UNIVERSIDAD NACIONAL DE SAN CRISTÓBAL DE HUAMANGA

FACULTAD DE INGENIERÍA DE MINAS, GEOLOGÍA Y CIVIL

ESCUELA PROFESIONAL DE INGENIERÍA CIVIL

TESIS:

Clasificación del peligro de accidentes de tránsito mediante redes neuronales artificiales y optimización del diseño geométrico según normas internacionales y DG-2018 en la vía Los Libertadores (PE-28A), Ayacucho

Para optar el título profesional de:
INGENIERO CIVIL

PRESENTADO POR:
Bach. Socrates ALLCCAHUAMAN ANDIA

ASESOR:
Dr. Ing. Hemerson LIZARBE ALARCÓN

AYACUCHO - PERÚ

2026

Resumen

La presente investigación desarrolló dos modelos de redes neuronales artificiales (RNA) para clasificar el nivel de riesgo de accidentes de tránsito en la Vía Los Libertadores (PE-28A), Ayacucho, combinados con un análisis normativo comparativo del diseño geométrico vial y un programa de alerta basado en geolocalización GPS en tiempo real.

El Modelo RNA N°01 (clasificación binaria: Riesgo Menor / Riesgo Mayor), entrenado con 202 registros históricos de accidentes y 23 variables del diseño geométrico, condiciones climáticas y factores viales conforme a la DG-2018, alcanzó un *accuracy* de 90,48 % y AUC-ROC de 0,9818 en el conjunto de evaluación. El Modelo RNA N°02 (clasificación multiclase: Riesgo Bajo / Medio / Alto), entrenado con 2,997 segmentos geométricos modelados en AutoCAD Civil 3D y expandidos a 5,394 registros mediante SMOTE, alcanzó un *accuracy* de 98,00 % y AUC-ROC de 0,9985, demostrando capacidad predictiva sobresaliente.

El análisis comparativo de 12 parámetros geométricos entre la DG-2018 y las normativas de Noruega, Islandia, Suecia, Reino Unido y Japón identificó deficiencias técnicas en la norma peruana, dando lugar a siete propuestas de mejora (P1–P7). Asimismo, se implementó un programa de alerta de accidentes como archivo HTML autocontenido, con rastreo GPS en tiempo real, cálculo de proximidad por la fórmula de Haversine y sistema de alerta multimodal (sonoro, vocal y vibración).

Los resultados demuestran que la integración de redes neuronales artificiales, análisis normativo y geolocalización constituye un enfoque efectivo para la predicción y reducción de accidentes de tránsito en carreteras nacionales del Perú.

Palabras clave: redes neuronales artificiales, predicción de accidentes de tránsito, diseño geométrico vial, DG-2018, normativas internacionales, geolocalización GPS, seguridad vial, PE-28A.

Abstract

This research developed two artificial neural network (ANN) models to predict traffic accidents on the Vía Los Libertadores (PE-28A), Ayacucho, Peru, combined with a comparative normative analysis of road geometric design and a real-time GPS-based accident alert program.

ANN Model No. 1 (binary classification: Minor Risk / Major Risk), trained on 202 historical accident records and 23 variables covering geometric design, climatic conditions, and road factors per the DG-2018 standard, achieved a test accuracy of 90.48 % and AUC-ROC of 0.9818. ANN Model No. 2 (multi-class classification: Low / Medium / High Risk), trained on 2,997 geometric road segments modeled in AutoCAD Civil 3D and expanded to 5,394 records using SMOTE, achieved a test accuracy of 98.00 % and macro AUC-ROC of 0.9985, demonstrating outstanding predictive performance.

A comparative analysis of 12 geometric parameters between DG-2018 and the standards of Norway, Iceland, Sweden, United Kingdom, and Japan identified technical deficiencies in the Peruvian standard, leading to seven improvement proposals (P1–P7). Additionally, a traffic accident alert program was implemented as a self-contained HTML file with real-time GPS tracking, Haversine-based proximity calculation, and a multimodal alert system (audio, voice, and vibration).

The results demonstrate that artificial neural networks, normative analysis, and geolocation technology is an effective approach for predicting and reducing traffic accidents on high-risk roads.

Keywords: artificial neural networks, traffic accident prediction, road geometric design, DG-2018, international standards, GPS geolocation, road safety, PE-28A.

Introducción

Los accidentes de tránsito constituyen la segunda causa de muerte a nivel mundial y uno de los principales problemas de salud pública en el Perú. Según el Ministerio de Transportes y Comunicaciones, en 2024 se registraron 86,757 siniestros viales a nivel nacional, con 3,002 personas fallecidas y 56,747 lesionadas (**MTC2024**). En la región Ayacucho, la situación es particularmente crítica: el 36,7 % de las muertes en 2023 se produjeron en carreteras, y a finales de 2024 se contabilizaron más de 40 fallecidos por accidentes de tránsito, incluyendo el vuelco del bus Expreso Molina Unión que dejó 25 muertos en el sector de Rumichaca de la Vía Los Libertadores (PE-28A) (**ONSV2024**).

La PE-28A, con 333,903 km de extensión entre los departamentos de Ica, Huancavelica y Ayacucho, es uno de los corredores viales con mayor índice de accidentabilidad del país. Sus características geométricas curvas cerradas, pendientes pronunciadas, zonas sin visibilidad y condiciones climáticas adversas la convierten en un corredor de alto riesgo que requiere intervenciones técnicas urgentes. Pese a los avances de la norma DG-2018, esta presenta deficiencias técnicas respecto a los estándares internacionales de países con los índices de mortalidad vial más bajos del mundo.

Frente a esta problemática, la presente investigación aplica un enfoque integrador que combina: (i) el desarrollo de dos modelos de redes neuronales artificiales para la predicción de accidentes de tránsito RNA N°01 (clasificación binaria, 202 registros, 23 variables) y RNA N°02 (clasificación multiclase, 2,997 segmentos geométricos modelados en Civil 3D); (ii) el análisis comparativo de la DG-2018 con las normativas de Noruega, Islandia, Suecia, Reino Unido y Japón; y (iii) la implementación de un programa de alerta de accidentes mediante el GPS en tiempo real con los 202 registros históricos georreferenciados de la PE-28A.

El trabajo se organiza en cinco capítulos: (I) planteamiento del problema , (II) marco teórico, (III) metodología y desarrollo, (IV) resultados y contrastación de hipótesis, y (V)conclusiones y recomendaciones. Se espera que los resultados contribuyan a la reducción de la siniestralidad en la PE-28A.

.

Dedicatoria:

Dedico la presente investigación a mi madre querida, Natalia Andía Quispe, por su apoyo incondicional en mis decisiones y trabajo duro para salir adelante. También, agradezco los sabios consejos de mis hermanos, que siempre me otorgan sus energías positivas para seguir adelante.

Agradecimientos

Al Dr. Ing. Hemerson Lizarbe Alarcón, asesor de la presente tesis, por sus sugerencias, recomendaciones, apreciaciones y por brindarme la información necesaria para la formulación del presente trabajo de investigación.

A los docentes de la Escuela de Formación Profesional de Ingeniería Civil de la Universidad Nacional de San Cristóbal de Huamanga por su contribución durante mi desarrollo académico y profesional.

UNIVERSIDAD NACIONAL DE SAN CRISTÓBAL DE HUAMANGA

Ayacucho, abril del 2026

Allccahuaman Andía, Sócrates

Índice General

Resumen ii
Abstract iii
Introducción iv
Dedicatoria v
Agradecimientos vi
Índice General vii
Índice de Figuras xi
Índice de Tablas xix
Glosarios xxii
Lista de Acrónimos xxiv
Símbolos xxv

Capítulo I
Planteamiento del Problema 2
1.1. Descripción del problema 2
1.1.1. Contexto Internacional 2
1.1.2. Contexto Nacional 3
1.1.3. Contexto Regional 6
1.2. Delimitación del problema 9
1.2.1. Espacial (geográfica) 9
1.2.2. Temporal 12
1.2.3. Temática y unidad de análisis 12
1.3. Formulación del problema 13
1.3.1. Problema general 13
1.3.2. Problemas específicos 14
1.4. Justificación e importancia 14
1.5. Limitaciones de la investigación 15
1.6. Objetivos 16
1.6.1. Objetivo general 16
1.6.2. Objetivos específicos 16

Capítulo II

Marco Teórico. 17
2.1. Antecedentes 17
2.1.1. Investigaciones internacionales 17
2.1.2. Investigaciones nacionales. 21
2.2. Bases teóricas 25
2.2.1. Redes Neuronales Artificiales (RNA) 25
2.2.2. Tipos de Redes Neuronales 29
2.2.3. Redes Neuronales Multicapa para Clasificación 32
2.2.4. Diseño geométrico de carreteras 40
2.2.5. Seguridad vial y accidentes de tránsito 42
2.2.6. Fuentes de datos para la investigación. 43
2.2.7. Estudio de accidentabilidad 50
2.2.8. Normas técnicas internacionales de diseño geométrico de vías 50
2.2.9. Herramientas computacionales aplicadas 58
2.3. Marco Conceptual 60
2.3.1. Fundamentos de las Redes Neuronales Artificiales y aprendizaje automático.60
2.3.2. Fundamentos del diseño geométrico de infraestructura vial 65

Capítulo III
Método de la Investigación 75
3.1. Método. 75
3.2. Orientación 75
3.3. Enfoque 75
3.4. Alcance 76
3.5. Diseño de investigación. 76
3.5.1. Diseño no experimental 76
3.5.2. Diseño comparativo 77
3.5.3. Diseño transversal 77
3.6. Población y muestra 77
3.6.1. Población 77
3.6.2. Muestra 80
3.7. Hipótesis 80
3.7.1. Hipótesis general 80
3.7.2. Hipótesis específicas 81
3.8. Operacionalización de variables, definición conceptual y operacional 81
3.8.1. Variables dependientes. 81
3.8.2. Variables independientes 83

3.8.3. Resumen general de variables por unidad de análisis 90
3.9. Técnicas e instrumentos 90
3.9.1. Técnicas de recolección de datos 90
3.9.2. Instrumentos de recolección de datos 93
3.10. Técnicas estadísticas para el procesamiento de la información 96
3.10.1. Procesamiento de datos: Modelo RNA N°1 96
3.10.2. Procesamiento de datos: Modelo RNA N°2 100
3.10.3. Procesamiento de datos: Análisis Normativo Comparativo 102
3.10.4. Software y herramientas utilizadas 103
3.11. Desarollo del trabajo de tesis. 104
3.11.1. Desarrollo del trabajo de tesis de la Red Neuronal N°01 104
3.11.2. Desarrollo del trabajo de tesis de la Red Neuronal N°02 167
3.11.3. Análisis normativo comparativo 217
3.11.4. Programa de alerta de accidentes de tránsito 235

Capítulo IV
Resultados 248
4.1. Análisis y contrastación de hipótesis. 248
4.1.1. Hipótesis general 248
4.1.2. Hipotesis 01: Modelo RNA N°01 248
4.1.3. Hipotesis 02: Modelo RNA N°02 254
4.1.4. Hipótesis 03: Análisis normativo comparativo 260
4.1.5. Hipótesis 04: Programa de alerta de accidentes 261

Capítulo V
Conclusiones y recomendaciones 264
5.1. Conclusiones 264
5.1.1. Conclusión general. 264
5.1.2. Conclusiones específicas 264
5.2. Recomendaciones 266
5.3. Trabajos futuros. 266

Capítulo VI
Anexos 268
6.1. Fotografias de la Via Los Libertadores (in situ) 268
6.2. Matriz de consistencia 272
6.3. Base de datos de RNA N°01. 273
6.4. Base de datos de RNA N°02. 277
6.5. Desarrollo de código de RNA N°01. 280

6.6. Desarrollo de código de RNA N°02.. 287
Referencias bibliográficas 294

Índice de Figuras

Figura 1 Muertes en accidentes de tráfico por cada cien mil hab. 2023 a nivel mundial. . 2
Figura 2 Histórico de siniestros de tránsito, 2010 - 2024 en el Perú. 3
Figura 3 Personas fallecidas en siniestros de tránsito, 2010 – 2024 en el Perú. 4
Figura 4 Personas lesionadas en siniestros de tránsito, 2010 – 2024 en el Perú. 5
Figura 5 Número de siniestros de tránsito, personas fallecidas y lesionadas, por región, 2024. 5
Figura 6 Personas fallecidas por siniestros de tránsito, según tipo de red vial, Ayacucho, 2021-2023. 6
Figura 7 Accidente de tránsito de Molina Unión, Ayacucho - 2023. 7
Figura 8 Accidente de tránsito de Molina Unión y Civa, Ayacucho - 2024. 8
Figura 9 Personas fallecidas en siniestros de tránsito, Ayacucho, 2021-2023. 8
Figura 10 Plano de la Carretera Vía Los Libertadores (PE-28A). 9
Figura 11 Progresiva 0+000km - ICA (dist. San Clemente) 10
Figura 12 Progresiva 333+90km - Ayacucho (Grifo. Ayacucho). 10
Figura 13 Departamentos que recorre la Vía Los Libertadores (PE-28A) 11
Figura 14 Provincias que recorre la Vía Los Libertadores (PE-28A) 11
Figura 15 Distritos que recorre la Vía Los Libertadores (PE-28A) 11

Figura 16 Redes neuronales inspirados en el cerebro humano. 26
Figura 17 Arquitectura de un Red Neuronal. 27
Figura 18 Red Neuronal Perceptron Simple. 30
Figura 19 Red Neuronal Perceptron Múltiple. 30
Figura 20 Redes neuronales Convulocionales. 31
Figura 21 Redes neuronales Recurrentes. 31
Figura 22 Redes neuronales de Base Radial. 32
Figura 23 Redes neuronales Generativas Adversarias. 32
Figura 24 Manual de diseño geométrico del Perú (DG-2018). 40
Figura 25 Elementos de la curva circular 41
Figura 26 Análisis de datos de los accidentes de tránsito. 43
Figura 27 Registro de accidentes de tránsito a nivel nacional 44
Figura 28 Ubicación de los accidentes de tránsito a nivel nacional 44
Figura 29 Estaciones metereológicas del SENAMHI 45
Figura 30 Registro de las condiciones climáticas de Estaciones metereológicas. 47
Figura 31 Obtención de datos en campo de la Via PE-28A. 48
Figura 32 Estado actual de la carretera PE-28A. 49

Figura 33 Estado actual de la carretera PE-28A. 49
Figura 34 Muertes en carretera por cada 100.000 habitantes de Noruega.. 52
Figura 35 Muertes en carretera por cada 100.000 habitantes de Islandia. 53
Figura 36 Muertes en carretera por cada 100.000 habitantes de Suecia. 54
Figura 37 Muertes en carretera por cada 100.000 habitantes de Reino Unido. 56
Figura 38 Muertes en carretera por cada 100.000 habitantes de Japon.. 57
Figura 39 Rangos de la Velocidad de Diseño en función a la clasificación de la carretera por demanda y orografía. 68
Figura 40 Distancia de visibilidad de parada con pendiente (metros) 69
Figura 41 Mínima distancia de visibilidad de adelantamiento 69
Figura 42 Longitudes de tramos en tangente 69
Figura 43 Peralte en zona rural (Tipo 3 ó 4). 70
Figura 44 Elementos de la curva circular 70
Figura 45 Relación entre radios consecutivos (Carretera de 2da y 3era clase) 71
Figura 46 Longitud mínima de curva de transición 72
Figura 47 Valores de sobreancho en función a “L” del tipo de vehículo de diseño 72
Figura 48 Pendientes máximas (%) 73
Figura 49 Sección transversal típica para carretera 73
Figura 50 Anchos mínimos de calzada en tangente 74
Figura 51 Ancho de bermas 74

Figura 52 Plano de la Carretera Vía Los Libertadores (PE-28A). 78
Figura 53 Progresiva 0+000km - Ica (distr. San Clemente). 78
Figura 54 Puente Rumichaca.. 79
Figura 55 Progresiva 333+903km - Ayacucho 79
Figura 56 Diagrama de flujo de la RNA N°01 - Parte 1. 105
Figura 57 Diagrama de flujo de la RNA N°01 - Parte 2. 106
Figura 58 Registro de accidentes de tránsito a nivel nacional 107
Figura 59 Base de datos de la RNA N°01:202 datos de accidentes. 108
Figura 60 Registro de accidentes por distrito 108
Figura 61 Registro de accidentes por departamento y provincia.. 109
Figura 62 Distrib. de la variable de salida: frecuencia absoluta (A) y proporción (B) . . 110
Figura 63 Histogramas y KDE: Mes, Día, Hora de Siniestro. 110
Figura 64 Histogramas y KDE por clase de variables: Clima Despejado, Clima Nublado, Clima Lluvioso.. 111
Figura 65 Histogramas y KDE: Temperatura Max (°C), Precipitación (mm/dia) y Curva111
Figura 66 Histogramas y: Peligro en pendiente, Radio de curva (m) y Peligro en curva. 112
Figura 67 Histogramas y KDE por clase de variables: Longitud de Curva (L), Longitud de subtangente (T) y Angulo de deflexión (°). 112

Figura 68 Histogramas y KDE: Longitud de cuerda (LC), Distancia Externa (E) y Distancia de ordenada media 113
Figura 69 Histogramas y KDE: Peligro por visibilida, Tramo en 'S' y Tramo en 'O'. . . 113
Figura 70 Histogramas y KDE: Long. tangente y Peligro estado del pavimento. 114
Figura 71 Boxplots de variables: Mes de siniestro, Temperatura max y Precipitación. . 115
Figura 72 Boxplots de variables: Radio de curva, Longitud de curva y Long. de subtangente. 115
Figura 73 Boxplots de variables: Angulo de deflexión, Long. de cuerda y Dist. externa (E). 115
Figura 74 Boxplots de variables: Dist. de la ordenada media y Long. de tangente. 116
Figura 75 Porcentaje de outliers por variable. Umbral: $[Q1 - 1.5 \cdot IQR,\ Q3 + 1.5 \cdot IQR]$ 116
Figura 76 Heatmap de correlaciones de Pearson entre las 23 variables de entrada. 117
Figura 77 Barras de correlación individual con la variable objetivo 118
Figura 78 Distrib. porcentual de Riesgo menor/mayor de la var. Peligro en pendiente. 120
Figura 79 Distri. porcentual de Riesgo menor/mayor de la var. Peligro en curva. 120
Figura 80 Distrib. porcentual de Riesgo menor/mayor de la var. Peligro por visibilidad. 121
Figura 81 Distrib. porcentual de Riesgo menor/mayor de la var. Peligro estado pavimento. 121
Figura 82 Análisis estadístico Mann-Whitney U de las 23 variables. 122
Figura 83 Diagr. de dispersión Radio de curva vs Long. de curva vs Long. de subtangente. 124
Figura 84 Diagr.de dispersión Radio de curva vs Angulo de deflexión vs Long. de cuerda. 124
Figura 85 Diagr. de disper. Radio de curva vs Dist. Externa vs Dist. de la ord. media. . 124
Figura 86 División estratificada 70/20/10% de la base de datos. 126
Figura 87 Normalización Z-Score: Mes, Día y Hora de Siniestro. 127
Figura 88 Normaliz. Z-Score: Clima Despejado, Clima Nublado y Clima Lluvioso. . . . 128
Figura 89 Normaliz. Z-Score:Temperatura Máx., Precipitación y Curva/Tramo Recto. . 128
Figura 90 Normaliz. Z-Score: Peligro Pendiente, Radio de Curva y Peligro en Curva. . 129
Figura 91 Normalización Z-Score de las variables: Long. (L), Subtangente (T) y Ángulo Deflexión. 129
Figura 92 Normaliz. Z-Score: Long. Cuerda (LC), Distancia Externa (E) y Ordenada Media (M) 130
Figura 93 Normaliz. Z-Score: Peligro Visibilidad, Tramo en "S" y Tramo Tangente "O" 130
Figura 94 Normaliz. Z-Score: Longitud Tangente y Peligro Estado del Pavimento. . . . 131
Figura 95 Arquitectura del Perceptrón Multicapa de la RNA N°01. 132
Figura 96 Accuracy vs. Alpha L2. 134
Figura 97 GAP de overfitting por valor de alpha. 135

Figura 98 Accuracy de entrenamiento y validación cruzada por arquitectura. 136
Figura 99 GAP de Overfitting por arquitectura. 136
Figura 100 GAP de Overfitting vs Número de parametros de la arquitectura. 137
Figura 101 Curva de convergencia: Perdida vs Epocas. 138
Figura 102 Accuracy en entrenamiento, validación y evaluación. 138
Figura 103 Pérdida (loss) vs. Épocas 139
Figura 104 Accuracy vs. Épocas 140
Figura 105 F1-Score por Clase en las Tres Particiones. 142
Figura 106 Precisión, Recall y F1-Score por Clase. 143
Figura 107 Resumen Global del F1-Score 144
Figura 108 Matriz de Confusión del Modelo. 146
Figura 109 Matriz de confusión normalizada (%). 147
Figura 110 Conteo por tipo de predicción. 148
Figura 111 Proporción de aciertos. 148
Figura 112 Curva ROC y área bajo la curva (AUC) del modelo RNA. 150
Figura 113 Curva de Aprendizaje: Diagnóstico de Sobreajuste. 152
Figura 114 Clase Real vs Frontera de Decisión. 154
Figura 115 Distribución de Predicciones y Confianza del Modelo. 154
Figura 116 Reducción media de Accuracy por variable (promedio ± desviación estándar, 20 repeticiones) 156
Figura 117 Ranking normalizado de importancia de variable 156
Figura 118 Precisión, Recall y F1-Score macro por partición. 158
Figura 119 Comparación de las 4 métricas por partición. 159
Figura 120 Diagrama de flujo de la creación del programa RNA N°01 160
Figura 121 Diagram de flujo de uso del programa RNA N°01 163
Figura 122 Paso 1: Abrir el programa 165
Figura 123 Paso 2: Ingresar variables temporales. 165
Figura 124 Paso 3: Ingresar variables geométricas. 166
Figura 125 Paso 4: Ingresar niveles de peligro. 166
Figura 126 Paso 5 y 6: Ejecutar e interpretar en caso Riesgo Menor. 166
Figura 127 Paso 5 y 6: Ejecutar e interpretar en caso Riesgo Mayor. 167
Figura 128 Diagrama de flujo del modelo RNA N°2 - parte 1. 168
Figura 129 Diagrama de flujo del modelo RNA N°2 - parte 2. 169
Figura 130 Modelamiento en el Civil 3D de la Via Los Libertadores. 170
Figura 131 Manual de diseño geométrico del Perú (DG-2018). 170
Figura 132 Modelamiento en el Civil 3D: Tramo llegada a Ayacucho. 171
Figura 133 Modelamiento en el Civil 3D: Tramo llegada a San Clemente. 171
Figura 134 Modelamiento en el Civil 3D: Tramo Pilpichaca. 171
Figura 135 Modelamiento en el Civil 3D: Tramo Huaytara. 172
Figura 136 Modelamiento en el Civil 3D: Datos geometricos exportados. 172

Figura 137 Base de datos RNA N° 2: 2997 registros. 172
Figura 138 Frecuencia y porcentaje por clase vs número de registros. 175
Figura 139 Proporción por clase de riesgo 175
Figura 140 Hitograma y KDE: Curva/Tramo Recto, Peligro en Pendiente, Radio de Curva. 176
Figura 141 Hitograma y KDE: Peligro en Curva, Long. de Curva y Long. Subtangente. 176
Figura 142 Hitograma y KDE: Ángulo Deflexión, Long. Cuerda y Dist. Externa.. 177
Figura 143 Hitograma y KDE: Ordenada Media, Peligro por Visib. y Tramo en S. 177
Figura 144 Hitograma y KDE: Tramo Tangente O, Long. Tangente, Peligro Pavimento. 178
Figura 145 Boxplots de variables de Radio de curva y longitud de curva. 179
Figura 146 Boxplots de variables de Longitud de subtangente y angulo de deflexión. . . . 179
Figura 147 Boxplots de variables de Longitud de cuerda y distancia externa. 180
Figura 148 Boxplots de variables de Distancia de ordenada media y longitud tangente. . 180
Figura 149 Análisis de outliers (Método IQR). 181
Figura 150 Matriz de correlación de Pearson. 182
Figura 151 Correlación individual de cada variable. 183
Figura 152 Variable de peligro en pendiente vs clase de accidente.. 184
Figura 153 Variable de peligro en curva vs clase de accidente. 184
Figura 154 Variable de peligro por visibilidad vs clase de accidente. 185
Figura 155 Variable de peligro por pavimento vs clase de accidente. 185
Figura 156 Diagr. de dispersión de Radio de curva vs long. de curva y long. de subtang. 186
Figura 157 Diagr. de dispersión de Radio de curva vs ángulo de deflexión y long. de cuerda. 186
Figura 158 Diagr. de dispersión de Radio de curva vs dist. externa y orden. media. 186
Figura 159 Significancia de variables - Test de Mann Whitney U. 187
Figura 160 División estratificada de la base de datos 70/20/10%. 188
Figura 161 Normalización Z-Score y KDE por clase de variables del 1-3 189
Figura 162 Normalización Z-Score y KDE por clase de variables del 4-6 190
Figura 163 Normalización Z-Score y KDE por clase de variables del 7-9 190
Figura 164 Normalización Z-Score y KDE por clase de variables del 10-12 191
Figura 165 Normalización Z-Score y KDE por clase de variables 13-15 191
Figura 166 Arquitectura de la red neuronal N°02: 15→128→64→32→16→3. Optimizador Adam · L2 α=0.05 Batch=32 Epocas=66. 192
Figura 167 Accuracy vs Alpha L2.. 193
Figura 168 GAP vs Alfa L2.. 194
Figura 169 Accuracy vs por arquitectura RNA. 195
Figura 170 GAP vs por arquitectura de la RNA. 195
Figura 171 GAP vs número de parametros de la RNA. 196
Figura 172 Curva de convergencia. 197
Figura 173 Accuracy por partición. 197

Figura 174 Pérdida vs épocas. 198
Figura 175 Accuracy vs épocas. 198
Figura 176 F1 Score por clase. 199
Figura 177 Precision, Recall y F1 por clase. 200
Figura 178 F1-Score global. 200
Figura 179 Matriz absoluta. 201
Figura 180 Matriz normalizada. 202
Figura 181 VP/VN/FP/FN por clase. 202
Figura 182 Curvas ROC y área bajo la curva (AUC). 203
Figura 183 Curva de Aprendizaje: Diagnóstico de Bias Varianza del Modelo RNA 204
Figura 184 Importancia de variables con intervalo de error. 205
Figura 185 Ranking normalizado de importancia de variables 206
Figura 186 Clase real vs frontera de decisión del RNA. 207
Figura 187 Distribución de predicciones del RNA. 208
Figura 188 Precisión, Recall y F1-Score macro por partición — RNA N°02. 209
Figura 189 Comparación de las 4 métricas por partición — RNA N°02. 209
Figura 190 Diagram de flujo de uso del programa RNA N°02 213
Figura 191 Paso 1: Abrir el programa 215
Figura 192 Paso 2, 3, 4: Ingresar variables geométricas. 215
Figura 193 Paso 5: Ingresar variables de condición de peligro. 216
Figura 194 Ejecutar e interpretar en caso Riesgo bajo. 216
Figura 195 Ejecutar e interpretar en caso Riesgo Medio. 216
Figura 196 Ejecutar e interpretar en caso Riesgo Alto. 217
Figura 197 Diagrama de flujo - Parte I. 236
Figura 198 Diagrama de flujo - Parte II. 237
Figura 199 Paso 1: Apertura del programa. 243
Figura 200 Paso 2: Exploración del mapa. 244
Figura 201 Paso 3: Configuración de parámetros (panel lateral). 244
Figura 202 Paso 3: Configuración de parámetros. 245
Figura 203 Paso 4: Activación del rastreo GPS. 245
Figura 204 Paso 5: Durante el recorrido. 246
Figura 205 Paso 6: Detención del rastreo y ubicación de accidentes de transito. 246
Figura 206 Datos de accidente de tránsito. 247
Figura 207 Programa ejecutado en el celular. 247

Figura 208 Arquitectura del Perceptrón Multicapa de la RNA N°01. 249
Figura 209 Comparación de las 4 métricas por partición. 250
Figura 210 Clase Real vs Frontera de Decisión. 251
Figura 211 Curva ROC y área bajo la curva (AUC) del modelo RNA. 251
Figura 212 Pérdida (loss) vs. Épocas 252
Figura 213 Accuracy vs. Épocas 252

Figura 214 Riesgo de accidentes por departamento y provincia. 253
Figura 215 Riesgo de accidentes por distrito 253
Figura 216 Paso 1: Abrir el programa 254
Figura 217 Arquitectura de la red neuronal N°02: 15→128→64→32→16→3. Optimizador Adam · L2 α=0.005 Batch=32 Epocas=66. 255
Figura 218 Comparación de las 4 métricas por partición — RNA N°02. 256
Figura 219 Clase real vs frontera de decisión del RNA. 257
Figura 220 Curvas ROC y área bajo la curva (AUC). 257
Figura 221 Pérdida vs épocas. 258
Figura 222 Accuracy vs épocas. 258
Figura 223 Riesgo de accidentes por departamento y provincia. 259
Figura 224 Riesgo de accidentes por distrito. 259
Figura 225 Paso 1: Abrir el programa 260
Figura 226 Paso 1: Apertura del programa. 262
Figura 227 Paso 2: Exploración del mapa. 263
Figura 228 Programa ejecutado en el celular - GPS. 263

Figura 229 Inicio de la carretera Via Los Libertadores - San clemente. 268
Figura 230 Final de la carretera Via Los Libertadores - Ayacucho. 268
Figura 231 Estado de la Via Los Libertadores - Distr. Humay. 269
Figura 232 Estado de la Via Los Libertadores - Distr. Huancano. 269
Figura 233 Estado de la Via Los Libertadores - Distr. Pilpichaca. 270
Figura 234 Estado de la Via Los Libertadores - Distr. Huaytara. 270
Figura 235 Estado de la Via Los Libertadores - Distr. Socos. 271
Figura 236 Estado de la Via Los Libertadores - Distr. Vinchos. 271
Figura 237 Importacion de librerias. 280
Figura 238 Importacion de base de datos. 280
Figura 239 Análisis estadistico. 280
Figura 240 Test de mann whittney y correlaciones. 281
Figura 241 Histogramas 281
Figura 242 Análsis de outilers por variable. 281
Figura 243 Correlación de Pearson. 282
Figura 244 Division de base de datos. 282
Figura 245 Normalización de datos. 282
Figura 246 Graficas de normalización. 283
Figura 247 Arquitectura de la RNA N°01. 283
Figura 248 Grafica de arquitectura de la RNA. 283
Figura 249 Graficas de perdida y Accuracy. 284
Figura 250 Matriz de confusión. 284
Figura 251 Curva Roc- AUC. 284

Figura 252 Curva de aprendizaje. 285
Figura 253 Importancia de variables. 285
Figura 254 Frontera de decisión. 285
Figura 255 Resumen de metricas de evaluación de la RNA. 286
Figura 256 Generación de archivo html. 286
Figura 257 Importación de librerias. 287
Figura 258 Importación de base de datos. 287
Figura 259 Estadistica de datos. 287
Figura 260 Estadistica de datos de Pearson. 288
Figura 261 Test estadistico de variables de Mann Whitney. 288
Figura 262 Grafica de histogramas. 288
Figura 263 Análisis de Outilers. 289
Figura 264 Matriz de correlación de pearson. 289
Figura 265 División de base de datos. 289
Figura 266 División de base de datos. 290
Figura 267 Arquitectura de la RNA N°02. 290
Figura 268 Grafica de arquitectura de la RNA. 290
Figura 269 Grafica de perdida vs epocas. 291
Figura 270 Gráfica de Accuracy vs epocas. 291
Figura 271 Gráfica de F1 - Score. 291
Figura 272 Matriz de confusión. 292
Figura 273 Curva ROC - AUC. 292
Figura 274 Importancia de variables. 292
Figura 275 Gráfica de frontera de decisión. 293
Figura 276 Resumen de metricas de la RNA. 293
Figura 277 Generación del programa html. 293

Índice de Tablas

Tabla 1 *Estaciones meteorológicas del SENAMHI de referencia para la Vía Los Libertadores (PE-28A).* 46

Tabla 2 *Operacionalización de las variables dependientes* 82
Tabla 3 *Variables independientes — Factor temporal* 83
Tabla 4 *Variables independientes — Factor climático y ambiental* 84
Tabla 5 *Variables independientes — Factor geométrico: alineamiento horizontal* 85
Tabla 6 *Variables independientes — Factor geométrico: alineamiento tangente* 86
Tabla 7 *Variables independientes — Factor de peligro: variables de alerta* 87
Tabla 8 *Variables independientes — Parámetros normativos comparados* 88
Tabla 9 *Resumen general de variables por unidad de análisis* 90
Tabla 10 *Técnicas de recolección de datos empleadas en la investigación* 92
Tabla 11 *Instrumentos de recolección de datos empleados en la investigación* 95
Tabla 12 *Esquema general del procesamiento de datos por unidad de análisis* 96
Tabla 13 *Criterios del score compuesto ponderado para el ranking de variables — RNA N°1* 97
Tabla 14 *Distribución del dataset por partición — Modelo RNA N°1* 98
Tabla 15 *Arquitecturas evaluadas y selección final del Modelo RNA N°1* 98
Tabla 16 *Hiperparámetros finales del Modelo RNA N°1* 99
Tabla 17 *Métricas de evaluación empleadas: Modelo RNA N°1* 100
Tabla 18 *Aplicación de SMOTE para el balanceo de clases — Modelo RNA N°2* 101
Tabla 19 *Distribución del dataset por partición del Modelo RNA N°2* 101
Tabla 20 *Arquitecturas evaluadas y selección final del Modelo RNA N°2* 101
Tabla 21 *Hiperparámetros finales del Modelo RNA N°2* 102
Tabla 22 *Métricas de evaluación empleadas — Modelo RNA N°2* 102
Tabla 23 *Criterios de clasificación del nivel de adecuación normativa de la DG-2018* . 103
Tabla 24 *Software y herramientas de procesamiento estadístico empleadas en la investigación* 103
Tabla 25 *Ficha técnica del Modelo RNA N°1* 104
Tabla 26 *Pares de variables con mayor correlación — Heatmap de Pearson* 118
Tabla 27 *Top-10 correlaciones de Pearson con la variable objetivo accidente de tránsito* 119
Tabla 28 *Distrib. porcentual de clases por nivel de variables de peligro* 122
Tabla 29 *Resumen del preprocesamiento de datos* 125
Tabla 30 *Resumen del diseño y entrenamiento del modelo RNA* 132
Tabla 31 *Descripción de cada capa de la arquitectura RNA* 133

Tabla 32 *Comparación de arquitecturas: rendimiento y nivel de sobreajuste* 137
Tabla 33 *Métricas de rendimiento del Modelo RNA N°1 en todas las particiones* 141
Tabla 34 *F1-Score por clase de riesgo según partición de datos* 142
Tabla 35 *Precisión, Recall y F1-Score por clase (conjunto de evaluación) (10%, n = 21)* 143
Tabla 36 *Resumen global del F1-Score macro y weighted por partición* 145
Tabla 37 *Síntesis del desempeño del clasificador binario RNA-MLP — métricas completas* 145
Tabla 38 *Métricas de clasificación por clase de la matriz de confusión: evaluación (n = 21)* 149
Tabla 39 *Síntesis del análisis de la matriz de confusión del modelo RNA-MLP* 149
Tabla 40 *Resultados de la Curva ROC por clase – conjunto de evaluación (n = 21)*. . . 151
Tabla 41 *Resultados de la curva de aprendizaje por tamaño de entrenamiento* 152
Tabla 42 *Resumen de la proyección PCA 2D* 155
Tabla 43 *Importancia por permutación de las 23 variables — conjunto de evaluación (n = 21)* 157
Tabla 44 *Precisión, Recall y F1-Score macro por partición del modelo RNA-MLP* 158
Tabla 45 *Comparación de las 4 métricas por partición del modelo RNA-MLP* 159
Tabla 46 *Parámetros extraídos del modelo entrenado e integrados en el programa RNA 1.html*. 161
Tabla 47 *Estructura de los tres bloques del archivo HTML generado* 162
Tabla 48 *Sub-pasos de la función `forwardPass()` y su equivalencia con Python* 162
Tabla 49 *Características técnicas del PROGRAMA RNA 1.html* 164
Tabla 50 *Procedimiento de uso del programa RNA 1.html para clasificación de tramos viales de la PE-28A* 164
Tabla 51 *Ficha técnica del modelo RNA N°2* 173
Tabla 52 *Resumen del análisis exploratorio de datos - Etapa I* 174
Tabla 53 *Resumen del preprocesamiento de datos - Etapa II* 188
Tabla 54 *Descripción de capas de la arquitectura RNA N°2*. 193
Tabla 55 *Métricas de rendimiento del modelo RNA N°2 en todas las particiones* 196
Tabla 56 *Distribución de registros por clase* 207
Tabla 57 *Aciertos y errores por clase* 208
Tabla 58 *Precisión, Recall y F1-Score macro por partición — RNA N°02* 209
Tabla 59 *Comparación de las 4 métricas por partición — RNA N°02* 210
Tabla 60 *Parámetros extraídos del modelo entrenado e integrados en el progrma RNA N°2* 210
Tabla 61 *Estructura de los tres bloques del archivo HTML generado — RNA N°02* . . . 211
Tabla 62 *Sub-pasos de la función `predecirRNA()` y su equivalencia con Python* 212
Tabla 63 *Características técnicas del programa RNA 2.html* 214
Tabla 64 *Procedimiento de uso del programa RNA 2.html* 214

Tabla 65 *Normativas internacionales de diseño geométrico analizadas* 218
Tabla 66 *Comparación de parámetros geométricos: Perú DG-2018 vs. Noruega N100* 219
Tabla 67 *Comparación de parámetros geométricos: Perú DG-2018 vs. Islandia Veghönnunarreglur* 221
Tabla 68 *Comparación de parámetros geométricos: Perú DG-2018 vs. Suecia VGU* . . 223
Tabla 69 *Comparación de parámetros geométricos: Perú DG-2018 vs. Reino Unido DMRB* 226
Tabla 70 *Comparación de parámetros geométricos: Perú DG-2018 vs. Japón RSO* . . . 229
Tabla 71 *Posición relativa del DG-2018 frente a las normativas internacionales analizadas* 231
Tabla 72 *Resumen de propuestas de mejora a la DG-2018 sustentadas en normativas internacionales* 234
Tabla 73 *Estructura de cada registro de accidente en la base de datos embebida* 238
Tabla 74 *Distribución anual de los 202 accidentes de tránsito — PE-28A (2020–2025)* 238
Tabla 75 *Tecnologías y librerías empleadas en el programa de alerta* 239
Tabla 76 *Parámetros de configuración del rastreo GPS* 240
Tabla 77 *Niveles de intensidad de la alerta sonora según distancia* 241
Tabla 78 *Parámetros configurables del programa de alerta* 242
Tabla 79 *Procedimiento de uso del programa de alerta de accidentes para recorridos en la carretera PE-28A* 242

Tabla 80 *Métricas de rendimiento del Modelo RNA N°01 en todas las particiones* 249
Tabla 81 *Métricas de rendimiento del Modelo RNA N°02 en todas las particiones* 255
Tabla 82 *Propuestas de mejora a la DG-2018 sustentadas en normativas internacionales* 261
Tabla 83 *Características técnicas verificadas del programa de alerta GPS* 262

Glosarios

Accidente de tránsito Evento súbito producido en la vía pública como consecuencia de la circulación vehicular, que provoca daños a personas o bienes materiales. La investigación utiliza 202 registros de accidentes de la PE-28A (2020–2025) del ONSV.

AUC-ROC Área bajo la Curva ROC. Mide la capacidad discriminativa del clasificador en todos los umbrales de decisión. Valores $\geq 0{,}90$ indican discriminación sobresaliente.

Berma Franja longitudinal paralela y adyacente a la calzada, que sirve de confinamiento de la capa de rodadura y como zona de seguridad para estacionamiento de emergencia.

Bombeo Inclinación transversal de la calzada en tramos en tangente, cuyo objetivo es evacuar las aguas superficiales de lluvia, expresada como porcentaje.

Carretera Camino para el tránsito de vehículos motorizados de por lo menos dos ejes, cuyas características geométricas deben cumplir las normas técnicas vigentes del Ministerio de Transportes y Comunicaciones.

Derecho de Vía Faja de terreno de ancho variable dentro del cual se encuentra comprendida la carretera, sus elementos constitutivos y las zonas de seguridad para el usuario.

Distancia de visibilidad Longitud continua hacia adelante de la carretera visible al conductor para ejecutar con seguridad las maniobras de parada o adelantamiento.

F1-Score Media armónica entre Precisión y Recall, $F1 = 2 \cdot (\text{Precisión} \times \text{Recall})/(\text{Precisión} + \text{Recall})$. El F1-Score macro promedia el resultado de cada clase sin ponderar por tamaño.

Función Softmax Función de activación de la capa de salida para clasificación multiclase, que convierte un vector de valores reales en una distribución de probabilidades tal que la suma es igual a 1.

Fórmula de Haversine Fórmula que calcula la distancia geodésica entre dos puntos sobre una esfera a partir de sus coordenadas geográficas, considerando la curvatura terrestre ($R = 6{,}371{,}000\,\text{m}$). Utilizada en el programa de alerta para medir la proximidad del conductor a los 202 registros de accidentes georreferenciados.

Georreferenciación Proceso de asignación de coordenadas geográficas (latitud y longitud, sistema WGS-84) a un registro o elemento, permitiendo su localización precisa sobre

mapas digitales.

Pendiente longitudinal Inclinación del eje de la carretera en el sentido del tránsito, expresada como porcentaje del desnivel entre dos puntos respecto a la distancia horizontal que los separa.

Peralte Inclinación transversal de la calzada en los tramos en curva, destinada a contrarrestar la fuerza centrífuga del vehículo, expresada como porcentaje.

Perceptrón Multicapa (MLP) Red neuronal *feedforward* con una o más capas ocultas entre la entrada y la salida. Es la arquitectura empleada en los Modelos RNA N°01 y N°02 de la presente investigación.

Radio de curvatura Radio de la curva circular que define la alineación horizontal de la carretera. El radio mínimo es el valor límite para circular a la velocidad de diseño con el peralte y coeficiente de fricción lateral máximos.

Red neuronal artificial (RNA) Modelo computacional inspirado en el sistema nervioso biológico, compuesto por capas de nodos interconectados que ajustan sus pesos durante el entrenamiento para aprender patrones a partir de datos.

Sección transversal Representación perpendicular al eje de la carretera que nomina y dimensiona los elementos que la conforman dentro del Derecho de Vía. Puede ser de tipo General o Especial.

Siniestralidad vial Indicador estadístico que expresa la frecuencia y gravedad de los accidentes de tránsito en una vía durante un período determinado.

SMOTE *Synthetic Minority Oversampling Technique*. Técnica de sobremuestreo que genera instancias sintéticas de las clases minoritarias por interpolación lineal, con el fin de balancear el dataset de entrenamiento.

Sobreajuste (*Overfitting*) Fenómeno en que un modelo aprende el ruido del conjunto de entrenamiento y pierde capacidad de generalización ante datos nuevos. Se detecta por un GAP elevado entre las métricas de entrenamiento y evaluación.

Velocidad de diseño Velocidad máxima seleccionada para el diseño de un tramo homogéneo de carretera, a la que se pueden circular con seguridad y comodidad bajo condiciones favorables.

Lista de Acrónimos

API	:	Interfaz de Programación de Aplicaciones (*Application Programming Interface*)
AUC	:	Área Bajo la Curva (*Area Under the Curve*)
DG-2018	:	Manual de Diseño Geométrico de Carreteras, Perú 2018
DMRB	:	Design Manual for Roads and Bridges, Reino Unido
DVP	:	Distancia de Visibilidad de Parada
GPS	:	Sistema de Posicionamiento Global (*Global Positioning System*)
IMD	:	Índice Medio Diario
MLP	:	Perceptrón Multicapa (*Multilayer Perceptron*)
MTC	:	Ministerio de Transportes y Comunicaciones
N100	:	Håndbok N100: Veg- og gateutforming, Noruega
ONSV	:	Observatorio Nacional de Seguridad Vial
PNP	:	Policía Nacional del Perú
PROVÍAS	:	Proyecto Especial de Infraestructura de Transporte Nacional
RNA	:	Red Neuronal Artificial
ROC	:	Característica Operativa del Receptor (*Receiver Operating Characteristic*)
RSO	:	Road Structure Ordinance — Dōro Kōzōrei, Japón
SENAMHI	:	Servicio Nacional de Meteorología e Hidrología del Perú
SMOTE	:	*Synthetic Minority Oversampling Technique*
VGU	:	Vägars och gators utformning, Suecia
WGS-84	:	*World Geodetic System 1984*

Símbolos

Acc	Exactitud global del modelo (*accuracy*) [%]
$\hat{y}$	Valor predicho por el modelo [adimensional]
α	Tasa de aprendizaje / parámetro de regularización L2 [adimensional]
AUC	Área bajo la curva ROC [adimensional]
b_j	Sesgo (*bias*) de la neurona j [adimensional]
b	Bombeo de la calzada [%]
d	Distancia geodésica entre dos puntos (Haversine) [m]
D_p	Distancia de visibilidad de parada [m]
$e_{\text{máx}}$	Peralte máximo [%]
F_1	F1-Score (media armónica de Precisión y Recall) [adimensional]
f	Coeficiente de fricción lateral [adimensional]
K	Número de clases del clasificador [adimensional]
φ	Latitud geográfica en radianes [rad]
$L_{\text{máx}}$	Longitud máxima de recta [m]
$L_{\text{mín}}$	Longitud mínima de recta [m]
λ	Longitud geográfica en radianes [rad]
$\mathscr{L}$	Función de pérdida (*loss function*) [adimensional]
N	Número total de registros del dataset [adimensional]
i	Pendiente longitudinal [%]
e	Peralte de la calzada [%]
w_{ij}	Peso sináptico entre la neurona i y la neurona j [adimensional]
R	Radio de curvatura horizontal [m]
R_T	Radio medio de la Tierra ($R_T = 6{,}371{,}000$ m) [m]
V_h	Velocidad de diseño horizontal [km/h]

CAPÍTULO I Planteamiento del Problema

1.1 Descripción del problema

1.1.1 Contexto Internacional

Los accidentes de tránsito constituyen una de las principales causas de mortalidad a nivel mundial, representando un desafío crítico para la salud pública y la ingeniería vial. En el *"Informe Mundial de Seguridad Vial 2024 del ITF"* **(International Transport Forum, 2024)**, se realiza un análisis comparativo de los resultados en función al diseño geométrico de cada uno de los países en cuanto a mortalidad en la vía. Para ello uno de los principales indicadores presentado es el número de fallecidos por accidente de tráfico por cada 100 000 habitantes. La figura siguiente permite observar claramente las diferencias entre países, destacándose Noruega, Islandia y Suecia como los más seguros, con tasas que rondan las 2 muertes por cada 100 000 habitantes. Los resultados son consistentes con el enfoque en políticas de seguridad vial rigurosas, con infraestructuras de calidad y con una cultura de cumplimiento normativa muy enraizada.

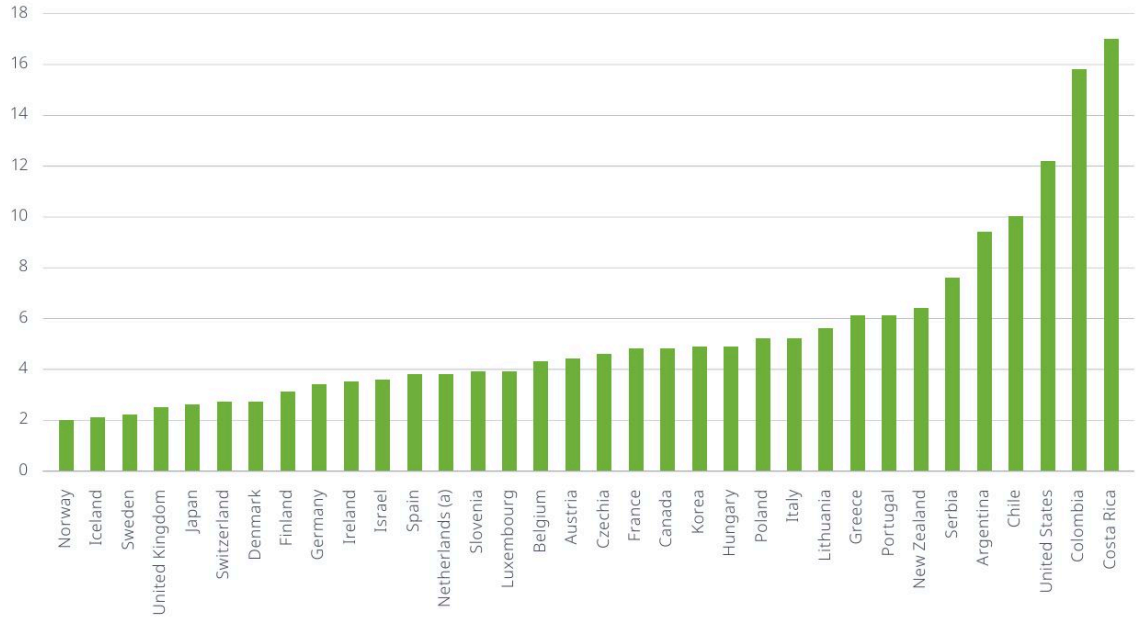


Figura 1 Muertes en accidentes de tráfico por cada cien mil hab. 2023 a nivel mundial.

Fuente: Road Safety Annual Report 2024 - International Transport Forum.

El Reino Unido, Japón, Suiza, Dinamarca y Finlandia son otros países de la Unión Europea que multiplican la situación claramente favorable los índices de por debajo de la cifra 4, siendo países de referencia en la gestión de la seguridad vial. Por el contrario, en el extremo

derecho de la gráfica se observan países con cifras muy superiores, como sería el caso de Colombia, de los Estados Unidos y en especial Costa Rica, que aporta más de 16 fallecimientos por cada 100,000 habitantes. La gráfica hace evidente las disparidades entre regiones relacionadas con la tasa de motorización, fizcalización, inversión en infraestructuras viales, y los niveles de calidad de la atención post-accidente.

1.1.2 Contexto Nacional

En el Perú, el diseño geométrico de las carreteras es factor determinante de seguridad vial, sobre todo en condiciones rurales y montañosas donde las condiciones topográficas críticas y carencia de estandarización técnica significan que el riesgo de siniestros se multiplica ampliamente. Aunque el Estado ha emprendido esfuerzos consistentes para mejorar el equipamiento vial, el país continúa teniendo graves limitaciones para aplicar criterios geométricos que sean adecuados para superar su diversidad territorial. Esa realidad se viene constatando en las estadísticas oficiales de siniestralidad vial.

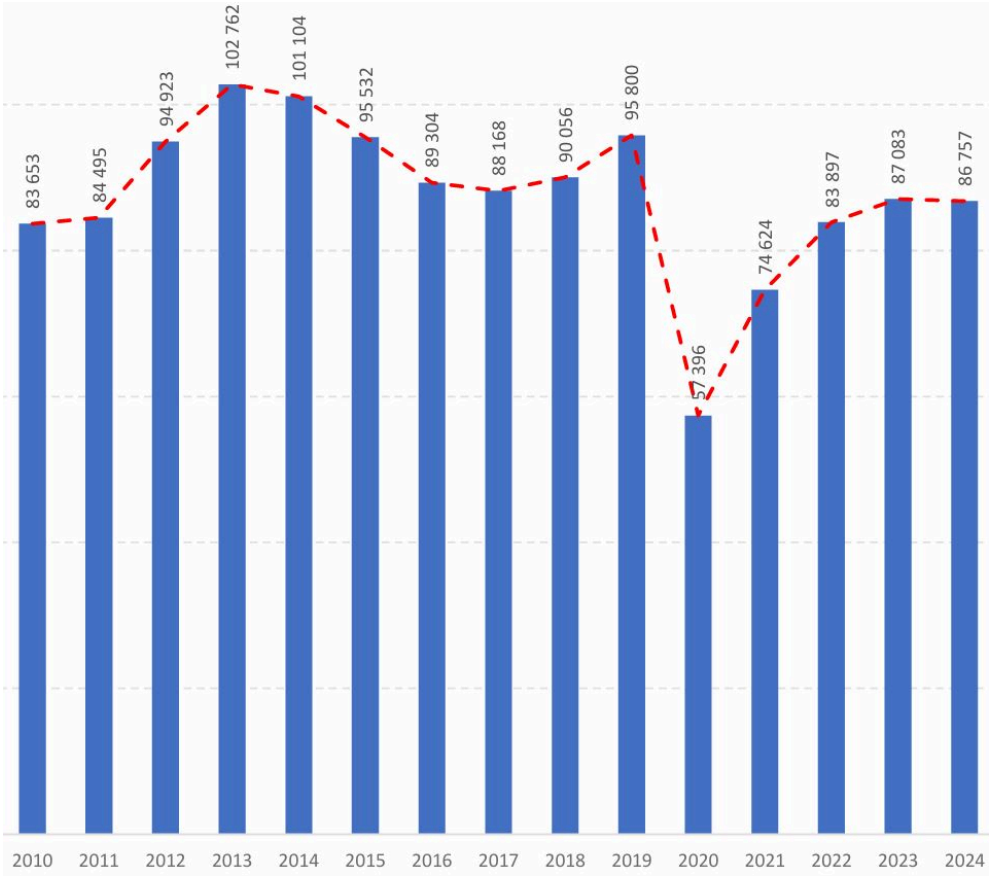


Figura 2 Histórico de siniestros de tránsito, 2010 - 2024 en el Perú.

Fuente: DIVEST/DIRTIC-PNP y Dirección de Seguridad Vial - MTC

De acuerdo con el **Boletín Anual de Siniestralidad Vial 2024**, editado por el Ministerio de Transportes y Comunicaciones (MTC), se contabilizaron 86,757 accidentes de tránsito de

forma nacional, resultando 3,002 muertos y 56,747 con lesiones **(MTC, 2024, p. 7)**. Dichos datos comprueban que, pese a haberse producido avances normativos y técnicos, el tema de la siniestralidad sigue estando crítico, principalmente en el diseño geométrico deficiente. Durante 2024, hubo un total de 3.002 muertes debido a accidentes de tráfico. Esta cifra refleja una disminución del 9.5 % en comparación con el año anterior 2023, cuando se reportaron 3,316 muertes. Aunque esta reducción es positiva, está claro que los incidentes de tráfico continúan siendo una grave preocupación que exige un cambio profundo y duradero, especialmente en áreas de alta concentración como Ayacucho **(MTC,2024, p. 10)**.

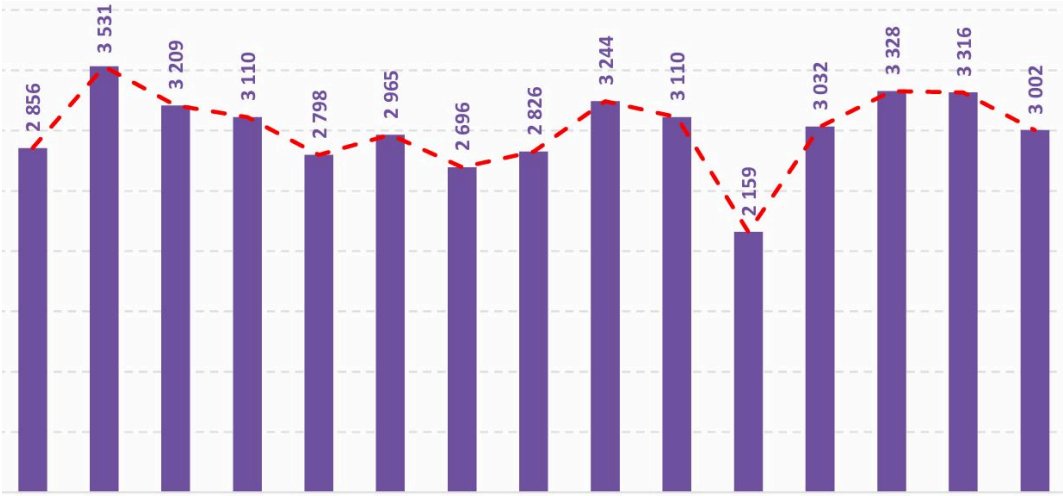


Figura 3 Personas fallecidas en siniestros de tránsito, 2010 – 2024 en el Perú.

Fuente: DIVEST/DIRTIC-PNP y Dirección de Seguridad Vial - MTC

En el periodo 2024, el número total de individus que sufrieron una condición de lesiones debido a siniestros de tránsito a nivel país fue de 56,747 lesionados. Esto significa que se experimentó una disminución del 2.2 % con respecto al año 2023, en el que se reporteran 58,000 lesionados, de acuerdo con las fuentes oficiales. Aunque es bajísimo, significa que se experimentó mejora moderada para incrementar la seguridad vial de país **(MTC,2024, p. 13).**

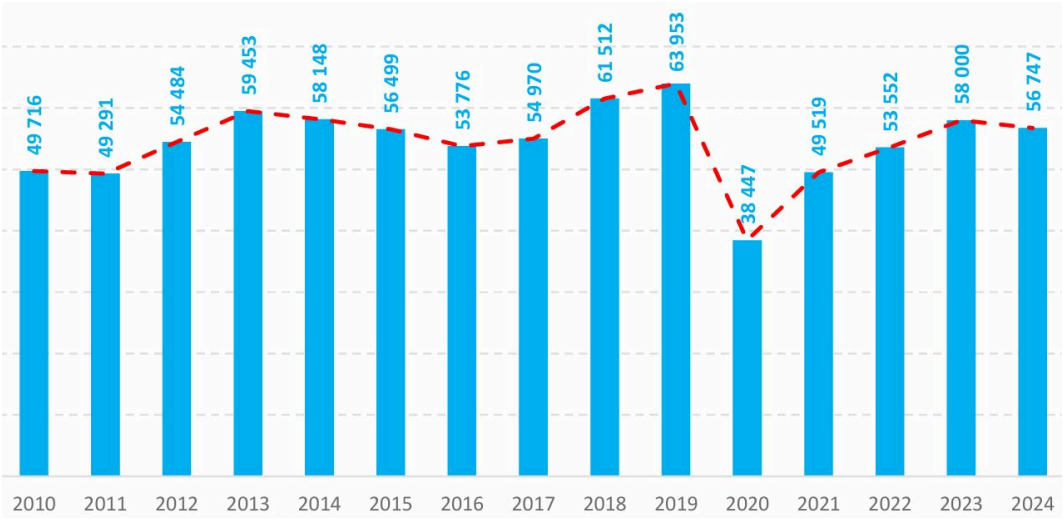


Figura 4 Personas lesionadas en siniestros de tránsito, 2010 – 2024 en el Perú.

Fuente: DIVEST/DIRTIC-PNP y Dirección de Seguridad Vial - MTC

Por lo tanto la siniestrabilidad de transito a nivel nacional, según la región durante el 2024, se muestra en la siguiente imagenes.

2024			
Región	Siniestros	Fallecidos	Lesionados
AMAZONAS	513	54	554
ANCASH	2 218	117	1 570
APURIMAC	703	130	510
AREQUIPA	5 604	187	3 051
AYACUCHO	549	77	475
CAJAMARCA	2 278	195	2 207
CALLAO	3 232	14	1 164
CUSCO	3 575	280	2 740
HUANCAVELICA	201	54	337
HUANUCO	1 161	76	1 124
ICA	2 281	93	1 491
JUNIN	3 124	218	3 259
LA LIBERTAD	4 091	120	3 593
LAMBAYEQUE	3 524	71	2 844
LIMA	43 786	570	23 261
LORETO	307	14	318
MADRE DE DIOS	404	60	329
MOQUEGUA	564	17	338
PASCO	381	57	489
PIURA	3 096	272	2 066
PUNO	1 287	138	997
SAN MARTIN	1 746	101	2 145
TACNA	1 160	49	674
TUMBES	264	15	261
UCAYALI	708	23	950
TOTAL	86 757	3 002	56 747

Figura 5 Número de siniestros de tránsito, personas fallecidas y lesionadas, por región, 2024.

Fuente: DIVEST/DIRTIC-PNP y Dirección de Seguridad Vial - MTC

1.1.3 Contexto Regional

La zona Ayacucho, donde se encuentra ubicada la Via Los Libertadores (PE-28A), contabilizó 549 accidentes de tránsito para el año 2024, con saldo de 77 muertos y 475 lesionados, números que aunque no son los más altos de todas las regiones, son alarmantes por ser muy altos porcentualmente, según el volumen de vehículos regionales **(MTC, 2024, p. 27)**. Dichos números indican el presence de factores estructurales más allá de la actitud de conductor para que estos accidentes se produzcan, resultando uno de ellos el diseño geométrico inadecuado.

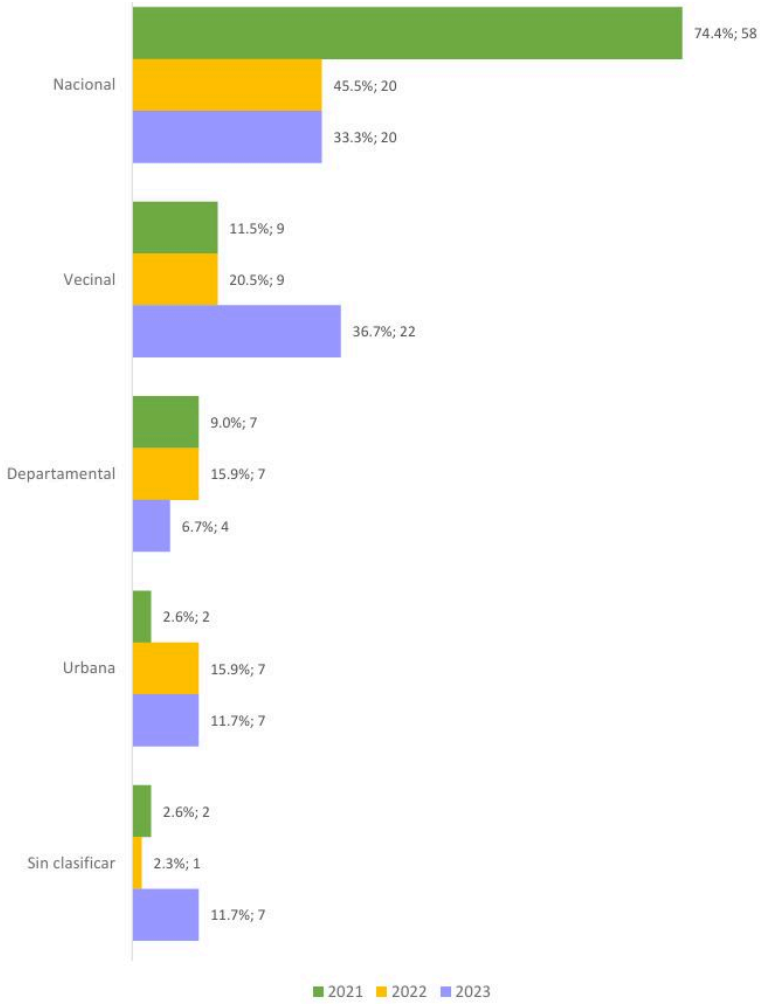


Figura 6 Personas fallecidas por siniestros de tránsito, según tipo de red vial, Ayacucho, 2021-2023.

Fuente: Observatorio Nacional de Seguridad Vial y Dirección de Seguridad Vial - MTC

De acuerdo al informe de la dirección de seguridad vial **"Informe de víctimas fatales en siniestros de tránsito e indentificación de puntos de alta siniestralidad"** manifiesta que en las carreteras de Ayacucho en gran mayoria presenta una serie de tramos con curvas cerradas, sin peralte, pendientes pronunciadas, anchos reducidos, y ausencia de visibilidad horizontal y vertical adecuada, factores que la convierten en un corredor de alto riesgo para el trans-

porte interprovincial y rural. En 2023, el 36.7 % de las muertes en Ayacucho se produjeron en carreteras vecinales, superando incluso a las muertes en vías nacionales y departamentales **(MTC, 2023, p. 12)**, y el 65 % de los fallecimientos fueron ocasionados por despistes, siniestros directamente relacionados con deficiencias geométricas **(MTC, 2023,p. 20)**.

Accidente de tránsito en Ayacucho en 2023 y 2024

El 18 de septiembre de 2023, el bus interprovincial de la compañía Molina, que salió de Ayacucho hacia Huancayo, se precipitó a un abismo de unos 200 metros frente al río Mantaro, con al menos 25 muertos. El autobús había sido parado por vegetación y así se salvó de una precipitación directa al río. El incidente puso de relieve carencias de infrastructura vial e inadecuado diseño geometrico de la carretera. Este hecho real da fuerza de convicción al llamado para renovar el proyecto geométrico viales para evitar que se produzcan accidentes semejantes.

Figura 7 Accidente de tránsito de Molina Unión, Ayacucho - 2023.

Fuente: Diario Perú21

A finales de 2024, Ayacucho contabilizó más de 40 muertos por accidente de tránsito, según el Observatorio Nacional de Seguridad Vial del MTC. Uno de los accidentes más duros se produjo el 16 de julio, al volcarse el bus de la compañía Expreso Molina Unión al fondo de un abismo, por el camino de Ccarhuaccpampa, sección Rumichaca (distrito de Vinchos, Huamanga), con 25 muertos. Otra se produjo el 14 de mayo, por el camino de Apacheta, también por la carretera de Libertadores, al volcar el bus de la compañía Civa, de Lima, que transportaba 34 pasajeros, con 17 fallecidos. Sus casos evidencian la necesidad de mejorar el diseño geométrico de carreteras y de la seguridad de las carreteras interprovinciales de la región.

Figura 8 Accidente de tránsito de Molina Unión y Civa, Ayacucho - 2024.

Fuente: Diario Jornada, Ayacucho.

Esto accidentes de tránsito es particularmente preocupante en Ayacucho, ya que el informe regional indica que el 35.9% de todas las muertes en la región en 2023 se debió a la falta de infraestructura adecuada y este porcentaje ha aumentado en comparación con el año anterior **(MTC, 2023, p. 25)**.

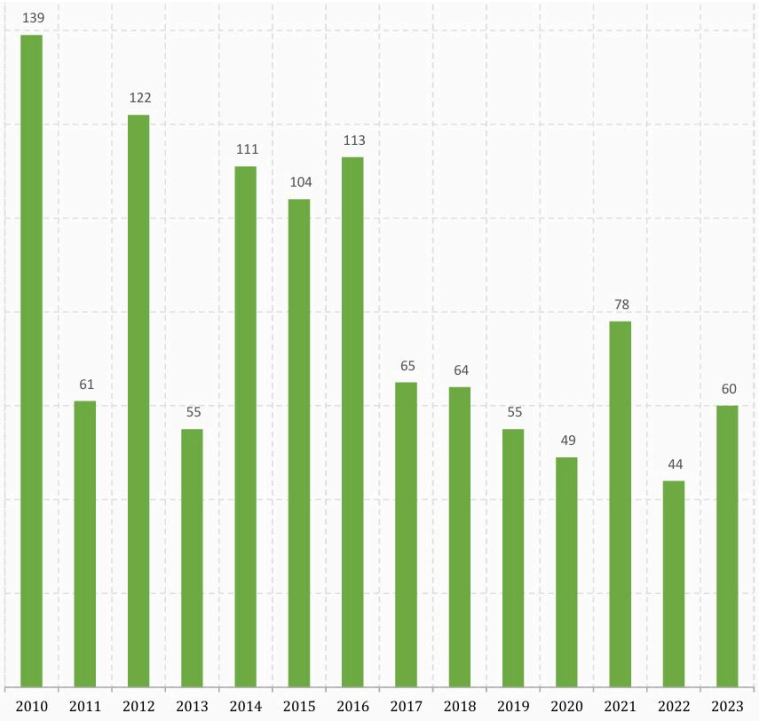


Figura 9 Personas fallecidas en siniestros de tránsito, Ayacucho, 2021-2023.

Fuente: Observatorio Nacional de Seguridad Vial y Dirección de Seguridad Vial – MTC.

Estas estadísticas confirman que el diseño geométrico no solo es un problema técnico, sino un factor de riesgo directo que aumenta la letalidad. Sin embargo, se deben interpretar junto a condiciones de entorno vial: un trazado geométrico incorrecto curvas cerradas sin curvas de peralte, curvas de radio mínimo no normativo o sin visibilidad, estado del vehiculo, condición

climatica, estado del pavimento, la visibilidad entre otras pueden inducir a una tragedia. En este sentido, se genera la necesidad de llevar acabo un análisis transversal de las normativas de diseño geométrico vial vigentes tanto en Perú como en otros países de interés de referencia, y para el tramo Vía Los Libertadores (PE-28A), con el propósito de determinar deficiencias geométricas importantes, sugerir parámetros mejorados mediante Redes Neuronales que ayudaran a predecir accidentes de tránsito y como consecuencia reducir los accidentes en la Via Los Libertadores (PE-28A).

1.2 Delimitación del problema

1.2.1 Espacial (geográfica)

La investigación se enmarca geográficamente en la **Vía Los Libertadores (PE-28A)**, esa ruta nacional que cumple un rol fundamental al conectar San Clemente, en el departamento de Ica, con la ciudad de Ayacucho, extendiéndose a lo largo de aproximadamente 333.90 km de recorrido. Es un corredor vial que atraviesa tres departamentos Ica, Huancavelica y Ayacucho y lo hace en condiciones extremas: Mal estado del pavimento, parametros geométricos muy desfasados, y en otros. los cuales a ocasionado múltiples accidentes de tránsito.

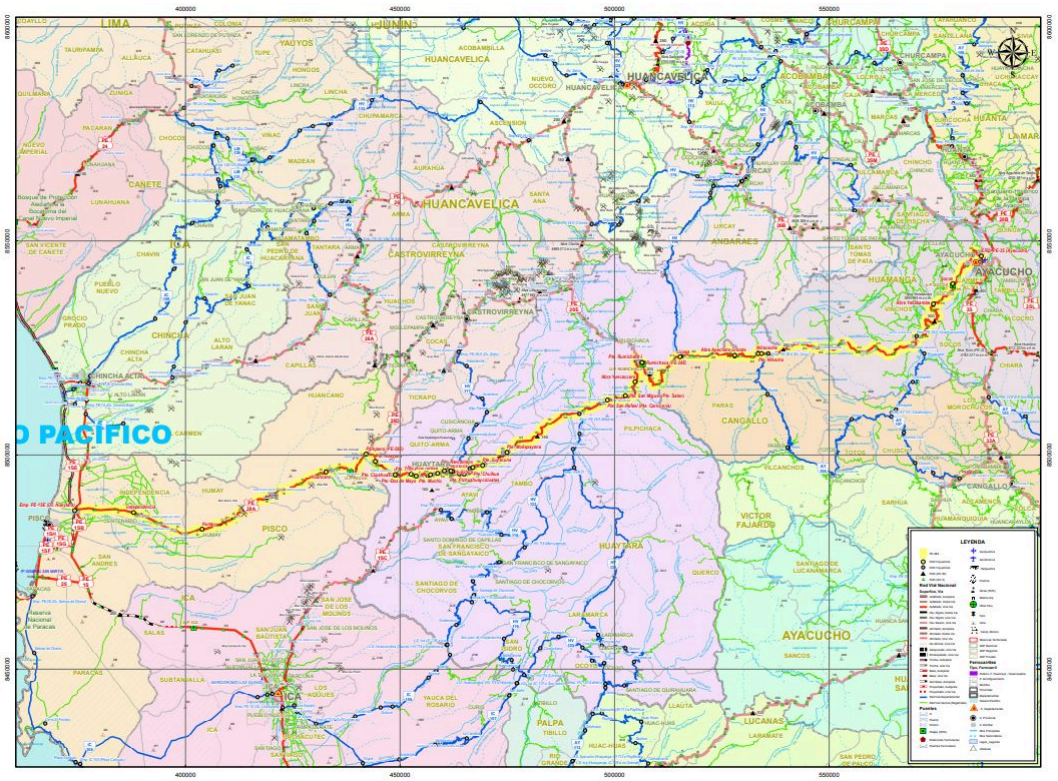


Figura 10 Plano de la Carretera Vía Los Libertadores (PE-28A)
Fuente: Ministerio de Transporte y Comunicaciones

Figura 11 Progresiva 0+000km - ICA (dist. San Clemente)
Fuente: Fotografía propia.

Figura 12 Progresiva 333+90km - Ayacucho (Grifo. Ayacucho)
Fuente:Fotografía propia.

Los departamento, provincia y distritos el cual comprende la **Vía Los Libertadores (PE-28A)** son las siguientes:

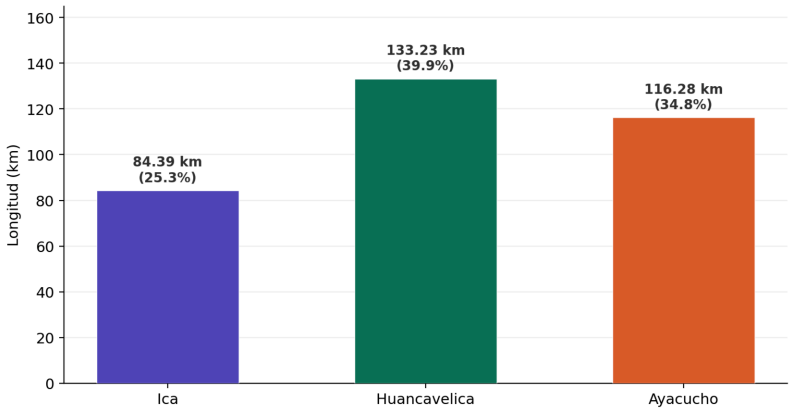


Figura 13 Departamentos que recorre la Vía Los Libertadores (PE-28A)
Fuente: Elaboración propia

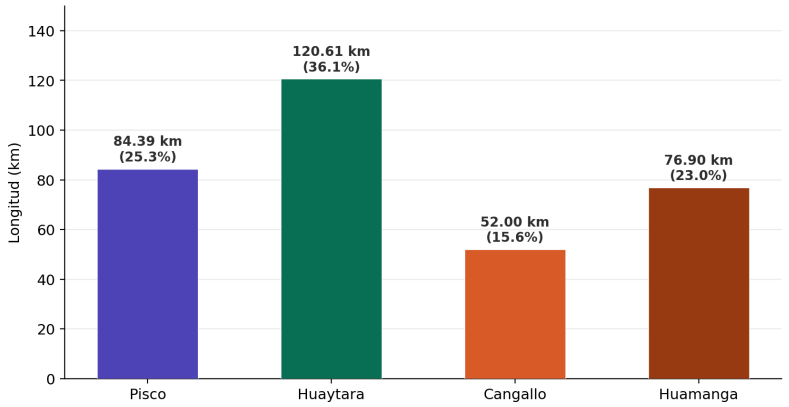


Figura 14 Provincias que recorre la Vía Los Libertadores (PE-28A)
Fuente: Elaboración propia

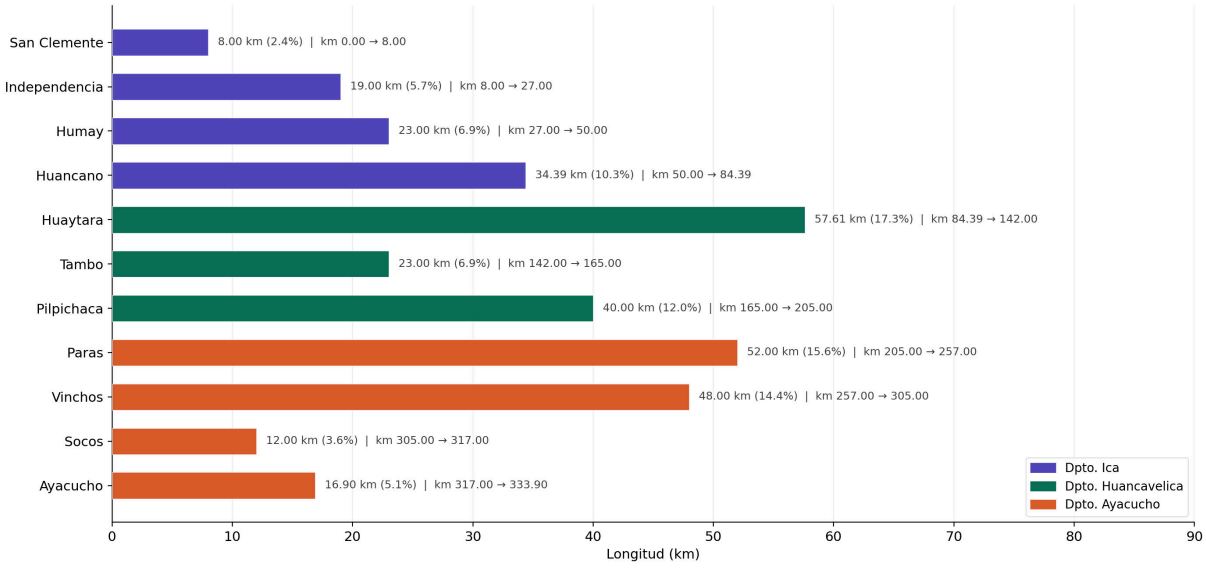


Figura 15 Distritos que recorre la Vía Los Libertadores (PE-28A)
Fuente: Elaboración propia

El foco del estudio está puesto, precisamente, en los puntos más críticos de esta vía, aquellos tramos donde la siniestralidad se concentra con mayor intensidad y donde las características del terreno, sumadas a factores climáticos y de diseño vial, configuran escenarios de alto riesgo. Identificar estos puntos no es un ejercicio meramente estadístico; es un paso necesario para entender qué está fallando y, sobre todo, dónde se deben priorizar las intervenciones que puedan salvar vidas.

1.2.2 Temporal

La investigación se enmarca en dos horizontes temporales claramente diferenciados: uno dedicado a la recopilación de los datos y otro al desarrollo del estudio en sí.

Hablemos primero de los datos históricos. El estudio recoge información desde el año 2020 hasta el 2025, lo que equivale a seis años de registros oficiales de siniestros de tránsito ocurridos a lo largo de la Vía Los Libertadores (PE-28A). Provienen de fuentes institucionales: el Ministerio de Transportes y Comunicaciones (MTC), la Policía Nacional del Perú (PNP) y el Observatorio Nacional de Seguridad Vial (ONSV), entidades que respaldan la solidez y confiabilidad de la información utilizada.

Se realizará dos modelos de Redes Neuronales Artificales:

- **El Modelo RNA N.°1 (Clasificación Binaria: Riesgo Menor/Mayor):** se construye a partir de un dataset de registros de siniestros de tránsito recopilados en la Vía Los Libertadores (PE-28A) durante el período 2020–2025.
- **El Modelo RNA N.°2 (Clasificación Multiclase: Riesgo Bajo/Medio/Alto):** se construye a partir de un dataset de registros, producto del procesamiento geométrico de la carretera en el software AutoCAD Civil 3D. Mediante el modelamiento del eje de la vía, el perfil longitudinal y las secciones transversales de la PE-28A

Por su parte, el desarrollo de la investigación se lleva a cabo durante el año académico 2025–2026, un período que comprende todas las etapas del proceso: desde la revisión bibliográfica y la recopilación y depuración de datos, pasando por el desarrollo y validación de los modelos RNA, el análisis comparativo de normativas internacionales y el modelamiento en Civil 3D y Vehicle Tracking.

1.2.3 Temática y unidad de análisis

A) Temática: la investigación abarca tres ejes fundamentales.

- El primer eje se centra en un análisis comparativo de normativas internacionales de diseño geométrico. Aquí se ponen frente a frente los parámetros de la DG-2018 del Perú (MTC, 2018) con las normativas de cinco países: Noruega, Islandia, Suecia, Reino Unido y Japón, precisamente los cinco con las tasas de mortalidad vial más bajas del mundo según el ITF (2024). La comparación abarca aspectos fundamentales como radios mínimos, peralte máximo, distancia de visibilidad de parada, pendientes longitudinales, anchos de carril y berma, curvas verticales y elementos especiales.
- El segundo eje apunta a la predicción de accidentes mediante Redes Neuronales Artificiales. Se desarrollan dos modelos predictivos con enfoques distintos pero complementarios: el Modelo RNA N.°1, de clasificación binaria (Riesgo Menor/Mayor) mediantes los datos historicos de los accidentes de tránsito de los años 2020 a 2025; y el Modelo RNA N.°2, de clasificación multiclase (Riesgo Bajo/Medio/Alto), que trabaja con un volumen considerablemente mayor con parametros geométricas extraídas directamente del modelamiento en Civil 3D y parametros ambientales. Ambos modelos desarolladas en Python.
- El tercer eje apunta en desarollar programas de predicción de accidentes de tránsito, resultados de la Redes Neuronales Artificales (RNA) para la evaluación de otras carreteras nacionales y ademas desarollar un programa de alerta de accidentes de tránsito para los conductures a partir de los registros de accidentes de tránsito de los 2020 a 2025.

B) La unidad de análisis: Es el tramo vial Via Los Libertadores (PE-28A) caracterizado por sus parámetros geométricos (radio, peralte, pendiente, DVP, sobreancho, y otros), su localización georreferenciada (progresiva kilométrica, coordenadas), condiciones climáticas y su historial de siniestralidad del período 2020–2025.

1.3 Formulación del problema

1.3.1 Problema general

¿En qué medida el desarrollo de dos modelos de redes neuronales artificiales, el análisis comparativo de la norma DG-2018 con normativas internacionales de diseño geométrico vial, y la implementación de un programa de alerta basado en geolocalización GPS en tiempo real, permiten clasificar la peligrosidad de los accidentes de tránsito y optimizar el diseño geométrico de la Vía Los Libertadores (PE-28A), Ayacucho?.

1.3.2 Problemas específicos

1. ¿Qué nivel de precisión alcanza el Modelo RNA N°01 (clasificación binaria: riesgo menor o riesgo mayor), entrenado con registros históricos de accidentes de tránsito de la Vía Los Libertadores (PE-28A), para predecir accidentes de tránsito utilizando variables del diseño geométrico, condiciones climáticas, características del siniestro y factores viales conforme a la norma DG-2018?.
2. ¿Qué nivel de precisión alcanza el Modelo RNA N°02 (clasificación multiclase: Riesgo Bajo, Riesgo Medio y Riesgo Alto), entrenado con registros de datos de la Vía Los Libertadores (PE-28A) modelados en AutoCAD Civil 3D, para predecir el nivel de riesgo de accidentes de tránsito en tramos viales?.
3. ¿En qué medida los parámetros del diseño geométrico de la norma DG-2018 presentan deficiencias respecto a las normativas internacionales de Noruega, Islandia, Suecia, Reino Unido y Japón, y qué mejoras pueden proponerse para optimizar la seguridad vial de la Vía Los Libertadores (PE-28A)?
4. ¿En qué medida un programa de alerta de accidentes de tránsito que integre registros históricos georreferenciados de la Vía Los Libertadores (PE-28A) y rastreo GPS en tiempo real, constituye una herramienta de seguridad vial viable para conductores?.

1.4 Justificación e importancia

La justificación de la investigación queda plenamente de manifiesto ante la urgente necesidad de poner freno a la elevada siniestralidad en la Vía Los Libertadores (PE-28A), una carretera clave en la región Ayacucho que presenta muchos peligros manifiestos por las deficientes geométricas que el actual diseño de la carretera tiene (basado en la norma DG-2018), el cual no refleja la complejidad real de las condiciones topográficas, las curvas cerradas que presenta la carretera, las pendientes escarpadas, las zonas sin visibilidad, condiciones climáticas, etc. En este sentido, la investigación se justifica en los siguientes puntos:

1. **Justificación técnica:** comparar los parametros de diseño de la norma DG-2018 (peruana) con las normas de los estándares internacionales de diseño de carreteras de aquellos países con menor siniestralidad como Noruega, Islandia y Suecia, en donde se extraerán los aspectos de mejora y se podrán aplicar en favor de una mayor seguridad vial y reducir cada uno de los factores de riesgo que intervienen en los accidentes de tráfico.

2. **Justificación tecnológica:** El estudio de la Vía Los Libertadores (PE-28A) también incorporará el uso de redes neuronales artificiales (RNA) como herramienta predictiva de accidentes de tránsito. Se desarrollan dos modelos de Redes Neuronales Artificiales (RNA) desarollados en el Python. El Modelo RNA N.°1 (clasificación binaria) y el Modelo RNA N.°2 (clasificación multiclase), demostrando que las redes neuronales es una herramienta viable y de precisión para predecir futuros accidentes de tránsito. Tambien se realizara el trazado del eje de la carretera PE-28A en el Software AutoCAD Civil 3D y se implementara el sistema de alerta para conductores.
3. **Justificación social:** La investigación busca contribuir en la predicción de futuros accidentes de transito, y ademas en favor de una infraestructura más segura, más eficiente y más sostenible para los usuarios de esta vía de que conecta 3 regiones, el desarrollo económico y reducir la siniestralidad en el espacio geográfico de la regiones de Ayacucho, Huancavelica e Ica.

1.5 Limitaciones de la investigación

a) **Disponibilidad limitada de registros de siniestros:** La base de datos de siniestros de tránsito de la Vía Los Libertadores (PE-28A) se circunscribe al período 2020–2025, obteniéndose únicamente 202 registros verificados para el Modelo RNA N.°1 (RNA de clasificación binaria). La limitación se debe a que, con anterioridad al año 2020, el Observatorio Nacional de Seguridad Vial del MTC y la policia nacional del Perú no dispone de una base de datos consolidada, georreferenciada y de acceso público para su uso en modelos de aprendizaje automático.

b) **Ausencia del levantamiento topográfico:** Por limitaciones económicas de la presente investigación de pregrado, no se realizó un levantamiento topográfico directo de los 333.90 km de la carretera Via Los Libertadores (PE-28A) mediante estación total o dron topográfico. No obstante, la carretera fue modelada en AutoCAD Civil 3D utilizando el modelo digital del terreno disponible de PROVÍAS Nacional y la información cartográfica, lo que permitió extraer los parámetros de diseño geométrico (radio de curva, pendiente longitudinal, distancia de visibilidad, sobreancho, longitud de curva y longitud tangente, radio entre otros datos),mediante el cual se logro la construcción de una base de datos de 2,997 registros para el Modelo RNA N.°2 (RNA de clasificación Multiclase).

1.6 Objetivos

1.6.1 Objetivo general

Clasificación de la peligrosidad de los accidentes de tránsito y optimizar el diseño geométrico de la Vía Los Libertadores (PE-28A), Ayacucho, mediante el desarrollo de dos modelos de redes neuronales artificiales, el análisis comparativo de la norma DG-2018 con normativas internacionales de diseño geométrico vial, y la implementación de un programa de alerta basado en geolocalización GPS en tiempo real, con el fin de contribuir a la reducción de los acccidentes de tránsito.

1.6.2 Objetivos específicos

a) Desarrollar y evaluar el Modelo RNA N°01 (clasificación binaria: riesgo menor o riesgo mayor), entrenado con 202 registros históricos de accidentes de tránsito de la Vía Los Libertadores (PE-28A), para predecir accidentes de tránsito utilizando variables del diseño geométrico, condiciones climáticas, características del siniestro y factores viales conforme a la norma DG-2018.
b) Desarrollar y evaluar el Modelo RNA N°02 (clasificación multiclase: riesgo bajo, riesgo medio y riesgo alto), entrenado con 2,997 registro de datos de la Vía Los Libertadores (PE-28A) modelados en AutoCAD Civil 3D, para predecir el nivel de riesgo de accidentes de tránsito en tramos viales.
c) Comparar los parámetros del diseño geométrico de la norma DG-2018 con las normativas internacionales de Noruega, Islandia, Suecia, Reino Unido y Japón, identificar las deficiencias del diseño actual de la Vía Los Libertadores (PE-28A) y proponer mejoras orientadas a optimizar su seguridad vial.
d) Desarrollar e implementar un programa de alerta de accidentes de tránsito para la Vía Los Libertadores (PE-28A), que integre registros históricos georreferenciados y rastreo GPS en tiempo real, como herramienta de seguridad vial para conductores.

CAPÍTULO II Marco Teórico

2.1 Antecedentes

2.1.1 Investigaciones internacionales

a) Según **(Codur & Tortum, 2015)**, manifesta en su artículo titulado **"An artificial neural network model for highway accident prediction: a case study of Erzurum, Turkey"** ("Un modelo de red neuronal artificial para la predicción de accidentes de carretera: un estudio de caso de Erzurum, Turquía") presenta un modelo de predicción de accidentes en las carreteras de la provincia de Erzurum, en Turquía, usando redes neuronales artificiales (ANN) que permite predecir los accidentes de tráfico ocurridos en la provincia de Erzurum, Turquía, a partir del análisis de 7.780 informes de accidentes ocurridos entre 2005 y 2012, de los cuales 7.285 han sido considerados tras la depuración de datos incorrectos o faltantes. Presentaron 8 variables que determinan la ocurrencia de los accidentes, donde el grado de curvatura vertical alcanzó un grado de determinación del 54.88%, seguido del tráfico diario promedio anual (AADT) y la curvatura horizontal con un grado de determinación próximo al 15% para cada uno de los factores citados. La red neuronal fue depurada utilizando el algoritmo Levenberg-Marquardt y obtuvo un buen rendimiento, con un coeficiente de correlación de 0.991, R2=0.982 y bajas tasas de error (MSE=4.11, RMSE=2.03). Las conclusiones apuntan a que las ANN son herramientas válidas y flexibles para la predicción de la mortalidad del tráfico, superando para ello a los métodos estadísticos probados. Los autores consideraron también la importancia de determinar los factores determinantes para mejorar la siniestralidad en el tráfico. Al final, los autores sugieren aumentar la base de datos y trabajar con algoritmos genéticos para ayudar a optimizar la red neuronal e intentar ayudar a la toma de decisiones de las políticas de tráfico que hay que desarrollar.

b) Según **(LuWang, 2023)**, manifesta en su artículo titulado **"Safety Evaluation for Highway Geometric Design Based on Spatial Path Properties"** ("Evaluación de la seguridad para el diseño geométrico de carreteras basada en las propiedades de la trayectoria espacial"). Este estudio fue desarrollado en China, donde los investigadores proponen una innovadora metodología para optimar la evaluación de seguridad vial basada en análisis tridimensional de las propiedades geométricas de las curvas, tal

como son: radio, curvatura acumulada, torsión espacial y longitud efectiva. A través de un acercamiento cuantitativo, el presente trabajo demuestra cómo curvas de radios pequeños, cambios bruscos en la ruta y transiciones geométricas inapropiadas muestran una fuerte correlación con el aumento en la tasa de accidentes por millón de vehículos-kilómetro.

Esta investigación Chino permite describir cada segmento de curva mediante modelos de correlación espacial, lo que constituye un eficaz instrumento predictivo del tramo crítico temprano. Por otra parte, se remarca que el rediseño de curvas en paisajes montañosos rurales, y las ubicadas en carriles de alta velocidad, puede hacer considerablemente disminuir la siniestralidad de las vías.

c) Según **(Alvarez, Fernandez, Gorliza, Mansilla & Molinero, 2020)**, manifesta en su artículo titulado **"Geometric road design factors affecting the risk of urban run off crashes. A case-control study"** ("Factores geométricos del diseño vial que afectan el riesgo de accidentes por escorrentía urbana. Un estudio de casos y controles") llevaron acabo un análisis de casos y controles para centrar su investigación en la ocurrencia de accidentes tipo "salida de vía" en contextos urbanos, para determinar qué elementos geométricos del diseño contribuyen al riesgo de siniestralidad significativamente. Su trabajo, realizado en Valladolid (España), encontró que determinados factores tales como radios de curvatura demasiado limitados, secciones rectas prolongadas no interrumpidas, carriles anchos no controlados por velocidad, y una pobre transición entre segmentos geométricos multiplican considerablemente las posibilidades para que se produzcan accidentes.

   Se basa en análisis estadístico y recomienda que estos factores geométricos se aborden desde las consideraciones iniciales al diseñar la infraestructura vial, y centrar las consideraciones en zonas de alta densidad automovilística. Especialmente, destaca que incoherencia entre geometría del trazado y comportamiento esperado del conductor resulta en un desencadenando directo para fallos de conducción.

   Esta experiencia trae beneficios al comparativismo de investigaciones como la presente tesis, al poder proyectar resultados urbanos hacia áreas rurales con las mismas condiciones de riesgo geométricas, como en la Via Los Libertadores (PE-28A), en donde se encuentran curvas sin visibilidad y longitudes rectas alargadas también.

d) Según **(Kriswardhana, Hasanuddin & Palestine, 2020)** se manifestaron en su

artículo titulado **"Modelling road traffic accident rate and road geometric parameters relationship"** ("Modelado de la relación entre la tasa de accidentes de tráfico y los parámetros geométricos de la carretera"). Este estudio se llevo en Indonesia, donde diseñaron un modelo estadístico para investigar cómo différents parámetros geométricos influían en las tasas de accidentes en redes de carreteras interurbanas de Indonesia. El artículo consideró los parámetros de diseño como radio de curvatura, super elevación, ancho de carril y libertad lateral mediante regresiones múltiples para llevar a cabo correlaciones significativas entre dichas variábles y las siniestralidades declaradas.

Uno de los resultados importantes fue una estricta correlación entre "libertad lateral" el ancho entre obstáculo adyacente y borde de calzada y accidentes más frecuentes. Se asumió que un mayor nivel de libertad lateral puede servir como un amortiguador de seguridad contra maniobras evasivas o desaparición de control en curvas, reduciendo mucho los incidentes en carreteras. Por otra parte, el análisis destaca la obligación de incluir estos factores en los procedimientos de diseño geométricos, fundamentalmente en secciones con curvas bruscas o condiciones topográficas complejas.

Esta orientación se puede aplicar ampliamente al caso de estudio de la ruta la Via Los Libertadores (PE-28A), ya que la existencia de curvas cerradas y bordes estrechos pone en riesgo gravemente la seguridad por carretera si no se re-proyectan bajo parámetros regulativos más exigentes y homogéneos.

e) Según **(Thangaraj & Sorimuthu, 2019)** publican en su articulo titulado **"Influence of Geometrical Design Elements on Traffic Safety"**("Influencia de los elementos de diseño geométrico en la seguridad vial"). Este estudio se llevo en la India, donde llevaron acabo un análisis exhaustivo de la literatura para decidir cómo se ven influenciados por elementos geométricos los factores de seguridad vial A su juicio, condiciones que fomentan un mayor riesgo por colisiones provienen de factores como un ancho escaso para carriles, una superamplificación no correcta y problemas en la visibilidad. Los autores destacan que, mediante una buena aplicación de los principios geométricos, junto con las normas utilizadas internacionalmente, se puede reducir sensiblemente los accidents El autor propusieron estrategias específicas de reconfiguración en secciones de alta siniestralidad, destacando especialmente para lograr parámetros homogéneos tanto en curvas como pendientes.

f) Según **(Islam, Teik, Hamid & Azarkerdar, 2019)** en su articulo titulado **"Relation-**

ship of Accident Rates and Road Geometric Design" ("Relación entre las tasas de accidentes y el diseño geométrico de las carreteras"). Este estudio se llevo en Malasia, donde realizaron un estudio cuantitativo sobre cómo los elementos geométricos del diseño, tales como el número de carriles, la distancia de visibilidad, la super-elevación, el ancho de la calzada y el radio de curva, son cruciales para influir en el índice de accidentes. El trabajo concluyó que curvas con radios bajos pendientes inclinadas se han relacionado estrechamente con un mayor siniestrismo. A través de regresión, demuestran cómo un avance en estos parámetros puede hacer desaparecer en un un 30% la probabilidad de accidentes. El presente trabajo reafirma la vigencia del papel de regular estándares geométricas óptimas como base para políticas seguridad vial.

g) Según **(Wang, Wang, Zhou, Jing & Cheng, 2018)** expresaron en su articulo titulado **"A Three-Dimensional Method for Safety Evaluation of Highway Geometric Design"** (Un método tridimensional para la evaluación de la seguridad del diseño geométrico de carreteras). Este estudio fue desarollado en el pais de China, donde ofrecieron un nuevo enfoque para valorar seguridad en las carreteras mediante un modelo tridimensional propiciando tanto la geometría horizontal como vertical del trazado, junto a la torsión espacial. A diferencia de las metodologías tradicionales, basadas en dos dimensions (planta y perfil) para estudiar pista en alineamiento, se trata de un acercamiento que combina el modelo "curvatura espacial continua", adaptado para una valoración más exacta de los cambios en dirección del vehículo en situaciones reales de funcionamiento

   El trabajo hizo un uso combinado de registros de accidentes junto con detalladas descripciones del perfil geométrico de múltiples segmentos de ruta para fijar una relación cantonibile entre variación en curvatura/torsión y número de accidentes por millón de vehículos-kilómetro. Los resultados revelaron que segmentos con cambios irregulares en curvatura o bruscas transiciones entre pendientes y curvas planas representan un mayor cargo cognitivo para el conductor, disminuyendo su capacidad para reaccionar y haciendo aumentar enormemente las posibilidades en siniestros.

   Este procedimiento se encontró particularmente eficaz en vías de topografía compleja, como territorios montañosos, ya que las variaciones en la geometría espacial son más comunes ahí. Según los resultados, se recomienda aplicar un diseño geométrico tridimensional más homogéneo, con transiciones suaves tanto horizontal como vertical,

que no solamente acelera la comodidad en la conducción, sino que también mejora la seguridad en las vías al reducir puntos críticos de conflicto.

h) Según **(Montoya Correa & Garrido Murillo, 2020)** en su tesis que lleva como título **"Criterios de diseño, seguridad y operación geométrico en el diseño vertical de carreteras. Análisis comparativo entre los manuales del invias (Colombia), ministerio de fomento (España) y AASHTO (EE.UU.)"**. En este estudio se lleva a cabo una comparación entre tres guías de diseño de vías en relación al componente geométrico vertical: el Manual de Diseño Geométrico de Carreteras 2008 del Instituto Nacional de Vías de Colombia, el A Policy on Geometric Design of Highways and Streets 2011 de la American Association of State Highway and Transportation Officials de Estados Unidos, y la Norma 3.1-IC Trazado de la Instrucción de Carreteras 2016 del Ministerio de Fomento de España. Para realizar un análisis comparativo serio, se tradujo al español el capítulo respectivo al diseño vertical del manual AASHTO. A partir de su traducción, se realizó un cuadro comparativo que permite hacer patentes los criterios técnicos respectivos referentes al diseño, seguridad y funcionamiento en trazado vertical según cada reglamento. Esta investigación apunta a desvelar las limitaciones existentes en el manual colombiano INVIAS para ofrecer propuestas de modificación en el diseño geométrico vertical para las redes viales colombianas.

i) Según **(Garzón, Escobar & Galindo, 2017)**, menciona en su artículo titulado **"Seguridad Vial en Carreteras"**. Ejemplo de implementación metodológica, una metodología general para su uso en Auditorías. "de Seguridad Vial" ASV, empleada en la ciudad de Manizales (Colombia), donde se realizaron como elementos de estudio, los lugares donde ocurrieron accidentes fatales durante el año 2015.

j) Según **(García Chávez, 2012)** en su Tesis **"Propuesta De Mejoramiento de la Seguridad Vial de una Carretera de Elevada Accidentalidad Utilizando Tecnologias ITS (Sistemas Inteligentes de Transporte)"** tenía como meta principal, determinar y elegir un segmento vial de la Red Carretera Federal de México adecuado para realizar un proyecto piloto de identificación, análisis, valoración e implementación de tecnologías ITS conexas a la seguridad, realizando el análisis de las razones de la siniestralidad en esa sección, la propuesta de tecnologías ITS para reducir los inconvenientes detectados en el análisis y la valoración La viabilidad económica de dicha propuesta.

2.1.2 Investigaciones nacionales

a) Según **(Quispe Condor, 2014)** publica en su tesis que lleva como título **"Diseño de una Red Neuronal para la Predicción de Accidentes de Tránsito en la Carretera Central Tramo: Tarma-La Merced, Junín 2022"** tuvo como objetivo crear una red neuronal artificial (RNA) que sirviera para predecir accidentes de tránsito a partir de aislar las variables geométricas de la carretera, entre ellas el radio de curva, el ángulo de deflexión, la pendiente longitudinal y la visibilidad. Para la fase de recopilación de datos históricos de accidentes y de las características geométricas de los lugares de accidente se optó por usar fichas de campo y aplicativos como Excel y NeuralTools. La base de datos utilizada contenía los puntos críticos donde se encontraban radios de curva de 20 a 50 metros, pendientes de 7 a 12 grados y visibilidad restringida.

   Se seleccionaron 27 puntos críticos para el entrenamiento que fueron clasificados utilizando los criterios del análisis de riesgo en tres niveles: bajo, medio y alto. Diversas arquitecturas de la RNA fueron tratadas mediante la variación del número de nodos de entrada y ocultos para reducir la tasa de error de la predicción. Las variables de entrada fueron tipo de curva (circular, elíptica, en "S", entre otras), radio, ángulo de deflección, pendiente y visibilidad, siendo el nivel de riesgo la variable salida. Los resultados obtenidos mostraron que las curvas constituidas por radios pequeños o medianos combinados con ángulos de deflección muy elevados (derechas) presentan alto riesgo, al igual que los tramos rectos con baja visibilidad. Asimismo, la RNA demostró ser una herramienta útil para anticipar las zonas críticas de las vías y para mejorar el diseño vial. Para ello, se recomienda aumentar la base de datos existente y aplicar el modelo en vías especialmente rurales con el fin de mejorar el diseño que optimiza la seguridad vial.

b) Según **(Chavez Castro, 2022)** publica en su tesis que lleva como título **"Estandarización de diseño geométrico a través del análisis de estándares de otros países de Sudamérica para actualizar el Manual de Carreteras - Diseño Geométrico DG 2018"**. El presente trabajo tiene como finalidad determinar unos parámetros geométricos que sirven para la concepción de caminos a fin de revisar y actualizar el manual de caminos - diseño geom étrico DG 2018, para ser usado en el Perú. De todo lo estudiado y ponderado en el contenido del diseño del transporte , la ampliación en curvas y el tiempo de reacción-percepción en relación con la distancia de visibilidad

de detención, poniendo en consideración los códigos en las normativas de regulaciones internacionales. Después de recolectar datos y hacer una comparación entre la normativas , se propuso modificar el manual peruano tomando como base principalmente las especificaciones del manual boliviano para vehículo de diseño y sobreadelgamiento en curvas. Al referirnos al tiempo percepción-reacción se propone mantener 2.5 segundos, en acuerdo con las regulaciones de Ecuador y Colombia, y desaparecer gráficos innecesarios incluidos en el DG 2018.

c) Según **(Maquera Maquera, 2023)** manifiesta en su tesis que lleva como título **"Evaluación del diseño geométrico de la carretera mo-518 tramo los Ángeles – Yacango, Moquegua, con la aplicación del Vehicle Tracking de Autocad CIVIL 3D"** se realizaron las investigaciones en tramo kilométrico 7+000 a 9+000, en sede Las Águilas, zona rural, provincia Mariscal Nieto, Moquegua, donde se encuentran curvas muy cerradas. Se buscan evaluar mediante simulación utilizando el software VEHICLE TRACKING el diseño geométrico para obtener condiciones óptimas de seguridad y reflejar mejores resultados, se utilizó una metodología descriptiva con Análisis estadístico multivariado ANOVA para comparar sobreancho y radios mínimos, determinando las curvas críticas defectuosas son seis, se compararon los parámetros actuales frente a las normas DG-2018 y AASHTO-2011, concluyendo que el diseño bajo la norma americana presenta mejores condiciones para la seguridad, se recomienda mejorar las curvas críticas diseño geométrico mediante los parámetros AASHTO-2011 para optimizar las condiciones de seguridad vial.

d) Según **(Cardenas Achahuanco, 2024)** en su tesis que lleva como título **"Propuesta de estandarización de diseño geométrico de trochas carrozables en la DG-2018 a través del análisis comparativo de normativas extranjeras"** presenta una propuesta de estandarización de parámetros para diseño geométrico para trochas carrozables, ya que las vías que equivalen al 20.98 % de las calles no asfaltadas en Perú no poseen especificaciones delineadas en el Manual de Carreteras DG-2018. El trabajo acoge un diseño deducido, utilizando herramientas como método Delphi, análisis comparativo de normativas extrajeras (Brasil y Chile) e información asistida por Civil 3D para diseño asistido por computadoras. Se consideraron parámetros importantes para trochas carrozables: IMDa < 200, radio mínimo 20 m, pendiente max 13 %, ancho mínimo carril 3 m y bermas 0.5 m. La validación mediante consenso experto mediante Delphi

y análisis costo-beneficio le permitió basar técnica y objetivamente su incorporación en futuras actualizaciones normativas.

e) Según **(Heredia Martinez & Perez Heredia, 2023)** en su tesis que lleva como título **"Propuesta de parámetros para el diseño geométrico de trochas carrozables bajo el enfoque de seguridad activa y normas internacionales, aplicado a la carretera Balconcillo-Cruce San Andrés, Cutervo"**. El estudio examina la falta de información sobre la norma peruana para el diseño geométrico de trochas carrozables, a pesar de ser una de las vías más transitadas en el país. Se entiende por enfoque de seguridad activa la sugerencia de parámetros básicos tales como radio mínimo, peralte, sobreancho, pendiente, curvas verticales, calzada, bombeo, bermas y distancias de visibilidad. Se consideran las dimensiones más seguras fundamentadas en antecedentes técnicos según la frecuencia de accidentes en caminos no pavimentados de bajo flujo vehicular. Se lleva a cabo un estudio comparativo entre las normativas internacionales (AASHTO, ARRB, MRRD) y el ya derogado manual peruano MDCNPBVT, dando como resultado los Parámetros Recomendados para Trochas Carrozables (PRTC). Estos se aplican al camino vecinal Balconcillo–Sahiundo–San Andrés (Cajamarca), donde se aprecian deficiencias en seguridad en el diseño original. Aunque costaba menos mover tierras, se sacrificaba la seguridad y se multiplicaba el mantenimiento. Finalmente, se deduce que, junto al uso de parámetros mínimos de seguridad, se debe hacer un análisis costo-beneficio en el diseño para conseguir un punto de equilibrio entre seguridad, costes de ejecución y mantenimiento.

f) Según **(Boza Pareja, 2018)** manifiesta en la tesis que lleva como título **"Análisis comparativo según los criterios de seguridad y operación a nivel de diseño geométrico de la ruta arco Tica Tica - Chinchero usando los manuales de diseño DG 2014 peruano, manual de diseño geométrico Ecuatoriano de carreteras 2003, y manual de diseño geométrico de carreteras Colombiano 2008 en la region del Cusco"**. La propuesta se realizó para diseñar geométricamente la alternativa de la vía entre Chinchero y Arco Tica Tica, utilizando criterios de seguridad y de operación según los manuales de diseño geométrico de Perú (DG-2014), Ecuador (2003) y Colombia (2008). Se revisaron rutas actuales y alternas, aplicando información del IMDA. Al realizar la comparación de las normas mostró que el manual ecuatoriano 2003 es uno de los más exigentes por criterio de seguridad y que el DG-2014 peruano resalta por

criterio de operación. Además, el DG-2014 y el manual ecuatoriano tienen dos tramos homogéneos, mientras que el colombiano tiene tres. En conclusión, se corroboró el supuesto que el ecuatoriano es más exigente en seguridad, y el peruano de operación, sugiriendo adoptar elementos más rigurosos para futuras actualizaciones normativas de diseño geométrico vial peruano.

g) Según **(Quispe Aquino & Poma Milian, 2019)** manifiesta en la tesis que lleva como título **"Norma de estandarización de diseño geométrico de trochas carrozables para modificar la norma del MTC a través del análisis de estándares de otros países"**. La presente tesis se centró principalmente en establecer una norma de estandarización para el diseño geométrico de vías carrozables, considerando manuales técnicos internacionales con la finalidad de optimizar y actualizar la normativa emitida por el MTC. El estudio adoptó un enfoque mixto (cuantitativo-cualitativo), lo que facilita la recolección, análisis e interpretación completa de información técnica y contextual. Se analizaron y definieron criterios de comparación entre varios manuales de diversas naciones con condiciones similares de geografía y tránsito, enfocándose en estos componentes esenciales de diseño geométrico: sección transversal (carril y berma), alineación horizontal (radio de curvatura mínimo, distancia de visibilidad de detención y de adelantamiento) y alineación vertical (pendiente máxima), con el objetivo de mejorar la operatividad y la seguridad vial de las carreteras. Por lo tanto, se realizó una propuesta normativa adaptada al volumen diario de tránsito promedio (IMD), estableciendo parámetros específicos para cada categoría: para un IMD de 0–50 veh/día, se sugiere una calzada de 4.50 m de ancho, radio mínimo de 15 m y pendiente máxima de 17%; para 50–100 veh/día, ancho de 5.50 m, radio de 25 m y pendiente máxima de 15%; y para 100–200 veh/día, ancho de 6.50 m, radio de 40 m y pendiente máxima de 12%. El estudio concluye que es factible y técnicamente imprescindible implementar un estándar específico para trochas carrozables, en línea con las condiciones topográficas y del tránsito rural en Perú.

2.2 Bases teóricas

2.2.1 Redes Neuronales Artificiales (RNA)

Las redes neuronales artificiales (RNA) son uno de los grupos de técnicas dentro del ámbito de la inteligencia artificial y del aprendizaje automático, que han tomado como modelo el

funcionamiento del cerebro humano. Estas técnicas se basan en una colección de nodos (neuronas artificiales) interconectados organizados en capas de tal manera que permiten modelar relaciones complejas entre entradas y salidas. Las redes neuronales artificiales, usualmente, aprenden a partir de datos históricos y pueden identificar patrones, no lineales, útiles para tareas de predicción o clasificación.

Figura 16 Redes neuronales inspirados en el cerebro humano.
Fuente: The University of Science Technology

Según **(Haykin, 2009)** define como:"Un procesador masivamente paralelo distribuido, constituido por unidades de procesamiento simples, que tiene una propensión natural para almacenar conocimiento experiencial y hacerlo disponible para su uso. Se asemeja al cerebro en dos aspectos: el conocimiento es adquirido por la red a partir de su entorno a través de un proceso de aprendizaje, y las fuerzas de conexión entre neuronas, conocidas como pesos sinápticos, se emplean para almacenar el conocimiento adquirido. (p. 2)"

En el caso del diseño geométrico vial, las redes neuronales pueden ser entrenadas con datos sobre características geométricas del trazado de las carreteras, y también con datos sobre accidentes, para predecir el riesgo vial, o para sugerir alternativas de rediseños más seguros, siguiendo la normativa de la técnica vigente.

a) **Neurona Artificial:** Las redes neuronales artificiales (RNA) son modelos computacionales que simulan la organización y el comportamiento del sistema nervioso biológico, conformados por conjuntos de neuronas artificiales distribuidas en capas interconectadas. Su principal capacidad radica en el procesamiento de información y en

el reconocimiento de patrones no lineales presentes en los datos, permitiendo aproximar funciones matemáticas complejas y generar predicciones a partir de variables de entrada. Cada neurona artificial opera mediante los siguientes componentes fundamentales:

- **Entradas:** Señales numéricas provenientes del conjunto de datos o transmitidas desde neuronas de capas anteriores.
- **Pesos sinápticos:** Parámetros ajustables que ponderan la contribución de cada señal de entrada al procesamiento de la neurona.
- **Función de agregación:** Operación que combina las entradas mediante una suma ponderada, integrando los pesos sinápticos y las señales recibidas.
- **Sesgo (*bias*):** Término independiente que desplaza el umbral de activación, otorgando mayor flexibilidad al modelo durante el aprendizaje.
- **Función de activación:** Transforma la salida de la función de agregación introduciendo no linealidad, lo que permite al modelo capturar relaciones complejas entre las variables.

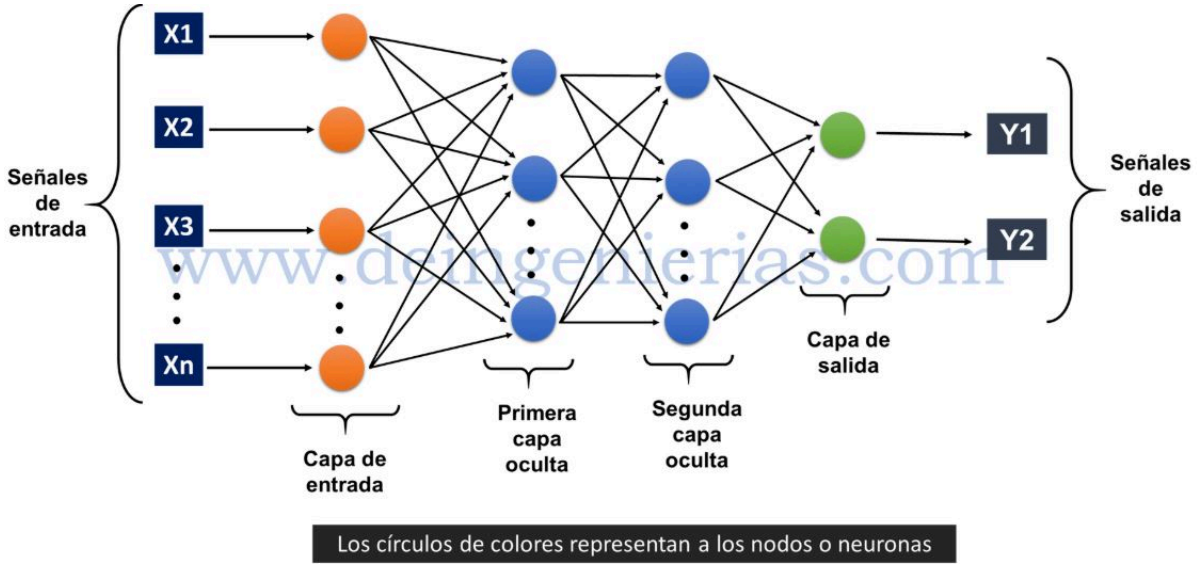


Figura 17 Arquitectura de un Red Neuronal.

Fuente: www.deingenierias.com

Según **(Bishop, 2006)** explica que ”los pesos sinápticos determinan la importancia relativa de cada entrada, y el sesgo (bias) permite desplazar la función de activación horizontalmente, proporcionando un grado adicional de flexibilidad al modelo” (p. 227). Sin la función de activación no lineal, una red de múltiples capas se reduciría algebraicamente a una transformación lineal simple, independientemente de su profundi-

dad, lo que limitaría drásticamente su capacidad representacional **(Goodfellow, Bengio & Courville, 2016)**.

Matemáticamente, la salida de una neurona puede expresarse como:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{2.1}$$

Donde:

- x_i son las entradas.
- w_i son los pesos.
- b es el bias.
- f es la función de activación.

b) **Funciones de activación:** Las funciones de activación constituyen elementos fundamentales en la arquitectura de una RNA, puesto que introducen no linealidad en el modelo y regulan la señal de salida de cada neurona hacia las capas subsiguientes (Goodfellow et al., 2016). Sin esta propiedad, la red se reduciría a una simple transformación lineal, incapaz de representar relaciones complejas entre variables. Entre las funciones de activación más empleadas en la literatura se tienen:

- **Sigmoide:** Definida como $f(x) = \frac{1}{1+e^{-x}}$, comprime cualquier valor real al intervalo $[0, 1]$, siendo útil en capas de salida para problemas de clasificación binaria.
- **Tangente hiperbólica (Tanh):** Expresada como $f(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$, mapea los valores al intervalo $[-1, 1]$, presentando salidas centradas en cero que favorecen la convergencia del entrenamiento.
- **ReLU (*Rectified Linear Unit*):** Definida como $f(x) = \text{máx}(0, x)$, anula las activaciones negativas y preserva las positivas sin modificación, lo que reduce el costo computacional y acelera la convergencia en redes profundas.
- **Leaky ReLU:** Variante de ReLU expresada como $f(x) = \text{máx}(\alpha x, x)$, donde α es un valor pequeño positivo. Permite un gradiente residual para entradas negativas, mitigando el problema de neuronas inactivas .
- **ELU (*Exponential Linear Unit*):** Extiende las ventajas de ReLU al incorporar una respuesta exponencial para valores negativos, reduciendo el sesgo en las activaciones y atenuando el problema de neuronas muertas.
- **Softmax:** Transforma un vector de valores reales en una distribución de probabilidad sobre K clases, garantizando que la suma de las salidas sea igual a la unidad.

Es la función estándar en la capa de salida para tareas de clasificación multiclase.

c) **Aprendizaje y Entrenamiento:** El entrenamiento de una RNA comprende dos etapas secuenciales e interdependientes que permiten ajustar los parámetros del modelo hasta minimizar el error de predicción:

- **Propagación directa (*Forward Propagation*):** Durante esta etapa, los datos de entrada recorren la red desde la capa de entrada hasta la capa de salida. En cada neurona se aplica la función de agregación seguida de la función de activación correspondiente, transformando progresivamente la señal hasta generar una predicción en la capa final (Goodfellow et al., 2016).
- **Retropropagación del error (*Backpropagation*):** Una vez obtenida la predicción, se cuantifica la discrepancia respecto al valor real mediante una función de pérdida. El gradiente de dicho error se propaga en sentido inverso a través de las capas de la red, permitiendo calcular la contribución de cada peso al error total. Mediante el algoritmo de descenso del gradiente, los pesos sinápticos se actualizan iterativamente con el objetivo de minimizar la función de pérdida (Goodfellow et al., 2016).

2.2.2 Tipos de Redes Neuronales

Las redes neuronales artificiales se clasifican en diversas arquitecturas según su topología, dirección del flujo de información y mecanismo de aprendizaje. Según Haykin **(Haykin, 2009)** establece que ”la arquitectura de una red neuronal está íntimamente ligada al algoritmo de aprendizaje utilizado para entrenarla, y la selección de la arquitectura apropiada es un paso crítico en el diseño del modelo” (p. 21). A continuación se describen los principales tipos de RNA:

a) **Perceptrón Simple:**

El perceptrón simple fue propuesto por **(Rosenblatt, 1958)** y representa la forma más elemental que puede adoptar una red neuronal: una capa de entradas conectada directamente a una capa de salida, sin nada en medio. Esa sencillez tiene un coste: solo puede clasificar correctamente patrones que sean linealmente separables, lo que limita mucho su uso en problemas reales. **(Minsky & Papert, 1969)** lo demostraron con rigor matemático en su obra Perceptrons, mostrando

que este modelo es incapaz de resolver algo tan básico como la función XOR, solo resuelven problemas basicos.

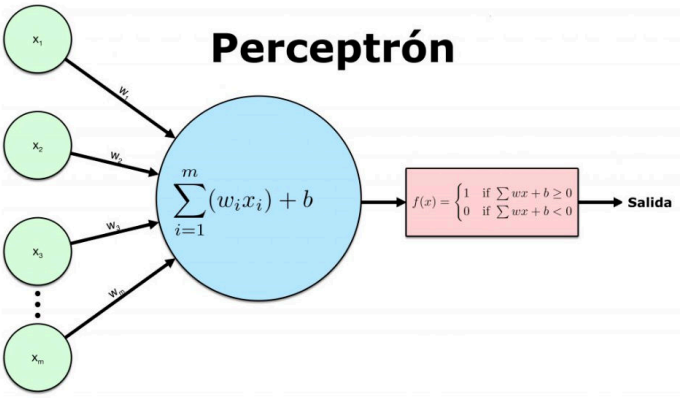


Figura 18 Red Neuronal Perceptron Simple.
Fuente: Blog de Jose Mariano Alvarez.

b) **Perceptrón Multicapa (MLP)**

El Perceptrón Multicapa descrito en detalle en la sección anterior supera las limitaciones del perceptrón simple mediante la incorporación de una o más capas ocultas con funciones de activación no lineales. **(Bishop, 2006)** señala que "la introducción de capas ocultas permite a la red aprender representaciones internas de los datos, transformando el espacio de entrada en un espacio donde las clases son linealmente separables" (p. 229).

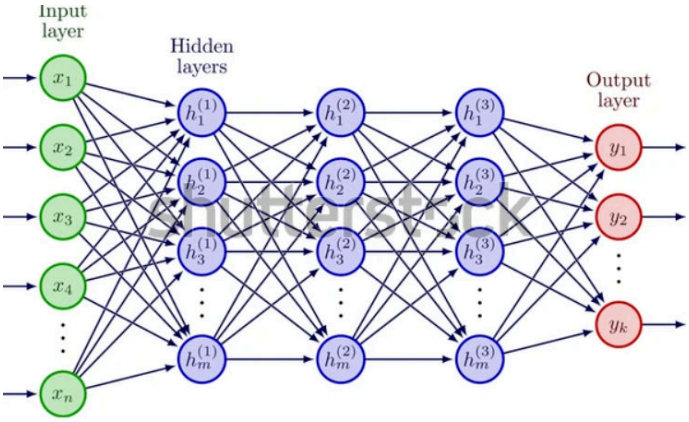


Figura 19 Red Neuronal Perceptron Múltiple.
Fuente: Shutterstock

c) **Redes Neuronales Convolucionales (CNN)**

Las Redes Neuronales Convolucionales (CNN, Convolutional Neural Networks), desarrolladas por **(LeCun, Bottou, Bengio & Haffner, 1998)**, están diseñadas específicamente para el procesamiento de datos con estructura espacial, como imágenes, mediante operaciones de convolución que extraen características locales de manera jerárquica. Según **(Goodfellow et al., 2016)** las definen como

”redes que emplean la operación matemática de convolución en lugar de la multiplicación matricial general en al menos una de sus capas” (p. 326).

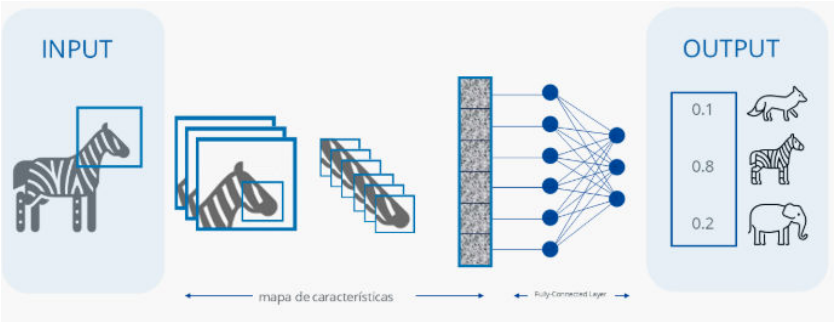


Figura 20 Redes neuronales Convulocionales.
Fuente: Ionos.

d) **Redes Neuronales Recurrentes (RNN) y LSTM** Las Redes Neuronales Recurrentes (RNN, Recurrent Neural Networks) incorporan conexiones de retroalimentación que permiten a la red mantener una ”memoria” de las entradas previas, lo que las hace idóneas para el procesamiento de datos secuenciales y series temporales. **(Haykin, 2009)** explica que ”las conexiones recurrentes permiten a la red exhibir comportamiento dinámico temporal, a diferencia de las redes feedforward que producen una respuesta estática” (p. 24).

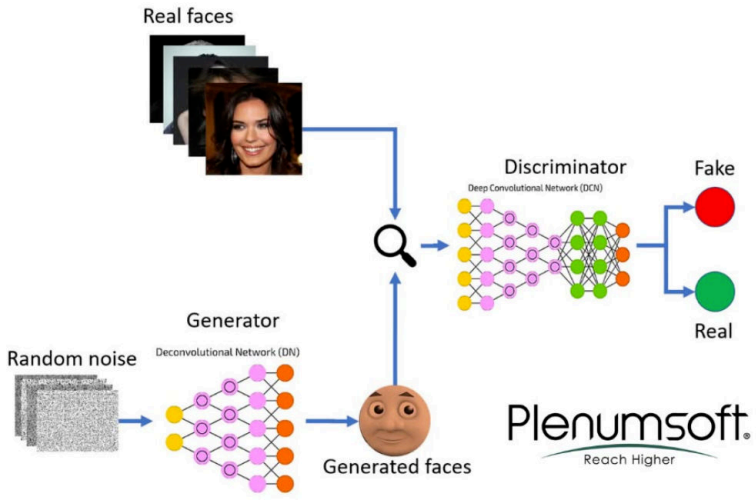


Figura 21 Redes neuronales Recurrentes.
Fuente: The University of Science Technology

e) **Redes de Base Radial (RBF)**

Las Redes de Base Radial (RBF, Radial Basis Function Networks) emplean funciones de base radial, típicamente gaussianas, como funciones de activación en la capa oculta, produciendo activaciones localizadas en el espacio de entrada. **(Bishop, 2006)** señala que ”las redes RBF realizan una transformación no lineal

del espacio de entrada seguida de una combinación lineal en la capa de salida, lo que permite un entrenamiento más rápido que los MLP convencionales" (p. 299).

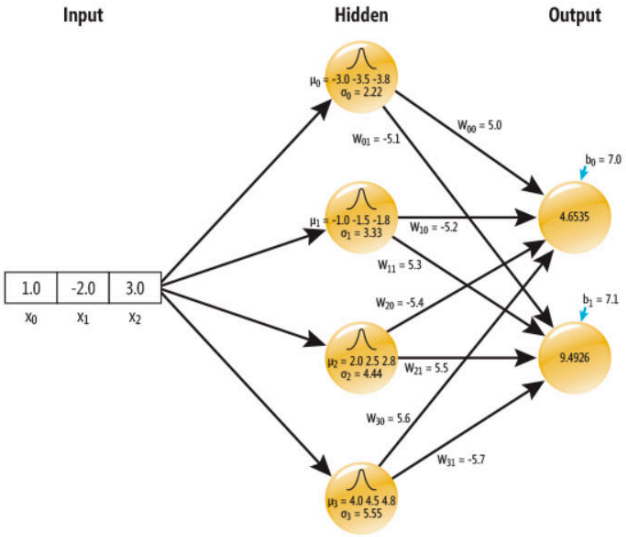


Figura 22 Redes neuronales de Base Radial
Fuente: James McCaffrey

f) **Redes Generativas Adversarias (GAN)**

Las Redes Generativas Adversarias (GAN, Generative Adversarial Networks), propuestas por **(Goodfellow et al., 2016)**, consisten en dos redes que compiten entre sí: un generador que produce datos sintéticos y un discriminador que intenta distinguir los datos sintéticos de los reales. Esta competencia adversaria produce datos sintéticos progresivamente más realistas.

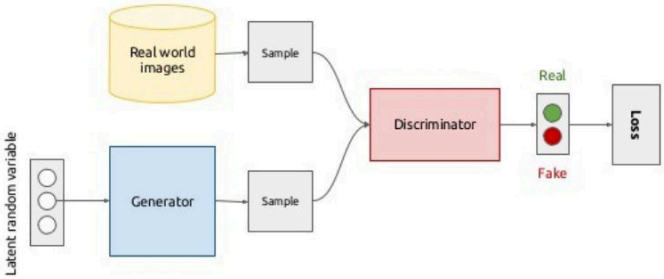


Figura 23 Redes neuronales Generativas Adversarias.
Fuente: The University of Science Technology

2.2.3 Redes Neuronales Multicapa para Clasificación

El Perceptrón Multicapa (*Multilayer Perceptron*, MLP) representa la evolución natural del perceptrón simple y constituye la arquitectura de red neuronal artificial más ampliamente utilizada en problemas de clasificación y regresión en ingeniería. Su principio fundamental radica en la capacidad de aprender representaciones no lineales de

los datos mediante la composición sucesiva de transformaciones paramétricas organizadas en capas (Goodfellow et al., 2016).

El Perceptrón Multicapa (*Multilayer Perceptron*, MLP) representa la evolución natural del perceptrón simple y constituye la arquitectura de red neuronal artificial más ampliamente utilizada en problemas de clasificación y regresión en ingeniería. Su principio fundamental radica en la capacidad de aprender representaciones no lineales de los datos mediante la composición sucesiva de transformaciones paramétricas organizadas en capas (Goodfellow et al., 2016).

2.2.3.1 Formulación del Problema de Clasificación

En un problema de clasificación supervisada, el objetivo del MLP es aprender una función discriminante $f : \mathbb{R}^n \to \{1, 2, \ldots, K\}$ que mapee un vector de características de entrada $\mathbf{x} \in \mathbb{R}^n$ hacia una de K clases predefinidas, a partir de un conjunto de entrenamiento etiquetado:

$$\mathscr{D} = \left\{(\mathbf{x}^{(i)}, y^{(i)})\right\}_{i=1}^{N}, \qquad y^{(i)} \in \{1, 2, \ldots, K\} \tag{2.2}$$

El MLP no aprende directamente la clase, sino la probabilidad posterior de pertenencia a cada clase condicionada a la entrada:

$$P(y = k \mid \mathbf{x};\, \theta\mathbf{0}), \qquad k = 1, 2, \ldots, K \tag{2.3}$$

La clase predicha es aquella con mayor probabilidad posterior (Bishop, 2006):

$$\hat{y} = \arg\max_{k}\ P(y = k \mid \mathbf{x};\, \theta\mathbf{0}) \tag{2.4}$$

2.2.3.2 Tipos de Clasificación y Configuración de la Capa de Salida

La arquitectura de la capa de salida del MLP se define en función del tipo de clasificación a resolver.

A) Clasificación Binaria ($K = 2$)

Se emplea una sola neurona de salida con función de activación **sigmoide**:

$$\hat{y} = \sigma\left(z^{(L)}\right) = \frac{1}{1 + e^{-z^{(L)}}} \in (0,\ 1) \tag{2.5}$$

La salida se interpreta como $P(y = 1 \mid \mathbf{x})$. La regla de decisión es:

$$\hat{y} = \begin{cases} 1 & \text{si } \sigma\left(z^{(L)}\right) \geq 0{,}5 \\ 0 & \text{si } \sigma\left(z^{(L)}\right) < 0{,}5 \end{cases} \tag{2.6}$$

La función de pérdida asociada es la **entropía cruzada binaria**:

$$\mathscr{L}_{\text{BCE}} = -\frac{1}{N}\sum_{i=1}^{N}\left[y^{(i)} \log \hat{y}^{(i)} + \left(1 - y^{(i)}\right) \log\left(1 - \hat{y}^{(i)}\right)\right] \tag{2.7}$$

B) Clasificación Multiclase ($K >= 3$).

Se emplean K neuronas de salida con función de activación **softmax**:

$$\hat{y}_k = \text{softmax}\left(z_k^{(L)}\right) = \frac{e^{z_k^{(L)}}}{\sum_{j=1}^{K} e^{z_j^{(L)}}}, \qquad k = 1, \ldots, K \tag{2.8}$$

Esta función garantiza que las salidas constituyan una distribución de probabilidad válida:

$$\sum_{k=1}^{K} \hat{y}_k = 1, \qquad \hat{y}_k \in (0,\ 1) \quad \forall\, k \tag{2.9}$$

La función de pérdida asociada es la **entropía cruzada categórica**:

$$\mathscr{L}_{\text{CCE}} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{ik} \log \hat{y}_{ik} \tag{2.10}$$

donde y_{ik} es la codificación *one-hot* de la clase verdadera del ejemplo i.

one-hot antes del entrenamiento. Si existen $K = 3$ clases, la codificación es:

$$y = 1 \;\Rightarrow\; \mathbf{y} = [1,\ 0,\ 0]^\top$$
$$y = 2 \;\Rightarrow\; \mathbf{y} = [0,\ 1,\ 0]^\top$$
$$y = 3 \;\Rightarrow\; \mathbf{y} = [0,\ 0,\ 1]^\top$$

Esta representación permite calcular la entropía cruzada categórica de forma vectorizada y es compatible con la activación softmax de la capa de salida (Goodfellow et al., 2016).

2.2.3.3 Flujo Completo de Clasificación en un MLP

El proceso completo, desde la entrada hasta la predicción de clase, puede resumirse en las siguientes etapas:

$$\underbrace{\mathbf{x}}_{\text{entrada}} \xrightarrow{\mathbf{W}^{(1)},\mathbf{b}^{(1)}} \underbrace{\mathbf{a}^{(1)}}_{\text{capa 1}} \xrightarrow{\mathbf{W}^{(2)},\mathbf{b}^{(2)}} \cdots \xrightarrow{\mathbf{W}^{(L)},\mathbf{b}^{(L)}} \underbrace{\hat{\mathbf{y}}}_{\text{softmax}} \xrightarrow{\arg\max} \underbrace{\hat{k}}_{\text{clase}} \tag{2.11}$$

Para cada capa oculta $l = 1, \ldots, L-1$:

$$\mathbf{z}^{(l)} = \mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)} \tag{2.12}$$
$$\mathbf{a}^{(l)} = \mathrm{ReLU}\left(\mathbf{z}^{(l)}\right) = \max\left(\mathbf{0},\ \mathbf{z}^{(l)}\right) \tag{2.13}$$

Para la capa de salida L:

$$\mathbf{z}^{(L)} = \mathbf{W}^{(L)}\mathbf{a}^{(L-1)} + \mathbf{b}^{(L)} \tag{2.14}$$
$$\hat{\mathbf{y}} = \mathrm{softmax}\left(\mathbf{z}^{(L)}\right) \tag{2.15}$$

2.2.3.4 Retropropagación Aplicada a Clasificación

A) Gradiente en la Capa de Salida.

Para la combinación **softmax + entropía cruzada categórica**, el gradiente tiene una forma notablemente simplificada (Goodfellow et al., 2016):

$$\delta \mathbf{0}^{(L)} = \hat{\mathbf{y}} - \mathbf{y} \tag{2.16}$$

Esta expresión indica que el error en la capa de salida es la diferencia entre la probabilidad predicha y la etiqueta real, lo que confiere estabilidad numérica al algoritmo de entrenamiento.

B) Gradiente en las Capas Ocultas.

Para cada capa oculta $l = L - 1, \ldots, 1$, usando activación ReLU:

$$\delta 0^{(l)} = \left(\mathbf{W}^{(l+1)}\right)^{\top} \delta 0^{(l+1)} \odot \mathbf{1}\left[\mathbf{z}^{(l)} > \mathbf{0}\right] \tag{2.17}$$

donde $\mathbf{1}[\mathbf{z}^{(l)} > \mathbf{0}]$ es la derivada de ReLU —vale 1 para activaciones positivas y 0 para negativas— y $\odot$ denota el producto de Hadamard elemento a elemento.

C) Actualización de Parámetros.

$$\mathbf{W}^{(l)} \leftarrow \mathbf{W}^{(l)} - \eta\, \delta 0^{(l)} \left(\mathbf{a}^{(l-1)}\right)^{\top} \tag{2.18}$$

$$\mathbf{b}^{(l)} \leftarrow \mathbf{b}^{(l)} - \eta\, \delta 0^{(l)} \tag{2.19}$$

2.2.3.5 Métricas de Evaluación en Clasificación

A diferencia de la regresión, la evaluación de un MLP clasificador requiere métricas específicas que capturen el rendimiento por clase y el comportamiento ante el desbalance de datos.

A) Matriz de Confusión.

La matriz de confusión representa la herramienta central para evaluar el desempeño de modelos de clasificación, dado que sintetiza de forma estructurada la relación entre los valores predichos por el modelo y las etiquetas reales del conjunto de datos (Minaee et al., 2021). A partir de sus elementos verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN) se derivan las siguientes métricas de evaluación:

Para K clases, la matriz de confusión $\mathbf{C} \in \mathbb{R}^{K \times K}$, donde C_{ij} representa el número de

ejemplos de la clase *i* predichos como clase *j*:

$$\mathbf{C} = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1K} \\ C_{21} & C_{22} & \cdots & C_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ C_{K1} & C_{K2} & \cdots & C_{KK} \end{pmatrix} \tag{2.20}$$

Los elementos de la diagonal C_{kk} corresponden a las clasificaciones correctas de cada clase.

B) Métricas Derivadas por Clase.

- **Exactitud global (*Accuracy*):** La exactitud (*Accuracy*) cuantifica la fracción de instancias correctamente clasificadas respecto al total de observaciones evaluadas (Brownlee, 2020b). No obstante, su interpretación debe realizarse en conjuntos de datos con distribuciones de clase desbalanceadas. Su expresión matemática es la siguiente:

$$\text{Exactitud} = \frac{VP + VN}{VP + VN + FP + FN} \tag{2.21}$$

  donde los términos que componen la expresión se definen como:

  - **VP (Verdaderos Positivos):** Instancias de clase positiva correctamente identificadas por el modelo.
  - **VN (Verdaderos Negativos):** Instancias de clase negativa correctamente identificadas por el modelo.
  - **FP (Falsos Positivos):** Instancias de clase negativa clasificadas erróneamente como positivas.
  - **FN (Falsos Negativos):** Instancias de clase positiva clasificadas erróneamente como negativas.

- **Precisión (*Precision*):** La precisión expresa la proporción de casos positivos predichos que corresponden efectivamente a instancias positivas reales, resultando especialmente relevante en contextos donde incurrir en falsos positivos

conlleva un costo elevado (Evidently AI, 2023).

$$\text{Precisión} = \frac{VP}{VP + FP} \tag{2.22}$$

- **Sensibilidad (*Recall*):** La sensibilidad, denominada también tasa de verdaderos positivos, evalúa la capacidad del modelo para detectar correctamente el total de instancias pertenecientes a la clase positiva (Google Developers, 2023). Un valor reducido de esta métrica indica que el modelo omite una proporción considerable de casos positivos reales.

$$\text{Sensibilidad} = \frac{VP}{VP + FN} \tag{2.23}$$

- **Puntuación F1 (*F1-Score*):** El F1-Score constituye la media armónica entre la precisión y la sensibilidad, integrando ambas métricas en un único indicador que penaliza los desequilibrios entre ellas (The Data Scientist, 2023). Su empleo resulta particularmente adecuado en problemas con distribuciones de clase desbalanceadas, donde optimizar exclusivamente la exactitud puede conducir a conclusiones erróneas sobre el rendimiento real del modelo.

$$F1\text{-}Score = 2 \times \frac{\text{Precisión} \times \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}} \tag{2.24}$$

C) Promedios Macro y Ponderado.

Cuando las clases están desbalanceadas situación frecuente en siniestralidad vial donde los accidentes graves son menos numerosos se utilizan (Sokolova & Lapalme, 2009):

Macro-promedio: promedio simple entre clases, penaliza igualmente el mal rendimiento en clases minoritarias:

$$F1_{\text{macro}} = \frac{1}{K} \sum_{k=1}^{K} F1_k \tag{2.25}$$

Promedio ponderado: promedio ponderado por el soporte de cada clase:

$$F1_{\text{pond}} = \frac{\sum_{k=1}^{K} n_k F1_k}{N} \tag{2.26}$$

donde n_k es el número de ejemplos reales de la clase k.

D) Curva ROC y AUC.

Para clasificación binaria o extensión *One-vs-Rest* en multiclase, la curva ROC grafica la tasa de verdaderos positivos frente a la tasa de falsos positivos para distintos umbrales de decisión. El área bajo la curva (AUC) resume el rendimiento discriminativo del modelo (Fawcett, 2006):

$$\text{AUC} \in [0{,}5\ ;\ 1{,}0] \tag{2.27}$$

Un valor AUC $= 1{,}0$ indica clasificación perfecta; AUC $= 0{,}5$ equivale a una clasificación aleatoria.

2.2.3.6 Desbalance de Clases en Clasificación Vial

En problemas de siniestralidad vial, el desbalance entre clases es una condición estructural: los accidentes con víctimas fatales o graves son naturalmente menos frecuentes que los leves o con daños materiales. Este desbalance introduce un sesgo sistemático en el MLP hacia las clases mayoritarias (He & Garcia, 2009). Las estrategias de tratamiento más utilizadas son:

A) SMOTE (*Synthetic Minority Oversampling Technique*).

Propuesto por Chawla, Bowyer, Hall y Kegelmeyer (2002), genera ejemplos sintéticos de las clases minoritarias interpolando entre ejemplos reales en el espacio de características:

$$\mathbf{x}_{\text{sint}} = \mathbf{x}_i + \lambda \left(\mathbf{x}_{nn} - \mathbf{x}_i\right), \qquad \lambda \sim \mathcal{U}(0,\ 1) \tag{2.28}$$

donde $\mathbf{x}_{nn}$ es un vecino más cercano de $\mathbf{x}_i$ perteneciente a la misma clase minoritaria.

B) Pesos de Clase (*Class Weights*).

Modifica la función de pérdida asignando mayor penalización a los errores en las clases minoritarias:

$$\mathscr{L}_{\text{pond}} = -\frac{1}{N}\sum_{i=1}^{N} w_{y^{(i)}} \sum_{k=1}^{K} y_{ik}\, \log \hat{y}_{ik} \tag{2.29}$$

donde $w_k \propto \dfrac{N}{K \cdot n_k}$ es el peso inversamente proporcional al tamaño de cada clase.

C) Submuestreo (*Undersampling*).

Reduce aleatoriamente los ejemplos de las clases mayoritarias hasta igualar la clase minoritaria, con el riesgo de perder información relevante (Lemaitre, Nogueira & Aridas, 2017).

2.2.4 Diseño geométrico de carreteras

El diseño geométrico de una carretera es el proceso técnico mediante el cual se determinan las características espaciales de la vía, definiendo su trazado en planta, perfil longitudinal y sección transversal. Estos parámetros se establecen en función de la velocidad de diseño el tipo de terreno el volumen de tráfico y la categoría funcional de la vía (MTC, DG-2018). El objetivo central del diseño geométrico es garantizar la seguridad, comodidad y economía del tránsito, minimizando la conflictividad entre los requerimientos del usuario y las limitaciones del terreno **(MTC, 2018)**.

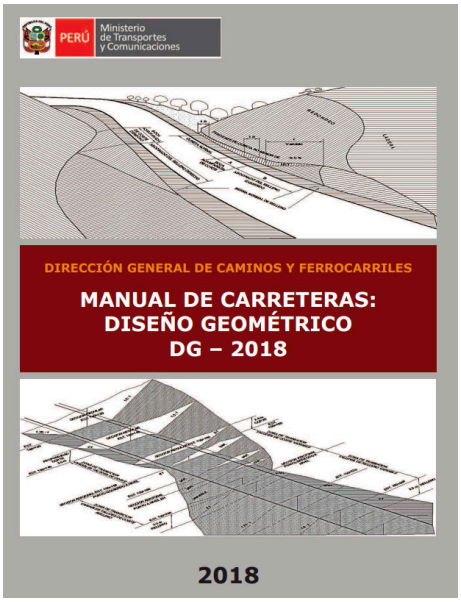


Figura 24 Manual de diseño geométrico del Perú (DG-2018).
Fuente: Ministerio de transportes y comunicaciones (MTC).

2.2.4.1 Alineamiento horizontal y elementos de curva

El alineamiento horizontal de una carretera está conformado por tramos rectos (tangentes) y curvas circulares, eventualmente complementadas con curvas de transición (espirales clothoide). Los parámetros geométricos son:

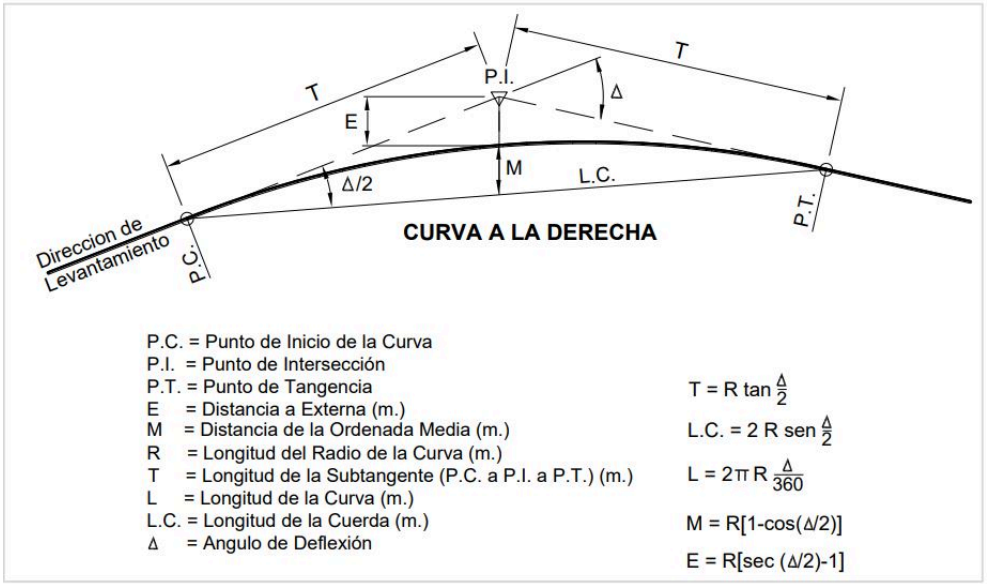


Figura 25 Elementos de la curva circular
Fuente: Manual de carreteras: Diseño geométrico DG 2018

- **Radio de curva (*R*):** distancia desde el centro de la curva hasta el eje de la calzada. Determina el grado de curvatura y condiciona directamente la velocidad segura de circulación y la fuerza centrífuga actuante sobre el vehículo. La DG-2018 es-

tablece radios mínimos en función de la velocidad de diseño y el peralte máximo adoptado.

- **Ángulo de deflexión (Δ):** ángulo central que define la magnitud del cambio de dirección en la curva. Valores elevados de ángulo de deflexión, especialmente combinados con radios pequeños, incrementan el riesgo de pérdida de control vehicular.
- **Longitud de curva (*L*):** longitud medida sobre el eje de la vía desde el principio de curva (PC) hasta el término de curva (TC). Curvas de longitud insuficiente generan problemas perceptivos en el conductor y reducen el tiempo de adaptación.
- **Longitud de subtangente (*T*):** distancia desde el vértice (PI) hasta el PC o TC. Interviene en la coordinación de alineamientos y en el cálculo de la longitud total de la curva.
- **Longitud de cuerda (*LC*):** distancia en línea recta entre el PC y el TC. Parámetro auxiliar para el replanteo topográfico de la curva.
- **Distancia externa (*E*):** distancia desde el PI hasta el punto medio de la curva, medida radialmente. Indica la desviación máxima del trazado respecto a la tangente.
- **Distancia de la ordenada media (*M*):** distancia desde el punto medio de la cuerda hasta el punto medio de la curva. Relevante para la determinación de la distancia de visibilidad en curvas.

2.2.4.2 Alineamiento vertical

El alineamiento vertical de una carretera define las variaciones de elevación a lo largo del eje vial mediante la sucesión de tramos de pendiente constante (rasante) enlazados por curvas verticales de transición. Según la DG-2018, el diseño del perfil longitudinal debe garantizar condiciones adecuadas de seguridad, comodidad y drenaje superficial, respetando los valores límite establecidos para cada categoría de carretera y tipo de orografía **(MTC, 2018)**.

2.2.4.3 Sección transversal

La sección transversal de la carretera define la geometría de la vía en un plano perpendicular al eje. Sus elementos principales son: la calzada (superficie de rodadura), las bermas, los taludes de corte y relleno, las cunetas y los elementos de drenaje. El ancho de calzada y la existencia de bermas son variables incluidas en la base de datos de la

presente investigación, reconocidas como factores que influyen en el comportamiento del conductor y en la probabilidad de accidente, especialmente en maniobras de emergencia. El peralte es la inclinación transversal de la calzada en las curvas, diseñado para contrarrestar la fuerza centrífuga actuante sobre el vehículo **(MTC, 2018)**.

2.2.5 Seguridad vial y accidentes de tránsito

La seguridad vial es el estado que permite asegurar la integridad física de los usuarios en una red de transporte. La siniestralidad vial, que se entiende como la frecuencia y severidad de un accidente de tránsito, depende de varias condiciones: el diseño de la vía, las condiciones climáticas, la señalización, y el comportamiento del conductor. **El Boletín de Siniestralidad Vial (MTC, 2024)** expone que Ayacucho mantiene elevados índices de siniestralidad , siendo las curvas cerradas, las pendientes pronunciadas y la nefasta visibilidad los motivos más comunes. Lo cual, además de evidenciarlo, nos propone una mejora del diseño geométrico de las vías vecinales de esta zona geográfica. Según **(AASHTO, 2018)**, la normativa geométrica adecuada, sumada a las tecnologías de simulación, permite prever escenarios de riesgo y proporcionar conclusiones a partir de la red de alternativas que se han diseñado.

2.2.6 Fuentes de datos para la investigación

2.2.6.1 Observatorio Nacional de Seguridad Vial (ONSV)

El Observatorio Nacional de Seguridad Vial (ONSV) es un órgano técnico adscrito al Ministerio de Transportes y Comunicaciones del Perú con una misión clara: recopilar, procesar, analizar y difundir la información estadística sobre siniestros de tránsito en todo el territorio nacional. No es una entidad más dentro del aparato estatal. Según el MTC (2018), el ONSV constituye la fuente oficial de datos de siniestralidad vial del Estado peruano, y su fortaleza radica en que integra información de múltiples actores institucionales: la Policía Nacional del Perú (PNP), la Superintendencia de Transporte Terrestre de Personas, Carga y Mercancías (SUTRAN), las Direcciones Regionales de Transporte y los establecimientos de salud del MINSA. Es decir, cruza miradas desde distintos frentes para construir un panorama lo más completo posible.

Figura 26 Análisis de datos de los accidentes de tránsito.

Fuente: Observatorio Nacional de Seguridad Vial (ONSV)

¿Qué registra exactamente su base de datos? Cada siniestro queda documentado con un nivel de detalle considerable: fecha y hora del evento, ubicación georreferenciada, departamento, provincia, distrito, vía y progresiva kilométrica, tipo de siniestro (choque, volcadura, despiste, atropello, caída de pasajero), causa probable, condiciones climáticas al momento del hecho, tipo de vehículo involucrado y, por supuesto, las consecuencias: número de fallecidos, lesionados y daños materiales. Cada registro cuenta una historia que no debería repetirse.

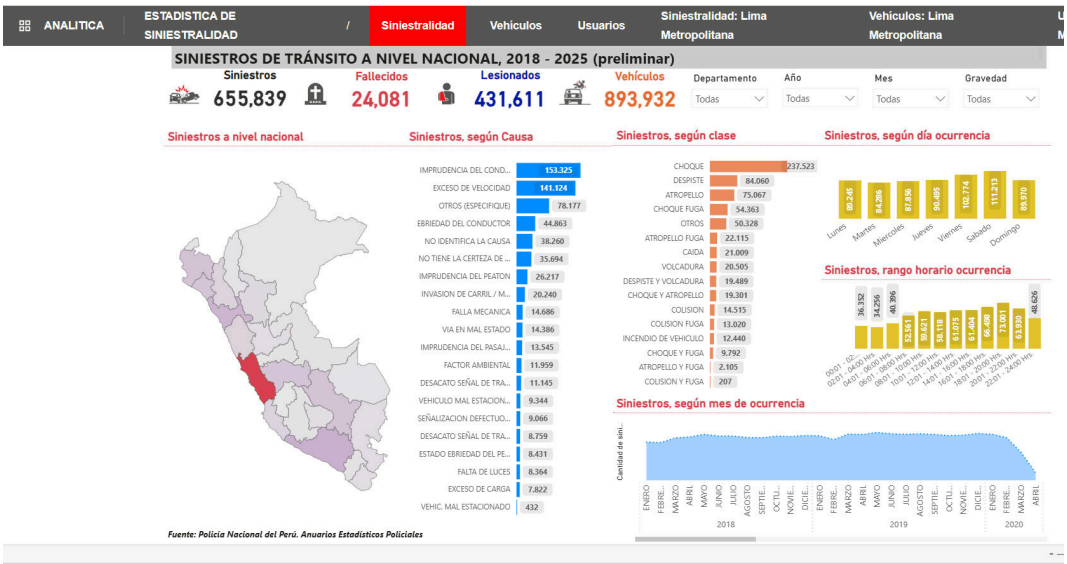


Figura 27 Registro de accidentes de tránsito a nivel nacional

Fuente: Observatorio Nacional de Seguridad Vial (ONSV)

En la presente investigación, los datos del ONSV correspondientes al período 2020–2025 constituyen la fuente primaria para la construcción de la base de datos con 202

registros que alimenta el Modelo RNA N.°1. Gracias a la georreferenciación de cada evento, fue posible vincular cada siniestro con las variables geométricas específicas del tramo donde ocurrió, conectando así la estadística con la geometría real de la vía.

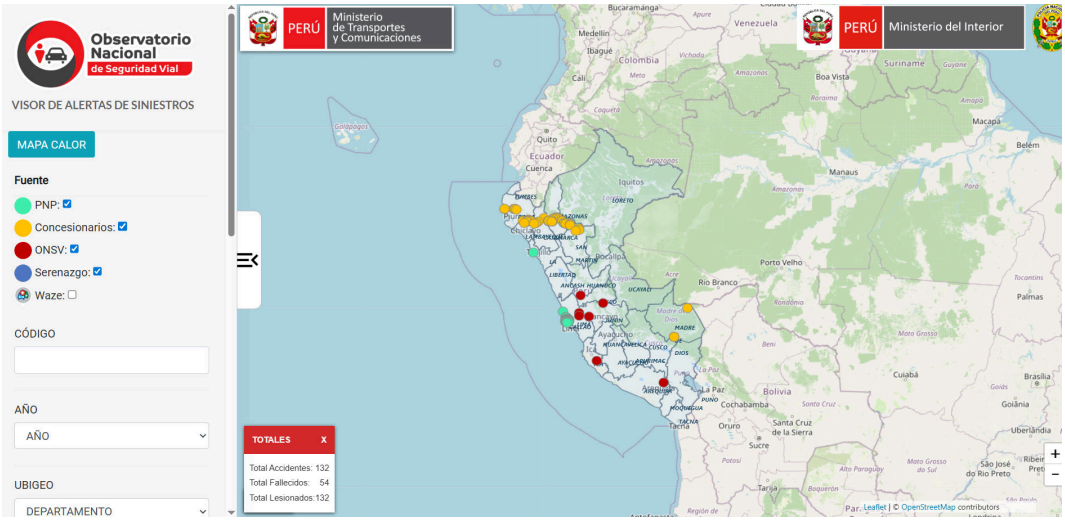


Figura 28 Ubicación de los accidentes de tránsito a nivel nacional

Fuente: Observatorio Nacional de Seguridad Vial (ONSV)

2.2.6.2 Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI)

El Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) es un organismo técnico especializado adscrito al Ministerio del Ambiente, responsable de la generación y difusión de información meteorológica, hidrológica y climática del territorio nacional. Según el SENAMHI (2024), la institución opera una red de estaciones meteorológicas distribuidas en todo el país, que registran las siguientes variables:

- Precipitación (mm)
- Temperatura (°C)
- Humedad relativa (%)
- Velocidad y dirección del viento (m/s)
- Presión atmosférica (hPa)
- Radiación solar (W/m^2)

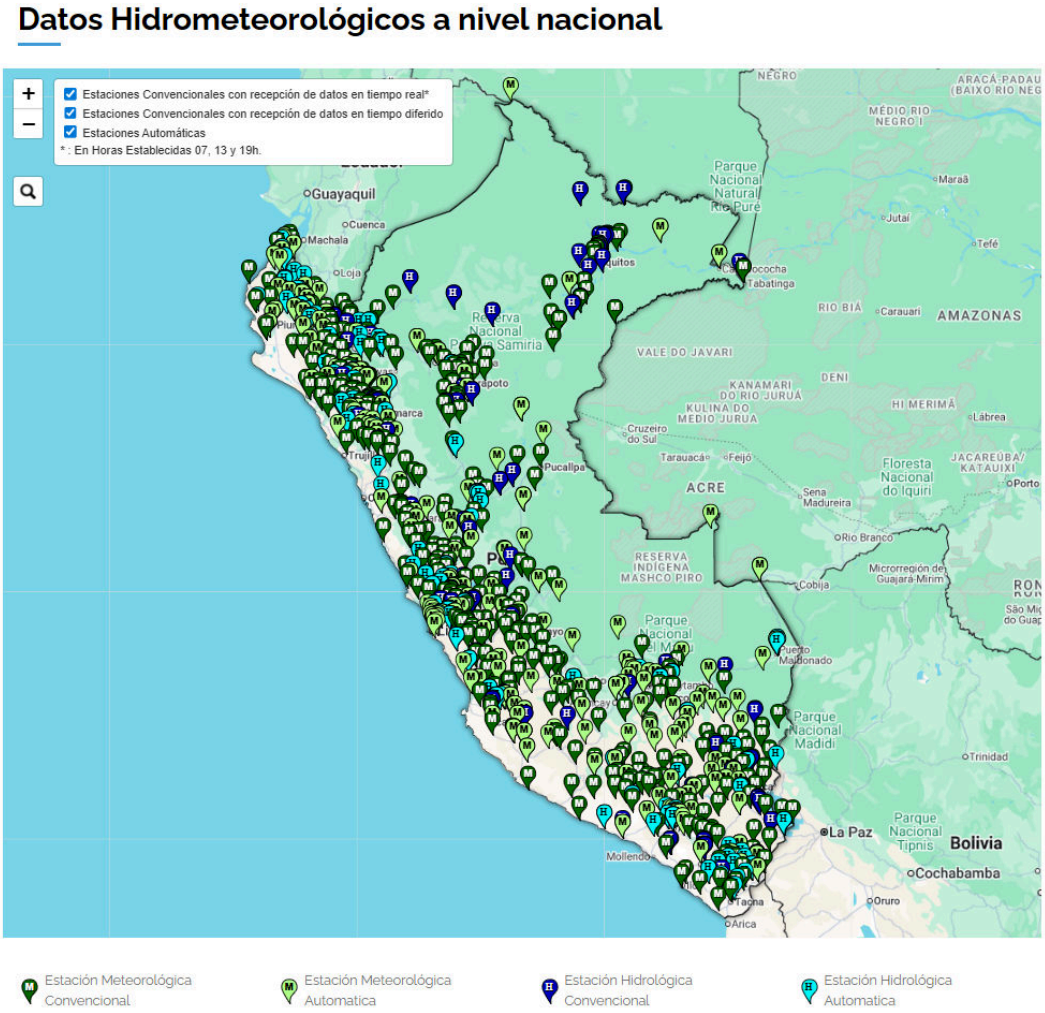


Figura 29 Estaciones metereológicas del SENAMHI
Fuente: Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI)

Estaciones meteorológicas de referencia. Para la presente investigación, los datos meteorológicos del SENAMHI se incorporaron como variables complementarias en el análisis exploratorio de datos del Modelo RNA N.°1, permitiendo evaluar la influencia de las condiciones climáticas en la siniestralidad de la PE-28A. En la Tabla 1 se presentan las estaciones de referencia para el corredor de estudio.

Tabla 1 *Estaciones meteorológicas del SENAMHI de referencia para la Vía Los Libertadores (PE-28A).*

Estación	Departamento	Progresiva aprox.
Hacienda Bernales	Ica	0 a 44km
Huancano	Ica	44 a 96km
Tambo	Huancavelica	96 a 169km
Santa Cruz de Hospicio	Ayacucho	169 a 274km
Santa Pedro de Cachi	Ayacucho	274 a 333.90km

Nota. Fuente: elaboración propia a partir de datos del SENAMHI (2024).

Estas estaciones proporcionan datos de precipitación acumulada mensual, temperatura mínima y máxima diaria, y frecuencia de eventos de neblina, variables que se vincularon a los registros de siniestros mediante correspondencia temporal y espacial.

Importancia de las variables meteorológicas en la siniestralidad vial. La incorporación de datos meteorológicos es fundamental porque las condiciones climáticas adversas lluvia, neblina, helada reducen significativamente la visibilidad y los coeficientes de fricción del pavimento, incrementando el riesgo de siniestros en carreteras de montaña. la combinación de geometría inadecuada y condiciones climáticas adversas multiplica exponencialmente la probabilidad de accidentes, especialmente en curvas con peralte insuficiente y tramos sin drenaje adecuado.

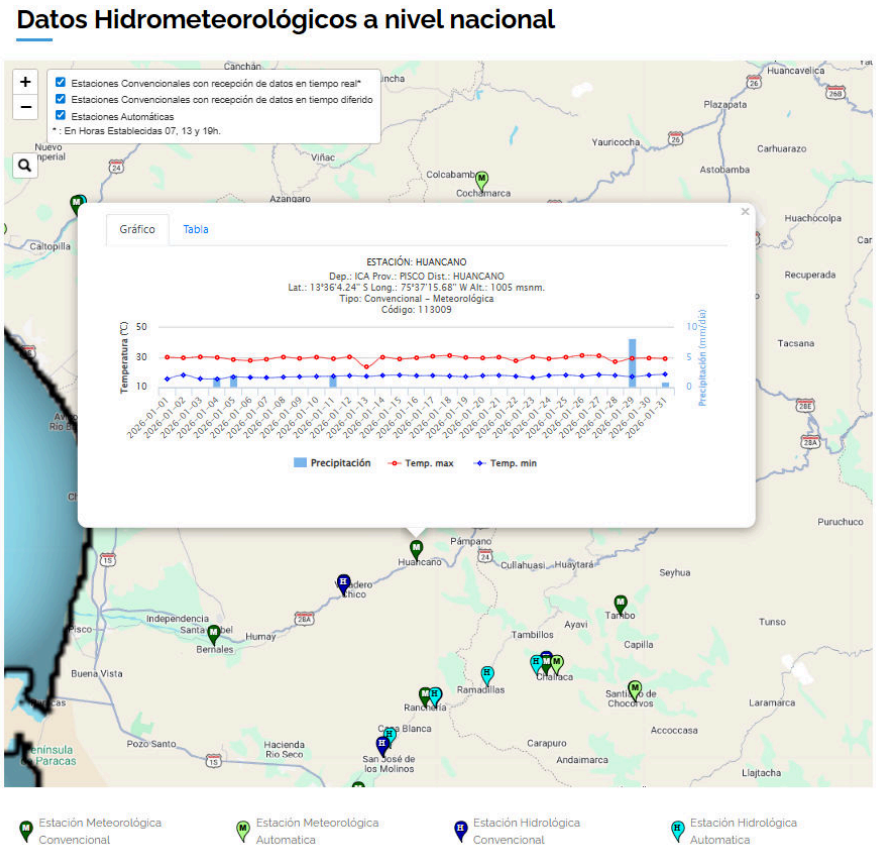


Figura 30 Registro de las condiciones climáticas de Estaciones metereológicas.
Fuente: Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI)

2.2.6.3 Recopilación de datos in situ o en campo

La recopilación de datos *in situ* constituye una etapa fundamental en toda investigación de seguridad vial, dado que permite obtener información directa de las condiciones reales de operación de la carretera que no siempre está disponible en bases de datos institucionales. Según (Garber & Hoel, 2015), "los estudios de campo en ingeniería

de transporte proporcionan la base empírica sobre la cual se construyen los modelos predictivos y se validan las hipótesis de diseño" (p. 78). En la presente investigación, la recopilación de datos *in situ* se realizó mediante los siguientes procedimientos complementarios:

a) Inspección visual y georreferenciación. Se realizaron recorridos vehiculares a lo largo de los 333.90 km de la PE-28A para identificar, fotografiar y georreferenciar tramos con deficiencias geométricas: curvas cerradas sin señalización, peralte insuficiente, bermas deterioradas y visibilidad comprometida. La georreferenciación se realizó mediante GPS y Google Earth Pro en coordenadas UTM (zona 18L).

Figura 31 Obtención de datos en campo de la Via PE-28A.
Fuente: Fotografia propia.

b) Registro fotográfico y videográfico. Se documentaron fotográficamente cinco aspectos clave del corredor:

- Estado del pavimento (baches, fisuras, material suelto).
- Geometría visible (curvas cerradas, pendientes, deficiencias de peralte).
- Señalización y dispositivos de seguridad (guardavías, delineadores, demarcación).
- Sitios donde ocurrieron los accidentes más graves del período 2020-2025.

Figura 32 Estado actual de la carretera PE-28A.
Fuente:Fotografía propia.

Figura 33 Estado actual de la carretera PE-28A.
Fuente:Fotografía propia.

2.2.7 Estudio de accidentabilidad

Con el fin de poder hacer una correcta referencia sobre la accidentalidad vial, hay que considerar al menos tres factores con respecto a los cuales, por lo menos uno de sus elementos de dichos factores debe estar presente en la exposición del análisis de los hechos que se está considerando y que son: el humano (conductor) , el dinámico (vehículo) y el estático (infraestructura de la vía). La elevada accidentalidad que existe en las vías, por razones como las expuestas, nos hace necesario un acucioso estudio que permita identificar la incidencia de cada uno de los factores que la componen. En los informes que suelen elaborarse, si bien en general admiten que la causa principal puede ser el conductor, es raro que se analice si dicho comportamiento erróneo del conductor proviene de la falta de un diseño geométrico adecuado de la vía, pudiendo ser este un correcto explicado que empecé a incidir en el comportamiento del conductor en muchas circunstancias **(M.T.C, Manual de seguridad vial, 2017)**.

2.2.8 Normas técnicas internacionales de diseño geométrico de vías

Las normas de diseño geométrico de las vías constituyen pautas técnicas que determinan los principios utilizados para la planificación, el diseño y la ejecución de las infraestructuras de carreteras, a fin de poder aportar las mejores condiciones de seguridad, funcionalidad, operatividad y confort del tránsito vehicular.

A nivel internacional se ha reconocido que aquellos países cuyos índices de mortalidad vial son inferiores a los otros países son aquellos que aplican normas de diseño geométrico elevados y que dan una importancia especial a la seguridad. En el *"Informe Mundial de Seguridad Vial 2024 del ITF"* **(International Transport Forum, 2024)**, sostiene que los cinco países que en 2023 se encontraban con el índice de mortalidad más bajo por muertes derivadas de la circulación son Noruega, Islandia, Suecia, Reino Unido y Japón, todos ellos con un índice inferior a 3 muertes por cada 100.000 habitantes.

En ese contexto, el presente estudio de tesis realizará una comparación y un análisis mediante redes neuronales entre las normas de diseño geométrico actualmente en uso en los cinco países con menor tasa de mortalidad vial **(Noruega, Islandia, Suecia,**

Reino Unido y Japón), con el objetivo de identificar criterios técnicos transferibles con el potencial de ser aplicados en los países con mayores niveles de siniestralidad vial, como el caso de Perú con la normativa **DG-2018**. Este análisis servirá de base para la formulación de propuestas de mejora de infraestructuras viales con la seguridad proactiva como elemento central.

2.2.8.1 Perú ("Manual de Carreteras - Diseño Geométrico DG-2018"):

El Manual de Carreteras "Diseño Geométrico DG-2018" , fue aprobado por el Reglamento Nacional de Gestión de Infraestructura Vial (DS N° 034-2008-MTC) constituyendo un documento técnico normativo de carácter obligatorio a nivel nacional; dicho manual es de aplicación para los organismos encargados de la gestión de infraestructura vial de los tres niveles de gobierno: nacional, regional y local **(MTC, 2018)**. El manual compila y sistematiza las técnicas, las fórmulas y los procedimientos necesarios para el diseño geométrico de la infraestructura vial, los cuales consideranán parámetros técnicos adecu á ndolo a la categoría y nivel de servicio del proyecto, y se encuadra en el resto de la normativa vigente que establece la gestión de infraestructura vial **(MTC, 2018)**.

2.2.8.2 Noruega ("N100, N200, N300, N400 y V712"):

Noruega es reconocido internacionalmente por su escaso índice de muertes por siniestros de tráfico, el cual alcanza aproximadamente a 2 muertes cada 100,000 habitantes en el 2023, según el *"Informe Mundial de Seguridad Vial 2024 del ITF"* **(International Transport Forum, 2024)**. Su éxito es el resultado de la implementación del enfoque denominado **"Visión Cero"** el cual considera inaceptables las muertes acumuladas por el tráfico. En el país se encuentran aplicadas normas de diseño geométrico moderno como la N100, que acentúa la seguridad a través de carreteras, zonas de velocidad reducida y la separación física de los carriles de circulación, entre otros. Por otra parte, se implementan controles de velocidad y alcoholemia bien establecidos, elevadísimos niveles de educación vial, y una infraestructura que se da para evitar errores humanos. Noruega combina todos estos beneficios con un sistema de atención médica muy eficiente y el uso de datos a alto nivel para prevenir los siniestros, configurándose así como un referente mundial en relación a la seguridad vial.

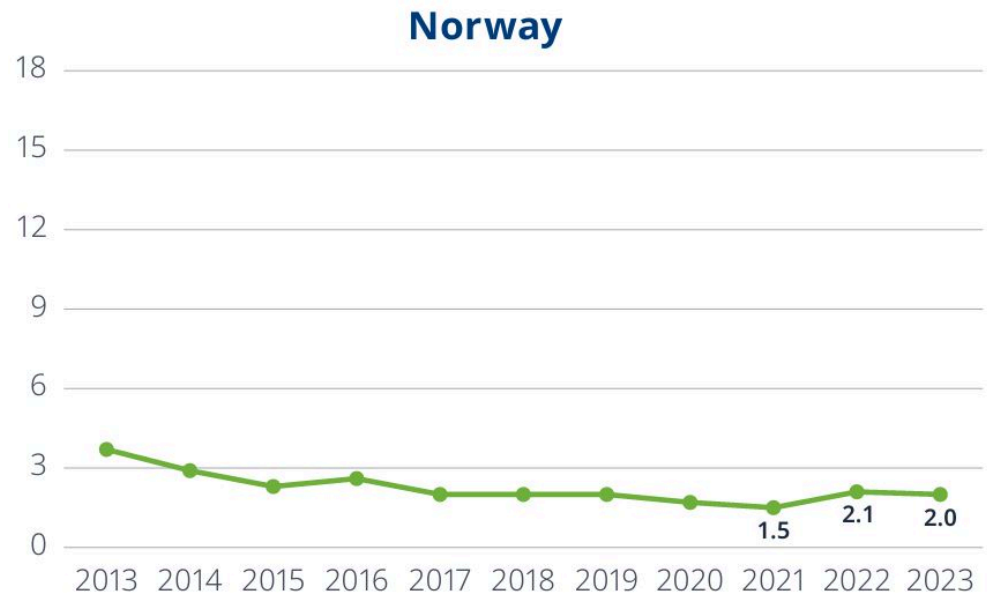


Figura 34 Muertes en carretera por cada 100.000 habitantes de Noruega.

Fuente: Road Safety Annual Report 2024 - International Transport Forum.

En la actualidad Noruega hace uso de un conjunto de normas técnicas que han ido elaborando desde Statens vegvesen (la Administración de Carreteras). La norma técnica principal sería la **"N100 - Veg og gateutforming"** (Diseño de vías y calles), donde se establecen los criterios geométricos a respetar en la construcción de carreteras urbanas y rurales seguras. Esta norma se ve complementada con otras en la misma línea como la **"N200-Geometrisk utforming av vegkryss"** (intersecciones), **"N300-Vegens linjeføring"** (alineación horizontal y vertical), **"N400-Vegens tverrprofil"** (perfil transversal), y el manual Håndbok **"V712-Vegnormaler"** (Manual general de diseño y construcción de carreteras), en la que se reúnen los estándares de diseño y construcción vial. Todo el conjunto de normas incorpora el enfoque **"Visión Cero"**, priorizando la seguridad a costa de la tolerancia al error humano en el diseño de la carretera **(Statens vegvesen, 2022)**.

2.2.8.3 Islandia ("Veglagna – Handbók um hönnun vega"):

En Islandia se ha producido un avance continuo, siendo uno de los países más seguros del mundo en la materia, con una cifra de fallecimientos por accidentes de tráfico en el rango de aprox. 2.1 muertes por cada 100,000 personas en 2023, según el *"Informe Mundial de Seguridad Vial 2024 del ITF"* **(International Transport Forum, 2024)**. Se trata de un resultado que se vincula directamente con el enfoque del país en este ámbito, integrando la seguridad vial de forma global en el país y que se encuentra

en consonancia con el marco de la prevención y el diseño de infraestructuras seguras, íntimamente relacionado con los principios de "Visión Cero", tal como se expone en el documento "CERO muertes. Los Diez Asuntos Clave de la Seguridad Vial", donde se expone la posición del país, que puede ser profundizada en el trabajo de que se hace eco el país sobre la materia.

Figura 35 Muertes en carretera por cada 100.000 habitantes de Islandia.

Fuente: Road Safety Annual Report 2024 - International Transport Forum.

Islandia cuenta con una serie de normas técnicas de carácter oficial para el diseño geométrico de las infraestructuras de carreteras, elaborados por la propia Administración de Carreteras y Costas (Vegagerðin). Estas normas están adaptadas a las especificidades de las condiciones climáticas y geográficas del país, como son los terrenos montañosos, la nieve, el hielo y la baja densidad de habitantes, lo que requiere de una infraestructura con un buen y seguro nivel de diseño. El documento maestro es el **"Veglagna – Handbók um hönnun vega" ("Construcción de carreteras – Manual de diseño de carreteras")** en el que se establecen los criterios de diseño geométrico de las carreteras, trayecto horizontal y vertical, radios de curvatura, pendiente, visibilidad y diseño vestido del perfil transversal **(Vegagerðin, 2022)**.

Además Islandia aplica límites de velocidad establecidos, una política de tolerancia cero con respecto a la conducción bajo la influencia del alcohol, el uso del cinturón de seguridad de forma obligatoria y la fiscalización automatizada en tramos geográficos críticos. El conocimiento vial se encuentra dentro del contenido escolar, la formación práctica para la obtención del permiso de conducir incluye la necesidad de practicar en

condiciones adversas.

2.2.8.4 Suecia ("Vägars och gators utformning (VGU)"):

Suecia es conocida a nivel internacional como el país precursor de la perspectiva "Visión cero", política que se instauró en el año 1997 y que fija como propósito que no se tolera una muerte o lesión grave en el tráfico. Gracias a esta política, Suecia ha conseguido conservar con constancia una de las más bajas tasas de mortalidad en las carreteras del mundo, registrando en 2023 en torno a 2.2 decesos por 100,000 habitantes, según el *"Informe Mundial de Seguridad Vial 2024 del ITF"* **(International Transport Forum, 2024)**.

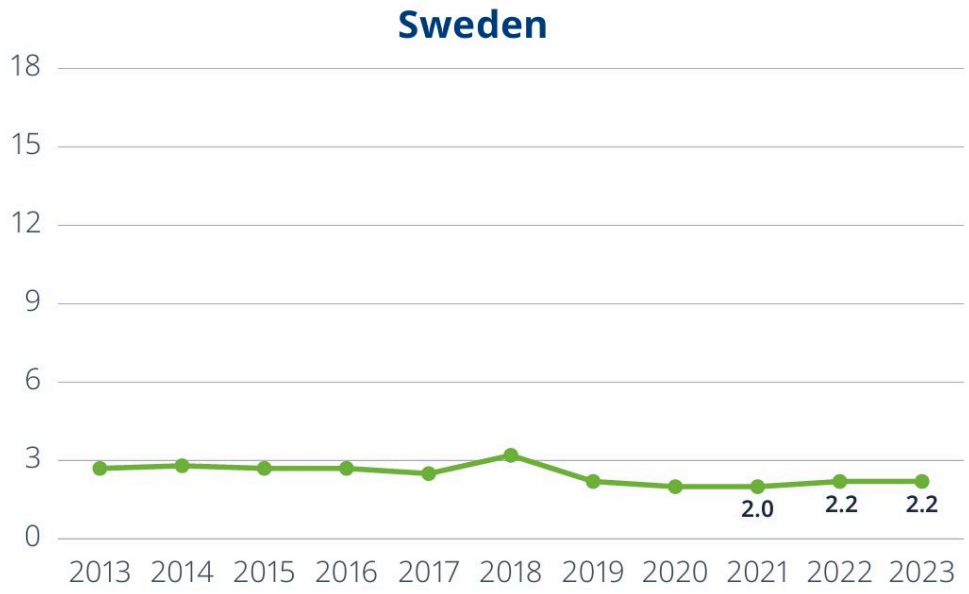


Figura 36 Muertes en carretera por cada 100.000 habitantes de Suecia.

Fuente: Road Safety Annual Report 2024 - International Transport Forum.

Una de las claves de su éxito radica en el diseño de la infraestructura centrado en la seguridad de todos los usuarios. Suecia hace uso de las especificaciones técnicas del manual **"Vägars och gators utformning(VGU)" ("Diseño de carreteras y calles")**, el cual fue emitido por la Administración Sueca de Transporte (Trafikverket), que establece criterios geométricos estrictos de las carreteras urbanas, al igual que de las rurales. Este documento determina valores tales como anchos de calzada, radios de curvatura, pendientes máximas, diseño de intersecciones, visibilidad y separación de usuarios vulnerables -ciclistas y peatones de forma segura.

Aparte de un diseño geométrico, Suecia refuerza la política de carretera con unos límites de velocidad muy estrictos (30 km/h como el límite en zonas urbanas), una

incoherente y disfuncional política de controles automáticos, una pobre tolerancia al alcohol (una sorpresa en un país como el sueco —0.2 g/L!), y programas de educación desde la escuela, una sólida estructura de monitorización de los siniestros por medio de la actuación con datos, espacios públicos que favorezcan el uso de rotondas en lugar de intersecciones semaforizadas y que hayan reducido la gravedad de los choques, etc., **(Trafikverket, 2020)**.

La aproximación sueca combina la ingeniería de las carreteras, la legislación, la instrucción y la fiscalización, para establecer un sistema de tránsito concebido para aceptar el error humano asumiendo mínimamente sus consecuencias. Esta visión holística ha ubicado a Suecia como el modelo mundial global en materia de seguridad de tránsito y ordenación urbana sostenible.

2.2.8.5 Reino Unido ("CD 109 – Highway Link Design"):

El Reino Unido ostenta uno de los índices de siniestralidad más impecables a nivel europeo y, de hecho, según el *"Informe Mundial de Seguridad Vial 2024 del ITF"* **(International Transport Forum, 2024)** la tasa de muertes se estimo en aproximadamente 2,5 muertes por cada 100.000 habitantes, una cifra que se encuentra entre las más notorias incluso mundialmente. Todo ello debido a un nivel de exigencia elevado basado en la planificación más amplia, en la regulación más estricta, en la mejora incesante de la seguridad de sus infraestructuras.

El diseño geométrico de las carreteras en el país se apoya en un conjunto de normas técnicas, conocidas como**"Design Manual for Roads and Bridges (DMRB)"**, elaborado por el National Highways, el cual se trabaja en los países de Inglaterra, Escocia, Gales e Irlanda del Norte. Se establecen parámetros para el trazado horizontal y vertical de las carreteras, secciones transversales, intersecciones, visibilidad, seguridad en las curvas y dispositivos de contención.

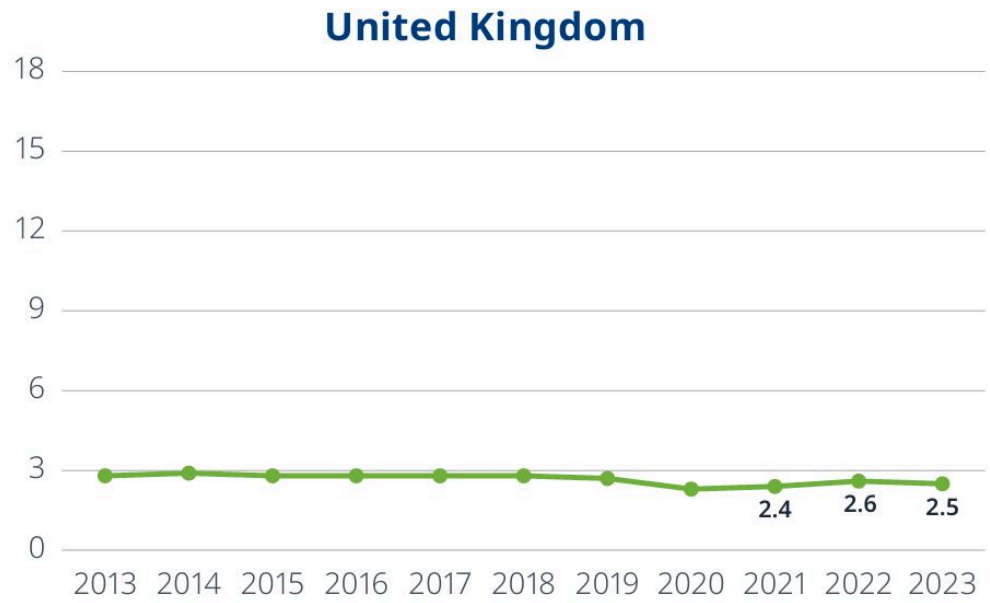


Figura 37 Muertes en carretera por cada 100.000 habitantes de Reino Unido.

Fuente: Road Safety Annual Report 2024 - International Transport Forum.

Entre los documentos del DMRB más relevantes laceramos el **"CD 109 - Highway Link Design"**, que establece la figura geométrica de las carreteras principales y las de alta velocidad, así como el del **"CD 116 - Geometric design of roundabouts"**, el cual regula el diseño de las rotondas, como un medio eficiente a la par que seguro de construcción de las intersecciones. El Reino Unido, además de desarrollar la infraestructura, implementa medidas complementarias complementarias como diferenciación de los límites de velocidad según el tipo de vía, educación continua mediante campañas educativas, alto cumplimiento con cinturón de seguridad **(National Highways, 2020)**.

2.2.8.6 Japon ("Design Standards for Road Structure - Dōro Kōzō Rei"):

Japón se encuentra entre los países de Asia que tienen mejor comportamiento en lo que se refiere a la seguridad vial, destacando que para el año 2023 presenta una tasa de mortalidad de unas 2,6 muertes por cada 100.000 personas, según el *"Informe Mundial de Seguridad Vial 2024 del ITF"* **(International Transport Forum, 2024)**. El dato reflejado se sostiene a lo largo de las últimas décadas gracias a unas estrictas políticas de tránsito, a los avances tecnológicos y a un fuerte clima institucional favorable a la reducción del número de siniestros viales.

Las características geométricas de las infraestructuras viales de Japón se encuentran

reguladas por las normas técnicas que son emanadas por el Ministerio de Tierras, Infraestructura, Transporte y Turismo (MLIT), destacando entre ellas el manual conocido como **"Design Standards for Road Structure"**, conocido en Japón bajo la denominación de **"Dōro Kōzō Rei"**, el cual fija los parámetros para el trazado geométrico de las carreteras, tales como alineación horizontal y vertical, radios mínimos de curvatura, anchos de calzada, pendientes, visibilidad y diseño de intersecciones, aplicables tanto en entornos todo urbano como en los de zonas rurales, en entornos donde se ve el modelo adecuadamente adaptado a las condiciones topográficas complejas del país como aquellas montañosas y las de regiones sísmicamente activas **(Ministry of Land, Infrastructure, Transport and Tourism, 2015)**.

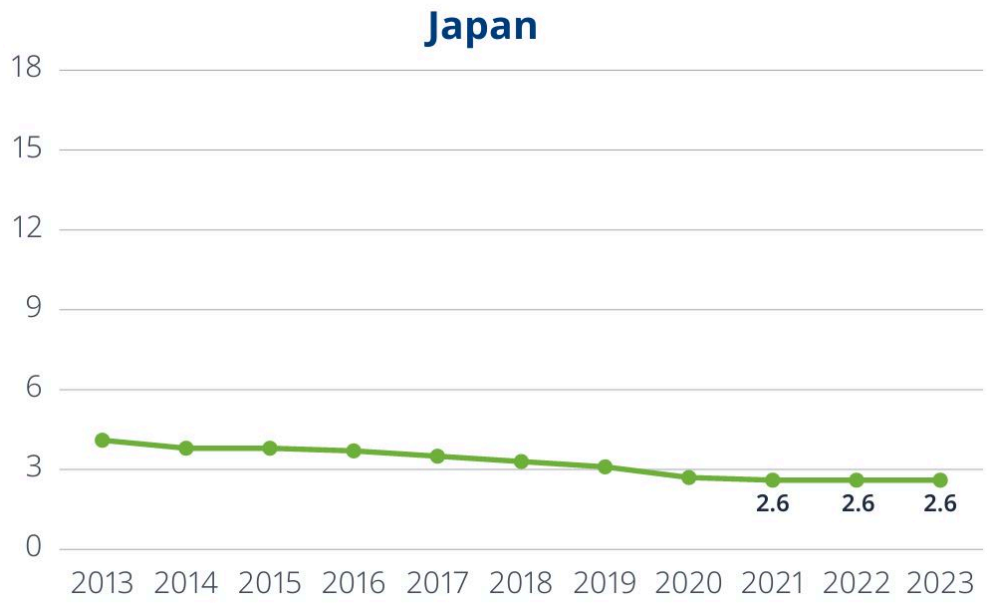


Figura 38 Muertes en carretera por cada 100.000 habitantes de Japon.

Fuente: Road Safety Annual Report 2024 - International Transport Forum.

La conjugación de normas técnicas precisas, un estricto control, las nuevas tecnologías (por ejemplo, el uso de vehículos con asistencia automatizada) y una cultura ciudadana que pone el énfasis en la prevención han contribuido al hecho de que Japón se sitúe como un referente regional en cuanto a la seguridad vial, ya que operando de este modo se reduce el efecto de los errores humanos y se prioriza la vida.

2.2.9 Herramientas computacionales aplicadas

2.2.9.1 Python y librerias de Machine learning

El desarrollo, entrenamiento y evaluación de los dos modelos de Redes Neuronales Artificiales se realizó íntegramente en JupyterLab, distribuido a través de Anaconda Navigator, una plataforma que pone a disposición un ecosistema completo de herramientas científicas para Python. La elección no fue arbitraria. JupyterLab tiene algo que lo hace especialmente valioso para una investigación como esta: permite construir notebooks interactivos donde el código ejecutable, las visualizaciones gráficas, las ecuaciones matemáticas y el texto explicativo conviven en un mismo documento. Todo queda registrado, todo es reproducible. Y eso, en ciencia, lo es todo porque garantiza que cualquier investigador pueda seguir el mismo camino y llegar a los mismos resultados. En cuanto al lenguaje de programación, se optó por Python. No sorprende. Python se ha ganado con creces el lugar de estándar de facto en el mundo del aprendizaje automático y la inteligencia artificial. Su sintaxis es clara, sus librerías especializadas son enormes en cantidad y calidad, y cuenta con una comunidad de soporte activa que respalda cada paso del desarrollo. Las librerías empleadas para construir los modelos se organizan en cinco categorías funcionales:

a) Modelado y aprendizaje automático: Se empleó **Scikit-learn**, la librería de referencia para aprendizaje automático en Python, utilizando los siguientes módulos:

- **MLPClassifier:** Implementación del Perceptrón Multicapa.
- **StandardScaler:** Normalización Z-Score de las variables de entrada, garantizando media cero y desviación estándar unitaria.
- **Train_test_split:** División estratificada del dataset en entrenamiento (70 %), validación (20 %) y prueba (10 %).
- **StratifiedKFold:** Validación cruzada estratificada con $K = 5$ *folds*.

b) Métricas de evaluación: Del módulo sklearn.metrics se emplearon:

- **Accuracy_score:** Exactitud global del modelo.
- **F1_score:** F1-Score con promedios *Macro* y *Weighted*.
- **Roc_auc_score:** AUC-ROC con esquema *One-vs-Rest* para clasificación multiclase.
- **Confusion_matrix:** Matriz de confusión absoluta y normalizada.

- **Classification_report:** Reporte detallado de precisión, *recall* y F1-Score por clase.
- **Brier_score_loss:** Calibración probabilística del modelo.
- **Permutation_importance:** Importancia de variables por permutación.

c) Manipulación y estructuración de datos:

- **Pandas:** Lectura, limpieza, filtrado y estructuración de los datasets en formato `DataFrame`.
- **NumPy:** Operaciones de álgebra lineal, cálculos matriciales y manipulación de *Arrays* multidimensionales.

d) Visualización de resultados:

- **Matplotlib:** Motor gráfico base para la generación de figuras de alta resolución (150–200 dpi).
- **Seaborn:** Visualizaciones estadísticas avanzadas como *heatmaps* de correlación, *boxplots* comparativos, curvas KDE y *scatter plots* con densidad.

e) Manejo de desbalance de clases: Se empleó la librería **Imbalanced-learn**, extensión de Scikit-learn especializada en técnicas de remuestreo para datasets desbalanceados:

- **SMOTE (*Synthetic Minority Over-sampling Technique*)**: Generación de muestras sintéticas.
- **RandomOverSampler:** Técnica alternativa de validación para comparar con los resultados de SMOTE.

2.2.9.2 Autocad Civil 3D aplicado al diseño de carreteras

AutoCAD Civil 3D es un software de cálculo y diseño altamente eficaz, el cual se utiliza en proyectos de diseño de sitios, urbanismo, carreteras, movimientos de tierra, cálculos topográficos, replanteos, etc. Concebido para ingenieros civiles, delineantes, diseñadores, técnicos, Civil 3D permite hacer proyectos de infraestructura vial, urbanística e hidráulica, manteniendo la coordinación de los equipos, probando alternativas de diseño, analizando el rendimiento de los proyectos y generando documentación técnica de alta calidad y coherencia , todo ello en el entorno de AutoCAD, que les resulta tan familiar. La principal característica de este programa es que Civil 3D, creado por Autodesk, vincula todos los componentes del diseño, de forma que los objetos están relacionados entre ellos, de modo que cualquier tipo de alteración efectuada se reflejará automáticamente en el modelo obteniéndose perfiles, secciones transversales,

superficies, tablas de datos y otros elementos relacionados, lo que aumenta de forma considerable la eficacia y la precisión del diseño.

2.3 Marco Conceptual

En la presente sección se definen y contextualizan los términos técnicos fundamentales empleados a lo largo de la investigación. A diferencia de las bases teóricas, donde se desarrollan en profundidad las teorías y modelos que sustentan el estudio, el marco conceptual tiene como propósito establecer con claridad el significado preciso de cada concepto clave, delimitando su alcance dentro del contexto de la predicción de accidentes de tránsito mediante redes neuronales artificiales y la optimización del diseño geométrico de la Vía Los Libertadores (PE-28A).

2.3.1 Fundamentos de las Redes Neuronales Artificiales y aprendizaje automático.

2.3.1.1 Inteligencia artificial y aprendizaje automático

La **Inteligencia Artificial (IA)** está dedicada al desarrollo de sistemas computacionales capaces de realizar tareas que normalmente requieren inteligencia humana. Dentro de este campo, el **Aprendizaje Automático** (*Machine Learning*) constituye un subcampo especializado; al respecto, Bishop (2006) señala que el aprendizaje automático "estudia algoritmos y modelos estadísticos que permiten a los sistemas mejorar automáticamente su rendimiento en una tarea determinada a partir de la experiencia". Un nivel más especializado lo representa el **Aprendizaje Profundo** (*Deep Learning*); según Goodfellow et al. (2016), este "permite a los sistemas computacionales aprender representaciones jerárquicas de los datos a través de redes neuronales con múltiples capas ocultas" desde características simples en las primeras capas hasta patrones complejos en las capas más profundas.

En cuanto a los paradigmas de aprendizaje, se distinguen dos enfoques principales según la disponibilidad de etiquetas en los datos. El **Aprendizaje Supervisado** es aquel en el que el modelo se entrena con datos etiquetados; de acuerdo con Bishop (2006), este paradigma implica "el uso de pares de entrada y salida conocidos, de modo que el algoritmo aprende a predecir la salida correspondiente a nuevas entradas no vis-

tas previamente". Por otro lado, el **Aprendizaje No Supervisado** refiere a un enfoque distinto; tal como indican Hastie, Tibshirani y Friedman (2009), en este paradigma "el modelo debe identificar por sí mismo patrones, agrupaciones o estructuras latentes en conjuntos de datos que carecen de etiquetas predefinidas", siendo de utilidad para tareas de segmentación y descubrimiento de conocimiento. En la presente investigación se emplea el aprendizaje supervisado, dado que los conjuntos de datos utilizados cuentan con etiquetas de clase previamente definidas.

2.3.1.2 Estructura y componentes de las redes neuronales artificiales

Una **Red Neuronal Artificial (RNA)** es un modelo computacional de procesamiento de información inspirado en la estructura y funcionamiento del sistema nervioso biológico compuesto por unidades de procesamiento denominadas neuronas artificiales, organizadas en capas interconectadas mediante pesos sinápticos que se ajustan iterativamente durante el proceso de aprendizaje para capturar patrones y relaciones en los datos (Haykin, 2009).

La unidad fundamental de toda RNA es la **Neurona Artificial**, que recibe un conjunto de entradas numéricas y las pondera mediante sus respectivos **Pesos Sinápticos** valores numéricos que determinan la importancia relativa de cada entrada en el cómputo de la neurona receptora y que se ajustan durante el entrenamiento (Haykin, 2009), les aplica una función de agregación (generalmente una suma ponderada), añade un término adicional denominado **Sesgo** (*Bias*) que permite desplazar horizontalmente la función de activación proporcionando un grado adicional al modelo (Bishop, 2006) y finalmente aplica una función de activación no lineal para producir una salida.

El **Perceptrón Multicapa (MLP)**, también denominado *Multilayer Perceptron*, es la arquitectura de red neuronal *feedforward* más ampliamente utilizada en problemas de clasificación y regresión en ingeniería. Está compuesto por una capa de entrada y una o más capas ocultas donde se realiza el procesamiento no lineal de la información, y una capa de salida que produce la predicción final, manteniendo conectividad total entre todas las neuronas de capas adyacentes (Bishop, 2006).

2.3.1.3 Funciones de activación

Las **Funciones de Activación** son funciones matemáticas no lineales aplicadas a la salida de cada neurona que introducen la no linealidad imprescindible para que la red

pueda modelar relaciones complejas entre las variables de entrada y la variable objetivo; sin esta no linealidad, una red de múltiples capas se reduciría algebraicamente a una simple transformación lineal, independientemente de su profundidad (Goodfellow et al., 2016).

La función **ReLU** (*Rectified Linear Unit*), se ha convertido en el estándar para capas ocultas de redes gracias a que resuelve el problema del gradiente desvaneciente que afectaba a funciones anteriores como la sigmoidea, y porque acelera la convergencia del entrenamiento al mantener un gradiente constante para valores positivos (Nair & Hinton, 2010). Por su parte, la función **Softmax** transforma un vector de valores reales (logits) en un vector de probabilidades normalizadas que suman 1.0, constituyendo la función de activación natural para la capa de salida en problemas de clasificación multiclase, ya que permite interpretar cada salida como la probabilidad de pertenencia a una clase (Bishop, 2006).

2.3.1.4 Entrenamiento y optimización

El proceso de aprendizaje de una RNA se fundamenta en el algoritmo de **Retropropagación** (*Backpropagation*), que calcula el gradiente de la función de pérdida respecto a cada peso de la red mediante la aplicación recursiva de la regla de la cadena del cálculo diferencial, propagando el error desde la capa de salida hacia las capas de entrada para determinar la dirección y magnitud del ajuste necesario en cada conexión sináptica (Rumelhart, Hinton & Williams, 1986). Este gradiente es utilizado por algoritmos de optimización como el **Descenso del Gradiente**, que actualiza iterativamente los parámetros del modelo en la dirección opuesta al gradiente de la función de pérdida con el objetivo de minimizar progresivamente el error de predicción (Goodfellow et al., 2016).

En esta investigación se emplea el **Optimizador Adam** (*Adaptive Moment Estimation*), un algoritmo avanzado que combina las ventajas de *Momentum* que acumula exponencialmente gradientes pasados para acelerar la convergencia en direcciones consistentes y *RMSProp* que adapta la tasa de aprendizaje individualmente para cada parámetro según la magnitud histórica de sus gradientes, logrando convergencia eficiente incluso con poca sintonización de hiperparámetros (Kingma & Ba, 2015).

La **Función de Pérdida** (*Loss Function*) cuantifica numéricamente la discrepancia entre las predicciones del modelo y los valores reales observados, y su minimización

constituye el objetivo central de todo el proceso de entrenamiento; para problemas de clasificación multiclase como los abordados en esta investigación, se emplea la entropía cruzada categórica como función de pérdida (Goodfellow et al., 2016). Una **Época** (*Epoch*) corresponde a una pasada completa del algoritmo de entrenamiento por la totalidad del dataset, de modo que el número total de épocas determina cuántas veces el modelo procesa todos los datos disponibles durante el aprendizaje (Géron, 2019).

2.3.1.5 Regularización y prevención del sobreajuste

El **Sobreajuste** (*Overfitting*) es un fenómeno indeseable en el que el modelo memoriza los datos de entrenamiento incluyendo su ruido estadístico, lo que produce un rendimiento artificialmente alto en los datos de entrenamiento pero un rendimiento significativamente inferior en datos no vistos, comprometiendo la capacidad de generalización del modelo (Goodfellow et al., 2016). Para prevenirlo, se emplean diversas técnicas de regularización.

La **Regularización L2** (*Weight Decay*) añade un término de penalización proporcional a la suma de los cuadrados de los pesos ($\lambda \sum w_i^2$) a la función de pérdida, lo que fuerza al modelo a encontrar soluciones con pesos más pequeños y, en consecuencia, representaciones más simples y generalizables (Bishop, 2006). El **Early Stopping** (Detención Temprana) es una técnica complementaria que monitorea continuamente la pérdida de validación durante el entrenamiento y detiene automáticamente el proceso cuando esta métrica deja de mejorar durante un número determinado de épocas consecutivas, denominado paciencia, evitando así que el modelo continúe ajustándose al ruido de los datos de entrenamiento (Goodfellow et al., 2016).

En la etapa de preprocesamiento, la **Normalización Z-Score** (*StandardScaler*) transforma cada variable para que posea media cero y desviación estándar unitaria mediante la fórmula $z = (x - \mu)/\sigma$, lo que garantiza que todas las variables contribuyan equitativamente al entrenamiento independientemente de sus escalas originales de medición (Géron, 2019).

Asimismo, la **División Estratificada** del dataset consiste en la partición de los datos en subconjuntos preservando las proporciones originales de cada clase en todos ellos, lo cual es fundamental para garantizar una evaluación imparcial del rendimiento del modelo (Hastie et al., 2009). Los tres subconjuntos resultantes cumplen funciones

diferenciadas y complementarias: el **Conjunto de Entrenamiento** (*Training Set*) se utiliza para ajustar los pesos de la red neuronal durante el aprendizaje; el **Conjunto de Validación** (*Validation Set*) se emplea para monitorear el rendimiento del modelo durante el entrenamiento y sintonizar los hiperparámetros sin introducir sesgo en la evaluación; y el **Conjunto de Evaluación** (*Test Set*) se reserva exclusivamente para la evaluación final del modelo, garantizando que nunca haya sido visto durante el entrenamiento ni la validación (Géron, 2019).

2.3.1.6 Métricas de evaluación de modelos de clasificación

La evaluación rigurosa de modelos de clasificación requiere el uso de múltiples métricas complementarias, dado que ninguna métrica individual captura completamente el rendimiento del modelo en todas sus dimensiones (Fawcett, 2006). La **Exactitud** (*Accuracy*) mide la proporción de predicciones correctas sobre el total de observaciones, calculándose como $Accuracy = (VP + VN)/N$, donde VP son los verdaderos positivos, VN los verdaderos negativos y N el total de observaciones; sin embargo, esta métrica puede resultar engañosa en datasets con desbalance significativo entre clases (Géron, 2019).

La **Precisión** (*Precision*) cuantifica la proporción de predicciones positivas que son realmente correctas ($Precision = VP/(VP+FP)$), mientras que la **Sensibilidad** (*Recall*) mide la proporción de casos positivos reales que el modelo identifica correctamente ($Recall = VP/(VP + FN)$), siendo esta última la métrica más crítica en aplicaciones de seguridad vial donde los falsos negativos tramos peligrosos no detectados tienen consecuencias potencialmente fatales (Fawcett, 2006). La **Especificidad** (*Specificity*) complementa al *recall* midiendo la proporción de casos negativos reales correctamente clasificados como negativos ($Specificity = VN/(VN + FP)$), permitiendo evaluar simultáneamente la capacidad del modelo para detectar positivos y descartar negativos (Hosmer, Lemeshow & Sturdivant, 2013).

El **F1-Score** es la media armónica de la precisión y el *recall*, calculada como $F1 = 2\times(P\times R)/(P+R)$, que proporciona un equilibrio entre ambas métricas en un solo indicador; en problemas multiclase, se calcula en variantes *Macro* (promedio no ponderado por clase) y *Weighted* (promedio ponderado por el soporte de cada clase) para ofrecer una evaluación equitativa del rendimiento en todas las categorías (Géron, 2019). El **AUC-ROC** (Área Bajo la Curva ROC) mide la capacidad discriminativa global del

clasificador a través de todos los umbrales de decisión posibles, donde valores superiores a 0.90 indican discriminación sobresaliente según la escala interpretativa de Hosmer et al. (2013).

La **Curva ROC** (*Receiver Operating Characteristic*) es la representación gráfica de la relación entre la Tasa de Verdaderos Positivos (TPR) y la Tasa de Falsos Positivos (FPR) para cada umbral de decisión (Fawcett, 2006), mientras que la **Curva Precision-Recall** muestra la relación entre precisión y *recall* para diferentes umbrales, resultando más informativa que la curva ROC cuando el dataset presenta alto desbalance entre clases (Saito & Rehmsmeier, 2015).

La **Matriz de Confusión** es una tabla que resume de forma completa el rendimiento de un clasificador al mostrar los verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos para cada clase, constituyendo la base a partir de la cual se derivan todas las métricas anteriores (Fawcett, 2006). La **Validación Cruzada K-Fold** es un método de evaluación robusto que particiona el dataset en K subconjuntos mutuamente excluyentes, entrenando K modelos independientes donde cada subconjunto actúa una vez como conjunto de validación; la media y desviación estándar de las métricas a través de los K folds proporcionan una estimación confiable de la estabilidad y reproducibilidad del modelo (Hastie et al., 2009).

Finalmente, la **Importancia por Permutación** es un método para evaluar la contribución individual de cada variable al rendimiento del modelo, que consiste en aleatorizar los valores de cada variable de forma independiente y medir la caída resultante en la exactitud del modelo; las variables cuya aleatorización produce mayor degradación del rendimiento son las más importantes para la predicción (Breiman, 2001).

2.3.2 Fundamentos del diseño geométrico de infraestructura vial

2.3.2.1 Clasificación de carretera y orografía

De acuerdo al Manual de carreteras: Diseño geométrico DG 2018 se clasifican en:

A.1) Según la demanda:

- **Autopistas de Primera Clase:** "Son carreteras con IMDA (Índice Medio Diario Anual) mayor a 6 000 veh/día, de calzadas divididas por medio de un separador central mínimo de 6.00 m; cada una de las calzadas debe contar con dos o más carriles de 3.60 m de ancho como mínimo" (MTC,2018, p.12).

- **Autopistas de Segunda Clase:** "Son carreteras con un IMDA entre 60 00 y 4 001 veh/día, de calzadas divididas por medio de un separador central que puede variar de 6.00 m hasta 1.00 m, en cuyo caso se instalará un sistema de contención vehicular; cada una de las calzadas debe contar con dos o más carriles de 3.60 m de ancho como mínimo" (MTC,2018, p.12).
- **Carreteras de Primera Clase:** "Son carreteras con un IMDA entre 4 000 y 2 001 veh/día, con una calzada de dos carriles de 3.60 m de ancho como mínimo" (MTC,2018, p.12).
- **Carreteras de Segunda Clase:** "Son carreteras con IMDA entre 2 000 y 400 veh/día, con una calzada de dos carriles de 3.30 m de ancho como mínimo" (MTC,2018, p.12).
- **Carreteras de Tercera Clase:** "Son carreteras con IMDA menores a 400 veh/día, con calzada de dos carriles de 3.00 m de ancho como mínimo" (MTC,2018, p.12).
- **Trochas Carrozables:** "Son vías transitables, que no alcanzan las características geométricas de una carretera, que por lo general tienen un IMDA menor a 200 veh/día. Sus calzadas deben tener un ancho mínimo de 4.00 m, en cuyo caso se construirá ensanches denominados plazoletas de cruce, por lo menos cada 500 m" (MTC,2018, p.13).

A.2) Según la orografía:

- **Terreno plano (tipo 1):** "Tiene pendientes transversales al eje de la vía, menores o iguales al 10% y sus pendientes longitudinales son por lo general menores de tres por ciento (3%)" (MTC,2018, p.14).
- **Terreno ondulado (tipo 2):** "Tiene pendientes transversales al eje de la vía entre 11% y 50% y sus pendientes longitudinales se encuentran entre 3% y 6 %, demandando un moderado movimiento de tierras, lo que permite alineamientos rectos, alternados con curvas de radios amplios, sin mayores dificultades en el trazo" (MTC,2018, p.14).
- **Terreno accidentado (tipo 3):** "Tiene pendientes transversales al eje de la vía entre 51% y el 100% y sus pendientes longitudinales predominantes se encuentran entre 6% y 8%, por lo que requiere importantes movimientos de tierras, razón por la cual presenta dificultades en el trazo" (MTC,2018, p.14).
- **Terreno escarpado (tipo 4):** "Tiene pendientes transversales al eje de la vía superiores al 100% y sus pendientes longitudinales excepcionales son superiores

al 8%, exigiendo el máximo de movimiento de tierras, razón por la cual presenta grandes dificultades en su trazo" (MTC,2018, p.14).

A.2) Según la función:

- **Carreteras de la Red Vial Nacional:**La Red Vial Nacional agrupa las carreteras principales que enlazan entre sí ciudades, regiones y fronteras del estado; su función es facilitar el transporte interregional y nacional de personas y mercancías. Estas carreteras deben reunir altos y exigentes niveles de calidad técnica y son administradas por el gobierno nacional, brindando así una infraestructura vial adecuada para grandes flujos de tráfico si se le compara fija y prolongadamente con otros medios de transporte **(MTC, 2018)**.
- **Carreteras de la Red Vial Departamental o Regional:** La Red Vial Departamental incorpora las carreteras que dan conexión desde las capitales regionales y provinciales hacia otras localidades en el interior del departamento. Tiene por finalidad la comunicaci ón de la región y como posibilitar el transporte desde los pueblos hacia los núcleos urbanos o viceversa. Son caminos que presentan características técnicos intermediarios y son administrados por los gobiernos de los departamentos **(MTC, 2018)**.
- **Carreteras de la Red Vial Vecinal o Rural:** Las Vías Rurales de Conexión están formadas por caminos y vías rurales que conectan comunidades, núcleos de población y el campo con la red de carreteras que depende del Departamento o la Naciona. Estas vías son afirmadas o vías sin pavimentar con estándares técnicos bajos, siendo arregladas por los gobiernos locales o las propias comunidades. La utilidad de estas vías está en facilitar la movilidad del lugar y el transporte de producción agrícola en el campo **(MTC, 2018)**.

2.3.2.2 Índice Medio Diario Anual (IMDA)

Representa la media aritmética de los volúmenes diarios de tráfico registrados en todos los días del año para un tramo determinado de la carretera. Conocerlo permite estimar núm é ricamente la importancia funcional de la carretera para el tramo evaluado y resulta de gran utilidad para la elaboración de estudios de viabilidad económica. Los valores de IMDA referidos a tramos concretos permiten al proyecto conocer las características y funcionales del diseño, clasificar la carretera y diseñar estrategias de mejora, rehabilitación y mantenimiento **(MTC, 2018)**.

2.3.2.3 Velocidad de diseño

La seguridad vial de los usuarios será la primera consideración a la hora de establecer la asignación de la velocidad de diseño . Como consecuencia , la velocidad que se autoriza a lo largo del trazado , va a ser coherente y constante , de forma que evita bruscos o cambios de velocidad muy frecuentes que contraríen la heurística de la maniobra . La continuidad en la velocidad de diseño asegura una conducción más predecible, disminuye el número de accidentes y facilita la eficiencia de la operación de la vía **(MTC, 2018)**.

CLASIFICACIÓN	OROGRAFÍA	VELOCIDAD DE DISEÑO DE UN TRAMO HOMOGÉNEO VTR (km/h)									
		30	40	50	60	70	80	90	100	110	120	130
Autopista de primera clase	Plano											
	Ondulado											
	Accidentado											
	Escarpado											
Autopista de segunda clase	Plano											
	Ondulado											
	Accidentado											
	Escarpado											
Carretera de primera clase	Plano											
	Ondulado											
	Accidentado											
	Escarpado											
Carretera de segunda clase	Plano											
	Ondulado											
	Accidentado											
	Escarpado											
Carretera de tercera clase	Plano											
	Ondulado											
	Accidentado											
	Escarpado											

Figura 39 Rangos de la Velocidad de Diseño en función a la clasificación de la carretera por demanda y orografía.

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.4 Distancia de visibilidad

Según **(MTC, 2018, p. 103)** "es la longitud continua hacia adelante de la carretera, que es visible al conductor del vehículo para poder ejecutar con seguridad las diversas maniobras a que se vea obligado o que decida efectuar. En los proyectos se consideran dos distancias de visibilidad":

- **Distancia de visibilidad de parada:**

  "Es la mínima requerida para que se detenga un vehículo que viaja a la velocidad de diseño, antes de que alcance un objetivo inmóvil que se encuentra en su trayectoria" **(MTC, 2018, p. 103)**.

Velocidad de diseño (km/h)	Pendiente nula o en bajada			Pendiente en subida		
	3%	6%	9%	3%	6%	9%
20	20	20	20	19	18	18
30	35	35	35	31	30	29
40	50	50	53	45	44	43
50	66	70	74	61	59	58
60	87	92	97	80	77	75
70	110	116	124	100	97	93
80	136	144	154	123	118	114
90	164	174	187	148	141	136
100	194	207	223	174	167	160
110	227	243	262	203	194	186
120	283	293	304	234	223	214
130	310	338	375	267	252	238

Figura 40 Distancia de visibilidad de parada con pendiente (metros)

Fuente: Manual de carreteras: Diseño geométrico DG 2018

- **Distancia de visibilidad de paso o adelantamiento:**

"Es la mínima que debe estar disponible, a fin de facultar al conductor del vehículo a sobrepasar a otro que viaja a una velocidad menor, con comodidad y seguridad, sin causar alteración en la velocidad de un tercer vehículo que viaja en sentido contrario y que se hace visible cuando se ha iniciado la maniobra de sobrepaso" **(MTC, 2018, p. 106)**.

VELOCIDAD ESPECÍFICA EN LA TANGENTE EN LA QUE SE EFECTÚA LA MANIOBRA (km/h)	VELOCIDAD DEL VEHÍCULO ADELANTADO (km/h)	VELOCIDAD DEL VEHÍCULO QUE ADELANTA, V (km/h)	MÍNIMA DISTANCIA DE VISIBILIDAD DE ADELANTAMIENTO D_A (m)	
			CALCULADA	REDONDEADA
20	-	-	130	130
30	29	44	200	200
40	36	51	266	270
50	44	59	341	345
60	51	66	407	410
70	59	74	482	485
80	65	80	538	540
90	73	88	613	615
100	79	94	670	670
110	85	100	727	730
120	90	105	774	775
130	94	109	812	815

Figura 41 Mínima distancia de visibilidad de adelantamiento

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.5 Tramos en tangente

"Las longitudes mínimas admisibles y máximas deseables de los tramos en tangente, en función a la velocidad de diseño" **(MTC, 2018, p. 127)**.

V (km/h)	L mín.s (m)	L mín.o (m)	L máx (m)
30	42	84	500
40	56	111	668
50	69	139	835
60	83	167	1002
70	97	194	1169
80	111	222	1336
90	125	250	1503
100	139	278	1670
110	153	306	1837
120	167	333	2004
130	180	362	2171

Figura 42 Longitudes de tramos en tangente

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.6 Curvas circulares

"Las curvas horizontales circulares simples son arcos de circunferencia de un solo radio que unen dos tangentes consecutivas, conformando la proyección horizontal de las curvas reales o espaciales" **(MTC, 2018, p. 127)**.

2.3.2.7 Relación del peralte, radio y velocidad específica de diseño

Con el permitir que el gráfico se determine el peralte y la radio a tener en cuenta para poder diseñar la curva en una velocidad de diseño determinada, es decir, en una velocidad determinada para la curva a diseñar, de forma que los vehículos que circulan por la curva tengan, ante su curvatura, condiciones de seguridad y de confort necesarias **(MTC, 2018)**.

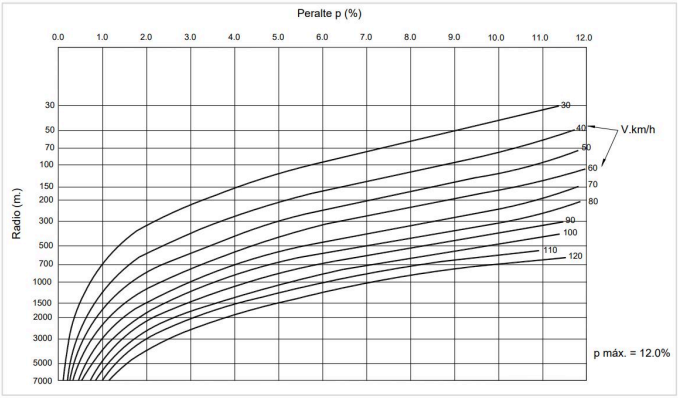


Figura 43 Peralte en zona rural (Tipo 3 ó 4)

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.8 Radios mínimos

"Los radios mínimos de curvatura horizontal son los menores radios que pueden recorrerse con la velocidad de diseño y la tasa máxima de peralte, en condiciones aceptables de seguridad y comodidad" **(MTC, 2018, p. 128)**.

Velocidad de diseño Km/h	Radio M
20	24
30	55
40	95
50	150
60	210
70	290
80	380
90	480

Figura 44 Elementos de la curva circular

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.9 Coordinación entre curvas circulares

"Para todo tipo de carretera, cuando se enlacen curvas circulares consecutivas sin tangente intermedia, así como mediante tangente de longitud menor o igual a 200 m, la relación de radios de las curvas circulares no sobrepasará los valores obtenidos a partir de la tabla siguiente" **(MTC, 2018, p. 134)**.

Radio Entrada (m)	Radio Salida (m)		Radio Entrada (m)	Radio Salida (m)	
	Máximo	Mínimo		Máximo	Mínimo
40	60	50	360	> 670	212
50	75	50	370	> 670	216
60	90	50	380	> 670	220
70	105	50	390	> 670	223
80	120	53	400	> 670	227
90	135	60	410	> 670	231
100	151	67	420	> 670	234
110	166	73	430	> 670	238
120	182	80	440	> 670	241
130	198	87	450	> 670	244
140	215	93	460	> 670	247
150	232	100	470	> 670	250
160	250	106	480	> 670	253
170	269	112	490	> 670	256
180	289	119	500	> 670	259
190	309	125	510	> 670	262
200	332	131	520	> 670	265
210	355	137	530	> 670	267
220	381	143	540	> 670	270
230	408	149	550	> 670	273
240	437	154	560	> 670	275
250	469	160	570	> 670	278
260	503	165	580	> 670	280

Figura 45 Relación entre radios consecutivos (Carretera de 2da y 3era clase)

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.10 Longitud de la curva de transicións

Se definen los valores mínimos para la longitud de la curva de transición, con la finalidad de asegurar una transición segura y agradable entre la recta y la curva circular, posibilitando la adaptación progresiva por parte del conductor de la velocidad y el peralte **(MTC, 2018)**.

Velocidad Km/h	Radio mín. m	J m/s^3	Peralte máx. %	A mín. m^2	Longitud de transición (L)	
					Calculada m	Redondeada m
30	24	0.5	12	26	28	30
30	26	0.5	10	27	28	30
30	28	0.5	8	28	28	30
30	31	0.5	6	29	27	30
30	34	0.5	4	31	28	30
30	37	0.5	2	32	28	30
40	43	0.5	12	40	37	40
40	47	0.5	10	41	36	40
40	50	0.5	8	43	37	40
40	55	0.5	6	45	37	40
40	60	0.5	4	47	37	40
40	66	0.5	2	50	38	40
50	70	0.5	12	55	43	45
50	76	0.5	10	57	43	45
50	82	0.5	8	60	44	45
50	89	0.5	6	62	43	45
50	98	0.5	4	66	44	45
50	109	0.5	2	69	44	45

Figura 46 Longitud mínima de curva de transición

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.11 Sobreancho

"Es el ancho adicional de la superficie de rodadura de la vía, en los tramos en curva para compensar el mayor espacio requerido por los vehículos. La necesidad de proporcionar sobreancho en una calzada, se debe a la extensión de la trayectoria de los vehículos y a la mayor dificultad en mantener el vehículo dentro del carril en tramos curvos" **(MTC, 2018, p. 31)**.

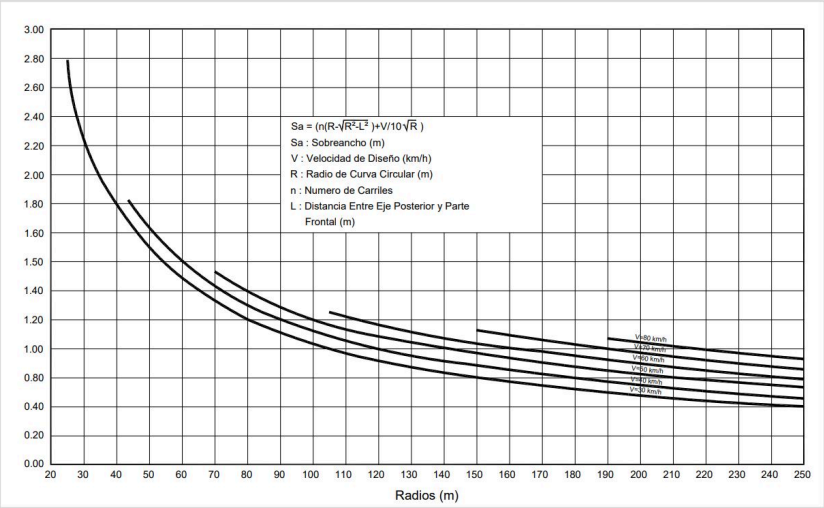


Figura 47 Valores de sobreancho en función a "L" del tipo de vehículo de diseño

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.12 Pendiente

"Es conveniente proveer una pendiente mínima del orden de 0.5%, a fin de asegurar en todo punto de la calzada un drenaje de las aguas superficiales" **(MTC, 2018, p. 32)**.

Demanda	Autopistas								Carretera				Carretera				Carretera			
Vehículos/día	> 6.000				6.000 - 4001				4.000-2.001				2.000-400				< 400			
Características	Primera clase				Segunda clase				Primera clase				Segunda clase				Tercera clase			
Tipo de orografía	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Velocidad de diseño: 30 km/h																			10.00	10.00
40 km/h																9.00	8.00	9.00	10.00	
50 km/h											7.00	7.00			8.00	9.00	8.00	8.00	8.00	
60 km/h					6.00	6.00	7.00	7.00	6.00	6.00	7.00	7.00	6.00	7.00	8.00	9.00	8.00	8.00		
70 km/h			5.00	5.00	6.00	6.00	6.00	7.00	6.00	6.00	7.00	7.00	6.00	6.00	7.00		7.00	7.00		
80 km/h	5.00	5.00	5.00	5.00	5.00	5.00	6.00	6.00	6.00	6.00	6.00		6.00	6.00			7.00	7.00		
90 km/h	4.50	4.50	5.00		5.00	5.00	6.00		5.00	5.00			6.00				6.00	6.00		
100 km/h	4.50	4.50	4.50		5.00	5.00	6.00		5.00				6.00							
110 km/h	4.00	4.00			4.00															
120 km/h	4.00	4.00			4.00															
130 km/h	3.50																			

Figura 48 Pendientes máximas (%)

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.13 Sección transversal

El diseño geométrico de la sección transversal está constituido por la descripción y definición de los elementos de la carretera y su descripción en una representación en un plano vertical que es perpendicular al alineamiento horizontal de la vía, y en función de conseguir la disposición y medida de los elementos que la hacen los que deben ser utilizados en cada uno de los puntos de la carretera, así como su disposición ahora con el terreno natural. teniendo como finalidad la estabilidad, funcionalidad y seguridad de la obra que se trata **(MTC, 2018)**.

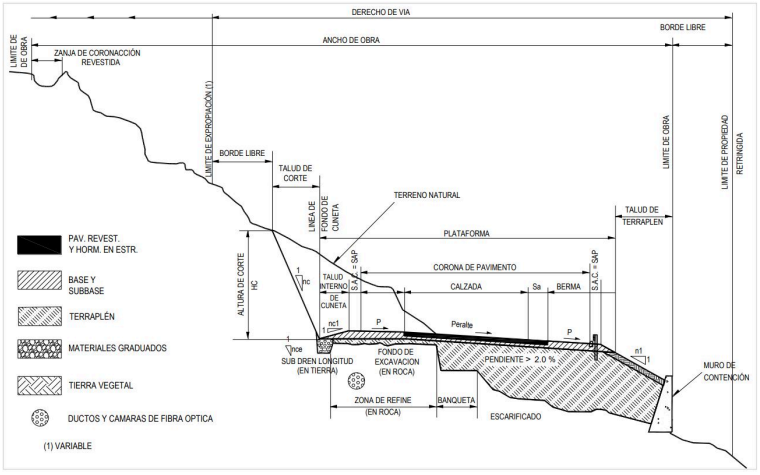


Figura 49 Sección transversal típica para carretera

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.14 Ancho de la calzada en tangente

"El ancho de la calzada en tangente, se determinará tomando como base el nivel de servicio deseado al finalizar el período de diseño. En consecuencia, el ancho y número de carriles se determinarán mediante un análisis de capacidad y niveles de servicio" **(MTC, 2018, p. 190)**.

Clasificación	Autopista								Carretera				Carretera				Carretera			
Tráfico vehículos/día	> 6,000				6,000 – 4,001				4,000-2.001				2,000-400				< 400			
Tipo	Primera Clase				Segunda Clase				Primera Clase				Segunda Clase				Tercera Clase			
Orografía	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Velocidad de diseño: 30km/h																			6.00	6.00
40 km/h																6.60	6.60	6.60	6.00	
50 km/h											7.20	7.20			6.60	6.60	6.60	6.60	6.00	
60 km/h					7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	6.60	6.60	6.60	6.60		
70 km/h			7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	6.60		6.60	6.60		
80 km/h	7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20	7.20		7.20	7.20			6.60	6.60		
90 km/h	7.20	7.20	7.20		7.20	7.20	7.20		7.20	7.20			7.20				6.60	6.60		
100 km/h	7.20	7.20	7.20		7.20	7.20	7.20		7.20				7.20							
110 km/h	7.20	7,20			7.20															
120 km/h	7.20	7.20			7.20															
130 km/h	7.20																			

Figura 50 Anchos mínimos de calzada en tangente

Fuente: Manual de carreteras: Diseño geométrico DG 2018

2.3.2.15 Ancho de las bermas

En la siguiente tabla se especifican los anchos de bermas, determinados según la clasificación de la vía, la velocidad de diseño y las características orográficas del terreno **(MTC, 2018)**.

Clasificación	Autopista								Carretera				Carretera				Carretera			
Tráfico vehículos/día	> 6.000				6.000 - 4001				4.000-2.001				2.000-400				< 400			
Características	Primera clase				Segunda clase				Primera clase				Segunda clase				Tercera Clase			
Tipo de orografía	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Velocidad de diseño: 30 km/h																			0.50	0.50
40 km/h																1.20	1.20	0.90	0.50	
50 km/h											2.60	2.60			1.20	1.20	1.20	0.90	0.90	
60 km/h					3.00	3.00	2.60	2.60	3.00	3.00	2.60	2.60	2.00	2.00	1.20	1.20	1.20	1.20		
70 km/h			3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	2.00	2.00	1.20		1.20	1.20		
80 km/h	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00		2.00	2.00			1.20	1.20		
90 km/h	3.00	3.00	3.00		3.00	3.00	3.00		3.00	3.00			2.00				1.20	1.20		
100 km/h	3.00	3.00	3.00		3.00	3.00	3.00		3.00				2.00							
110 km/h	3.00	3.00			3.00															
120 km/h	3.00	3.00			3.00															
130 km/h	3.00																			

Figura 51 Ancho de bermas

Fuente: Manual de carreteras: Diseño geométrico DG 2018

CAPÍTULO III
Método de la Investigación

3.1 Método

El método de investigación empleado es el **deductivo** porque se evaluará el fenomeno observado, es decir va de lo general a lo particular (Borja, 2012). Se parte de teorías consolidadas en inteligencia artificial y normativas internacionales de diseño geométrico vial como marco general, a partir de las cuales se formulan hipótesis verificables sobre la relación entre las variables geométricas de la PE-28A, ambientales, humanas, vehiculares que influyen en su siniestralidad. Estas hipótesis se contrastan empíricamente mediante modelos de redes neuronales artificiales entrenados con datos reales del período 2020–2025, permitiendo deducir conclusiones específicas sobre los tramos críticos y las propuestas de mejora para la vía en estudio.

3.2 Orientación

Según (Borja, 2012), ”la investigación aplicada consiste en buscar, conocer, interpretar, colaborar y modificar una realidad problemática”. Según su orientación, la investigación es de tipo **aplicada**, dado que hace uso del conocimiento científico disponible en inteligencia artificial específicamente las redes neuronales artificiales y de los estándares internacionales de diseño geométrico vial, con el propósito de resolver un problema concreto: la elevada siniestralidad registrada en la Vía Los Libertadores PE-28A durante el período 2020–2025. Los resultados obtenidos modelos predictivos, identificación de tramos críticos y propuestas de mejora geométrica tienen aplicación directa por parte de los organismos competentes como PROVÍAS Nacional y el MTC.

3.3 Enfoque

El enfoque de la investigación es **cuantitativo** permite medir, modelar, comparar y analizar datos objetivos y comprobables, aspecto que resulta crucial en trabajos relacionados con la seguridad vial. Este enfoque permite trabajar un conjunto de variables técnicas del diseño geométrico, como radios de curva, pendientes longitudinales, peraltes, velocidades de diseño, anchos de la calzada y frecuencia de accidentes, las cuales

son fácilmente evaluables mediante herramientas como AutoCAD Civil 3D. El enfoque cuantitativo también tiene la ventaja de permitir la combinación con modelos matemáticos avanzados, como las redes neuronales artificiales, que procesan grandes volúmenes de datos, descubren patrones ocultos entre conjunto de datos y son capaces de predecir la aparición de accidentes en función de un conjunto de variables geométricas y operativas de un determinado diseño vial. Sobre este aspecto, **(Hernandez, Fernandez & Baptista, 2014)** sostiene: *"El enfoque cuantitativo se fundamenta en la recolección de datos para probar hipótesis, con base en la medición numérica y el análisis estadístico, con el fin de establecer patrones de comportamiento y probar teorías"* **(p. 4).**

3.4 Alcance

El alcance de la presente investigación es de tipo **explicativo**. En una primera etapa se describe y caracteriza el estado geométrico actual de la Vía Los Libertadores (PE-28A) y su registro histórico de siniestros viales durante el período 2020–2025, lo que corresponde a un nivel descriptivo. Posteriormente, mediante el entrenamiento de modelos de redes neuronales artificiales con 23 variables de naturaleza multicausal —geométricas, ambientales, vehiculares y temporales —, se identifican y cuantifican las relaciones entre dichas variables y la ocurrencia y severidad de los accidentes, alcanzando un nivel correlacional. Finalmente, al establecer las causas predominantes de la siniestralidad en los tramos críticos y proponer mejoras al diseño geométrico sustentadas en normativas internacionales, la investigación alcanza el nivel explicativo **(Hernandez et al., 2014)**.

3.5 Diseño de investigación

3.5.1 Diseño no experimental

El diseño es no experimental ya que las variables no se manipulan a propósito, sino que se observan y analizan como realmente están situadas dentro de su contexto real mediante simulaciones digitales. **(Hernandez et al., 2014)** afirman que: *"dentro de los diseños no experimentales nunca se manipulan de manera intencional las variables independientes; se observan los fenómenos exactamente como suceden en su ambiente natural"* **(p. 151).**

3.5.2 Diseño comparativo

Dado el carácter comparativo del diseño, aquí se comprueban todos los parámetros geométricos establecidos en las normas extranjeras de diseño de carretera , con el fin de ver cómo cada uno influye en la seguridad vial. Es el simulador de varios escenarios por medio de los mismos tramos aplicándose distintaas normas internacionales, y así contrastar los efectos que tienen sobre la maniobrabilidad y las posibilidades de accidente . Según **(Sampieri, Collado & Lucio, 2014)** sostiene: *"El diseño comparativo busca establecer similitudes y diferencias entre dos o más variables, grupos o enfoques, para generar conclusiones más fundamentadas"* **(p. 135).**

3.5.3 Diseño transversal

Según **(Sampieri et al., 2014)** menciona que: *"Los estudios transversales se realizan en un solo momento del tiempo, proporcionando una fotografía del fenómeno estudiado"* **(p. 154).** Lo que comprende al presente trabajo de investigación, dado que se realiza la recolección y análisis de datos en un solo tiempo (año 2020 al 2025), corresponde a un diseño transversal. Lo que incluye la revisión de normas como tambien el análisis del tramo vial mediante programas especializados y ademas el registro de siniestros viales recientes.

3.6 Población y muestra

3.6.1 Población

La población de estudio comprende tres unidades de análisis diferenciadas según los objetivos de la investigación:

- **Primera población:** La totalidad de registros de siniestros viales oficialmente documentados por el Observatorio Nacional de Seguridad Vial (ONSV) en la Vía Los Libertadores PE-28A durante el período 2020–2025.
- **Segunda población:** El conjunto de tramos geométricos que conforman los 333.903 km de la Vía Los Libertadores PE-28A, modelados digitalmente mediante el software Civil 3D a partir del modelo digital del terreno (DTM), abarcando 3 departamentos, 4 provincias y 11 distritos.

- **Tercera población:** El conjunto de normativas internacionales de diseño geométrico vial vigentes, consideradas para el análisis comparativo con la normativa peruana DG-2018.

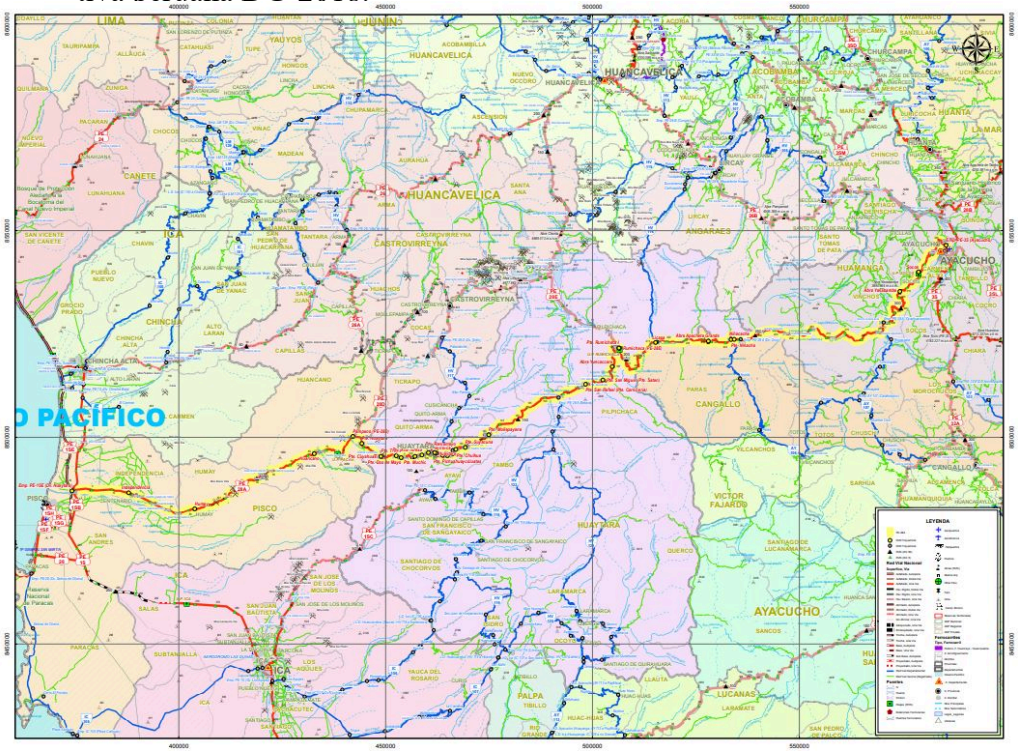


Figura 52 Plano de la Carretera Vía Los Libertadores (PE-28A)
Fuente: Ministerio de Transporte y Comunicaciones

Figura 53 Progresiva 0+000km - Ica (distr. San Clemente)
Fuente:Fotografía propia.

Figura 54 Puente Rumichaca.
Fuente:Fotografía propia.

Figura 55 Progresiva 333+903km - Ayacucho
Fuente:Fotografía propia.

3.6.2 Muestra

La determinación de la muestra se realizó de manera diferenciada para cada unidad de análisis:

- **Muestra para el Modelo RNA N°1 (clasificación binaria):** Conformada por 202 registros de accidentes de tránsito del período 2020–2025, con 23 variables por registro de naturaleza multicausal geométricas, ambientales, vehiculares y temporales. Se excluyeron los registros anteriores al año 2020 por presentar insuficiente consolidación estadística para su uso en modelos de aprendizaje automático. El tipo de muestreo es no probabilístico por criterio.
- **Muestra para el Modelo RNA N°2 (clasificación multiclase de riesgo):** Conformada por 2,997 registros de tramos geométricos de la totalidad de la vía PE-28A, con 15 variables geométricas extraídas mediante Civil 3D otros variables de los factores incidentes, garantizando una distribución equilibrada entre los niveles Riesgo Bajo, Riesgo Medio y Riesgo Alto.
- **Muestra para el análisis normativo comparativo:** Conformada por 6 normativas internacionales de diseño geométrico vial DG-2018 (Perú) y de los paises de Noruega, Suecia, Islandia, Reino Unido, y las Norma de Japón. La selección responde a paises que menor tasas de accidentes de tránsito.
- **Muestra para el progrma de alerta de accidentes de tránsito:** Conformada por 202 registros de accidentes de tránsito del período 2020–2025 de la Via Los Libertadores.

3.7 Hipótesis

3.7.1 Hipótesis general

La implementación de dos modelos de redes neuronales artificiales para la clasificación de la peligrosidad de accidentes de tránsito, combinada con el análisis comparativo de la norma DG-2018 con normativas internacionales de diseño geométrico vial y la implementación de un programa de alerta basado en geolocalización, contribuirá a la predicción y reducción de los accidentes de tránsito en la Vía Los Libertadores (PE-28A), Ayacucho.

3.7.2 Hipótesis específicas

a) El Modelo RNA N°01 (clasificación binaria: riesgo Menor o riesgo mayor), entrenado con 202 registros históricos de accidentes de tránsito de la Vía Los Libertadores (PE-28A) y variables del diseño geométrico, condiciones climáticas, características del siniestro y factores viales conforme a la norma DG-2018, alcanza un nivel de precisión alto para la predicción de accidentes de tránsito.
b) El Modelo RNA N°02 (clasificación multiclase: riesgo bajo, riesgo medio y riesgo alto), entrenado con 2,997 registros de datos geométricos de la Vía Los Libertadores (PE-28A) modelados en AutoCAD Civil 3D, alcanza un nivel de precisión alto para predecir el nivel de riesgo de accidentes de tránsito en tramos viales.
c) La comparación de los parámetros del diseño geométrico de la norma DG-2018 con las normativas internacionales de Noruega, Islandia, Suecia, Reino Unido y Japón permite identificar deficiencias en el diseño actual de la Vía Los Libertadores (PE-28A) y formular propuestas de mejora orientadas a optimizar su seguridad vial.
d) El desarrollo e implementación de un programa de alerta de accidentes de tránsito que integre los registros históricos georreferenciados de la PE-28A y rastreo GPS constituye una herramienta de seguridad vial viable para conductores en la Vía Los Libertadores (PE-28A).

3.8 Operacionalización de variables, definición conceptual y operacional

La presente investigación contempla tres unidades de análisis diferenciadas: el Modelo RNA N°1 (clasificación binaria de accidentes), el Modelo RNA N°2 (clasificación multiclase de riesgo geométrico) y el análisis normativo comparativo de la DG-2018 frente a cinco normativas internacionales. A continuación se presenta la operacionalización completa de las variables dependientes e independientes correspondientes a cada unidad.

3.8.1 Variables dependientes

Tabla 2 *Operacionalización de las variables dependientes*

N°	Variable	Modelo / Unidad	Definición conceptual	Definición operacional	Valores	Escala
1	Accidente de tránsito	RNA N°1 Binario	Evento vial no deseado que produce daños materiales, lesiones o muertes en la PE-28A, resultado de la interacción entre factores geométricos, climáticos y temporales	Registro binario de presencia de accidente, obtenido de la base de datos ONSV 2020–2025 con 202 registros reales	0 = Riesgo Menor 1 = Riesgo Mayor	Ordinal y cardinal
2	Accidente de tránsito	RNA N°2 Multiclase	Grado de peligrosidad de un tramo geométrico de la PE-28A determinado por sus características de trazo, clasificado en niveles que expresan la probabilidad de siniestralidad	Clasificación multiclase del tramo geométrico obtenida del Modelo RNA N°2, entrenado con 2,997 registros	0 = Riesgo Bajo 1 = Riesgo Medio 2 = Riesgo Alto	Ordinal y cardinal
3	Accidente de tránsito	Análisis Normativo	Grado en que los parámetros de diseño geométrico de la DG-2018 se ajustan a los estándares internacionales de países con los índices más bajos de siniestralidad	Clasificación del parámetro normativo comparado respecto al estándar internacional de referencia, determinando brechas técnicas que fundamentan las propuestas de mejora	Conforme Parcialmente conforme No conforme	Nominal

Continúa en la siguiente página

(Continuación de la Tabla 2)

N°	Variable	Modelo / Unidad	Definición conceptual	Definición operacional	Valores	Escala
4	Accidente de tránsito	Programa de Alerta	Proximidad del conductor a zonas históricas de siniestralidad en la PE-28A.	Distancia en metros calculada por la fórmula de Haversine entre la posición GPS del conductor y el punto de los 202 registros de acccidentes.	Fuera de zona de peligro Zona de peligro	Ordinal

Fuente: Elaboración propia.

3.8.2 Variables independientes

3.8.2.1 Factor temporal (exclusivo Modelo RNA N°1)

Tabla 3 *Variables independientes — Factor temporal*

N°	Variable	Definición conceptual	Definición operacional	Valores reales	Escala
1	Mes de siniestro	Mes calendario en que ocurrió el accidente, asociado a la estacionalidad climática y flujo vehicular en la PE-28A	Valor entero del mes del registro ONSV, agrupado en franjas estacionales para el análisis de patrones temporales	1 al 12 (meses del año)	Ordinal
2	Día de siniestro	Día de la semana en que ocurrió el evento vial, relacionado con patrones de movilidad y comportamiento del conductor	Valor entero del día registrado en el parte ONSV	1 al 7 (días de la semana)	Ordinal
3	Hora de siniestro	Franja horaria en que ocurrió el accidente, asociada a condiciones de visibilidad y densidad de tráfico	Valor entero de la franja horaria registrada en el parte ONSV	1 = Madrugada 2 = Mañana 3 = Tarde 4 = Noche	Ordinal

Fuente: Elaboración propia.

3.8.2.2 Factor climático y ambiental (exclusivo Modelo RNA N°1)

Tabla 4 *Variables independientes — Factor climático y ambiental*

N°	Variable	Definición conceptual	Definición operacional	Valores reales	Escala
4	Clima despejado	Condición atmosférica sin precipitaciones ni nubosidad significativa al momento del siniestro, que favorece la visibilidad y la adherencia neumático-calzada	Variable binaria codificada como 1 si el clima era despejado al momento del accidente y 0 en caso contrario	0 = No 1 = Sí	Cardinal
5	Clima nublado	Condición de nubosidad moderada o densa que reduce la visibilidad sin precipitaciones activas al momento del evento vial	Variable binaria codificada como 1 si existía nubosidad al momento del accidente y 0 en caso contrario	0 = No 1 = Sí	Cardinal
6	Clima lluvioso	Presencia activa de precipitaciones al momento del siniestro que reduce la visibilidad y la fricción neumático-calzada	Variable binaria codificada como 1 si existía lluvia activa al momento del accidente y 0 en caso contrario	0 = No 1 = Sí	Cardinal
7	Temperatura máxima (°C)	Temperatura máxima registrada en la zona del siniestro durante el día del evento, asociada al comportamiento del pavimento y del conductor	Valor numérico continuo de temperatura en grados Celsius registrado para el día y zona del accidente	Numerico °C	Cardinal
8	Precipitación (mm/día)	Volumen de precipitación pluvial acumulada en el día del accidente en la zona de ocurrencia, que condiciona el estado superficial de la calzada	Valor numérico continuo de precipitación diaria en milímetros correspondiente al día y zona del registro	Numerico mm/día	Cardinal

Fuente: Elaboración propia.

3.8.2.3 Factor geométrico: Alineamiento horizontal (Modelos RNA N°1 y N°2)

Tabla 5 *Variables independientes — Factor geométrico: alineamiento horizontal*

N°	Variable	Definición conceptual	Definición operacional	Valores reales	Escala
9	Curva o tramo recto	Clasificación del tramo vial donde ocurrió el siniestro, distinguiendo entre alineamiento curvo y tangente	Variable binaria codificada como 1 si el evento ocurrió en curva horizontal y 0 si ocurrió en tramo recto	0 = Tramo recto 1 = Curva	Cardinal
10	Radio de curva (m)	Parámetro geométrico que define la curvatura del alineamiento horizontal, determinante de la fuerza centrífuga actuante sobre el vehículo en movimiento	Valor numérico en metros del radio de la curva horizontal del tramo, extraído del modelo Civil 3D con DTM de PROVÍAS Nacional	Número ”m”	Cardinal
11	Longitud de curva (m)	Extensión del tramo curvo medida sobre el eje de la vía entre sus puntos de inicio y término	Valor numérico en metros de la longitud del arco de curva horizontal correspondiente al tramo	Número ”m”	Cardinal
12	Longitud de subtangente (m)	Distancia horizontal desde el vértice de la curva hasta el punto de inicio o fin de la misma, relacionada con la geometría de transición	Valor numérico en metros de la subtangente de la curva horizontal del tramo	Número ”m”	Cardinal
13	Ángulo de deflexión (°)	Ángulo de cambio de dirección entre dos alineamientos rectos consecutivos que definen la curva horizontal	Valor numérico en grados sexagesimales del ángulo de deflexión de la curva horizontal del tramo	Número ”°”	Cardinal
14	Longitud de cuerda (m)	Distancia rectilínea entre los puntos de inicio y fin de la curva horizontal, indicador de la amplitud del cambio de dirección	Valor numérico en metros de la cuerda de la curva horizontal del tramo	Número ”m”	Cardinal
15	Distancia externa (m)	Distancia entre el vértice de la curva y el punto medio del arco, indicador de la agudeza de la curvatura horizontal	Valor numérico en metros de la distancia externa de la curva horizontal del tramo	Número ”m”	Cardinal

Continúa en la siguiente página

(Continuación de la Tabla 5)

N°	Variable	Definición conceptual	Definición operacional	Valores reales	Escala
16	Ordenada media (m)	Distancia entre el punto medio de la cuerda y el punto medio del arco de la curva, parámetro complementario de la geometría horizontal	Valor numérico en metros de la ordenada media de la curva horizontal del tramo	Número ”m”	Cardinal

Fuente: Elaboración propia.

3.8.2.4 Factor geométrico: Alineamiento tangente (Modelos RNA N°1 y N°2)

Tabla 6 *Variables independientes — Factor geométrico: alineamiento tangente*

N°	Variable	Definición conceptual	Definición operacional	Valores reales	Escala
17	Tramo tangente en “S”	Configuración geométrica de dos curvas consecutivas en sentidos opuestos separadas por una tangente corta, que genera condición de peligro por inversión de curvatura	Variable ordinal que indica la presencia y nivel de peligrosidad del tramo tangente en configuración S	0 = Ausente 1–3 = Nivel de peligro creciente	Cardinal
18	Tramo tangente en “O”	Configuración geométrica de dos curvas consecutivas en el mismo sentido separadas por una tangente corta, que genera condición de visibilidad reducida	Variable ordinal que indica la presencia y nivel de peligrosidad del tramo tangente en configuración O	0 = Ausente 1–3 = Nivel de peligro creciente	Cardinal
19	Longitud de tangente (m)	Extensión del tramo recto entre dos curvas horizontales consecutivas, determinante de la percepción de velocidad y comportamiento del conductor	Valor numérico en metros de la longitud del tramo tangente de la ubicación del evento	Número ”m”	Cardinal

Fuente: Elaboración propia.

3.8.2.5 Factor de peligro: Variables de alerta (Modelos RNA N°1 y N°2)

Tabla 7 *Variables independientes — Factor de peligro: variables de alerta*

N°	Variable	Definición conceptual	Definición operacional	Valores reales	Escala
20	Peligro en pendiente	Condición de riesgo asociada a una pendiente longitudinal que supera los límites recomendados por la DG-2018 para el tipo de terreno de la PE-28A	Variable binaria que indica si el tramo presenta pendiente longitudinal en condición de peligro según parámetros DG-2018	0 = Sin peligro 1 = Con peligro	Ordinal
21	Peligro en curva	Nivel de riesgo asociado a los parámetros geométricos de la curva horizontal — radio, peralte, velocidad — respecto a los límites normativos DG-2018	Variable ordinal que clasifica el nivel de peligro geométrico de la curva horizontal del tramo	0 = Sin peligro 1 = Bajo 2 = Medio 3 = Alto	Ordinal
22	Peligro por visibilidad	Condición de riesgo derivada de una distancia de visibilidad insuficiente respecto a los mínimos establecidos por la DG-2018 para la velocidad de diseño del tramo	Variable ordinal que clasifica el nivel de déficit de visibilidad del tramo	0 = Adecuada 1 = Déficit parcial 2 = Déficit crítico	Ordinal
23	Peligro estado del pavimento	Condición de deterioro superficial del pavimento que reduce la fricción neumático-calzada e incrementa el riesgo de siniestro	Variable binaria que indica si el tramo presenta pavimento en estado de peligro al momento del registro	0 = Buen estado 1 = Deteriorado	Ordinal

Fuente: Elaboración propia.

3.8.2.6 Parámetros normativos comparados (Análisis normativo — 6 países)

Tabla 8 *Variables independientes — Parámetros normativos comparados*

N°	Parámetro	Definición conceptual	Definición operacional	Indicador	Escala
24	Radio mínimo de curvatura horizontal	Valor mínimo del radio de curva horizontal permisible para una velocidad de diseño dada, garantizando la estabilidad del vehículo sin superar el peralte máximo	Valor en metros del radio mínimo establecido por cada normativa para velocidad equivalente al tramo PE-28A, contrastado entre los 6 países	Metros (m) por normativa	Cardinal
25	Peralte máximo	Inclinación transversal máxima permisible de la calzada en tramos curvos para contrarrestar la fuerza centrífuga	Valor porcentual máximo del peralte establecido por cada normativa, contrastado entre los 6 países	Porcentaje (%) por normativa	Cardinal
26	Longitud mínima de tangente entre curvas	Extensión mínima del tramo recto requerida entre dos curvas consecutivas para garantizar la percepción segura del cambio de dirección	Valor en metros de la longitud mínima de tangente entre curvas establecida por cada normativa	Metros (m) por normativa	Cardinal
27	Pendiente máxima longitudinal	Valor máximo de inclinación del eje vial permisible para el tipo de terreno y velocidad de diseño, determinante del control vehicular y distancia de frenado	Valor porcentual máximo de pendiente longitudinal establecido por cada normativa para terreno accidentado y escarpado equivalente al de la PE-28A	Porcentaje (%) por normativa	Cardinal
28	Pendiente mínima longitudinal	Valor mínimo de inclinación del eje vial requerido para garantizar el drenaje superficial adecuado de la calzada	Valor porcentual mínimo de pendiente establecido por cada normativa	Porcentaje (%) por normativa	Cardinal
29	Radio mínimo de curva vertical	Valor mínimo del radio de la parábola de acuerdo vertical que garantiza la distancia de visibilidad de parada requerida	Valor en metros del radio mínimo de curva vertical cóncava y convexa establecido por cada normativa	Metros (m) por normativa	Cardinal

Continúa en la siguiente página

(Continuación de la Tabla 8)

N°	Parámetro	Definición conceptual	Definición operacional	Indicador	Escala
30	Ancho mínimo de calzada	Dimensión transversal mínima de la superficie destinada a la circulación vehicular, determinante de la seguridad en cruce y adelantamiento	Valor en metros del ancho mínimo de calzada establecido por cada normativa para carretera de dos carriles equivalente a la PE-28A	Metros (m) por normativa	Cardinal
31	Ancho mínimo de berma	Dimensión mínima de la franja longitudinal paralela a la calzada destinada a paradas de emergencia y protección del borde de la vía	Valor en metros del ancho mínimo de berma establecido por cada normativa	Metros (m) por normativa	Cardinal
32	Talud de corte máximo	Inclinación máxima permisible de la superficie lateral de excavación según el tipo de material del terreno	Relación H:V del talud de corte establecida por cada normativa para terreno rocoso y suelto	Relación H:V por normativa	Cardinal
33	Velocidad de diseño	Velocidad máxima de circulación segura en un tramo homogéneo, parámetro base de todos los elementos geométricos	Valor en km/h de la velocidad de diseño establecida por cada normativa para carretera primaria en terreno accidentado equivalente a la PE-28A	km/h por normativa	Cardinal
34	Distancia mínima de visibilidad de parada	Longitud mínima de vía que el conductor debe percibir para detener el vehículo ante un obstáculo en condiciones normales	Valor en metros de la distancia mínima de visibilidad de parada establecida por cada normativa para la velocidad de diseño equivalente	Metros (m) por normativa	Cardinal
35	Distancia mínima de visibilidad de adelantamiento	Longitud mínima de vía libre requerida para que un vehículo pueda sobrepasar a otro con seguridad	Valor en metros de la distancia mínima de visibilidad de adelantamiento establecida por cada normativa	Metros (m) por normativa	Cardinal

Fuente: Elaboración propia.

3.8.3 Resumen general de variables por unidad de análisis

Tabla 9 *Resumen general de variables por unidad de análisis*

Unidad de análisis	Variable dependiente	Variables independientes	N° total	Instrumento
RNA N°1 Binario	Accidente de tránsito (0/1)	Diseño geométrico	**23**	Base ONSV 2020–2025 / Civil 3D / SENAMHI
RNA N°2 Multiclase	Accidente de tránsito (0/1/2)	Diseño geométrico	**15**	Civil 3D / Google Earth
Análisis Normativo	Accidente de tránsito	Diseño geométrico (12) × 6 normativas	**12 × 6**	Fichas de análisis documental normativo
Programa de Alerta GPS	Accidente de tránsito	Coordenadas GPS del conductor (lat, lon), distancia a los 202 registros georreferenciados	**202**	Base ONSV 2020–2025 / Geolocalización

Fuente: Elaboración propia.

3.9 Técnicas e instrumentos

Para el desarrollo de la presente investigación se emplearon cinco técnicas de recolección de datos, seleccionadas en función de la naturaleza de cada unidad de análisis y de los objetivos específicos planteados.

3.9.1 Técnicas de recolección de datos

3.9.1.1 Análisis documental

Técnica orientada a la revisión, extracción y sistematización de información contenida en documentos oficiales. Se aplicó en tres niveles diferenciados:

- Revisión de los registros oficiales de siniestros viales emitidos por el Observatorio Nacional de Seguridad Vial (ONSV) para la PE-28A, correspondientes al período 2020–2025.
- Revisión y descarga de datos meteorológicos históricos de las estaciones climatológicas del Servicio Nacional de Meteorología del Perú (SENAMHI), correspondientes a las zonas de influencia de la vía PE-28A en los departamentos de Ica, Huancavelica y Ayacucho, incluyendo temperatura diaria (°C), precipitación acumulada diaria (mm/día) y otros.

- Revisión y extracción sistemática de parámetros geométricos de las seis normativas internacionales de diseño geométrico vial: DG-2018 (Perú), Håndbok N100 (Noruega), Vägars och gators utformning (Suecia), Veghönnunarreglur (Islandia), Design Manual for Roads and Bridges DMRB (Reino Unido) y Road Structure Ordinance - Dōro Kōzōrei (Japón).

3.9.1.2 Observación de campo

Técnica de registro visual directo empleada para el reconocimiento de las condiciones físicas de la PE-28A. Los recorridos se realizaron a lo largo de los tramos de estudio, registrando mediante fotografías y videos el estado del pavimento, señalización vial, características topograficas, configuraciones geométricas críticas y condiciones del entorno vial. La información recopilada complementó y validó los datos geométricos obtenidos del modelo Civil 3D, sin que se realizara intervención física alguna sobre la vía.

3.9.1.3 Modelamiento digital

Técnica de caracterización geométrica computacional empleada para la extracción de las variables geométricas de la PE-28A mediante el software Civil 3D, utilizando como insumo principal el modelo digital del terreno (DTM) en Google Earth, y compatibilizado con el plano de la via de acuerdo a la SINAC (Sistema nacional de carreteras). Esta técnica permitió procesar la totalidad de los 333.903 km de la vía y extraer las variables geométricas del Modelo RNA N°2, sin necesidad de un levantamiento topográfico convencional en campo, dado que el DTM oficial constituyó la fuente primaria de datos geométricos.

3.9.1.4 Análisis estadístico

Técnica empleada para el procesamiento, selección y validación de variables. Se aplicó en dos niveles:

- Prueba no paramétrica de Mann-Whitney U para la selección de las variables estadísticamente significativas del Modelo RNA N°1, permitiendo reducir y depurar el conjunto inicial de variables de entrada.
- Técnica SMOTE (*Synthetic Minority Oversampling Technique*) para el balanceo de clases del Modelo RNA N°2, de los 2,997 datos entre las clases de riesgo Bajo,

Medio y Alto.

3.9.1.5 Modelamiento computacional con redes neuronales artificiales

Técnica de aprendizaje automático supervisado empleada para el entrenamiento, validación y evaluación de los dos modelos predictivos RNA. Se utilizó el algoritmo MLPClassifier de la librería Scikit-learn bajo el entorno Python/JupyterLab, con arquitecturas diferenciadas para el Modelo N°1 (clasificación binaria) y el Modelo N°2 (clasificación multiclase). La evaluación de ambos modelos se realizó mediante las métricas de exactitud (*accuracy*), AUC-ROC, matriz de confusión, precisión, sensibilidad y F1-score.

3.9.1.6 Resumen de las técnicas de recolección de información

Tabla 10 *Técnicas de recolección de datos empleadas en la investigación*

N°	Técnica	Unidad de análisis aplicada
1	Análisis documental	RNA N°1: revisión de registros oficiales ONSV 2020–2025 y datos meteorológicos SENAMHI. Análisis Normativo: extracción de parámetros geométricos de las normativas de 6 países
2	Observación de campo	Validación y complementación de datos geométricos mediante recorridos en la PE-28A, registrando estado del pavimento, señalización, topografia y configuraciones geométricas críticas
3	Modelamiento digital	RNA N°2: extracción de las 15 variables geométricas de los 333.903 km de la PE-28A mediante Civil 3D y el Gogle Earth.
4	Análisis estadístico	RNA N°1: prueba Mann-Whitney U para selección de variables significativas. RNA N°2: técnica SMOTE para balanceo de clases de los 2,997 datos.
5	Modelamiento computacional con redes neuronales artificiales	RNA N°1: clasificación binaria. RNA N°2: clasificación multiclase. Ambos modelos con algoritmo MLPClassifier de Scikit-learn bajo Python/JupyterLab
6	Desarrollo de programa	Programa de alerta de accidentes de tránsito: implementación de archivo HTML con los 202 registros históricos georreferenciados de la PE-28A (2020-2025), rastreo GPS en tiempo real mediante la geolocalización.

Fuente: Elaboración propia.

3.9.2 Instrumentos de recolección de datos

3.9.2.1 Registros oficiales de siniestros viales — ONSV

Partes oficiales de accidentes de tránsito emitidos por el Observatorio Nacional de Seguridad Vial (ONSV) para la Vía Los Libertadores PE-28A. Cada registro contiene información sobre las condiciones del evento vial que constituyen las variables de entrada del Modelo RNA N°1.

- **Fuente:** Observatorio Nacional de Seguridad Vial (ONSV).
- **Período de cobertura:** 2020–2025.
- **Total de registros:** 202 accidentes reales registrados en la PE-28A.
- **Variables extraídas:** Las variables temporales mes, día y hora del siniestro y la variable dependiente accidente de tránsito (0 = Riesgo Menor / 1 = Riesgo Mayor) fueron obtenidas directamente de los registros oficiales del ONSV. Los parámetros geométricos curva o tramo recto; radio de curva; longitud de curva; longitud de subtangente; ángulo de deflexión; longitud de cuerda; distancia externa; ordenada media; longitud de tangente; tramo tangente en configuración S; tramo tangente en configuración fueron extraídos del modelo digital de la PE-28A procesado en Civil 3D con Google Earth, asignados a cada registro de accidente según su ubicación georreferenciada en la vía.

3.9.2.2 Datos meteorológicos de estaciones SENAMHI

Registros históricos de variables climáticas obtenidos de las estaciones meteorológicas del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) ubicadas en la zona de influencia de la PE-28A. Los datos climatológicos fueron cruzados con los registros de siniestros ONSV por fecha y zona geográfica, permitiendo construir las cinco variables climáticas del Modelo RNA N°1.

- **Fuente:** SENAMHI - Servicio Nacional de Meteorología e Hidrología del Perú
- **Estaciones consideradas:** Estaciones meteorológicas de la zona de influencia de la PE-28A en los departamentos de Ica, Huancavelica y Ayacucho
- **Período de cobertura:** 2020 - 2025
- **Variables extraídas:**
  - Temperatura máxima diaria (°C).
  - Precipitación acumulada diaria (mm/día).

– Condición climática despejado, nublado y lluvioso.

3.9.2.3 Modelamiento de geométrico de la vía en Civil 3D

Instrumento de extracción digital de parámetros geométricos generado a partir del modelo de la PE-28A procesado en Civil 3D con Google Earth. Constituye la fuente principal de datos del Modelo RNA N°2.

- **Fuente:** Civil 3D + Google Earth + Mapa de la carretera PE-28A.
- **Cobertura:** 333.903 km de la PE-28A (3 departamentos, 4 provincias, 11 distritos).
- **Total de registros:** 2,997 tramos geométricos originales, expandidos a 5,394 mediante SMOTE.
- **Variables extraídas:** Curva o tramo recto; peligro en pendiente; radio de curva; peligro en curva; longitud de curva; subtangente ; ángulo de deflexión; longitud de cuerda; distancia externa; ordenada media; peligro por visibilidad; tramo tangente en configuración S; tramo tangente en configuración O; longitud de tangente; peligro en estado del pavimento.

3.9.2.4 Análisis documental normativo comparativo

Instrumento de elaboración propia diseñado para la extracción y comparación sistemática de los parámetros geométricos de las seis normativas internacionales consideradas. Cada ficha recoge el valor del parámetro establecido por cada normativa para condiciones equivalentes a las de la PE-28A, permitiendo identificar las brechas técnicas respecto a la DG-2018 y fundamentar las propuestas de mejora.

- **Fuente:** Normativas oficiales de 6 países.
- **Normativas:** DG-2018 (Perú), Håndbok N100 (Noruega), Vägars och gators utformning (Suecia), Veghönnunarreglur (Islandia), Design Manual for Roads and Bridges DMRB (Reino Unido) y Road Structure Ordinance - Dōro Kōzōrei (Japón).
- **Parámetros comparados:** Radio mínimo de curvatura horizontal, peralte máximo, longitud mínima de tangente entre curvas, pendiente máxima longitudinal, pendiente mínima longitudinal, radio mínimo de curva vertical, ancho mínimo de calzada, ancho mínimo de berma, talud de corte máximo, velocidad de diseño,

distancia mínima de visibilidad de parada y distancia mínima de visibilidad de adelantamiento.

3.9.2.5 Registro fotográfico y videográfico de campo

Registro visual sistemático de las condiciones físicas de la PE-28A obtenido mediante fotografías y videos capturados durante recorridos. Constituyó un instrumento de validación y complementación de los datos geométricos del Civil 3D, con énfasis en el estado del pavimento, señalización vial, taludes y configuraciones geométricas críticas de los tramos de mayor siniestralidad.

- **Fuente:** Trabajo de campo directo en la PE-28A.
- **Cobertura:** Tramos críticos y zonas de alta siniestralidad identificadas en la vía.
- **Formato:** Fotografías digitales y videos de recorrido.

3.9.2.6 Resumen de instrumentos por unidad de análisis

Tabla 11 *Instrumentos de recolección de datos empleados en la investigación*

N°	Instrumento	Unidad de análisis	Fuente	Registros / Cobertura
1	Registros oficiales de siniestros viales	RNA N°1 - Binario	ONSV	202 accidentes reales registrados en la PE-28A, período 2020–2025
2	Datos meteorológicos de estaciones climatológicas	RNA N°1 - Binario	SENAMHI	Temperatura máxima diaria (°C) y precipitación acumulada (mm/día) en los departamentos de Ica, Huancavelica y Ayacucho, período 2020–2025
3	Ficha de análisis geométrico en Civil 3D	RNA N°2 - Multiclase	DTM PROVÍAS Nacional	2,997 tramos geométricos de los 333.903 km de la carretera PE-28A
4	Ficha de análisis documental normativo comparativo	Análisis Normativo	Normativas oficiales de 6 países	12 parámetros geométricos comparados entre DG-2018 (Perú), Noruega, Suecia, Islandia, Reino Unido y Japón
5	Registro fotográfico y videográfico de campo	Validación de campo	Trabajo de campo PE-28A	Tramos críticos y zonas de alta siniestralidad identificadas a lo largo de la PE-28A

Fuente: Elaboración propia.

3.10 Técnicas estadísticas para el procesamiento de la información

El procesamiento de la información se organizó en tres bloques según cada unidad de análisis. Para cada bloque se describen las técnicas estadísticas empleadas, su fundamento metodológico y su aplicación específica en la investigación. Los valores numéricos obtenidos de la aplicación de estas técnicas se presentan y analizan en detalle en el Capítulo IV de Resultados.

Tabla 12 *Esquema general del procesamiento de datos por unidad de análisis*

Unidad de análisis	Entrada	Técnicas aplicadas	Salida
RNA N°1: Binario	202 registros / 23 variables (ONSV + SENAMHI)	EDA → Mann-Whitney U → Z-Score → División 70/20/10 → MLP	Riesgo Menor (0) / Riesgo Mayor (1)
RNA N°2: Multiclase	2,997 registros / 15 variables (Civil 3D)	EDA → SMOTE → Z-Score → División 70/20/10 → MLP	Riesgo Bajo (0) / Medio (1) / Alto (2)
Análisis Normativo	12 parámetros × 6 normativas	Comparación sistemática → Índice de brecha → Clasificación de adecuación	Nivel de adecuación DG-2018
Programa de Alerta	202 registros georreferenciados (ONSV 2020–2025) + coordenadas GPS en tiempo real	Georreferenciación.	Alerta sonora y de vibración al conductor

Fuente: Elaboración propia.

3.10.1 Procesamiento de datos: Modelo RNA N°1

3.10.1.1 Análisis exploratorio de datos (EDA)

Estadística descriptiva. Se empleó para caracterizar la distribución inicial de las 23 variables de entrada del dataset ONSV-SENAMHI (202 registros). Se calcularon frecuencias absolutas y relativas para variables nominales y ordinales, y medidas de tendencia central (media, mediana) y dispersión (desviación estándar, rango intercuartílico — IQR) para variables cardinales como temperatura máxima, precipitación, radio de curva y longitud de tangente. Se realizó además la verificación de valores nulos por columna como paso previo a la limpieza del dataset.

Análisis de valores atípicos — Método IQR. Se identificaron outliers en cada variable

bajo el criterio $[Q1 - 1.5 \cdot IQR,\ Q3 + 1.5 \cdot IQR]$. Su propósito fue evaluar el impacto de valores extremos en el proceso de *backpropagation*, dado que los outliers pueden distorsionar los gradientes y dificultar la convergencia del modelo.

3.10.1.2 Selección de variables — Prueba de Mann-Whitney U

Prueba no paramétrica aplicada para determinar qué variables presentan diferencias estadísticamente significativas entre los grupos Riesgo Menor (0) y Riesgo Mayor (1). Las hipótesis planteadas fueron:

- H_0: No existe diferencia en la distribución de la variable entre ambos grupos
- H_1: Existe diferencia estadísticamente significativa ($p < 0.05$)

Se seleccionaron para el entrenamiento únicamente las variables que rechazaron H_0 ($p < 0.05$). Esta prueba fue complementada con correlación de Pearson, correlación de Spearman y *effect size* de Cohen's *d*, generando un score compuesto ponderado para el ranking final de variables:

$$\text{Score} = 0.35\times\text{Sig(Mann-Whitney)}+0.30\times|r_{\text{Pearson}}|+0.20\times|\rho_{\text{Spearman}}|+0.15\times|d_{\text{Cohen}}| \tag{3.1}$$

Tabla 13 *Criterios del score compuesto ponderado para el ranking de variables — RNA N°1*

Criterio	Peso	Fundamento
Significancia Mann-Whitney U	35%	Diferencia estadística entre grupos Riesgo Menor y Riesgo Mayor
Correlación de Pearson $\|r\|$	30%	Relación lineal entre la variable independiente y la variable dependiente
Correlación de Spearman $\|\rho\|$	20%	Relación monótona entre la variable independiente y la variable dependiente
Effect size Cohen's *d*	15%	Magnitud de la diferencia entre los grupos de ambas clases
Total	**100%**	

Fuente: Elaboración propia.

3.10.1.3 Preprocesamiento: Normalización y división del dataset

Normalización Z-Score (StandardScaler). Se aplicó estandarización Z-Score a todas las variables de entrada ($\mu = 0$, $\sigma = 1$), ajustada exclusivamente sobre el conjunto de entrenamiento para evitar fuga de información (*data leakage*). Esta técnica garantiza que variables con diferentes unidades y magnitudes metros, grados, °C, mm ingresen a la RNA en igual escala, evitando que variables de mayor rango dominen el proceso de aprendizaje.

División estratificada del dataset. El dataset fue dividido en tres particiones con estratificación por clase (*stratify=y*), preservando la proporción de clases en cada sub-conjunto:

Tabla 14 *Distribución del dataset por partición — Modelo RNA N°1*

Partición	Porcentaje	Registros	Uso
Entrenamiento (Train)	70%	141	Ajuste de pesos de la RNA
Validación (Val)	20%	40	Monitoreo del aprendizaje
Evaluación final (Test)	10%	20	Evaluación independiente del modelo
Total	**100%**	**202**	

Fuente: Elaboración propia.

3.10.1.4 Entrenamiento del Modelo RNA N°1

Se empleó el algoritmo MLPClassifier (*Multilayer Perceptron Classifier*) de la librería Scikit-learn. Para la selección de la arquitectura óptima se evaluaron sistemáticamente 7 configuraciones de capas ocultas y 8 valores del parámetro de regularización L2 (*alpha* $\in$ [0.0001, 0.001, 0.005, 0.01, 0.05, 0.1, 0.3, 0.5]), seleccionando los valores que minimizaron el sobreajuste (*overfitting*) sin caer en subajuste (*underfitting*).

Tabla 15 *Arquitecturas evaluadas y selección final del Modelo RNA N°1*

Arquitectura	Capas ocultas	Resultado	Seleccionada
(32, 16)	2	Subajuste (*underfitting*)	No
(64, 32)	2	Exactitud moderada	No
(64, 32, 16)	3	Mejora parcial	No
(64, 48, 32, 16, 8)	**5**	**Mejor balance train/val**	**Sí**
(128, 64, 32)	3	Sobreajuste leve	No
(128, 64, 32, 16)	4	Sobreajuste moderado	No
(128, 64, 32, 16, 8)	5	Sobreajuste alto	No

Fuente: Elaboración propia.

Tabla 16 *Hiperparámetros finales del Modelo RNA N°1*

Hiperparámetro	Valor y justificación
Arquitectura	23 → 64 → 48 → 32 → 16 → 8 → 2 (topología piramidal escalonada — 5 capas ocultas)
Función de activación	ReLU: Evita el problema del gradiente desvaneciente
Optimizador	Adam: Convergencia adaptativa y estable
Regularización L2 (*alpha*)	0.05: Valor óptimo seleccionado tras búsqueda sistemática
Batch size	64: Gradientes más estables que mini-batch pequeño
Tasa de aprendizaje	Adaptativa: Ajuste automático según convergencia
Early stopping	Activado con restauración de mejores pesos
Máximo de iteraciones	400 épocas con parada anticipada

Fuente: Elaboración propia.

3.10.1.5 Validación y evaluación del Modelo RNA N°1

Se aplicaron las siguientes técnicas de validación y diagnóstico:

- **Validación cruzada estratificada 5-Fold (StratifiedKFold):** divide el dataset en 5 particiones, entrenando en 4 y evaluando en 1 de forma rotativa. Proporciona 5 estimaciones independientes del rendimiento, evaluando accuracy, F1-macro y F1-weighted en cada fold para verificar la estabilidad y generalización del modelo.
- **Curva de aprendizaje (diagnóstico Bias-Variance):** varía el tamaño del conjunto de entrenamiento de 10% a 100%, diagnosticando sobreajuste (gap amplio entre curvas train y validación), subajuste (ambas curvas estancadas) o buen ajuste (curvas convergentes hacia valores altos).
- **Curva de pérdida (*loss curve*):** registra la *cross-entropy loss* por época para entrenamiento y validación, diagnosticando la convergencia del modelo y la presencia de sobreajuste.
- **Importancia de variables por permutación:** barajea aleatoriamente los valores de cada variable (20 repeticiones) y mide la caída de exactitud del modelo. Una caída grande indica alta importancia de la variable; una caída cercana a cero indica escaso aporte predictivo.
- **Fronteras de decisión:** visualiza la frontera de clasificación de la RNA para cada variable versus la variable Top-1 del ranking estadístico, mediante una malla 2D de predicciones.

Tabla 17 *Métricas de evaluación empleadas: Modelo RNA N°1*

Métrica	Fórmula	Propósito
Exactitud (*Accuracy*)	$(TP + TN)$/Total	Proporción global de predicciones correctas
AUC-ROC	Área bajo curva ROC	Capacidad discriminante del modelo entre clases
Precisión (*Precision*)	$TP/(TP + FP)$	Calidad de las predicciones positivas
Sensibilidad (*Recall*)	$TP/(TP + FN)$	Capacidad de detectar los positivos reales
F1-Score	$2 \times (P \times R)/(P + R)$	Balance entre Precisión y Recall
Matriz de confusión	Binaria 2×2	Diagnóstico completo de aciertos y errores

Fuente: Elaboración propia.

3.10.2 Procesamiento de datos: Modelo RNA N°2

3.10.2.1 Análisis exploratorio de datos (EDA)

Estadística descriptiva y análisis de desbalance de clases. Se analizó la distribución de los 2,997 registros geométricos entre las tres clases de riesgo (Bajo, Medio, Alto), identificando el desbalance de clases como el problema estadístico central. Un dataset desbalanceado genera accuracy artificialmente alta pero F1-macro bajo, ya que el modelo tiende a ignorar las clases minoritarias durante el entrenamiento.

Análisis de valores atípicos — Método IQR. Aplicado sobre las 15 variables geométricas del dataset para documentar el porcentaje de outliers por variable y evaluar su impacto en el entrenamiento multiclase, considerando que la normalización StandardScaler reduce pero no elimina el efecto de los valores extremos.

Análisis de variables de peligro. Se realizó un análisis específico de las cuatro variables de peligro categóricas mediante distribución de frecuencias por nivel de peligro y clase de riesgo, confirmando su capacidad predictiva para el modelo multiclase.

3.10.2.2 Balanceo de clases — Técnica SMOTE

SMOTE (*Synthetic Minority Oversampling Technique*) genera muestras sintéticas de las clases minoritarias interpolando entre registros reales existentes en el espacio de características. Se aplicó para equilibrar las tres clases de riesgo antes del entrenamiento, garantizando que el modelo no favorezca ninguna clase durante el proceso de aprendizaje:

Tabla 18 *Aplicación de SMOTE para el balanceo de clases — Modelo RNA N°2*

Clase	Registros originales	Registros tras SMOTE	Observación
Riesgo Bajo (0)	Variable	Balanceado	Clase de referencia
Riesgo Medio (1)	Variable	Balanceado	+ Registros sintéticos
Riesgo Alto (2)	Variable	Balanceado	+ Registros sintéticos
Total	**2,997**	**5,394**	**+2,397 registros sintéticos**

Fuente: Elaboración propia.

3.10.2.3 Preprocesamiento:Normalización y división del dataset

Se aplicó la misma normalización Z-Score del Modelo RNA N°1, ajustada sobre el conjunto de entrenamiento expandido por SMOTE. La división estratificada se realizó sobre los 5,394 registros:

Tabla 19 *Distribución del dataset por partición del Modelo RNA N°2*

Partición	Porcentaje	Registros	Uso
Entrenamiento (Train)	70%	2,097	Ajuste de pesos de la RNA
Validación (Val)	20%	600	Monitoreo del aprendizaje
Evaluación final (Test)	10%	300	Evaluación independiente del modelo
Total	**100%**	**2,997**	

Fuente: Elaboración propia.

3.10.2.4 Entrenamiento del Modelo RNA N°2

Se empleó el mismo algoritmo MLPClassifier con arquitectura optimizada para clasificación multiclase de 3 clases. Se evaluaron 7 arquitecturas alternativas y 8 valores de alpha, seleccionando la arquitectura de 4 capas ocultas con mejor balance entre entrenamiento y validación:

Tabla 20 *Arquitecturas evaluadas y selección final del Modelo RNA N°2*

Arquitectura	Capas ocultas	Resultado	Seleccionada
(32, 16)	2	Subajuste multiclase	No
(64, 32)	2	Exactitud moderada	No
(64, 32, 16)	3	Mejora parcial	No
(128, 64, 32, 16)	**4**	**Mejor balance train/val**	**Sí**
(64, 48, 32, 16, 8)	5	Sobreajuste leve	No
(128, 64, 32, 16, 8)	5	Sobreajuste moderado	No

Fuente: Elaboración propia.

Tabla 21 *Hiperparámetros finales del Modelo RNA N°2*

Hiperparámetro	Valor y justificación
Arquitectura	15 → 128 → 64 → 32 → 16 → 3 (expansión inicial con reducción progresiva)
Función de activación	ReLU — evita el problema del gradiente desvaneciente
Optimizador	Adam — convergencia adaptativa y estable
Regularización L2 (*alpha*)	0.05 — valor óptimo seleccionado tras búsqueda sistemática
Tasa de aprendizaje	Adaptativa — ajuste automático según convergencia
Early stopping	Activado con restauración de mejores pesos
Pesos de clase	Balanceados — compensación del desbalance original pre-SMOTE
Máximo de iteraciones	400 épocas con parada anticipada

Fuente: Elaboración propia.

3.10.2.5 Validación y evaluación del Modelo RNA N°2

Se aplicaron las mismas técnicas de validación del Modelo RNA N°1, adaptadas al contexto multiclase con métricas calculadas por clase (Riesgo Bajo, Riesgo Medio, Riesgo Alto):

Tabla 22 *Métricas de evaluación empleadas — Modelo RNA N°2*

Métrica	Fórmula	Propósito
Exactitud (*Accuracy*)	$(TP + TN)$/Total	Proporción global de predicciones correctas
AUC-ROC multiclase	Área bajo curva ROC	Capacidad discriminante entre las 3 clases de riesgo
Precisión por clase	$TP/(TP + FP)$	Calidad de predicciones por nivel de riesgo
Sensibilidad por clase	$TP/(TP + FN)$	Detección correcta por nivel de riesgo
F1-Score por clase	$2 \times (P \times R)/(P + R)$	Balance Precisión-Recall por nivel de riesgo
Matriz de confusión 3×3	Multiclase	Diagnóstico de errores entre niveles de riesgo
Curva de pérdida	*Cross-entropy*	Convergencia del modelo por época

Fuente: Elaboración propia.

3.10.3 Procesamiento de datos: Análisis Normativo Comparativo

3.10.3.1 Análisis comparativo de parámetros geométricos

Se realizó la comparación sistemática de 12 parámetros geométricos de la DG-2018 frente a los estándares de 5 normativas internacionales. Para cada parámetro se obtuvo

el valor establecido por cada normativa para condiciones equivalentes a las de la PE-28A, generando una matriz de parámetros de 12×6.

3.10.3.2 Clasificación del nivel de adecuación

Cada parámetro fue clasificado según el siguiente criterio de adecuación:

Tabla 23 *Criterios de clasificación del nivel de adecuación normativa de la DG-2018*

Clasificación	Criterio	Implicancia
Adecuado	El valor DG-2018 cumple o supera el estándar internacional de referencia	No requiere modificación
Parcialmente adecuado	El valor DG-2018 se aproxima al estándar con una brecha menor al 20%	Requiere revisión y ajuste menor
No adecuado	El valor DG-2018 presenta una brecha superior al 20% respecto al estándar	Requiere propuesta de mejora inmediata

Fuente: Elaboración propia.

3.10.4 Software y herramientas utilizadas

Tabla 24 *Software y herramientas de procesamiento estadístico empleadas en la investigación*

N°	Software / Herramienta	Uso específico en la investigación
1	Python / JupyterLab	Entorno de desarrollo de ambos modelos RNA
2	Scikit-learn (MLPClassifier)	Algoritmo MLP para entrenamiento de ambos modelos RNA
3	Scikit-learn (StandardScaler)	Normalización Z-Score de variables de entrada
4	Scikit-learn (StratifiedKFold)	Validación cruzada estratificada 5-Fold
5	Scikit-learn (learning_curve)	Curva de aprendizaje y diagnóstico Bias-Variance
6	Scikit-learn (permutation_importance)	Importancia de variables por permutación
7	Scikit-learn (PCA)	Reducción dimensional — Modelo RNA N°2
8	Imbalanced-learn (SMOTE)	Balanceo de clases — Modelos RNA N°1 y N°2
9	SciPy (mannwhitneyu)	Prueba estadística Mann-Whitney U — RNA N°1
10	Pandas / NumPy	Procesamiento, limpieza y transformación de datos
11	Matplotlib / Seaborn	Visualización de resultados y figuras diagnósticas
12	Civil 3D (Autodesk)	Modelamiento geométrico de la PE-28A
13	Microsoft Excel	Organización y preprocesamiento inicial de datos

Fuente: Elaboración propia.

3.11 Desarollo del trabajo de tesis

3.11.1 Desarrollo del trabajo de tesis de la Red Neuronal N°01

El presente apartado describe el proceso completo de desarrollo del Modelo de Red Neuronal Artificial N°1 (RNA N°1) para la clasificación binaria del nivel de riesgo de accidentes de tránsito en la Vía Los Libertadores PE-28A. El proceso se organizó en cuatro etapas secuenciales: Diagrama de flujo, Análisis Exploratorio de Datos, Preprocesamiento de datos, Diseño, optimización y entrenamiento de la RNA y Evaluación e interpretación del Modelo RNA N°1.

Tabla 25 *Ficha técnica del Modelo RNA N°1*

Parámetro	**Valor**
Institución	Universidad Nacional de San Cristóbal de Huamanga
Escuela Profesional	Ingeniería Civil
Vía de estudio	Los Libertadores (PE-28A) — Ayacucho, Perú
Dataset	DATOS_SINIESTROS.csv — 202 registros reales (ONSV 2020–2025)
Variables de entrada	23 variables (geométricas, climáticas, temporales y de peligro)
Variable objetivo	ACCIDENTE DE TRÁNSITO (binaria: 0 = Riesgo Menor / 1 = Riesgo Mayor)
Arquitectura RNA	MLP: 23→64→48→32→16→8→2 (5 capas ocultas, 6,906 parámetros)
Activación / Salida	ReLU (capas ocultas) | Sigmoid (capa de salida)
Optimizador	Adam | Regularización L2 α=0.05 | Batch=64 | Early Stopping pac=20
Épocas ejecutadas	48 / 200 (máximo)
Accuracy — Test	90.48% ✓ (umbral excelencia ≥90%)
AUC-ROC — Macro	0.9818 ✓ (umbral excelencia ≥0.90)
F1-Score Macro — Test	0.9045 ✓ (umbral excelencia ≥0.90)
CV 5-Fold Accuracy	0.886 ± 0.033
Software	Python / JupyterLab / Scikit-learn / Pandas / Matplotlib
Año	2026

Fuente: Elaboración propia.

3.11.1.1 Diagrama de flujo

Diagrama de flujo

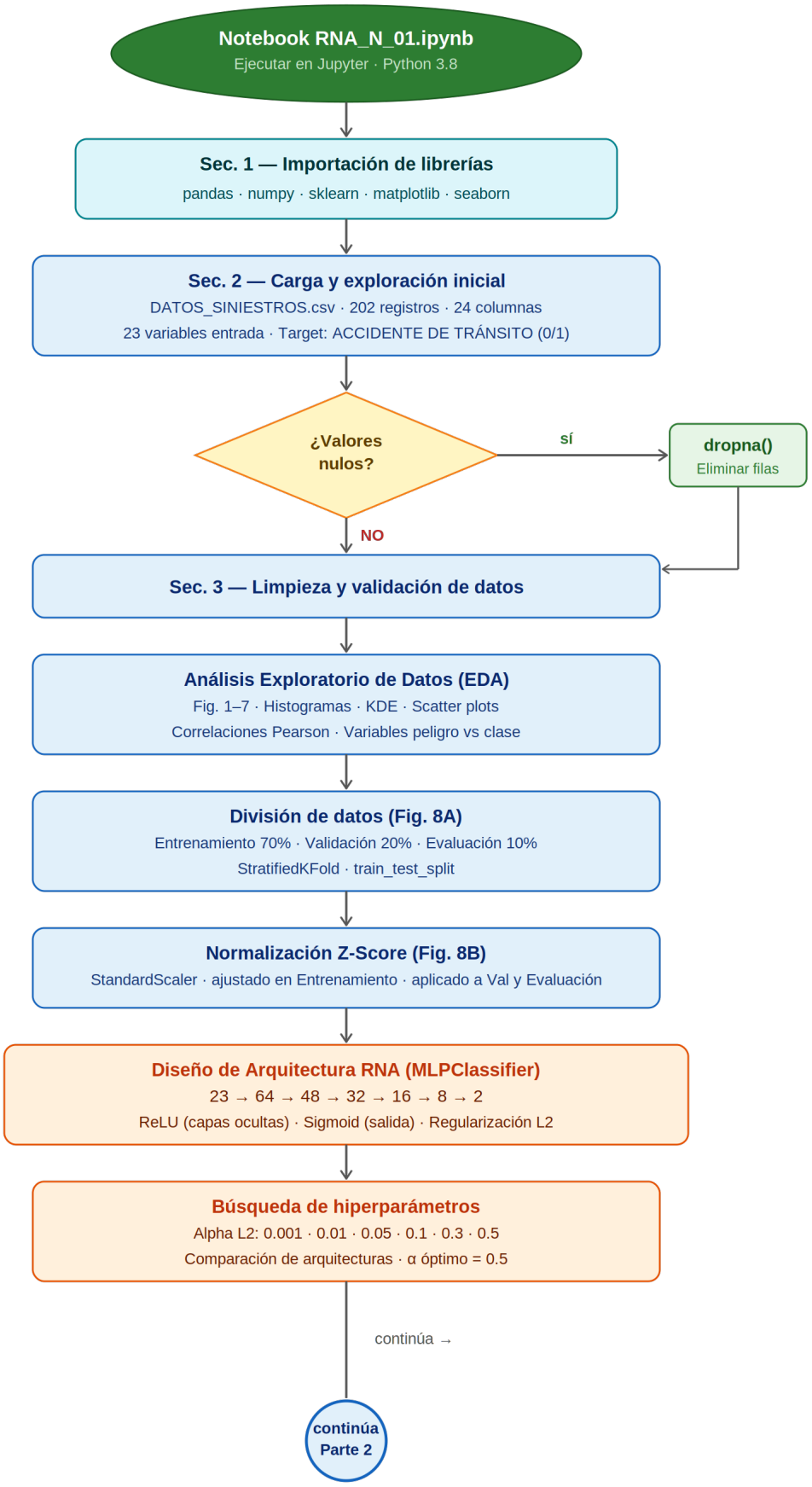


Figura 56 Diagrama de flujo de la RNA N°01 - Parte 1
Fuente: Elaboración propia

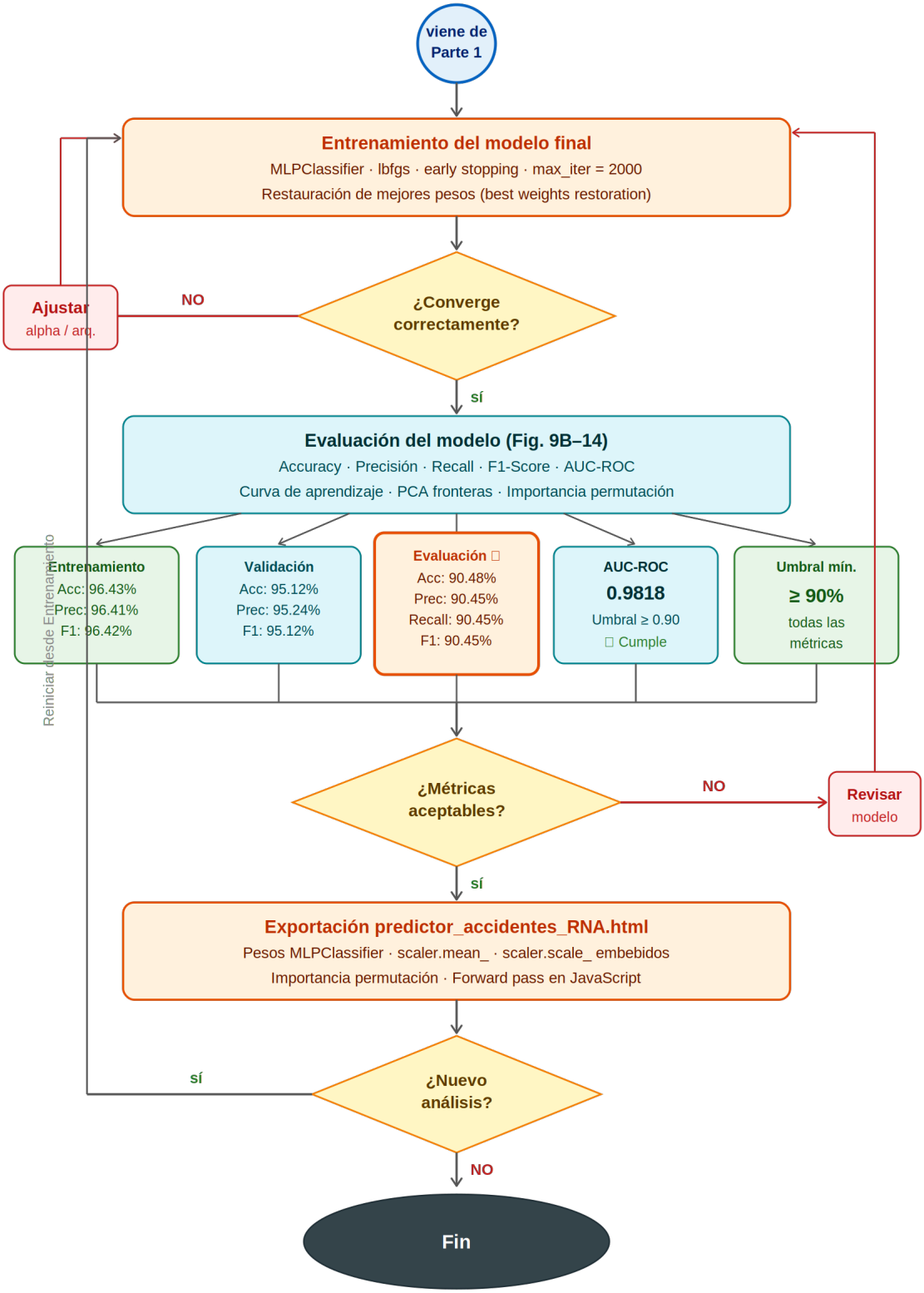


Figura 57 Diagrama de flujo de la RNA N°01 - Parte 2
Fuente: Elaboración propia

3.11.1.2 Registro de datos de la RNA N°01

La base de datos empleada para la RNA N°01 fue construida a partir de los registros oficiales de siniestros viales proporcionados por la Oficina Nacional de Seguridad Vial, correspondientes al tramo ICA-AYACUCHO de la carretera PE-28A (Vía Los Libertadores), para el período 2020–2025, cuyas características se detallan a continuación:

- **Fuente principal:** Registros de accidentes de tránsito de la ONSV, período 2020-2025.
- **Vía:** Carretera PE-28A (Vía Los Libertadores), tramo ICA-AYACUCHO.
- **Total:** 202 registros de siniestros georreferenciados.
- **Variables:** 23 variables de entrada para la RNA mediante análisis de importancia y correlación.
- **Variable objetivo:** Clasificación binaria del nivel de riesgo, riesgo menor (0) y riesgo mayor (1).
- **Fuentes complementarias:** Datos meteorológicos de temperatura y precipitación obtenidos de SENAMHI.
- **Instituciones involucradas:** PROVÍAS Nacional, Policía Nacional del Perú y Ministerio de Transportes y Comunicaciones.

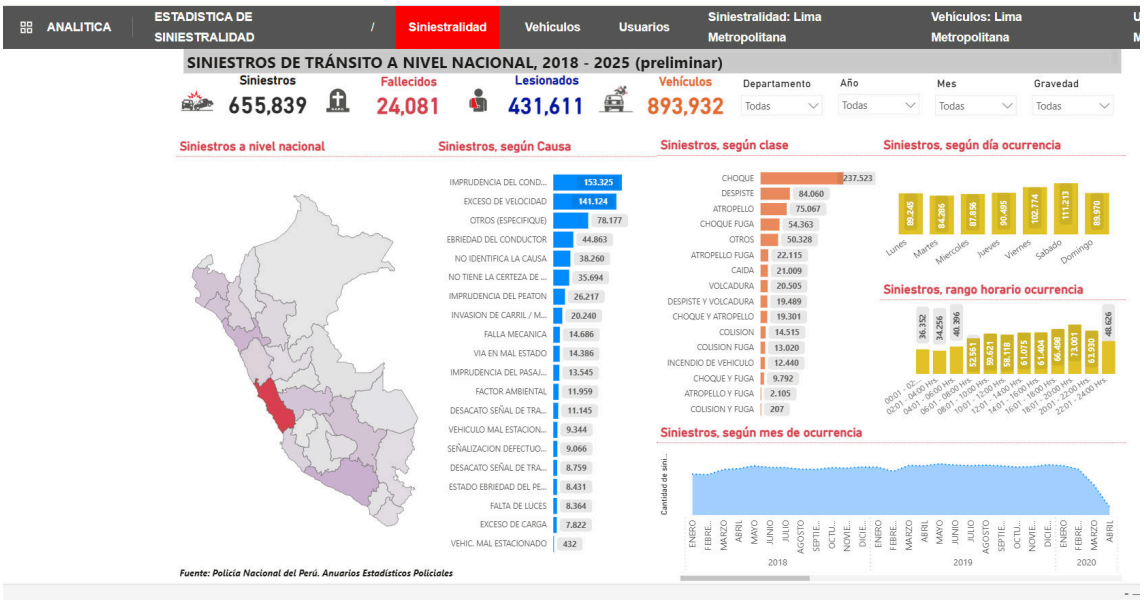
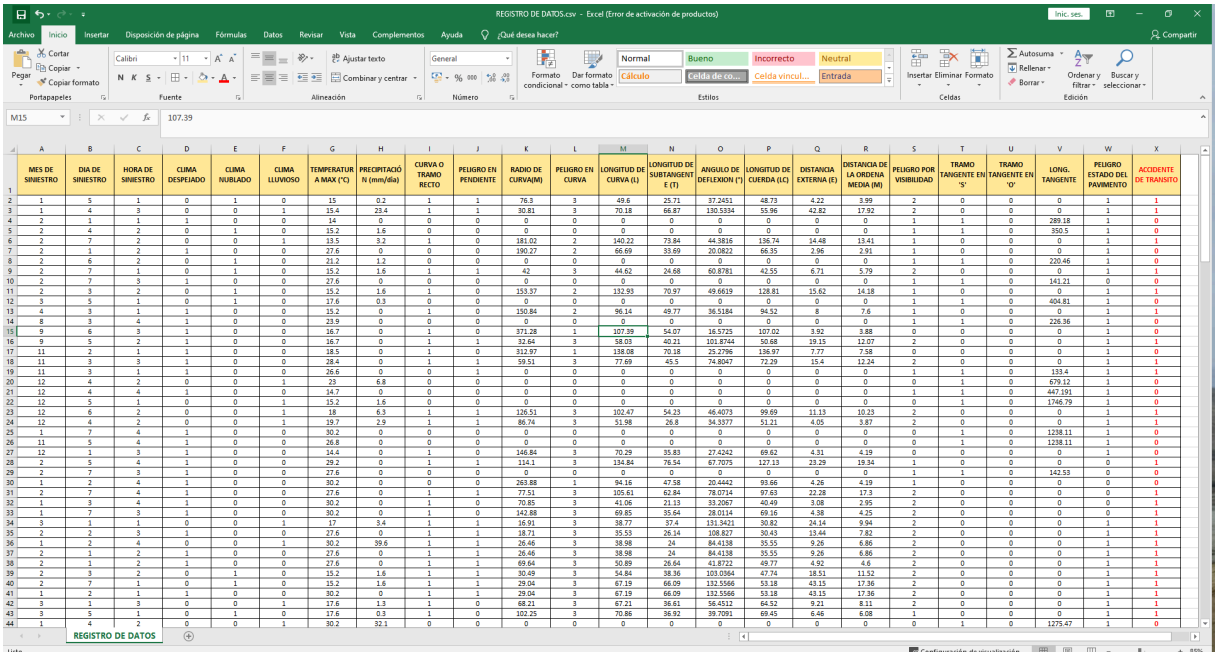


Figura 58 Registro de accidentes de tránsito a nivel nacional

Fuente: Observatorio Nacional de Seguridad Vial (ONSV)

.



MES DE SINIESTRO	DIA DE SINIESTRO	HORA DE SINIESTRO	CLIMA DESPEJADO	CLIMA NUBLADO	CLIMA LLUVIOSO	TEMPERATURA MAX (°C)	PRECIPITACIÓN (mm/día)	CURVA O TRAMO RECTO	PELIGRO EN PENDIENTE	RADIO DE CURVA(M)	PELIGRO EN CURVA	LONGITUD DE CURVA (L)	LONGITUD DE SUBTANGENTE (T)	ANGULO DE DEFLEXION (°)	LONGITUD DE CUERDA (LC)	DISTANCIA EXTERNA (E)	DISTANCIA DE LA ORDENA MEDIA (M)	PELIGRO POR VISIBILIDAD	TRAMO TANGENTE EN 'S'	TRAMO TANGENTE EN 'O'	LONG. TANGENTE	PELIGRO ESTADO DEL PAVIMENTO	ACCIDENTE DE TRANSITO
1	5	1	0	1	0	15	0.2	1	1	76.3	3	49.6	25.71	37.2451	48.73	4.22	3.99	2	0	0	0	1	1
1	4	3	0	0	1	15.4	23.4	1	1	30.81	3	70.18	66.87	130.5334	55.96	42.82	17.92	2	0	0	0	1	1
2	1	1	1	0	0	14	0	0	0	0	0	0	0	0	0	0	0	1	1	0	289.18	1	0
2	4	2	0	1	0	15.2	1.6	0	0	0	0	0	0	0	0	0	0	1	1	0	350.5	1	0
2	7	2	0	0	1	13.5	3.2	1	0	181.02	2	140.22	73.84	44.3816	136.74	14.48	13.41	1	0	0	0	1	1
2	1	2	1	0	0	27.6	0	0	0	190.27	2	66.69	33.69	20.0822	66.35	2.96	2.91	1	0	0	0	1	0
2	6	2	0	1	0	21.2	1.2	0	0	0	0	0	0	0	0	0	0	1	1	0	220.46	1	0
2	7	1	0	1	0	15.2	1.6	1	1	42	3	44.62	24.68	60.8781	42.55	6.71	5.79	2	0	0	0	1	1
2	7	3	1	0	0	27.6	0	0	0	0	0	0	0	0	0	0	0	1	1	0	141.21	0	0
2	3	2	0	1	0	15.2	1.6	1	0	153.37	2	132.93	70.97	49.6619	128.81	15.62	14.18	1	0	0	0	1	1
3	5	1	0	1	0	17.6	0.3	0	0	0	0	0	0	0	0	0	0	1	1	0	404.81	1	0
4	3	1	1	0	0	15.2	0	1	0	150.84	2	96.14	49.77	36.5184	94.52	8	7.6	1	0	0	0	1	1
8	3	4	1	0	0	23.9	0	0	0	0	0	0	0	0	0	0	0	1	1	0	226.36	1	0
9	6	3	1	0	0	16.7	0	1	0	371.28	1	107.39	54.07	16.5725	107.02	3.92	3.88	0	0	0	0	1	0
9	5	2	1	0	0	16.7	0	1	1	32.64	3	58.03	40.21	101.8744	50.68	19.15	12.07	2	0	0	0	1	1
11	2	1	1	0	0	18.5	0	1	0	312.97	1	138.08	70.18	25.2796	136.97	7.77	7.58	0	0	0	0	1	0
11	3	3	1	0	0	28.4	0	1	1	59.51	3	77.69	45.5	74.8047	72.29	15.4	12.24	2	0	0	0	1	1
11	3	1	1	0	0	26.6	0	0	1	0	0	0	0	0	0	0	0	1	1	0	133.4	1	1
12	4	2	0	0	1	23	6.8	0	0	0	0	0	0	0	0	0	0	0	1	0	679.12	1	0
12	4	4	1	0	0	14.7	0	0	0	0	0	0	0	0	0	0	0	1	1	0	447.191	1	0
12	5	1	0	0	1	15.2	1.6	0	0	0	0	0	0	0	0	0	0	0	1	0	1746.79	1	0
12	6	2	0	0	1	18	6.3	1	1	126.51	3	102.47	54.23	46.4073	99.69	11.13	10.23	2	0	0	0	1	1
12	4	2	0	0	1	19.7	2.9	1	1	86.74	3	51.98	26.8	34.3377	51.21	4.05	3.87	2	0	0	0	1	1
1	7	4	1	0	0	30.2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1238.11	1	0
11	5	4	1	0	0	26.8	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1238.11	1	0
12	1	3	1	0	0	14.4	0	1	0	146.84	3	70.29	35.83	27.4242	69.62	4.31	4.19	0	0	0	0	1	0
2	5	4	1	0	0	29.2	0	1	1	114.1	3	134.84	76.54	67.7075	127.13	23.29	19.34	1	0	0	0	0	1
2	7	3	1	0	0	27.6	0	0	0	0	0	0	0	0	0	0	0	1	1	0	142.53	0	0
1	2	4	1	0	0	30.2	0	0	0	263.88	1	94.16	47.58	20.4442	93.66	4.26	4.19	1	0	0	0	0	0
2	7	4	1	0	0	27.6	0	1	1	77.51	3	105.61	62.84	78.0714	97.63	22.28	17.3	2	0	0	0	0	1
1	3	4	1	0	0	30.2	0	1	0	70.85	3	41.06	21.13	33.2067	40.49	3.08	2.95	2	0	0	0	0	1
1	7	3	1	0	0	30.2	0	1	0	142.88	3	69.85	35.64	28.0114	69.16	4.38	4.25	2	0	0	0	0	1
3	1	1	0	0	1	17	3.4	1	1	16.91	3	38.77	37.4	131.3421	30.82	24.14	9.94	2	0	0	0	1	1
2	2	3	1	0	0	27.6	0	1	1	18.71	3	35.53	26.14	108.827	30.43	13.44	7.82	2	0	0	0	1	1
1	2	4	0	0	1	30.2	39.6	1	1	26.46	3	38.98	24	84.4138	35.55	9.26	6.86	2	0	0	0	1	1
2	1	2	1	0	0	27.6	0	1	1	26.46	3	38.98	24	84.4138	35.55	9.26	6.86	2	0	0	0	1	1
2	1	2	1	0	0	27.6	0	1	1	69.64	3	50.89	26.64	41.8722	49.77	4.92	4.6	2	0	0	0	1	1
2	3	2	0	1	0	15.2	1.6	1	1	30.49	3	54.84	38.36	103.0364	47.74	18.51	11.52	2	0	0	0	1	1
2	7	1	0	1	0	15.2	1.6	1	1	29.04	3	67.19	66.09	132.5566	53.18	43.15	17.36	2	0	0	0	1	1
1	2	1	1	0	0	30.2	0	1	1	29.04	3	67.19	66.09	132.5566	53.18	43.15	17.36	2	0	0	0	1	1
3	1	3	0	0	1	17.6	1.3	1	0	68.21	3	67.21	36.61	56.4512	64.52	9.21	8.11	2	0	0	0	1	1
3	5	1	0	1	0	17.6	0.3	1	0	102.25	3	70.86	36.92	39.7091	69.45	6.46	6.08	1	0	0	0	1	1
1	4	2	0	0	1	30.2	32.1	0	0	0	0	0	0	0	0	0	0	0	1	0	1275.47	1	0

Figura 59 Base de datos de la RNA N°01:202 datos de accidentes.
Fuente: Elaboración propia.

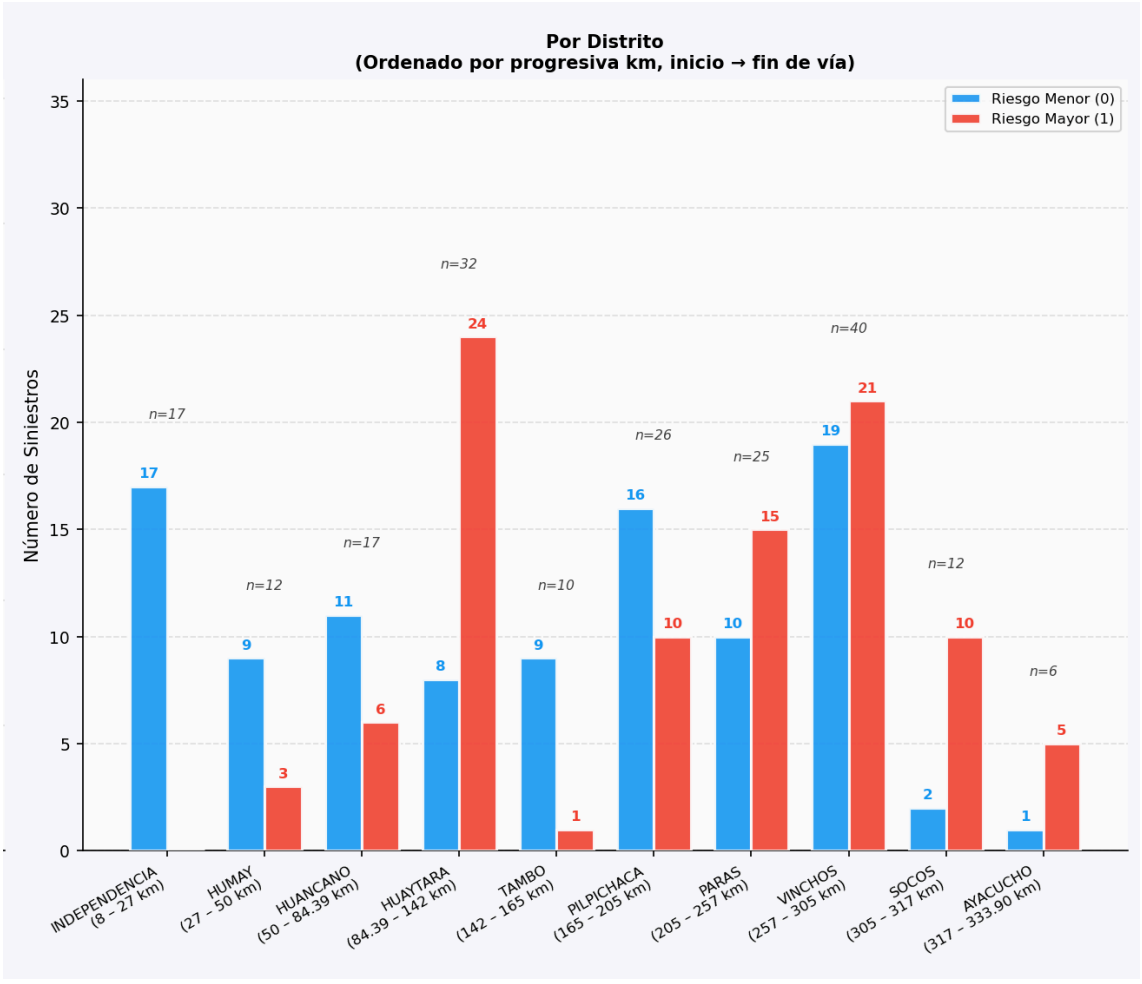


Figura 60 Registro de accidentes por distrito
Fuente: Elaboración propia.

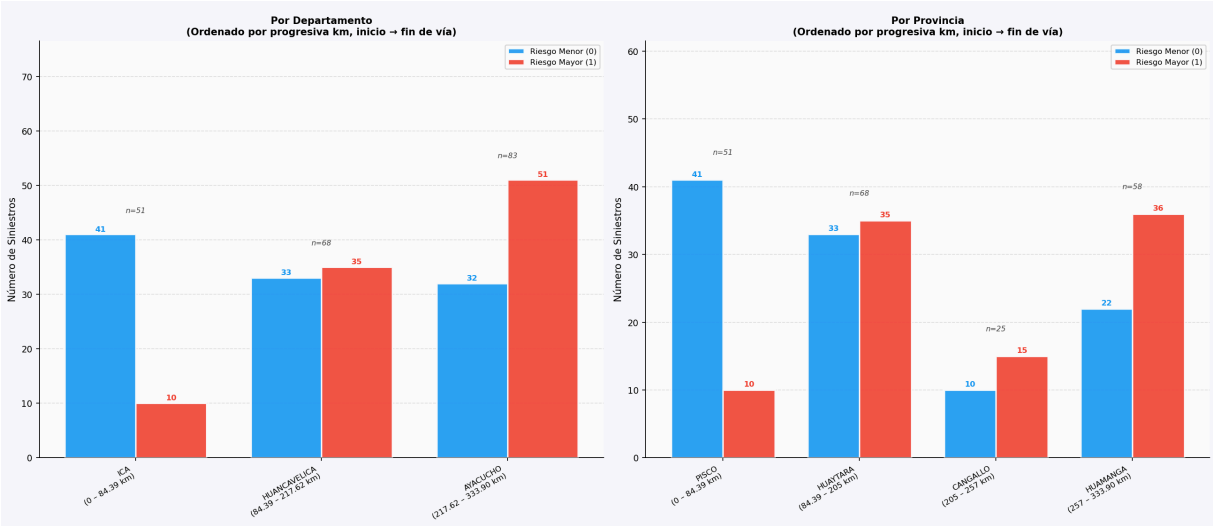


Figura 61 Registro de accidentes por departamento y provincia.
Fuente: Elaboración propia.

3.11.1.3 Etapa I: Análisis exploratorio de datos (EDA)

El análisis exploratorio de datos tuvo como propósito caracterizar el dataset de 202 registros de siniestros viales, evaluar la calidad de los datos y determinar el poder discriminativo de cada una de las 23 variables de entrada respecto a la variable dependiente (accidentes de tránsito). Este análisis guió todas las decisiones posteriores de preprocesamiento y arquitectura del modelo (Montgomery & Runger, 2018).

A) Distribución de la variable objetivo

La Figura siguiente se muestra la distribución de la variable dependiente accidentes de tránsito a través de tres paneles complementarios: frecuencia absoluta (A), proporciones relativas (B) y ratio de desbalance (C). El resultado revela un balance casi perfecto entre clases: RIESGO MENOR (n=106, 52.5%) y RIESGO MAYOR (n=96, 47.5%), con un ratio de 1.10:1. Este balance elimina el riesgo de sesgo de clase y permite que las métricas de accuracy sean directamente representativas del rendimiento real del modelo sin distorsiones por clase mayoritaria (Brownlee, 2020a).

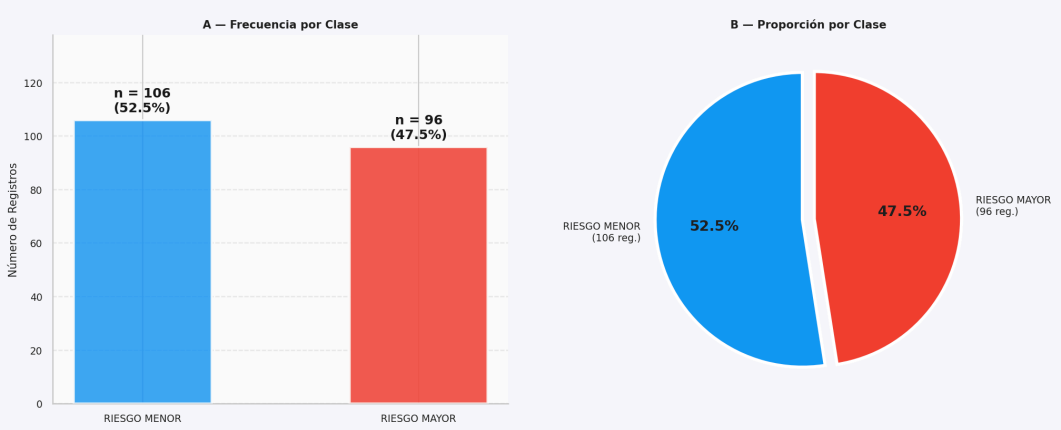


Figura 62 Distrib. de la variable de salida: frecuencia absoluta (A) y proporción (B)
Fuente: Elaboración propia

B) Histogramas y KDE por clase (23 variables)

Las figuras muestran histogramas globales (con media y mediana) y curvas KDE por clase (riesgo menor/mayor) para las 23 variables de entrada. Variables como peligro en curva, ángulo de deflexión, peligro por visibilidad, tipo de tramo (curva/recto) y peligro en pendiente presentan bajo solapamiento entre clases, lo que indica alta capacidad discriminatoria y las convierte en predictores clave. En cambio, variables temporales (mes, día, hora) y peligro por estado del pavimento muestran curvas KDE casi superpuestas, reflejando escaso poder discriminativo.

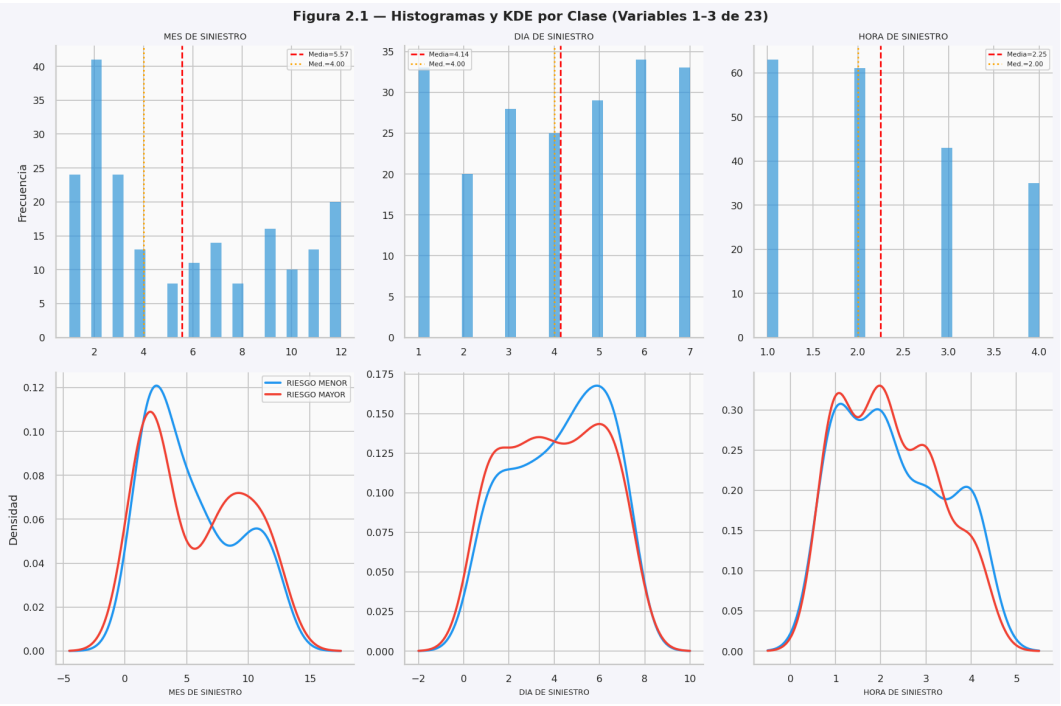


Figura 63 Histogramas y KDE: Mes, Día, Hora de Siniestro.
Fuente: Elaboración propia

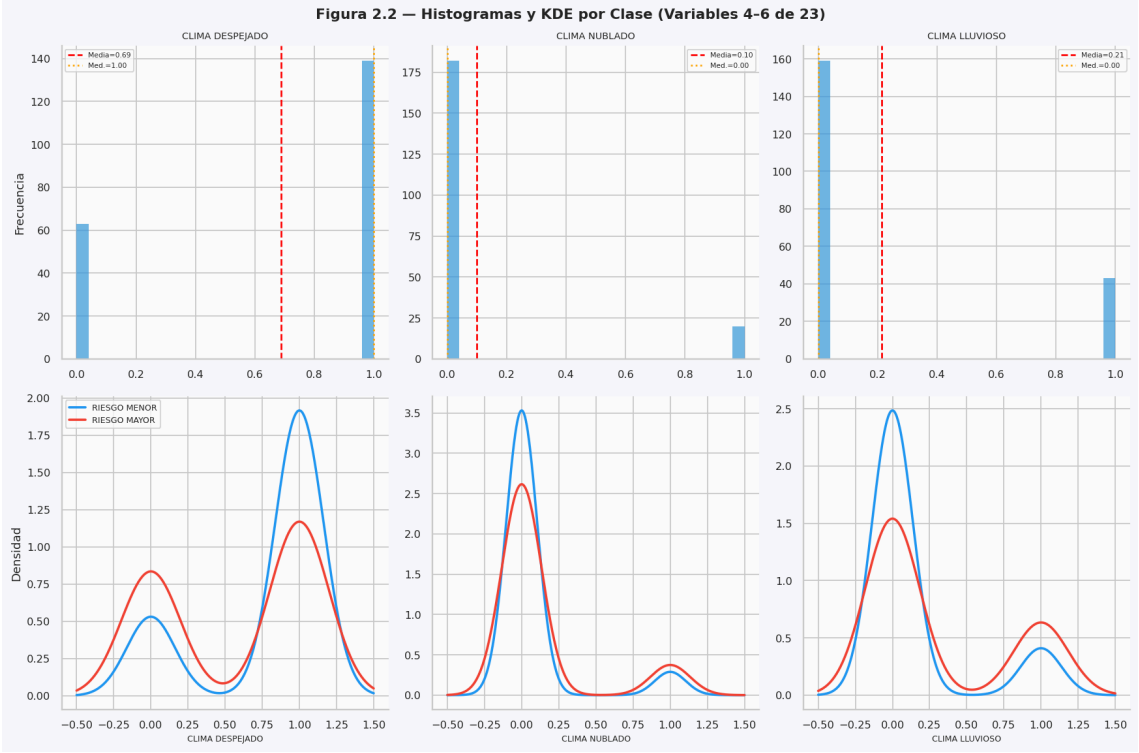


Figura 64 Histogramas y KDE por clase de variables: Clima Despejado, Clima Nublado, Clima Lluvioso.

Fuente: Elaboración propia

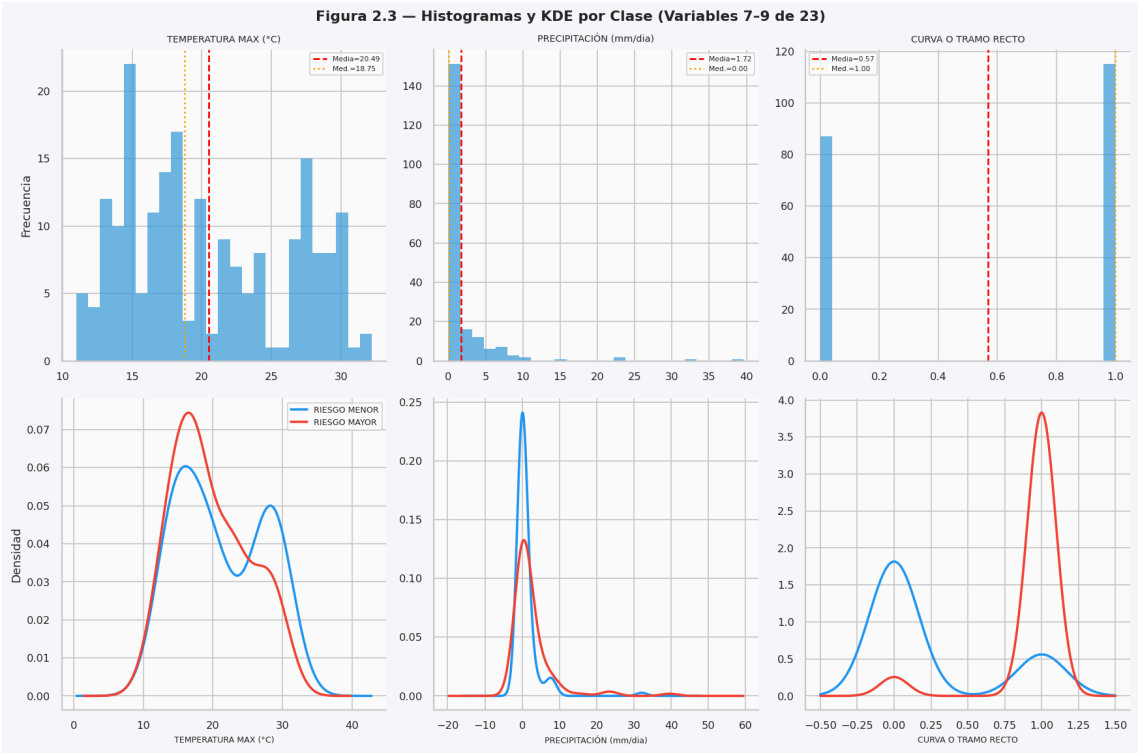


Figura 65 Histogramas y KDE: Temperatura Max (°C), Precipitación (mm/dia) y Curva.

Fuente: Elaboración propia

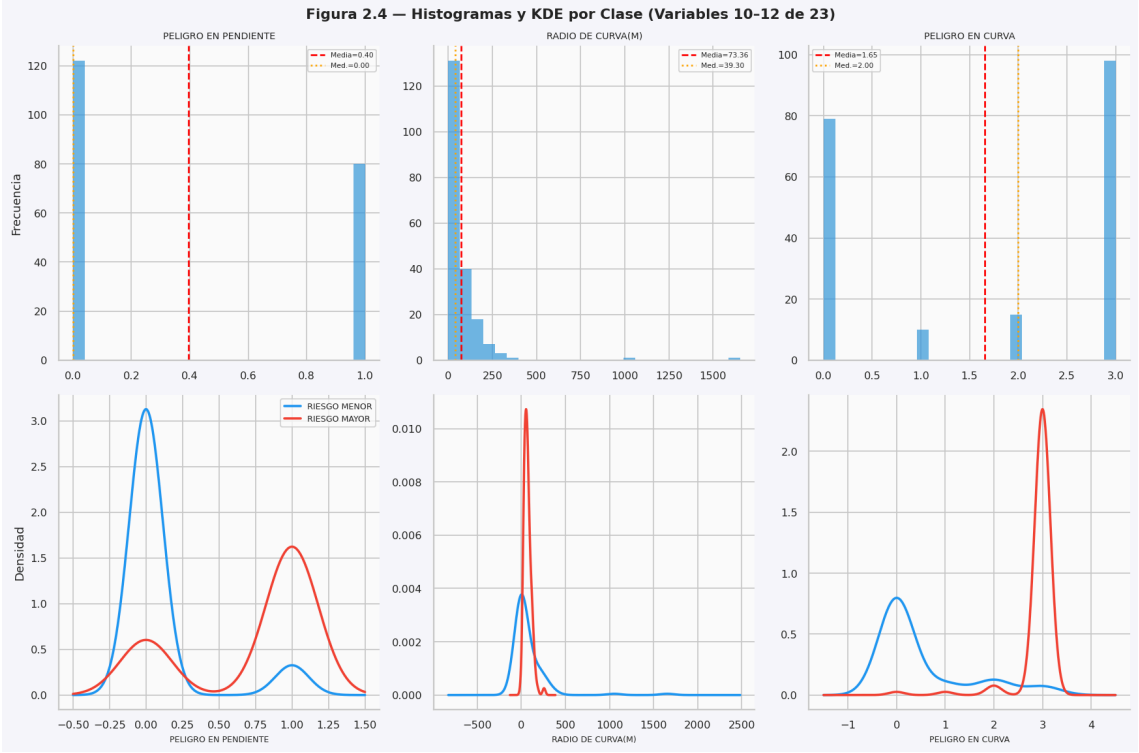


Figura 66 Histogramas y: Peligro en pendiente, Radio de curva (m) y Peligro en curva.
Fuente: Elaboración propia

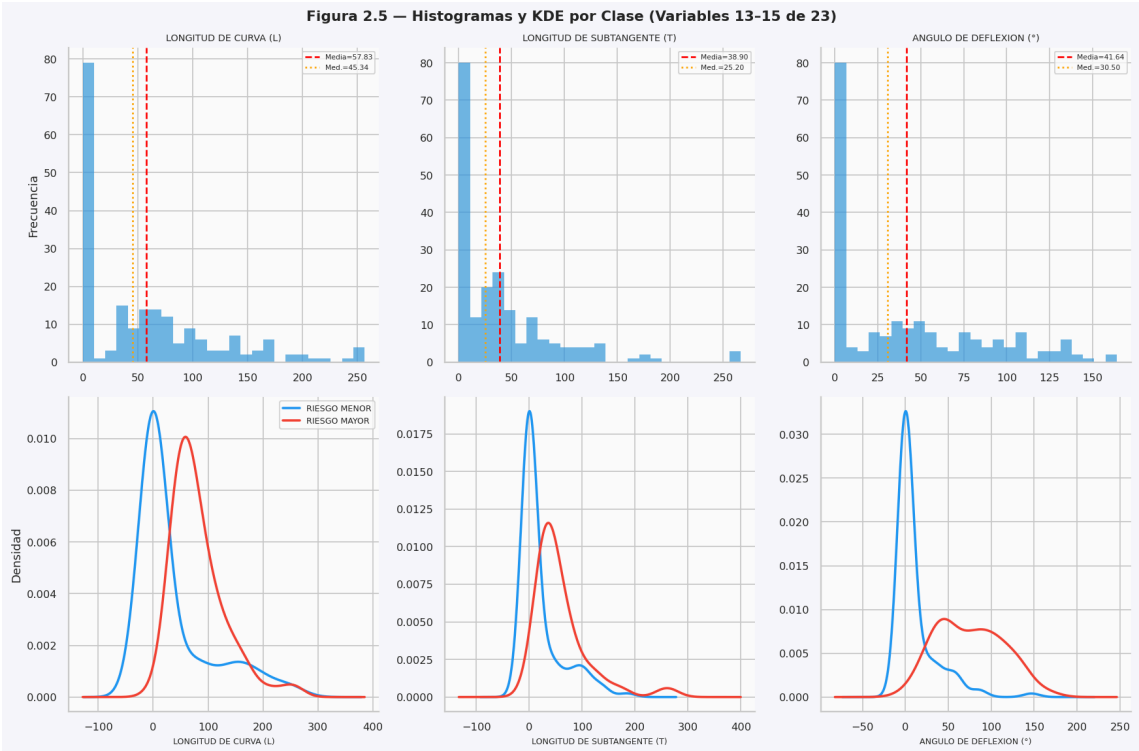


Figura 67 Histogramas y KDE por clase de variables: Longitud de Curva (L), Longitud de subtangente (T) y Angulo de deflexión (°).
Fuente: Elaboración propia

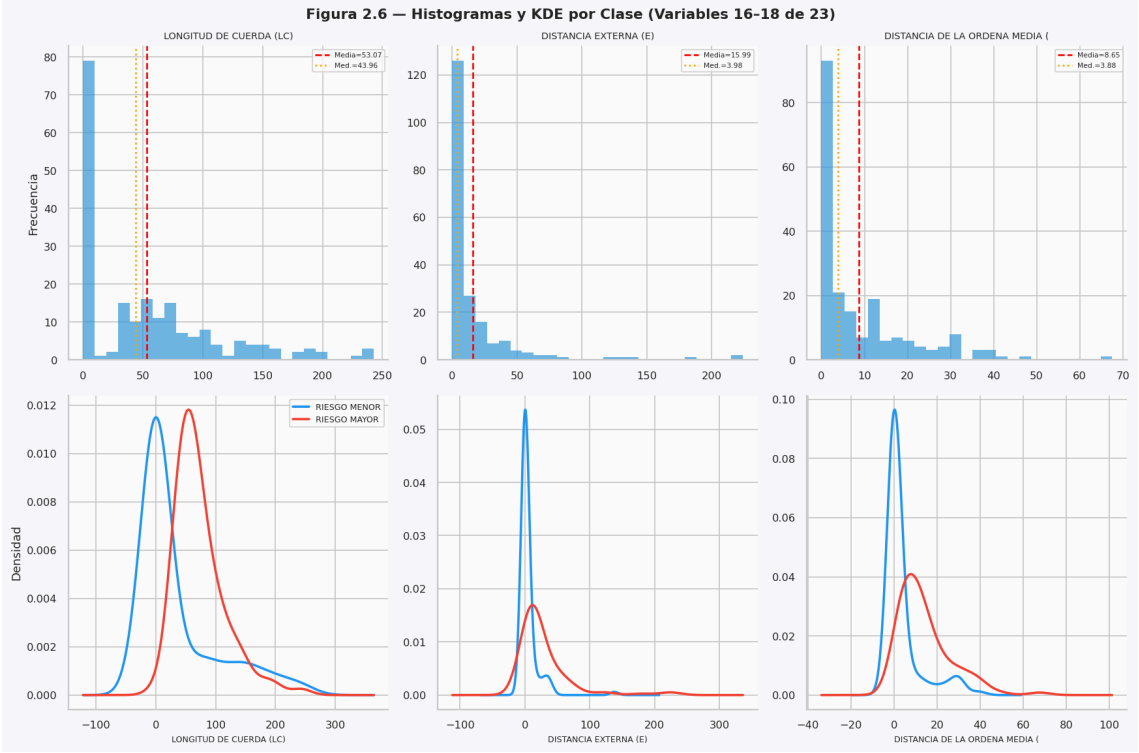


Figura 68 Histogramas y KDE: Longitud de cuerda (LC), Distancia Externa (E) y Distancia de ordenada media

Fuente: Elaboración propia

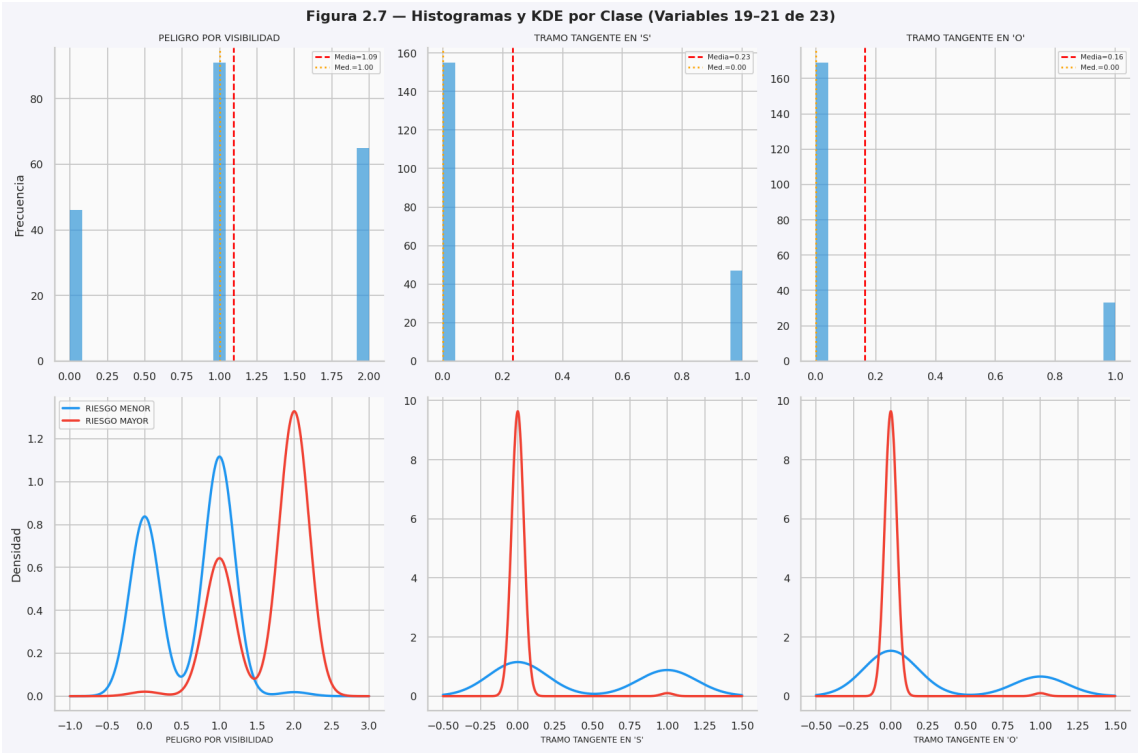


Figura 69 Histogramas y KDE: Peligro por visibilida, Tramo en ’S’ y Tramo en ’O’.

Fuente: Elaboración propia

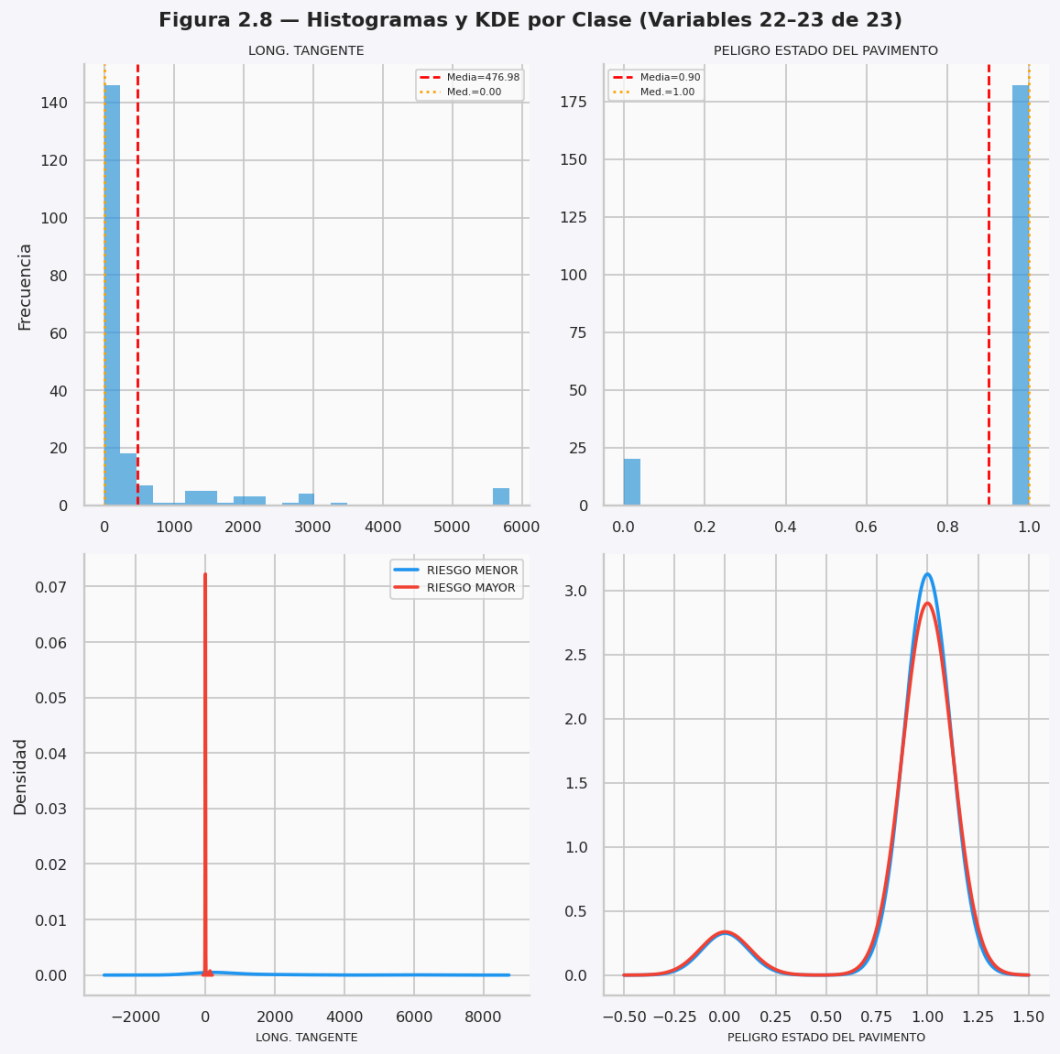


Figura 70 Histogramas y KDE: Long. tangente y Peligro estado del pavimento.
Fuente: Elaboración propia

C) Boxplots por clase de riesgo

Las figuras siguientes presentan los boxplots de las variables continuas separados por clase de riesgo, donde el rombo azul marino indica la media aritmética y las cajas representan el rango intercuartílico Q1–Q3. El ángulo de deflexión exhibe la mayor separación entre medianas: la mediana de riesgo mayor es aproximadamente 5 veces superior a la de riesgo menor, confirmando que las curvas cerradas generan mayor riesgo de accidente grave. Longitud de curva (L), subtangente (T), longitud de cuerda (LC) y distancia externa (E) también presentan medianas claramente superiores en riesgo mayor. Temperatura máxima y mes de siniestro no muestran diferencias notables.

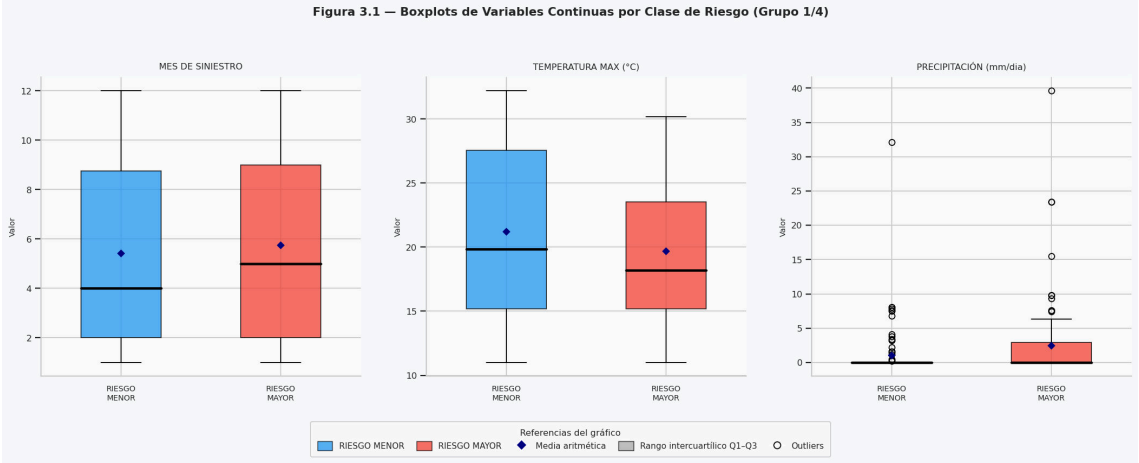


Figura 71 Boxplots de variables: Mes de siniestro, Temperatura max y Precipitación.
Fuente: Elaboración propia

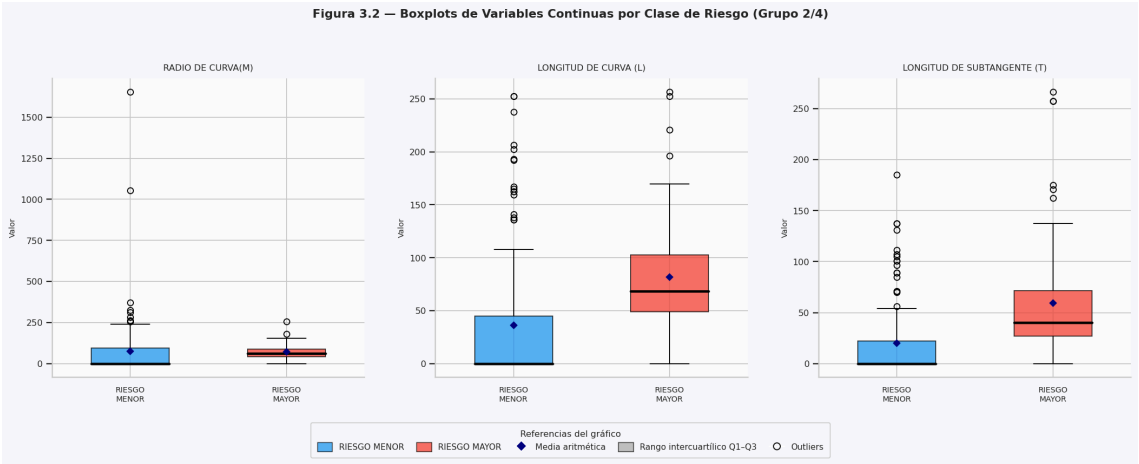


Figura 72 Boxplots de variables: Radio de curva, Longitud de curva y Long. de subtangente.
Fuente: Elaboración propia

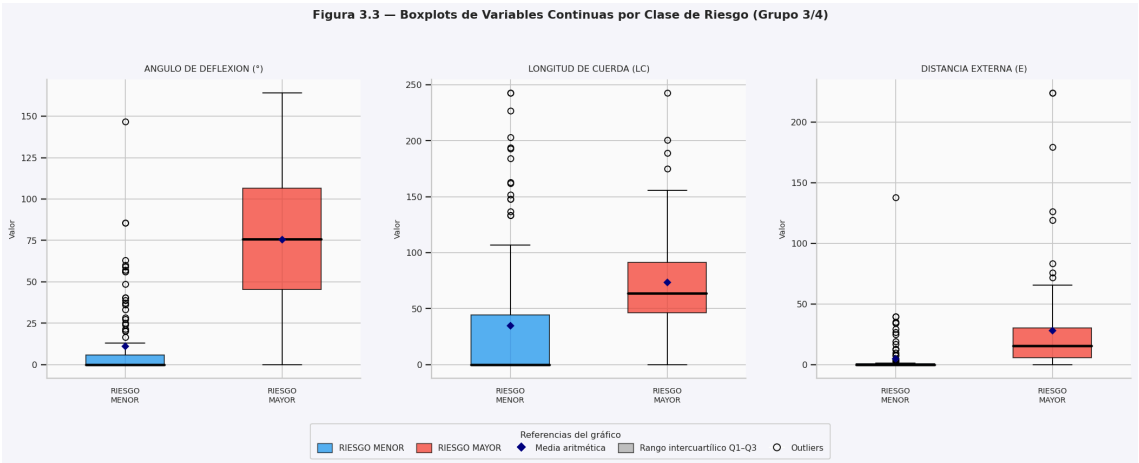


Figura 73 Boxplots de variables: Angulo de deflexión, Long. de cuerda y Dist. externa (E).
Fuente: Elaboración propia

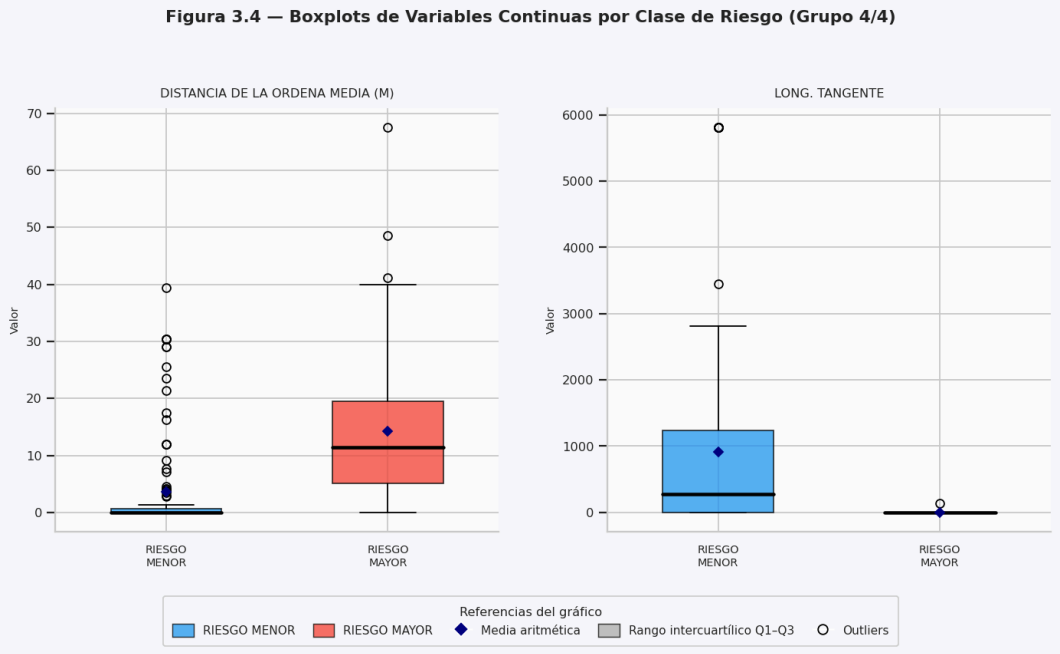


Figura 74 Boxplots de variables: Dist. de la ordenada media y Long. de tangente.
Fuente: Elaboración propia

D) Análisis de outliers (Método IQR)

La figura siguiente presenta el porcentaje de outliers por variable (panel A) y la tabla resumen IQR (panel B), bajo el criterio $[Q1 - 1.5 \cdot IQR,\ Q3 + 1.5 \cdot IQR]$ (Tukey, 1977).

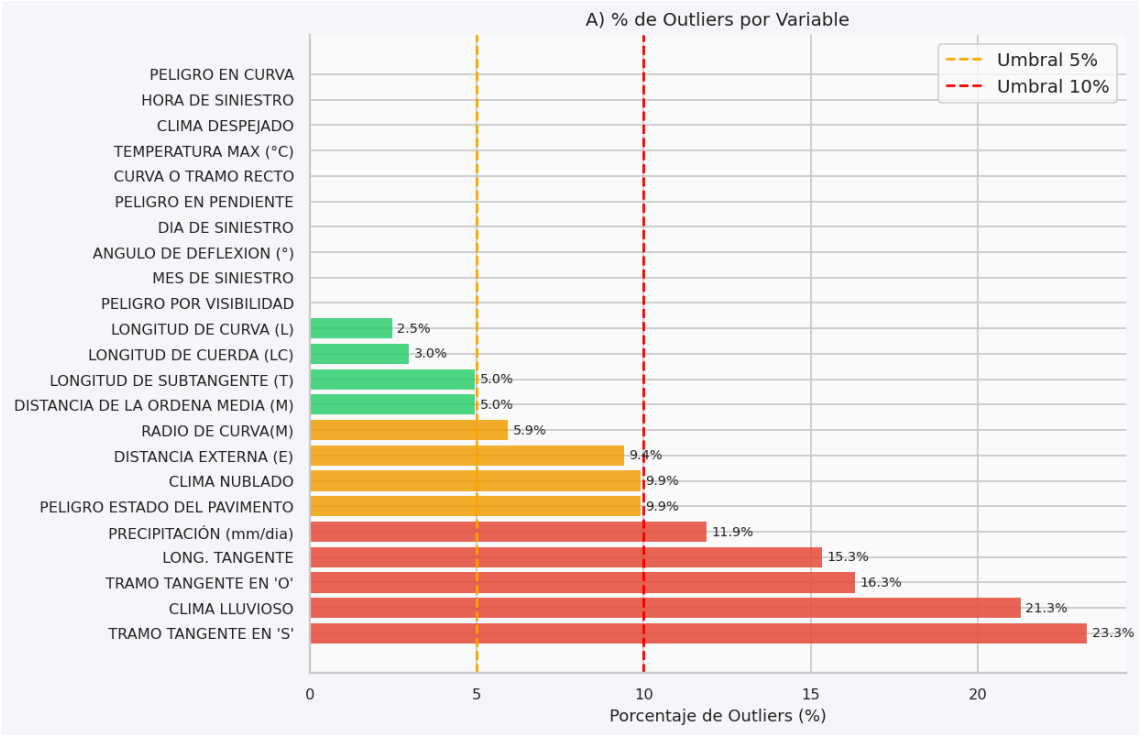


Figura 75 Porcentaje de outliers por variable. Umbral: $[Q1 - 1.5 \cdot IQR,\ Q3 + 1.5 \cdot IQR]$.
Fuente: Elaboración propia

Las primeras cuatro tienen IQR=0 al ser variables binarias; sus "outliers" son valores 1 en distribuciones mayoritariamente nulas. Los valores extremos en variables geométricas corresponden a condiciones reales excepcionales que deben preservarse. La normalización Z-Score mitiga su impacto en el entrenamiento.

E) Correlaciones de Pearson entre variables

La Figura siguiente presenta el heatmap de correlaciones de Pearson entre las 23 variables de entrada y la variable objetivo y las barras de correlación individual.

Heatmap de Correlaciones

El hallazgo más importante del heatmap es el bloque rojo intenso en la zona central-inferior, donde las variables geométricas de curva se correlacionan fuertemente entre sí, evidenciando multicolinealidad significativa. La tabla siguiente resume los pares de variables con mayor correlación detectada.

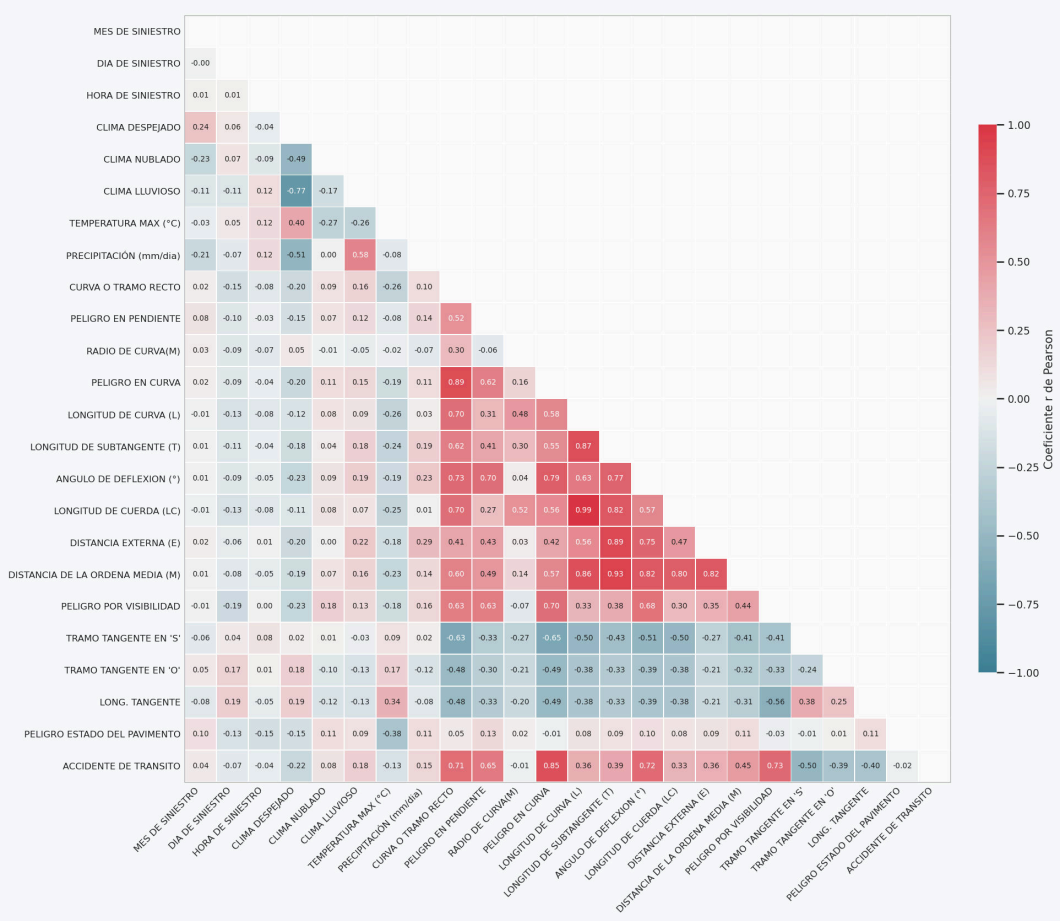


Figura 76 Heatmap de correlaciones de Pearson entre las 23 variables de entrada.
Fuente: Elaboración propia

Tabla 26 *Pares de variables con mayor correlación — Heatmap de Pearson*

Par de Variables	r de Pearson	Interpretación
Longitud de Cuerda (LC) ↔ Longitud de Subtangente (T)	0.99	Casi perfectamente colineales
Distancia Ordena Media ↔ Longitud de Subtangente (T)	0.93	Correlación muy alta
Peligro en Curva ↔ Curva o Tramo Recto	0.89	Correlación alta
Ángulo de Deflexión ↔ Longitud de Curva (L)	0.73	Correlación alta
Clima Despejado ↔ Clima Lluvioso	−0.77	Alta negativa (variables complementarias)

Nota: Valores obtenidos mediante el coeficiente de correlación de Pearson *r* aplicado a las 23 variables de entrada del Modelo RNA N°1.
Fuente: Elaboración propia.

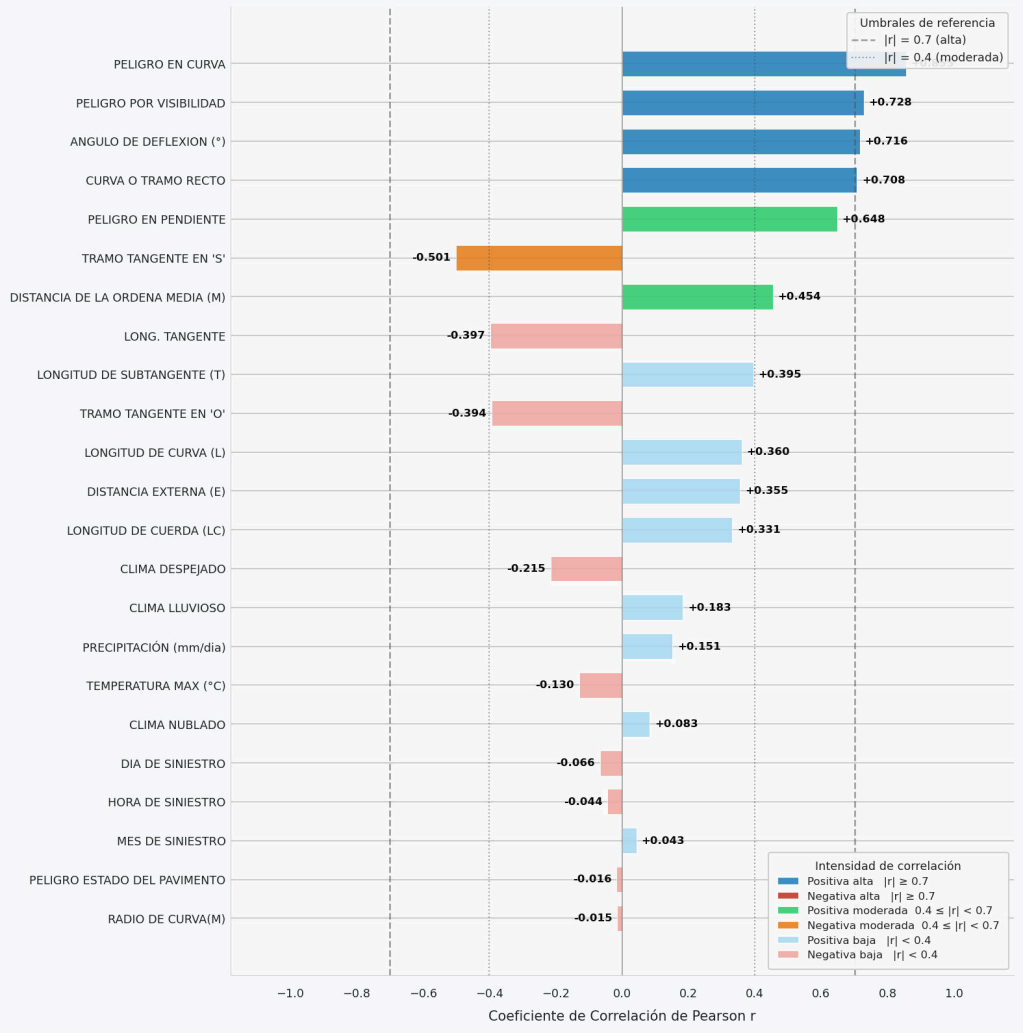


Figura 77 Barras de correlación individual con la variable objetivo
Fuente: Elaboración propia

Las variables con mayor correlación individual con accidente de tránsito se agrupan en tres niveles:

- **Predictores fuertes** ($|r| \geq 0.70$): peligro en curva ($r = +0.855$), peligro por visibilidad ($r = +0.728$), ángulo de deflexión ($r = +0.716$) y curva o tramo recto ($r = +0.708$), confirmando que la geometría horizontal es el factor dominante de accidentalidad en la vía PE-28A.
- **Predictores moderados** ($0.40 \leq |r| < 0.70$): peligro en pendiente ($r = +0.648$), tramo tangente en 'S' ($r = -0.501$), distancia de la ordenada media ($r = +0.454$) y long. tangente ($r = -0.397$), donde las correlaciones negativas indican que los tramos en tangente mitigan el riesgo de accidente.
- **Predictores débiles** ($|r| < 0.40$): las variables climáticas (clima lluvioso, precipitación, temperatura) y temporales (mes, día, hora) presentan correlaciones lineales bajas; sin embargo, su influencia es de naturaleza no lineal e interactiva con la geometría, efecto que la RNA captura en sus capas ocultas.

Tabla 27 *Top-10 correlaciones de Pearson con la variable objetivo accidente de tránsito*

Rank	Variable	r de Pearson	Interpretación
1	PELIGRO EN CURVA	0.855	Correlación muy alta positiva - predictor principal
2	PELIGRO POR VISIBILIDAD	0.728	Alta positiva - curvas con visibilidad comprometida
3	ÁNGULO DE DEFLEXIÓN (°)	0.716	Alta positiva - curvas más cerradas = mayor riesgo
4	CURVA O TRAMO RECTO	0.708	Alta positiva - presencia de curva vs. tangente
5	PELIGRO EN PENDIENTE	0.648	Moderada-alta positiva - pendiente pronunciada
6	DIST. DE LA ORDENADA MEDIA M	0.454	Moderada positiva - parámetro geométrico de curva
7	TRAMO TANGENTE EN S	−0.501	Moderada negativa - tangente mitiga el riesgo
8	LONG. TANGENTE	−0.397	Baja-moderada negativa
9	LONG. SUBTANGENTE T	0.395	Baja-moderada positiva
10	LONGITUD DE CURVA L	0.360	Baja-moderada positiva

Fuente: Elaboración propia.

F) Variables de peligro vs. clase (barras 100% apiladas)

La figura siguiente muestra la distribución porcentual de riesgo menor/mayor por nivel de las cuatro variables categóricas de peligro. Peligro en curva nivel 3 (máximo) concentra el 93% de registros de riesgo mayor, y peligro por visibilidad nivel 2 alcanza el 98% de riesgo mayor. Por el contrario, peligro estado pavimento no presenta poder discriminativo (≈50/50 en ambos niveles). Estos resultados validan la coherencia física del dataset con el conocimiento experto en ingeniería vial.

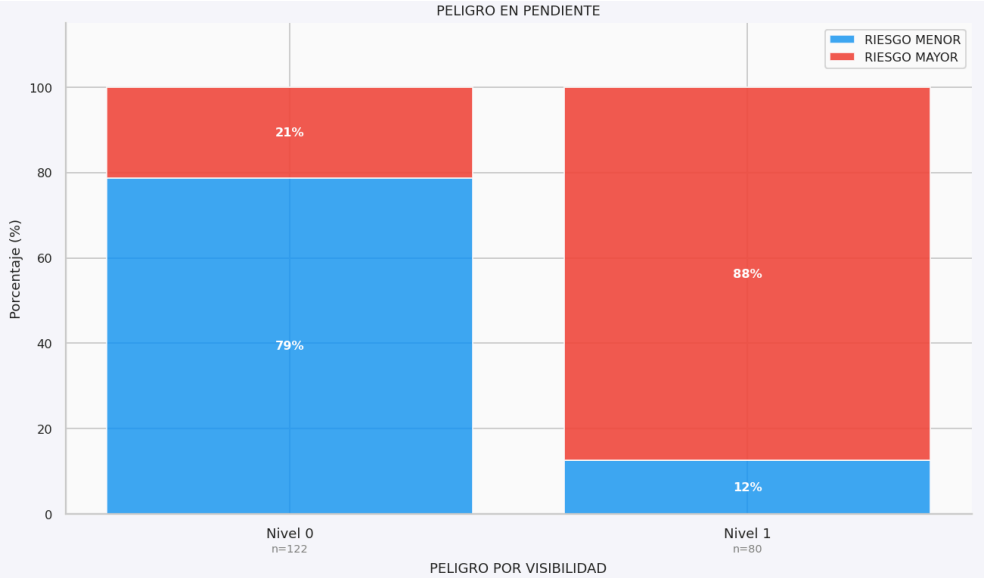


Figura 78 Distrib. porcentual de Riesgo menor/mayor de la var. Peligro en pendiente.
Fuente: Elaboración propia

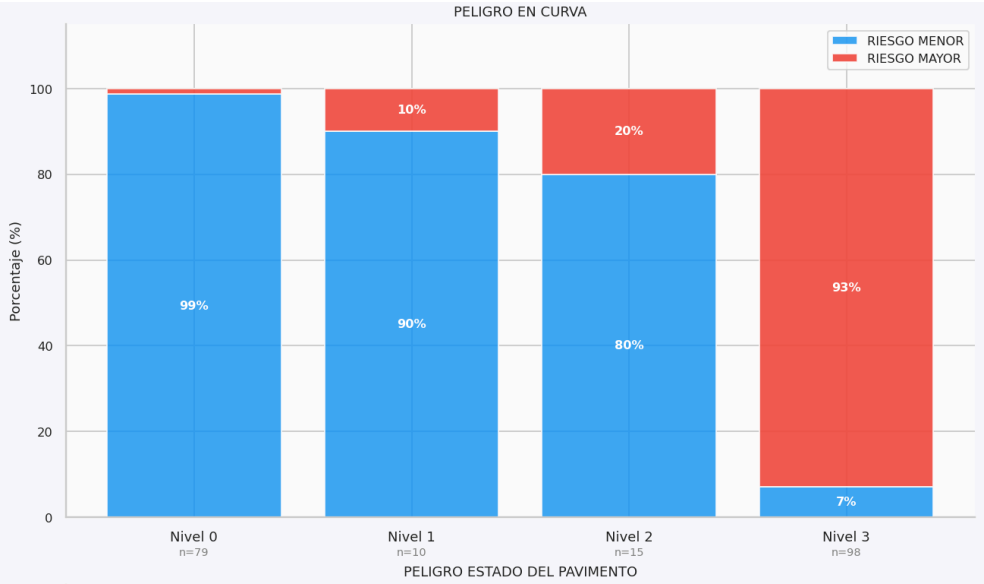


Figura 79 Distri. porcentual de Riesgo menor/mayor de la var. Peligro en curva.
Fuente: Elaboración propia

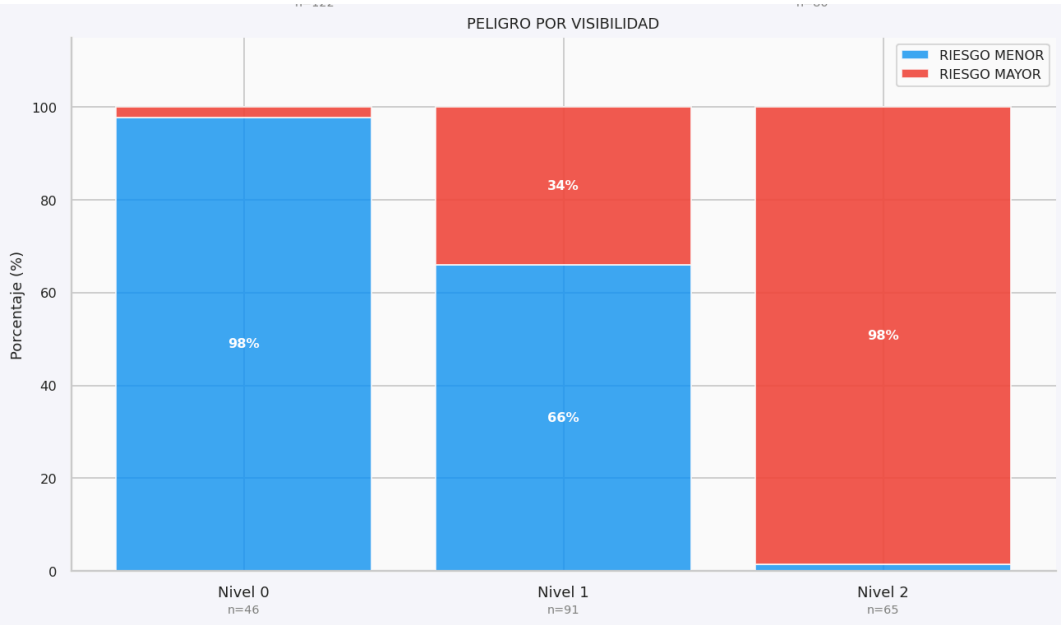


Figura 80 Distrib. porcentual de Riesgo menor/mayor de la var. Peligro por visibilidad.
Fuente: Elaboración propia

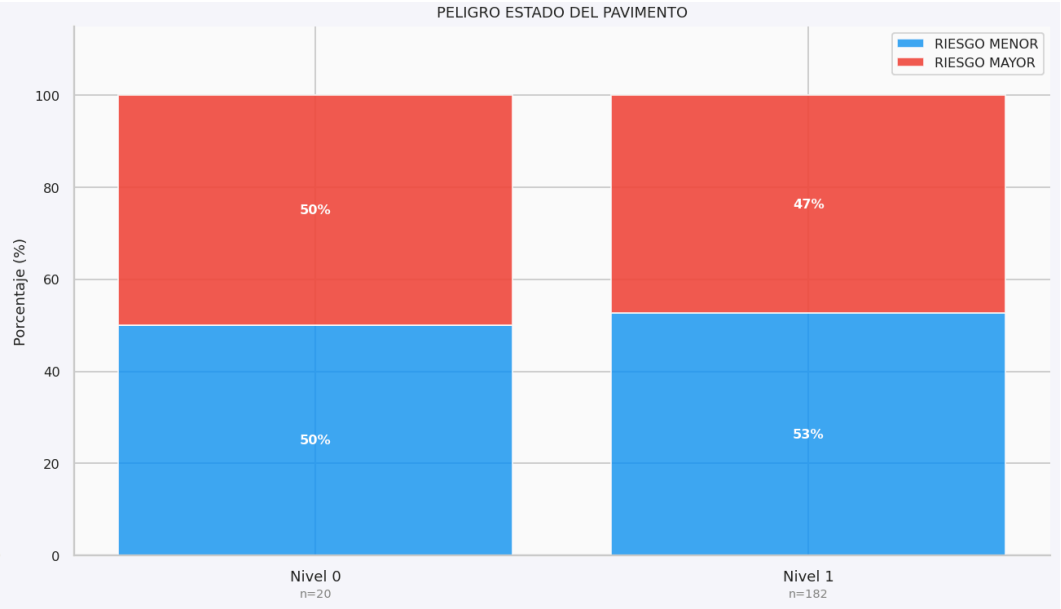


Figura 81 Distrib. porcentual de Riesgo menor/mayor de la var. Peligro estado pavimento.
Fuente: Elaboración propia

Tabla 28 *Distrib. porcentual de clases por nivel de variables de peligro*

Variable	Nivel	n	% R. MENOR	% R. MAYOR	Conclusión
PELIGRO EN PENDIENTE	0 (sin peligro)	122	79%	21%	Tramos sin peligro son mayormente seguros
PELIGRO EN PENDIENTE	1 (con peligro)	80	12%	88%	Pendiente peligrosa eleva riesgo en +67 pp
PELIGRO EN CURVA	0 (sin peligro)	79	99%	1%	Tramos rectos libres de riesgo mayor
PELIGRO EN CURVA	3 (máximo)	98	7%	93%	Curvas máximas: 93% son RIESGO MAYOR
PELIGRO POR VISIB.	0	46	98%	2%	Sin obstáculos: riesgo mínimo
PELIGRO POR VISIB.	2	65	2%	98%	Visibilidad comprometida = riesgo casi certero
PELIGRO PAVIMENTO	0 y 1	202	≈50%	≈50%	Sin poder discriminativo — distribución aleatoria

Fuente: Elaboración propia.

G) Análisis estadístico Mann-Whitney U

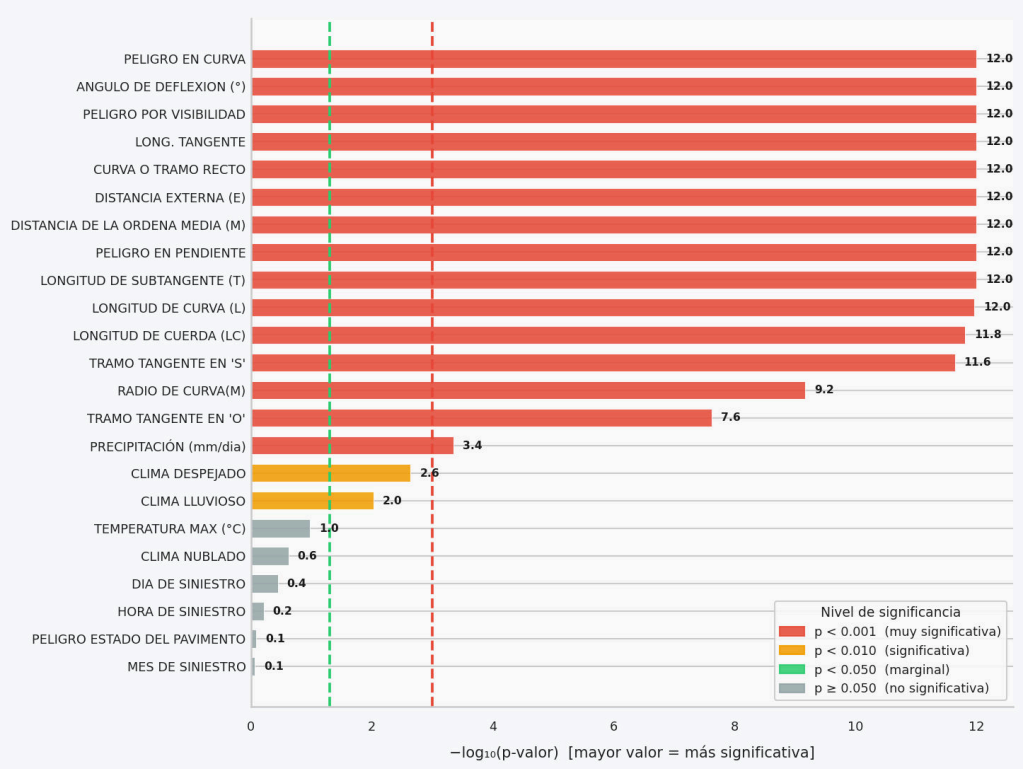


Figura 82 Análisis estadístico Mann-Whitney U de las 23 variables
Fuente: Elaboración propia

El Test de Mann-Whitney U es una prueba estadística no paramétrica que compara dos grupos independientes y determina si sus valores son significativamente diferentes o

producto del azar. El Test de Mann-Whitney U valida estadísticamente que 19 de las 23 variables (83%) presentan diferencias reales entre clases de riesgo, siendo la geometría vial el factor dominante con *p*-valores infinitesimales. Este resultado justifica la selección de variables realizada para el entrenamiento del Modelo RNA N°1 y confirma que el dataset contiene información discriminante suficiente para que la red neuronal aprenda a clasificar correctamente los niveles de accidentalidad en la carretera PE-28A.

- **Variables geométricas como factor dominante.** Los diez primeros lugares del ranking corresponden exclusivamente a parámetros geométricos de la PE-28A con $p < 10^{-12}$, confirmando que la geometría horizontal y vertical constituye el factor crítico de riesgo en el tramo.
- **Variables climáticas de influencia secundaria.** Precipitación, Clima Despejado y Clima Lluvioso resultan significativas ($p < 0.010$), con menor fuerza discriminante que las geométricas, aunque su inclusión es pertinente pues modulan el comportamiento del conductor y la adherencia del pavimento.
- **Variables temporales no discriminantes linealmente.** Mes, Día y Hora de Siniestro no superan $p = 0.05$, sin embargo, su efecto puede ser no lineal e interactivo, capturado por las capas ocultas de la RNA y no detectable mediante el Test Mann-Whitney U.

H) Diagrama de Dispersión de variables geométricas

Los seis scatter plots confirman un patrón estructural consistente: radios de curva pequeños (< 250 m) concentran la mayor densidad de puntos rojos (RIESGO MAYOR), mientras que los puntos azules (RIESGO MENOR) se dispersan hacia radios mayores. Este comportamiento es uniforme en todos los pares analizados Radio vs Longitud de Curva, vs Ángulo de Deflexión, vs Subtangente, vs Cuerda, vs Distancia Externa y vs Ordena Media lo que confirma que el Radio de Curva actúa como variable geométrica eje de la accidentalidad en la PE-28A. La separación entre clases no sigue una frontera lineal sino curva y compleja, lo que valida el uso de una arquitectura RNA-MLP con capas ocultas para aprender estas fronteras de decisión.

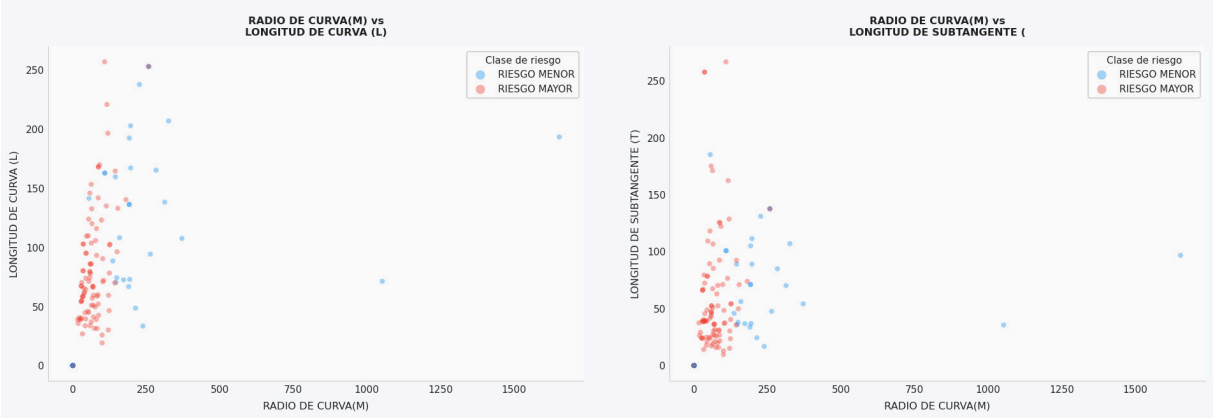


Figura 83 Diagr. de dispersión Radio de curva vs Long. de curva vs Long. de subtangente.
Fuente: Elaboración propia

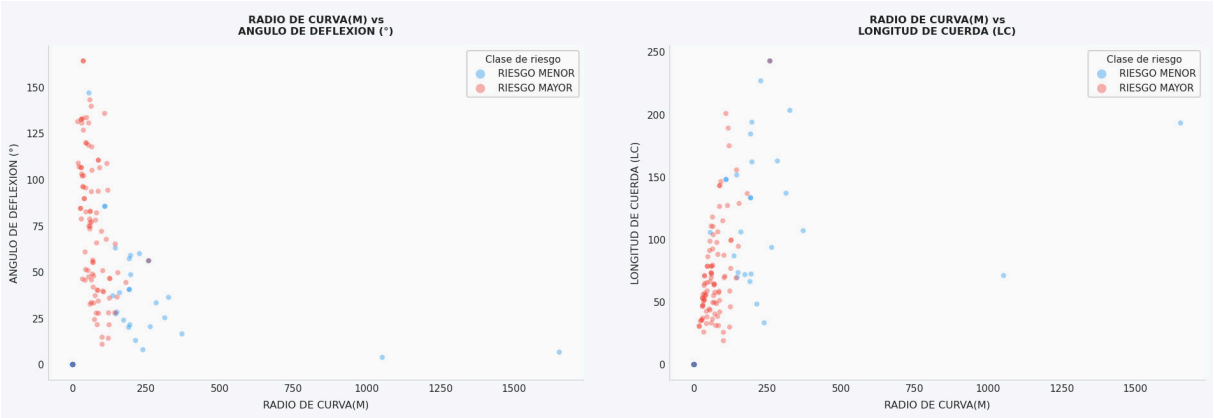


Figura 84 Diagr.de dispersión Radio de curva vs Angulo de deflexión vs Long. de cuerda.
Fuente: Elaboración propia

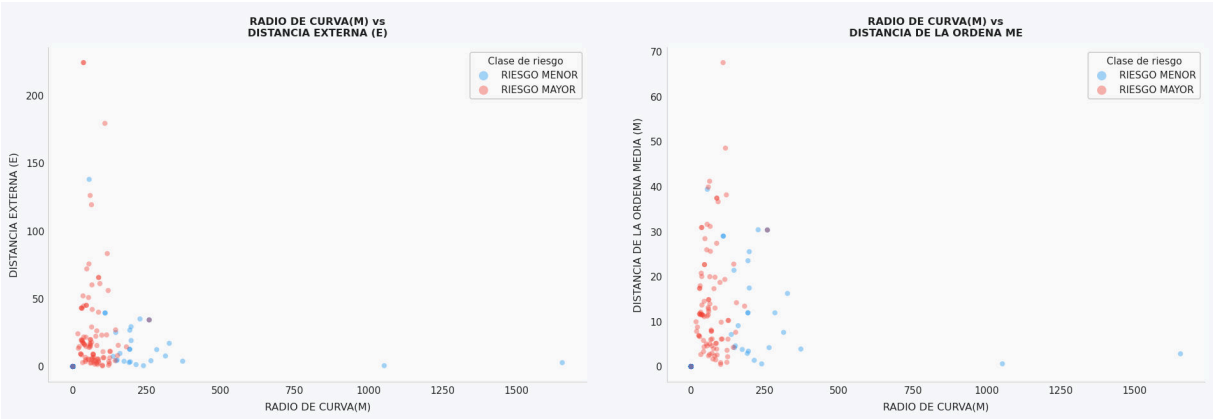


Figura 85 Diagr. de disper. Radio de curva vs Dist. Externa vs Dist. de la ord. media.
Fuente: Elaboración propia

En síntesis, los scatter plots demuestran que cuando el radio de curva es pequeño, todas las variables geométricas adoptan simultáneamente valores críticos, generando condi-

ciones de riesgo geométrico compuesto que la RNA debe identificar como patrón de accidentalidad en la carretera PE-28A.

3.11.1.4 Etapa II: Preprocesamiento de datos

El preprocesamiento de datos tuvo como objetivo transformar la base de datos en un formato adecuado para el entrenamiento eficiente de la RNA, garantizando la ausencia de fuga de información (*data leakage*) entre particiones y la escala uniforme de todas las variables de entrada. Esta etapa comprendió tres procesos secuenciales: división estratificada del dataset, normalización Z-Score y verificación visual de la separación entre clases tras la normalización.

La Tabla siguiente resume los tres procesos aplicados, sus parámetros clave y los resultados obtenidos.

Tabla 29 *Resumen del preprocesamiento de datos*

Pasos	Proceso	Parámetros Clave	Resultado
Paso 1	División estratificada	70% Entrenamiento (140 registros) 20% Validación (41 registros) 10% Evaluación (21 registros)	Proporción de clases preservada en las 3 particiones: 52% RIESGO MENOR / 48% RIESGO MAYOR
Paso 2	Normalización Z-Score	`fit()` aplicado solo en Train `transform()` en Val y Test Fórmula: $z = (x - \mu)/\sigma$	$\mu \approx 0, \sigma \approx 1$ en todas las particiones. Ausencia de *data leakage* garantizada
Paso 3	Grafica Z-Score × variable y KDE.	$\Delta\mu$ = diferencia de medias normalizadas entre clases: $\Delta\mu = \bar{z}_{\text{MAY}} - \bar{z}_{\text{MEN}}$	Máximo: PELIGRO EN CURVA $\Delta\mu = 1.67$ Mínimo: PELIGRO PAVIMENTO $\Delta\mu = 0.03$

Nota: El ajuste Z-Score se calculó exclusivamente con los datos de entrenamiento para evitar fuga de información hacia las particiones de validación y prueba. $\Delta\mu$ cuantifica la separación entre clases tras la normalización; valores altos indican mayor capacidad de esa variable para ayudar a la RNA a clasificar correctamente.
Fuente: Elaboración propia.

A) División estratificada 70/20/10% de la base de datos

La base de datos se cuenta con 202 registros se dividió en tres particiones: 70% para entrenamiento (140 registros), 20% para validación (41 registros) y 10% para la evaluación (21 registros). Se empleó división estratificada para preservar la proporción original de clases (52% RIESGO MENOR / 48% RIESGO MAYOR) en cada parti-

ción, evitando sesgos de muestreo que pudieran invalidar la evaluación del modelo.

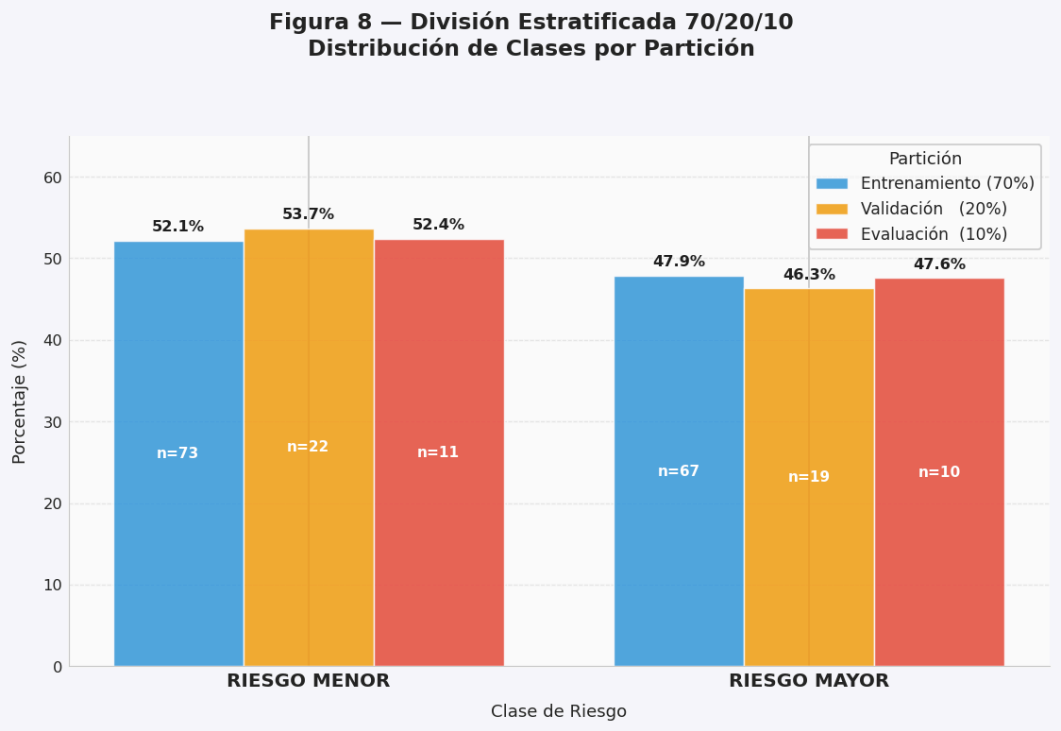


Figura 86 División estratificada 70/20/10% de la base de datos.
Fuente: Elaboración propia

B) Normalización Z-Score variable por variable (23 variables)

Se aplicó la estandarización Z-Score a todas las variables numéricas de entrada mediante la expresión:

$$z = \frac{x - \mu}{\sigma} \tag{3.2}$$

donde x es el valor original, μ es la media y σ es la desviación estándar. Como resultado, todas las variables quedan con media $\mu \approx 0$ y desviación estándar $\sigma \approx 1$. Para evitar *data leakage*, el ajuste (`fit()`) se realizó exclusivamente sobre el conjunto de entrenamiento, aplicando posteriormente la transformación (`transform()`) a las particiones de validación y prueba con los parámetros obtenidos del entrenamiento.

Verificación visual Z-Score. Se calculó la diferencia de medias normalizadas entre clases ($\Delta\mu$) para cada variable, definida como:

$$\Delta\mu = \bar{z}_{\text{RIESGO MAYOR}} - \bar{z}_{\text{RIESGO MENOR}} \tag{3.3}$$

Los resultados mostraron que PELIGRO EN CURVA registró el mayor $\Delta\mu = 1.67$, confirmando su alta capacidad discriminativa entre clases, mientras que PELIGRO ESTADO DEL PAVIMENTO presentó el menor $\Delta\mu = 0.03$, indicando escasa diferenciación entre grupos tras la normalización. Lo que significa mientras $\Delta\mu$ sea mayor, la variable SEPARA BIEN las clases antes de entrar a la RNA. El cual se graficará con Kernel Density Estimation (KDE) o Estimación de Densidad por Núcleo, el cual es una técnica estadística que suaviza un histograma para convertirlo en una curva continua que representa cómo se distribuyen los datos.

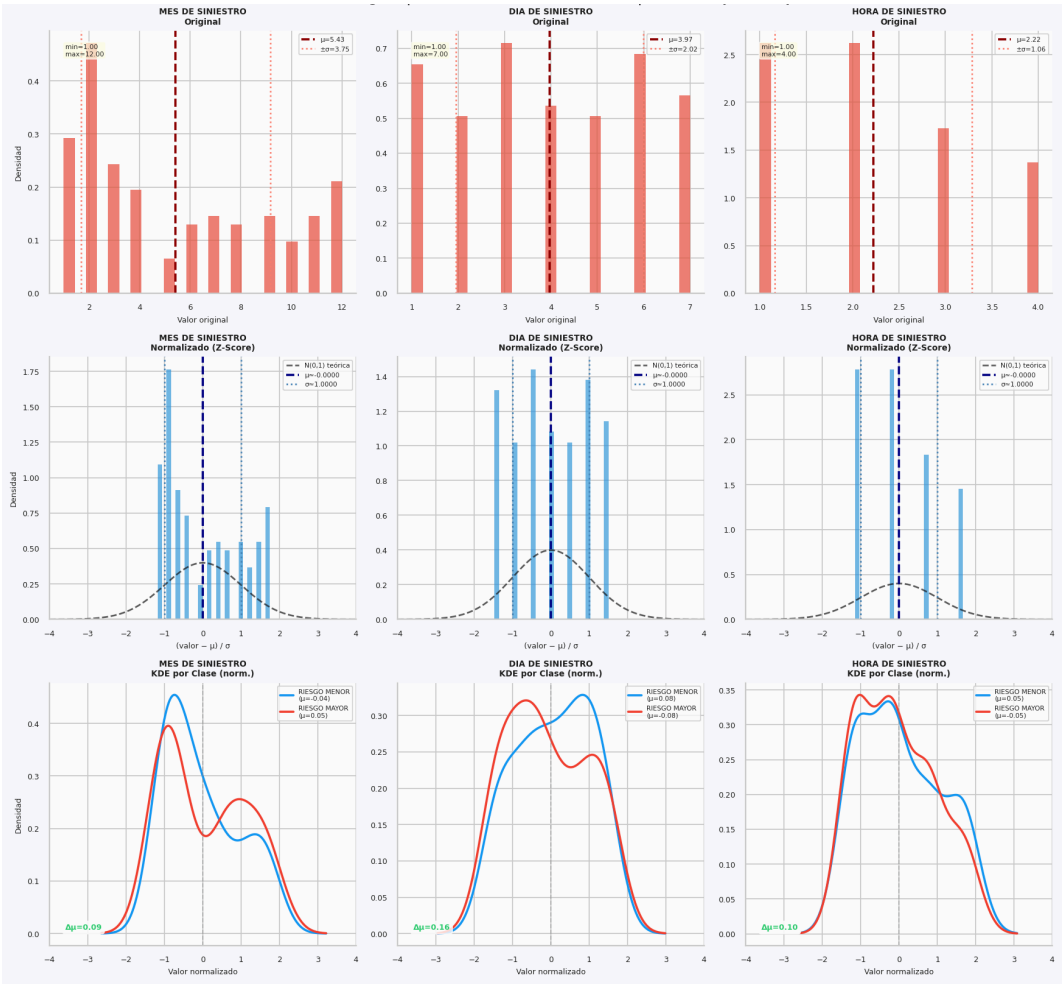


Figura 87 Normalización Z-Score: Mes, Día y Hora de Siniestro.
Fuente: Elaboración propia

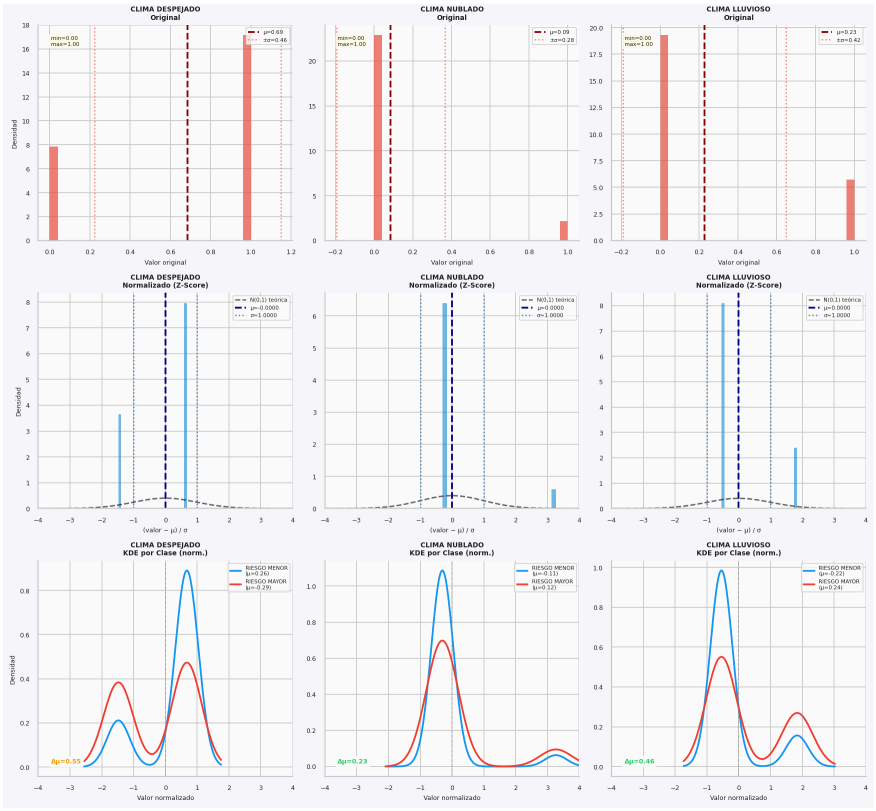


Figura 88 Normaliz. Z-Score: Clima Despejado, Clima Nublado y Clima Lluvioso.
Fuente: Elaboración propia

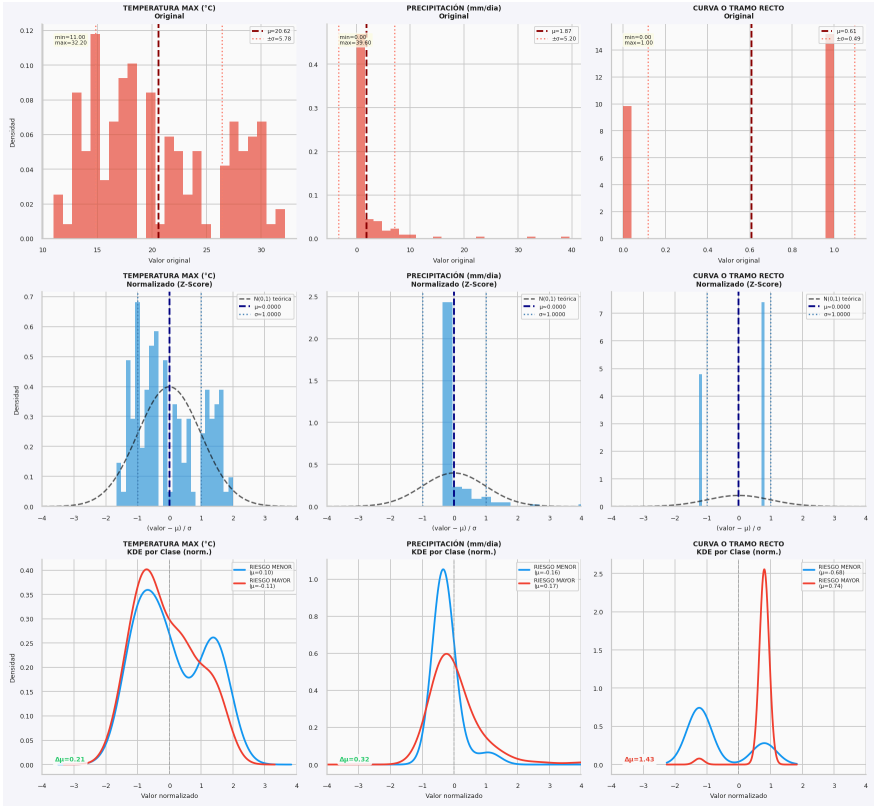


Figura 89 Normaliz. Z-Score:Temperatura Máx., Precipitación y Curva/Tramo Recto.
Fuente: Elaboración propia

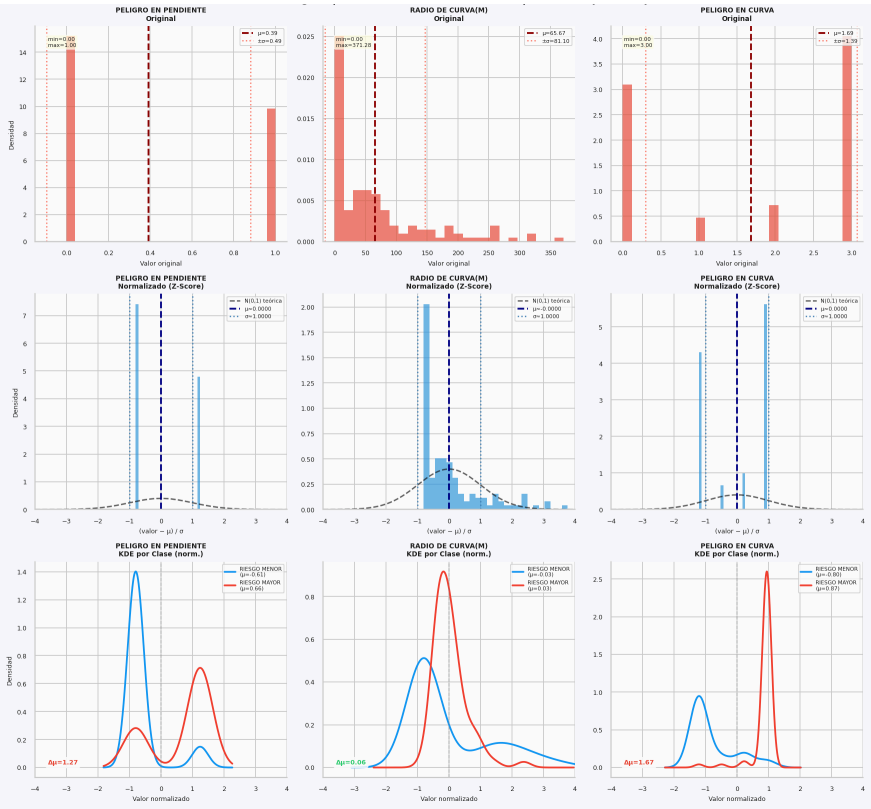


Figura 90 Normaliz. Z-Score: Peligro Pendiente, Radio de Curva y Peligro en Curva.
Fuente: Elaboración propia

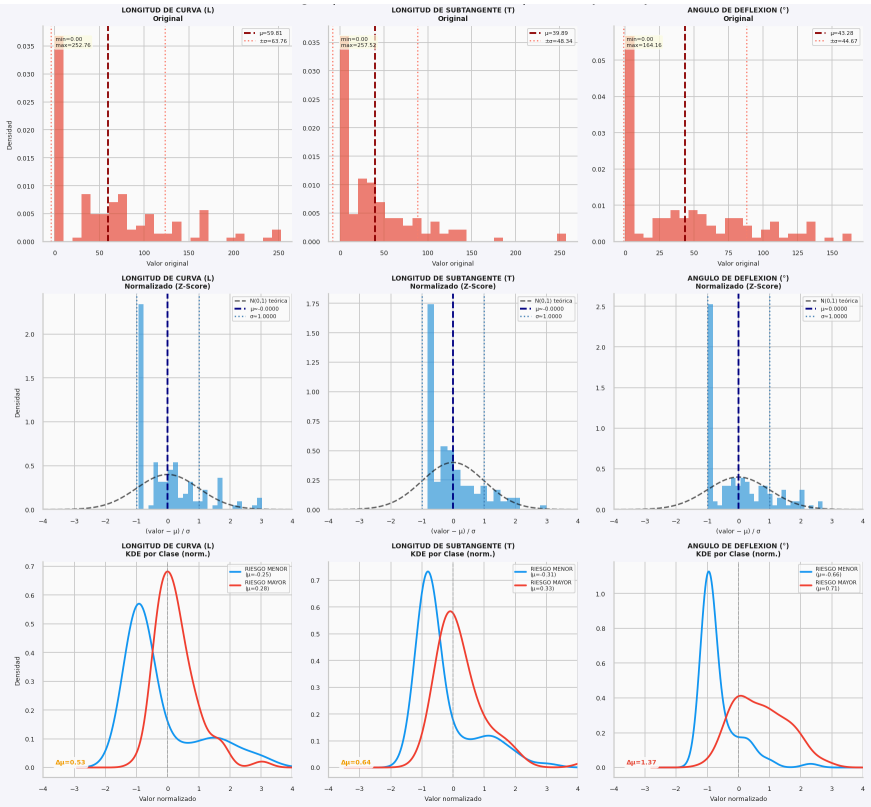


Figura 91 Normalización Z-Score de las variables: Long. (L), Subtangente (T) y Ángulo Deflexión.
Fuente: Elaboración propia

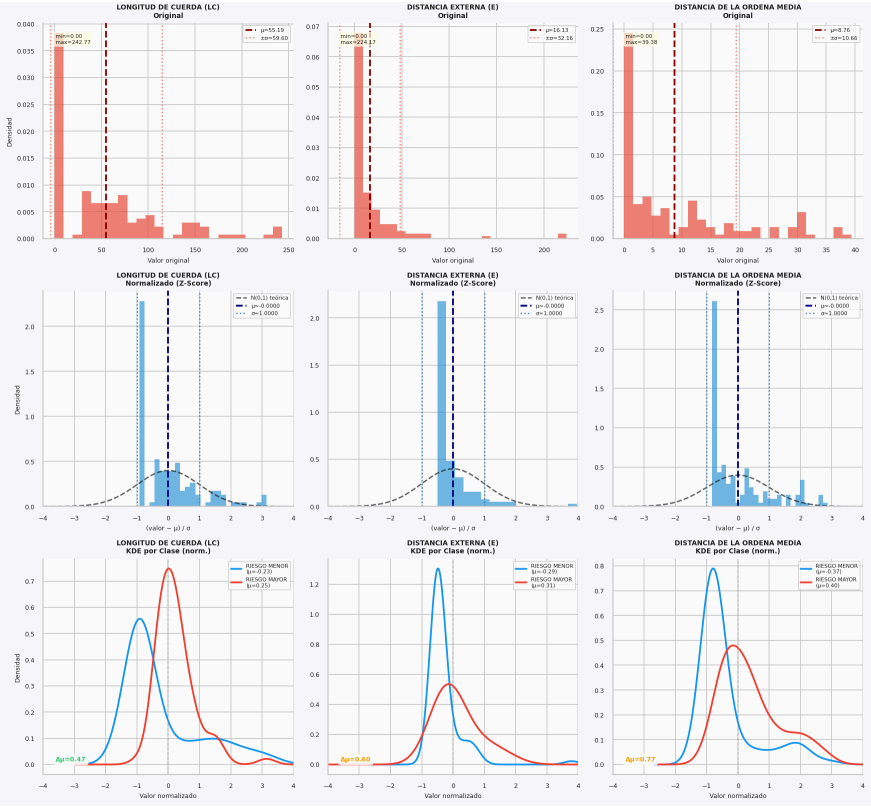


Figura 92 Normaliz. Z-Score: Long. Cuerda (LC), Distancia Externa (E) y Ordenada Media (M)
Fuente: Elaboración propia

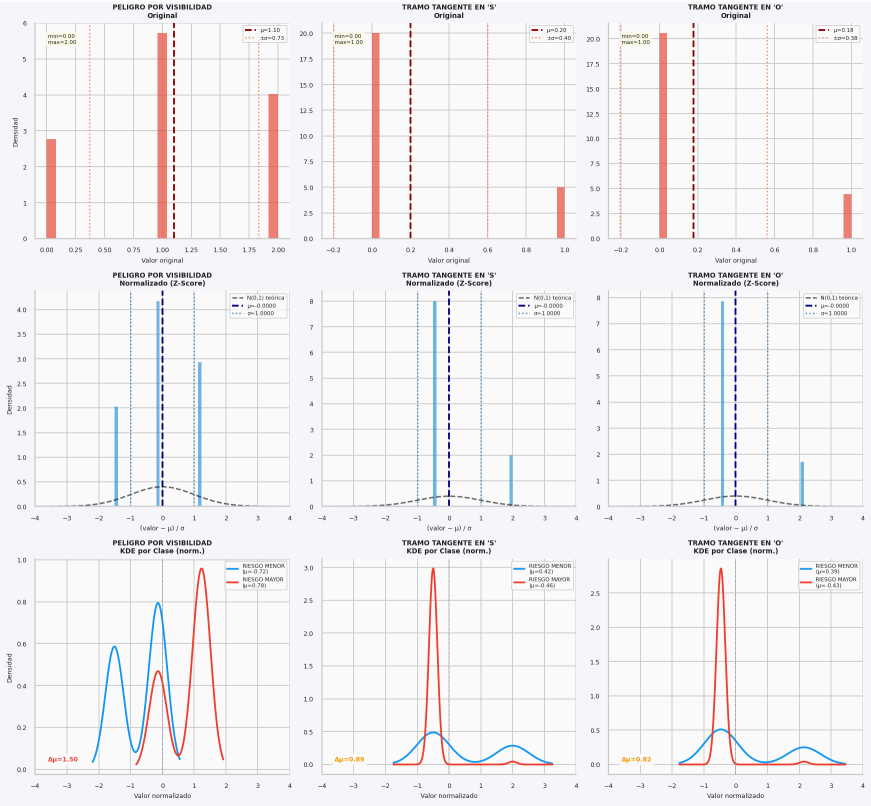


Figura 93 Normaliz. Z-Score: Peligro Visibilidad, Tramo en ”S” y Tramo Tangente ”O”
Fuente: Elaboración propia

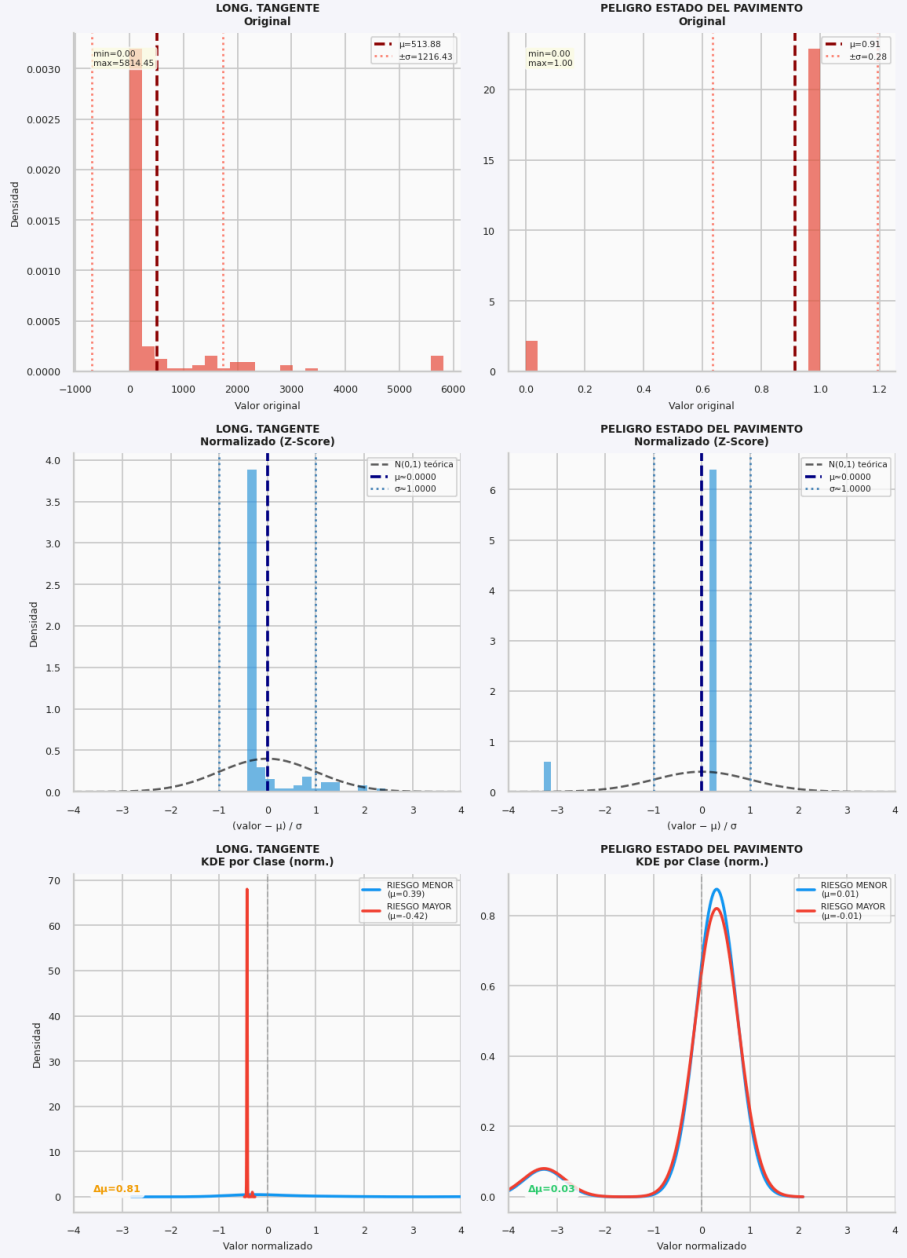


Figura 94 Normaliz. Z-Score: Longitud Tangente y Peligro Estado del Pavimento.
Fuente: Elaboración propia

3.11.1.5 Etapa III: Diseño, optimización y entrenamiento de la RNA

Esta etapa comprendió el diseño de la arquitectura óptima del Perceptrón Multicapa (MLP), la búsqueda sistemática del parámetro de regularización L2, la comparación de arquitecturas candidatas, el diagnóstico de sobreajuste y el entrenamiento definitivo del modelo .

Tabla 30 *Resumen del diseño y entrenamiento del modelo RNA*

Item	Proceso	Parámetros / Resultado	Decisión Tomada
1	Arquitectura MLP definitiva	23→64→48→32→16→8→2	Topología piramidal: abstracción progresiva de patrones
2	Búsqueda alpha L2	8 valores: 0.0001 → 0.5	Alpha óptimo=0.05 → GAP overfitting mínimo
3	Comparación 7 arquitecturas	(32,16) a (128,64,32,16,8)	Seleccionada (64,48,32,16,8): GAP=0.064, CV≈0.88
4	Diagnóstico de overfitting	Curvas Entr./Valid. 48 épocas	Mejor modelo en época 29; GAP<0.20 siempre
5	Convergencia del entrenamiento	Pérdida 0.81→0.11 en 48 épocas	Descenso suave en 3 fases; sin oscilaciones
6	Pérdida y precisión vs. épocas	Valid._loss<Entr._loss (efecto L2)	Precisión Val máxima=97.56% en época 11

Fuente: Elaboración propia.

A) Arquitectura del Modelo RNA N°1 (MLP)

La Figura siguiente muestra la topología definitiva del Perceptrón Multicapa: 23→64→48→32→16→8→2, con 5 capas ocultas y 6,906 parámetros totales. La arquitectura garantiza una abstracción progresiva de los patrones: la primera capa oculta (64 neuronas) extrae características primarias; las capas intermedias (48, 32, 16) combinan y abstraen representaciones progresivamente más compactas; la quinta capa oculta (8 neuronas) genera la representación final pre-clasificación; y la capa de salida (2 neuronas con Softmax) produce las probabilidades (RIESGO MENOR=0) y (RIESGO MAYOR=1).

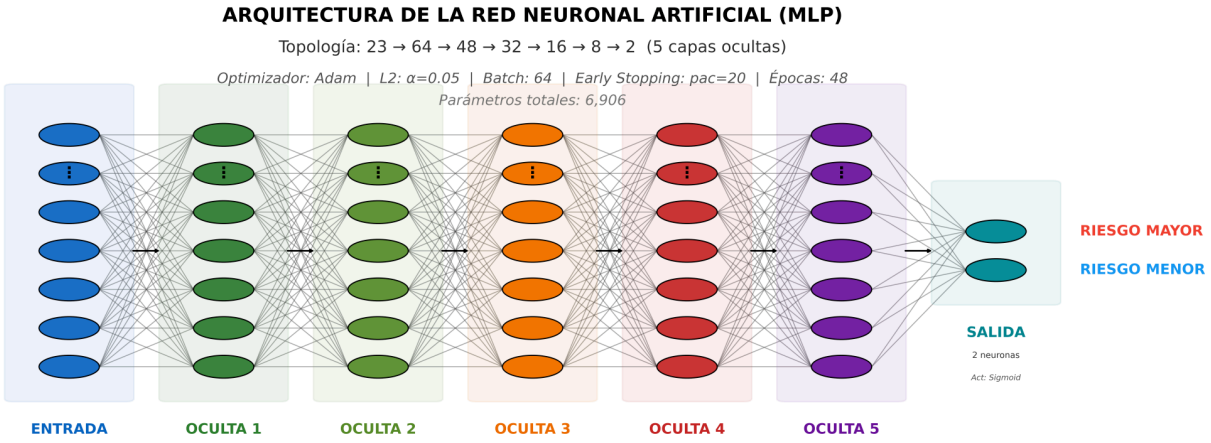


Figura 95 Arquitectura del Perceptrón Multicapa de la RNA N°01
Fuente: Elaboración propia

Tabla 31 *Descripción de cada capa de la arquitectura RNA*

Capa	Tipo	Neuronas	Activación	Función
Entrada	Capa de Entrada	23	-	Una neurona por cada variable de entrada normalizada mediante Z-Score
Oculta 1	Capa Completamente Conectada (Dense)	64	ReLU	Detección de patrones primarios en las 23 variables de entrada: geometría, clima y temporalidad
Oculta 2	Capa Completamente Conectada (Dense)	48	ReLU	Combinación de patrones detectados y reducción progresiva de dimensionalidad
Oculta 3	Capa Completamente Conectada (Dense)	32	ReLU	Abstracción de patrones de riesgo de nivel intermedio
Oculta 4	Capa Completamente Conectada (Dense)	16	ReLU	Síntesis y compresión de las representaciones aprendidas
Oculta 5	Capa Completamente Conectada (Dense)	8	ReLU	Representación compacta final previa a la clasificación
Salida	Capa de Salida Completamente Conectada	2	Sigmoid	Genera la probabilidad de cada clase: P(RIESGO MENOR) + P(RIESGO MAYOR) = 1.0

Nota: ReLU (Rectified Linear Unit): función de activación $f(x) = \max(0, x)$ que introduce no linealidad y evita el problema del gradiente desvaneciente. Sigmoid normaliza las salidas como probabilidades que suman 1.0. Dense (Completamente Conectada): cada neurona está conectada con todas las neuronas de la capa anterior.
Fuente: Elaboración propia.

B) Efecto de la regularización L2 sobre el overfitting

La Figura evalúa el impacto de la regularización L2 sobre el overfitting del Modelo RNA N°1, analizando ocho valores de alpha: $\alpha \in \{0.0001, 0.001, 0.005, 0.01, 0.05, 0.1, 0.3, 0.5\}$.

¿Qué es la regularización L2? Añade un término de penalización a la función de error durante el entrenamiento:

$$\mathscr{L}_{\text{total}} = \mathscr{L}_{\text{error}} + \lambda \sum w^2 \tag{3.4}$$

donde λ controla la intensidad de la penalización y w son los pesos sinápticos. Al penalizar pesos grandes, L2 impide que la RNA memorice el conjunto de entrenamiento

favoreciendo la RNA prediga mejor futuro datos nuevos.

El overfitting se cuantifica mediante el **GAP de overfitting**, GAP (la diferencia entre qué tan bien aprende el modelo con datos conocidos vs datos nuevos):

$$\text{GAP} = \text{Accuracy}_{\text{Entrenamiento}} - \text{Accuracy}_{\text{CV (Validación cruzada)}} \tag{3.5}$$

Se estableció un umbral de alerta en GAP = 0.10; valores superiores indican sobreajuste inaceptable.

- **Accuracy vs Alpha:** El Entrenamiento Accuracy (línea roja) se mantiene estable entre 0.93–0.95 en todos los valores evaluados. El CV Accuracy (línea azul, validación cruzada 5-Fold) permanece alrededor de 0.88–0.89, con ambas curvas convergiendo en $\alpha = 0.5$.

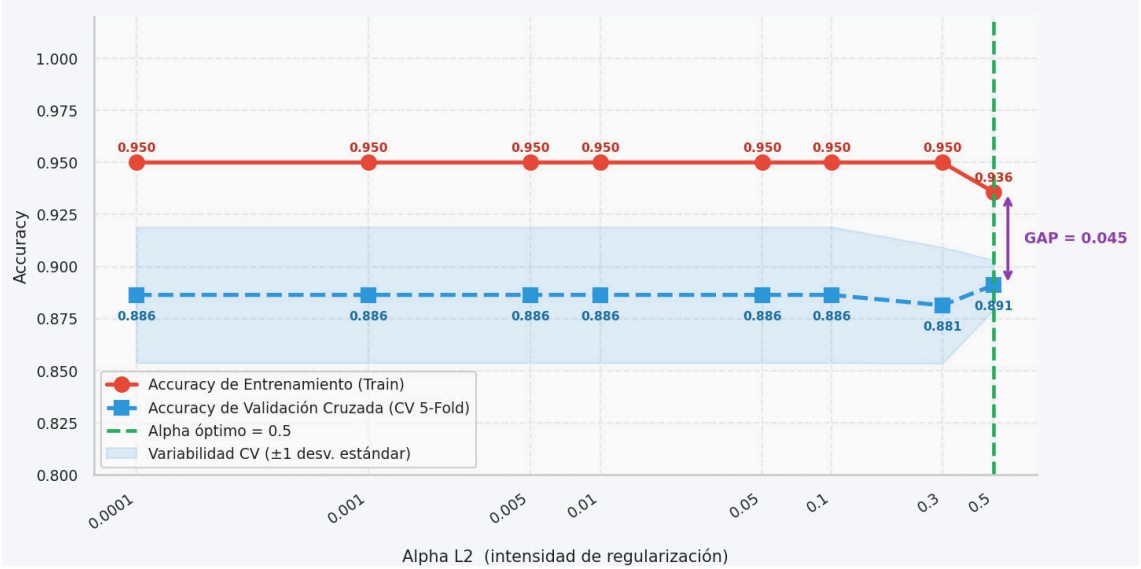


Figura 96 Accuracy vs. Alpha L2.
Fuente: Elaboración propia

- **GAP de Overfitting:** Para $\alpha \in \{0.0001 \text{ a } 0.1\}$, el GAP se mantiene constante en 0.064 — nivel aceptable. En $\alpha = 0.3$ sube levemente a 0.069, mientras que en $\alpha = 0.5$ alcanza el valor mínimo de **0.045** (barra verde), confirmándolo como el alpha óptimo.

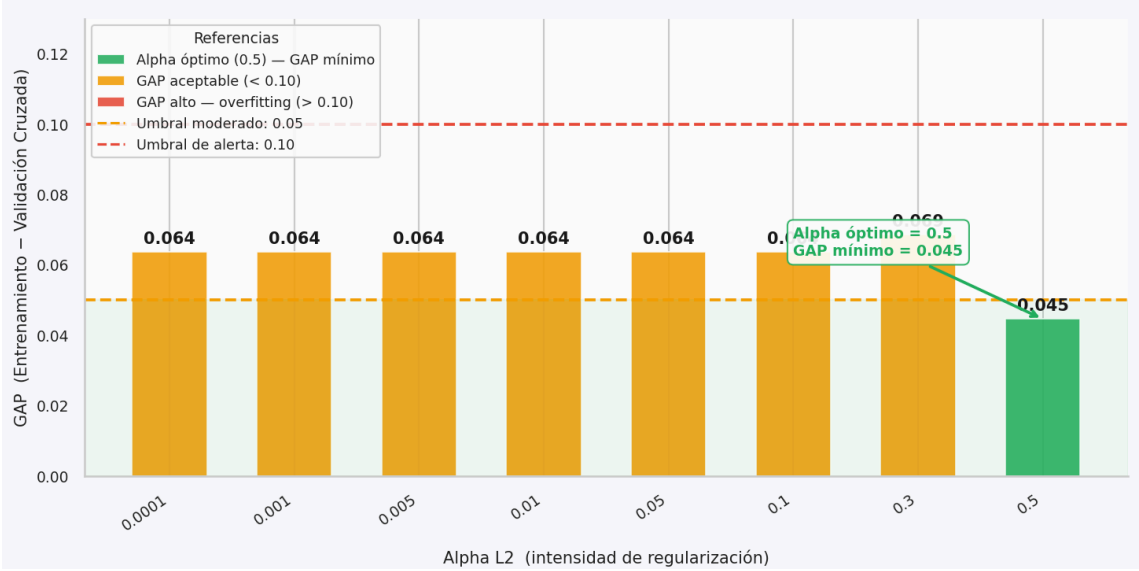


Figura 97 GAP de overfitting por valor de alpha.
Fuente: Elaboración propia

Conclusión: Se seleccionó $\alpha = 0.5$ como valor óptimo de regularización L2 por producir el menor GAP de overfitting (GAP = 0.045), por debajo del umbral de alerta de 0.10. Esta estrategia garantiza que las métricas del Modelo RNA N°1 prediga mejor para datos nuevos.

C) Comparación de arquitecturas candidatas

Se compara seis arquitecturas RNA evaluando tres criterios: accuracy de entrenamiento, accuracy de validación cruzada (CV 5-Fold) y GAP de overfitting, con el objetivo de seleccionar la arquitectura que mejor generaliza sin memorizar los datos de entrenamiento.

Resultados por grafica:

- **Accuracy de entrenamiento y validación cruzada por arquitectura:** Todas las arquitecturas mantienen Train Accuracy entre 0.93–0.97 y CV Accuracy entre 0.87–0.91, confirmando capacidad de aprendizaje consistente. La arquitectura seleccionada (64→48→32→16→8) logra CV = 0.89 con la menor brecha respecto al entrenamiento.

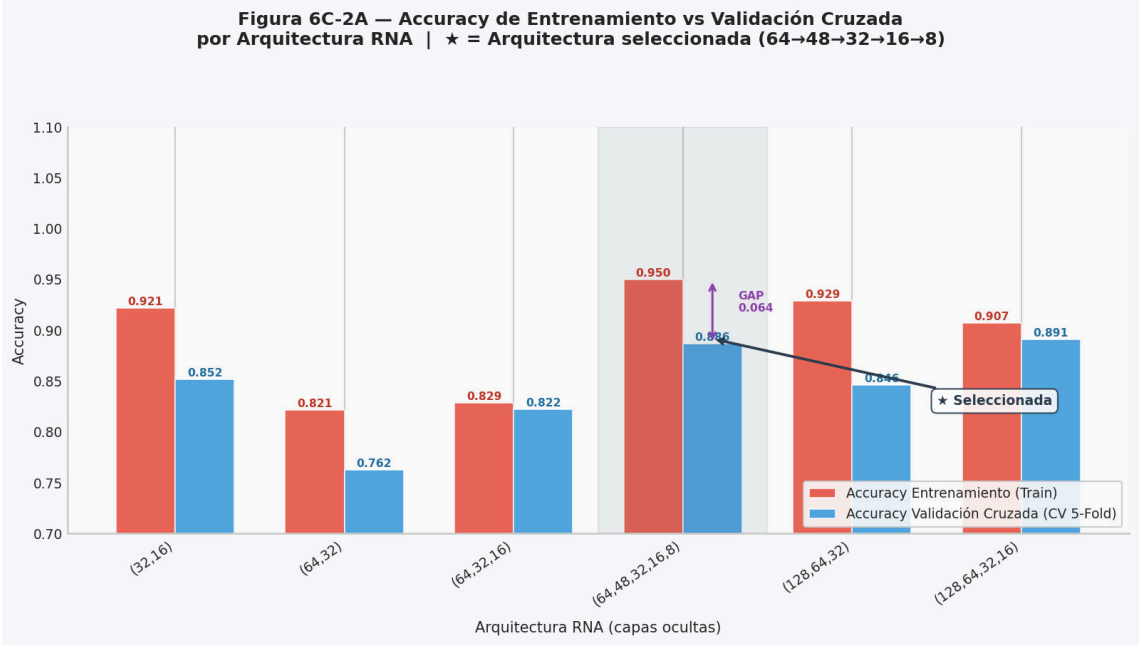


Figura 98 Accuracy de entrenamiento y validación cruzada por arquitectura.
Fuente: Elaboración propia

- **GAP de Overfitting por arquitectura:** La arquitectura seleccionada registra el menor GAP entre las evaluadas, por debajo del umbral de alerta de 0.10, indicando ausencia de sobreajuste inaceptable.

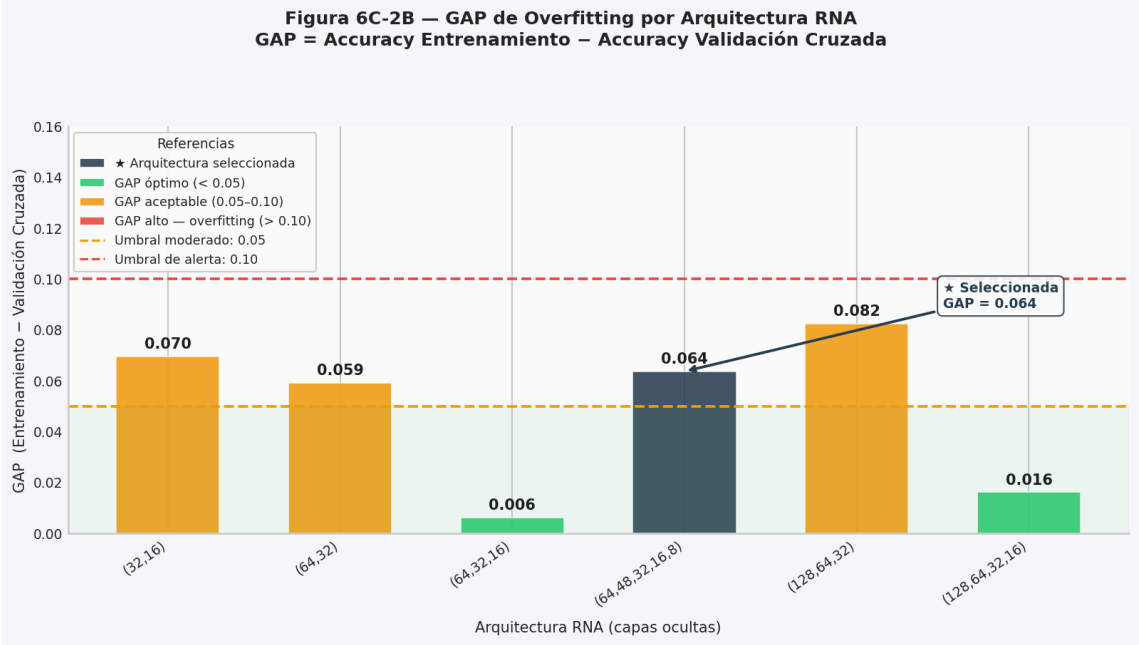


Figura 99 GAP de Overfitting por arquitectura.
Fuente: Elaboración propia

- **GAP de Overfitting vs Número de parametros de la arquitectura:** Arquitecturas más grandes como (128,64,32,16) no reducen el GAP respecto a la seleccionada, confirmando que mayor número de parámetros no garantiza mejor generalización.

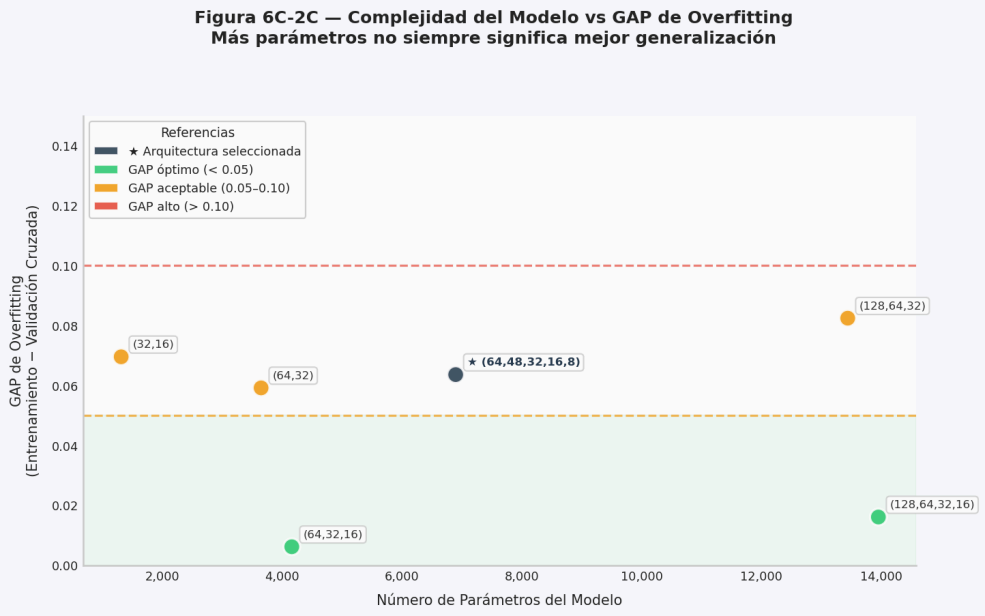


Figura 100 GAP de Overfitting vs Número de parametros de la arquitectura.
Fuente: Elaboración propia

Tabla 32 *Comparación de arquitecturas: rendimiento y nivel de sobreajuste*

Arquitectura	Acc. Entre-namiento	Acc. CV	GAP Entr.-CV	Seleccionada
(32, 16)	0.925	0.855	+0.070	No
(64, 32)	0.830	0.830	+0.059	No
(64, 32, 16)	0.830	0.830	+0.006	No
(64, 48, 32, 16, 8) ⋆	**0.950**	**0.880**	**+0.064**	**SÍ**
(128, 64, 32)	0.925	0.880	+0.082	No
(128, 64, 32, 16)	0.905	0.890	+0.016	No

Fuente: Elaboración propia.

Conclusión: Se seleccionó la arquitectura (64→48→32→16→8) por ofrecer el mejor equilibrio entre accuracy de validación y control del overfitting, con una reducción progresiva de neuronas por capa que favorece la abstracción jerárquica de patrones de accidentalidad en la vía PE-28A.

D) Convergencia del entrenamiento y accuracy por partición

La Figuras siguiente evalúa el proceso de entrenamiento del Modelo RNA N°1:

Curva de Convergencia (48 épocas): La pérdida (*Cross-Entropy Loss*) desciende progresivamente en tres fases: descenso acelerado (épocas 1–10, de 0.81 a 0.30), descenso moderado (épocas 10–25, de 0.30 a 0.15) y convergencia estable (épocas 25–48, pérdida final $\mathscr{L} = 0.110$). La ausencia de oscilaciones confirma la estabilidad del op-

timizador Adam.

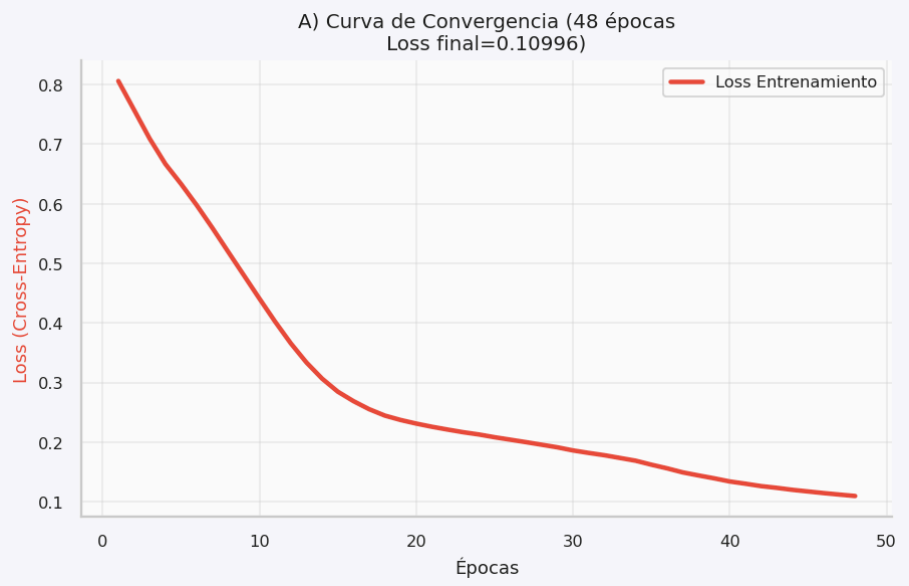


Figura 101 Curva de convergencia: Perdida vs Epocas.
Fuente: Elaboración propia

Accuracy por Partición: Los resultados obtenidos en cada partición se presentan a continuación:

- Entrenamiento (70%, $n = 140$): Accuracy = 96.43%
- Validación (20%, $n = 41$): Accuracy = 95.12%
- Evaluación (10%, $n = 21$): Accuracy = 90.48%

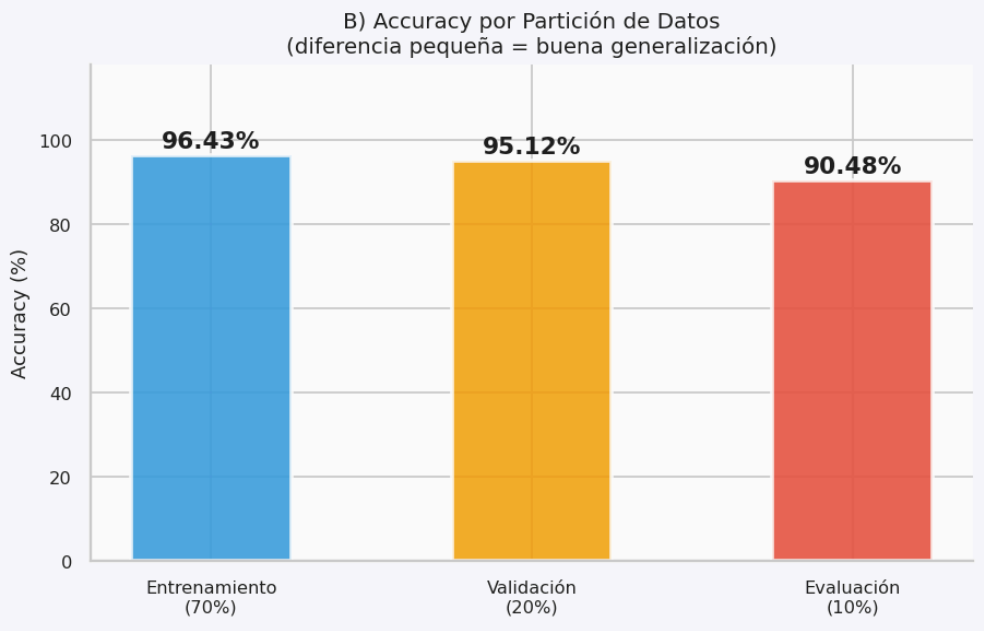


Figura 102 Accuracy en entrenamiento, validación y evaluación.
Fuente: Elaboración propia

La diferencia entre entrenamiento y validación es de apenas $\Delta = 1.31\%$, y entre entre-

namiento y evaluación de $\Delta = 5.95\%$, ambas dentro del rango aceptable para un base de datos de 202 registros. Estos resultados confirman la ausencia de overfitting significativo y validan que la regularización L2 ($\alpha = 0.5$) combinada con Early Stopping controló eficazmente el sobreajuste del modelo.

E) Pérdida y Accuracy vs. épocas (Entrenamiento vs. Validación)

La Figuras siguientes muestra las curvas de pérdida y Accuracy para entrenamiento y validación por época. La pérdida de validación es sistemáticamente inferior a la de entrenamiento por efecto de la penalización L2.

Resultados por grafica:

- **Gráfica de Pérdida vs. Épocas**: Muestra el error del modelo en cada época. Ambas curvas bajan progresivamente, el cual el modelo aprende. La curva naranja (validación) alcanza su punto mínimo en la época 29 ($val_loss = 0.0608$), ese es el mejor momento del modelo. Después sube levemente porque empieza a sobreajustarse. El early stopping detectó eso, esperó 20 épocas más sin mejora, paró en época 48 y restauró los pesos de época 29. Los pesos de época 48 fueron descartados.

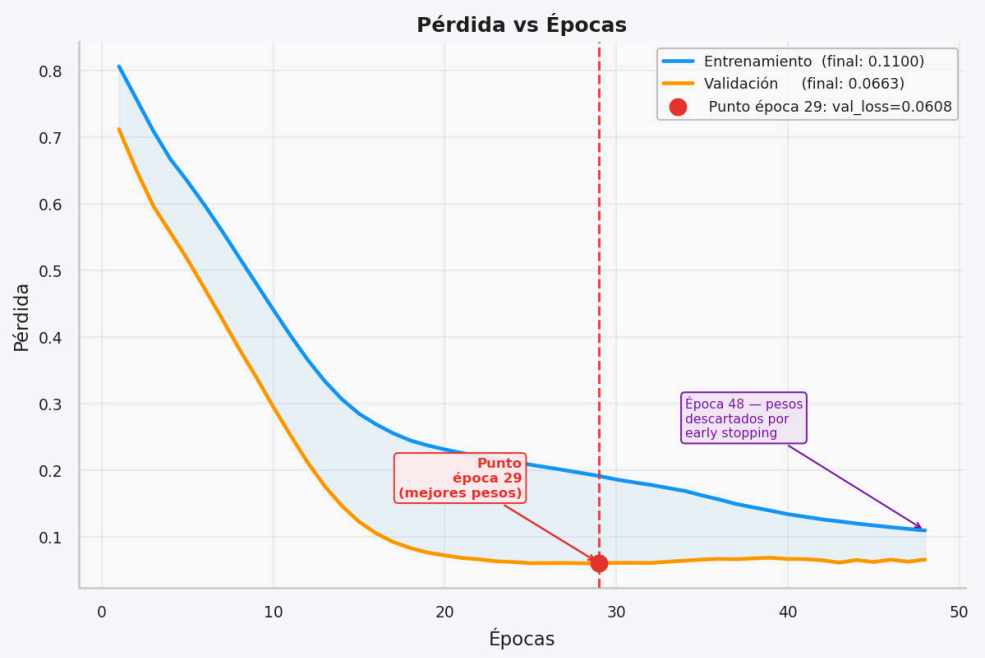


Figura 103 Pérdida (loss) vs. Épocas
Fuente: Elaboración propia

- **Accuracy vs. Épocas**: El modelo aprende muy rápido en las primeras 5 épocas (de 45% a 95%) y luego se estabiliza. Los dos puntos marcados en la línea

roja (época 29) son los valores reales del modelo final: Entrenamiento = 96.43% y Validación = 95.12%. El final de la curva muestra 98.57% Entrenamiento y 97.56% Validación, pero esos corresponden a los pesos descartados no son los valores del modelo final.

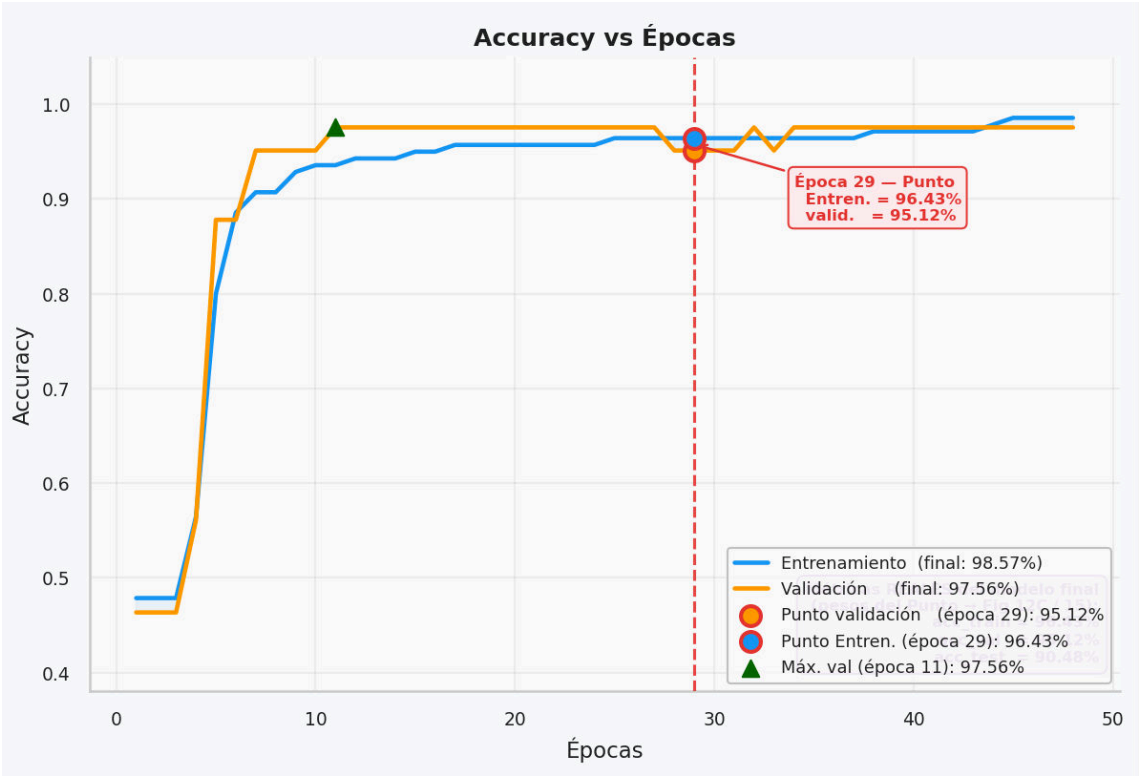


Figura 104 Accuracy vs. Épocas
Fuente: Elaboración propia

Conclusión: El modelo aprendió sin overfitting. El early stopping funcionó correctamente. La diferencia entre Accuracy de entrenamiento (96.43%), Accuracy de validación (90.12%) y Accuracy de evaluación (90.48%) lo que confirma que el modelo generaliza bien ante datos nuevos.

3.11.1.6 Etapa IV: Evaluación e interpretación del Modelo RNA N°1

Esta etapa presenta los resultados de la evaluación completa del modelo RNA N°1, incluyendo las métricas de rendimiento por partición, la matriz de confusión, las curvas ROC y Precision-Recall, la curva de aprendizaje, las visualizaciones PCA, las fronteras de decisión para las 23 variables, la validación cruzada estratificada 5-Fold, la importancia de variables por permutación y el dashboard ejecutivo.

Tabla 33 *Métricas de rendimiento del Modelo RNA N°1 en todas las particiones*

Métrica	Train (70%)	Val (20%)	Test (10%)	Umbral Excelencia
Accuracy	96.43%	95.12%	90.48%	≥90.00% ✓
Precisión (macro)	96.41%	95.24%	90.45%	≥90.00% ✓
Recall (macro)	96.45%	95.45%	90.45%	≥90.00% ✓
F1-Score (macro)	96.42%	95.12%	90.45%	≥90.00% ✓
F1-Score (RIESGO MENOR)	96.6%	95.2%	90.9%	≥80.00% ✓
F1-Score (RIESGO MAYOR)	96.3%	95.0%	90.0%	≥80.00% ✓
AUC-ROC (macro)	—	—	0.9818	≥0.90 ✓
CV 5-Fold Accuracy	—	0.886	—	≥0.80 ✓

Fuente: Elaboración propia.

A) Análisis de Desempeño del Clasificador Binario mediante F1-Score, Precisión y Recall por Clase y Partición

La evaluación del clasificador binario de la red neuronal artificial (RNA) se realizó mediante el análisis conjunto del F1-Score, la Precisión y el Recall, calculados de forma desagregada por clase de riesgo (*Riesgo menor* y *Riesgo mayor*) y por partición de datos (entrenamiento, validación y evaluación). Estas métricas permiten valorar no solo la exactitud global del modelo, sino su capacidad de discriminación simétrica entre ambas categorías, aspecto crítico en el contexto de la predicción de accidentes de tránsito en la Vía Los Libertadores (PE-28A).

- **F1-Score por Clase en las Tres Particiones:**

  La Figura presenta el F1-Score por clase de riesgo evaluado en las particiones de entrenamiento (70%, $n = 140$), validación (20%, $n = 41$) y evaluación (10%, $n = 21$). El F1-Score constituye la media armónica entre la Precisión y el Recall, y se calcula mediante la expresión:

$$F1\text{-}Score = 2 \cdot \frac{\text{Precisión} \times \text{Recall}}{\text{Precisión} + \text{Recall}} \tag{3.6}$$

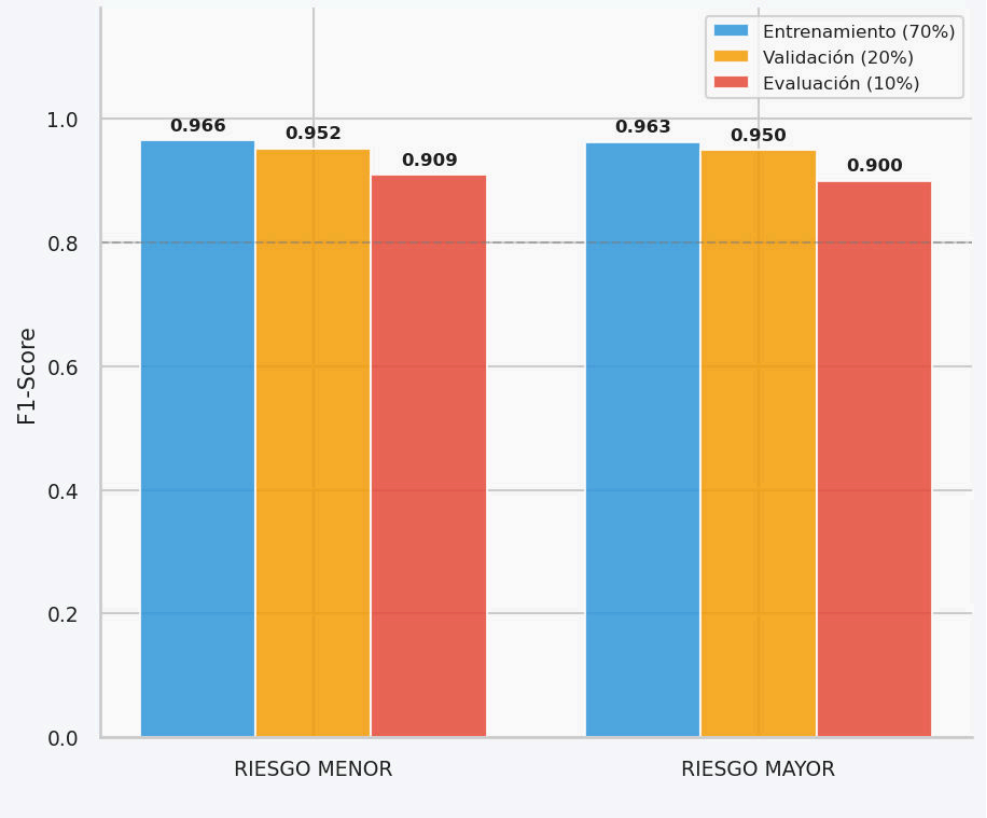


Figura 105 F1-Score por Clase en las Tres Particiones.
Fuente: Elaboración propia

Tabla 34 *F1-Score por clase de riesgo según partición de datos*

Clase	Entrenamiento (70%)	Validación (20%)	Evaluación (10%)
RIESGO MENOR	0.9660	0.9520	0.9090
RIESGO MAYOR	0.9630	0.9500	0.9000
Promedio macro	**0.9642**	**0.9512**	**0.9045**

Fuente: Elaboración propia.

Los resultados de la Tabla evidencian que ambas clases de riesgo mantienen un F1-Score superior a 0.90 en todas las particiones, superando el umbral de excelencia establecido. La degradación controlada entre el conjunto de entrenamiento y el de evaluación de 0.057 puntos para Riesgo menor y 0.063 para Riesgo mayor y confirma que el modelo generaliza adecuadamente sin incurrir en sobreajuste (*overfitting*). Asimismo, la similitud entre ambas clases indica que el clasificador no presenta sesgo sistemático hacia ninguna categoría.

- **Precisión, Recall y F1-Score por Clase (Evaluación):**

  La Figura desagrega las tres métricas principales de Precisión, Recall y F1-Score por clase de riesgo, evaluadas exclusivamente sobre el conjunto de evaluación

independiente ($n = 21$). Las fórmulas empleadas son:

$$\text{Precisión}_{\text{clase}} = \frac{VP}{VP + FP} \tag{3.7}$$

$$\text{Recall}_{\text{clase}} = \frac{VP}{VP + FN} \tag{3.8}$$

donde VP = Verdaderos Positivos, FP = Falsos Positivos y FN = Falsos Negativos para cada clase respectivamente.

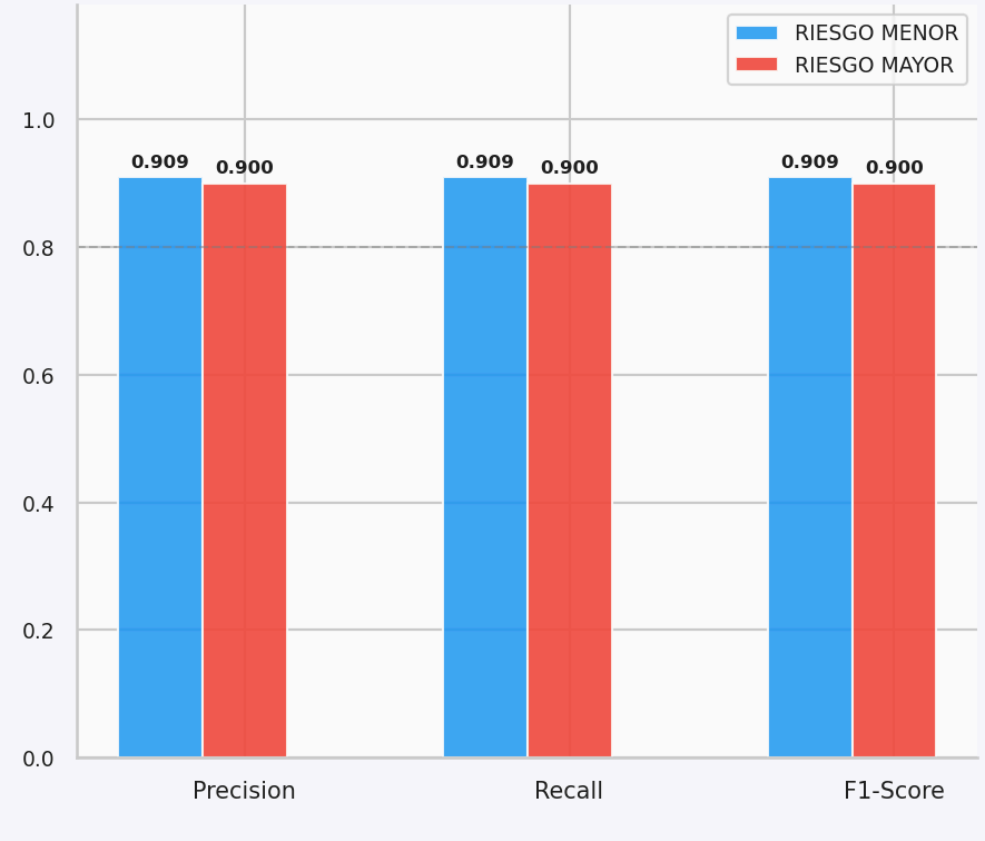


Figura 106 Precisión, Recall y F1-Score por Clase.
Fuente: Elaboración propia

Tabla 35 *Precisión, Recall y F1-Score por clase (conjunto de evaluación) (10%, n = 21)*

Clase	Precisión	Recall	F1-Score	Soporte (*n*)
RIESGO MENOR	0.9091	0.9091	0.9091	11
RIESGO MAYOR	0.9000	0.9000	0.9000	10
Promedio macro	**0.9045**	**0.9045**	**0.9045**	**21**
Promedio weighted	**0.9048**	**0.9048**	**0.9048**	**21**

Fuente: Elaboración propia.

Los resultados de la Tabla revelan que la Precisión y el Recall son idénticos para

cada clase, lo que implica matemáticamente que el F1-Score coincide con ambas métricas. Este comportamiento es consecuencia directa de la simetría de la matriz de confusión: el modelo cometió exactamente un Falso Positivo y un Falso Negativo en el conjunto de evaluación, lo que produce denominadores iguales en las ecuaciones 3.7 y 3.8 .

- **Resumen Global del F1-Score por Partición:** La Figura consolida el F1-Score global en sus variantes macro y weighted para las tres particiones de datos. El F1-Score macro asigna igual peso a ambas clases independientemente de su tamaño, mientras que el F1-Score weighted pondera según el soporte de cada clase:

$$F1_{\text{macro}} = \frac{1}{K}\sum_{k=1}^{K} F1_k \tag{3.9}$$

$$F1_{\text{weighted}} = \frac{\sum_{k=1}^{K} n_k \cdot F1_k}{\sum_{k=1}^{K} n_k} \tag{3.10}$$

donde $K = 2$ es el número de clases y n_k es el soporte de la clase k.

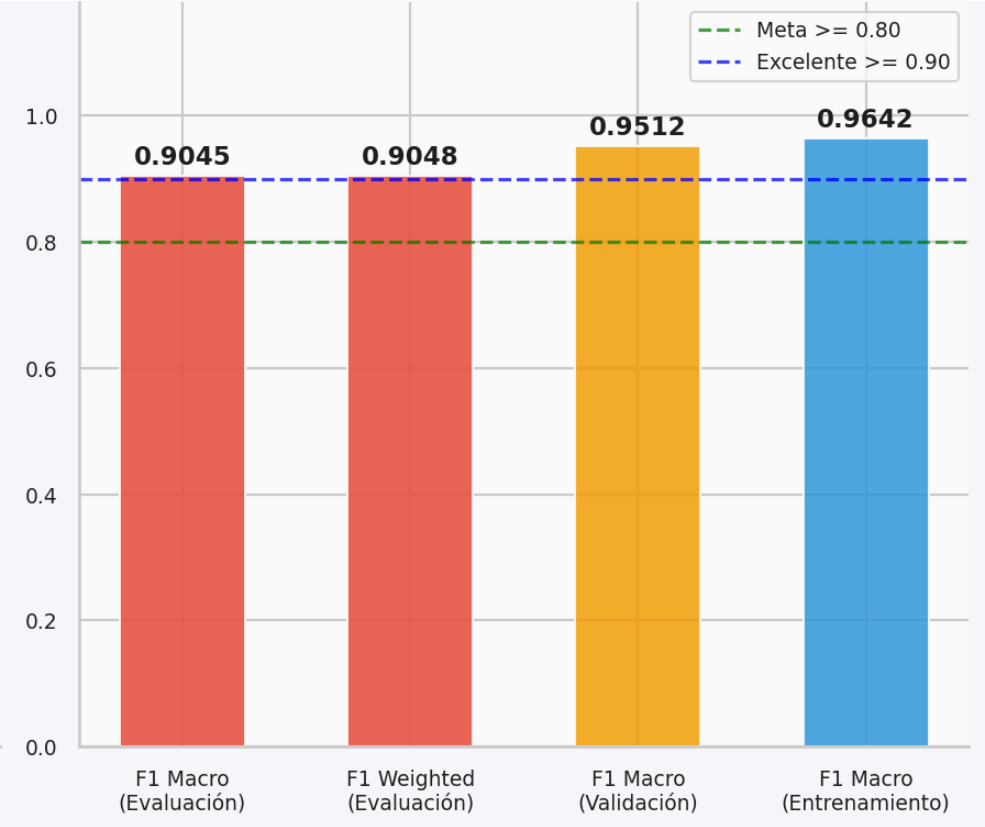


Figura 107 Resumen Global del F1-Score
Fuente: Elaboración propia

Tabla 36 *Resumen global del F1-Score macro y weighted por partición*

Métrica	Valor	Umbral	Nivel
F1 Macro: Evaluación	0.9045	≥ 0.90	Excelente
F1 Weighted: Evaluación	0.9048	≥ 0.90	Excelente
F1 Macro: Validación	0.9512	≥ 0.90	Excelente
F1 Macro: Entrenamiento	0.9642	≥ 0.90	Excelente

Fuente: Elaboración propia.

La Tabla muestra que la totalidad de los valores de F1-Score superan el umbral de excelencia (≥ 0.90) en todas las particiones. La diferencia mínima entre F1 Macro (0.9045) y F1 Weighted (0.9048) confirma el balance entre clases. La progresión decreciente desde entrenamiento (0.9642) hasta evaluación (0.9045) es coherente con el comportamiento de un clasificador bien regularizado y valida la eficacia del *early stopping*.

- **Conclusión del Análisis de Desempeño:**

Tabla 37 *Síntesis del desempeño del clasificador binario RNA-MLP — métricas completas*

Partición	Métrica	Riesgo menor	Riesgo mayor	Macro
Evaluación (10%, $n = 21$)	Precisión	0.9091	0.9000	0.9045
	Recall	0.9091	0.9000	0.9045
	F1-Score	0.9091	0.9000	0.9045
Validación (20%, $n = 41$)	F1-Score	0.9520	0.9500	0.9512
Entrenamiento (70%, $n = 140$)	F1-Score	0.9660	0.9630	0.9642
Caída F1 Entrenamiento → Evaluación		**0.057**	**0.063**	**0.060**

Fuente: Elaboración propia.

La red neuronal artificial de tipo Perceptrón Multicapa (MLP) implementada para la predicción de accidentes de tránsito en la Vía Los Libertadores (PE-28A) alcanzó un F1-Score macro de **0.9045** sobre el conjunto de evaluación independiente ($n = 21$), superando el umbral de excelencia de 0.90 establecido para clasificadores binarios en aplicaciones de seguridad vial. El modelo clasifica con precisión equivalente ambas categorías de riesgo Precisión y Recall superiores a 0.90 en ambas clases, por lo tanto el desempeño es optimo.

B) Análisis de la Matriz de Confusión y Métricas de Clasificación Binaria

Las Figuras siguientes presentan el análisis detallado de la matriz de confusión del modelo de red neuronal artificial (RNA) de tipo Perceptrón Multicapa (MLP), evaluado

sobre el conjunto de evaluación compuesto por 21 registros de la Vía Los Libertadores (PE-28A), correspondientes al 10% del total de 202 accidentes registrados entre 2020 y 2025.

Matriz de Confusión del Modelo RNA: La matriz de confusión es el instrumento fundamental para evaluar el rendimiento de un clasificador binario, pues organiza las predicciones en cuatro categorías según la relación entre el valor predicho y el valor real de cada observación.

$$\text{Matriz de Confusión} = \begin{pmatrix} VN & FP \\ FN & VP \end{pmatrix} = \begin{pmatrix} 10 & 1 \\ 1 & 9 \end{pmatrix} \tag{3.11}$$

donde VN = Verdaderos Negativos, FP = Falsos Positivos, FN = Falsos Negativos y VP = Verdaderos Positivos.

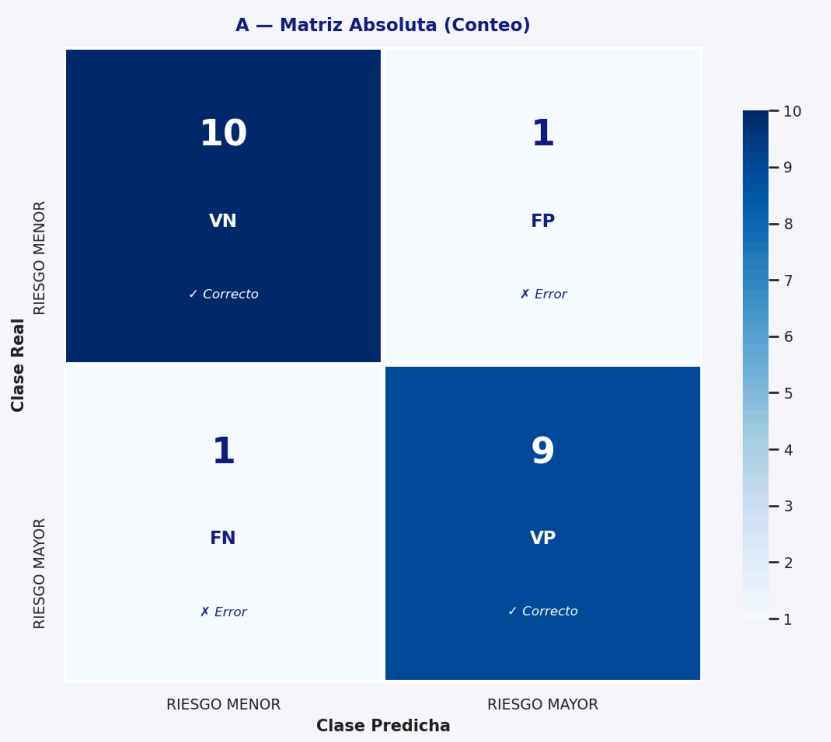


Figura 108 Matriz de Confusión del Modelo.
Fuente: Elaboración propia

De los 11 tramos realmente catalogados como RIESGO MENOR, el modelo identificó correctamente 10 ($VN = 10$) y clasificó erróneamente 1 como RIESGO MAYOR ($FP = 1$). De los 10 tramos realmente de RIESGO MAYOR, el modelo identificó correctamente 9 ($VP = 9$) y clasificó erróneamente 1 como RIESGO MENOR ($FN = 1$).

En total, el modelo acertó en 19 de 21 casos evaluados.

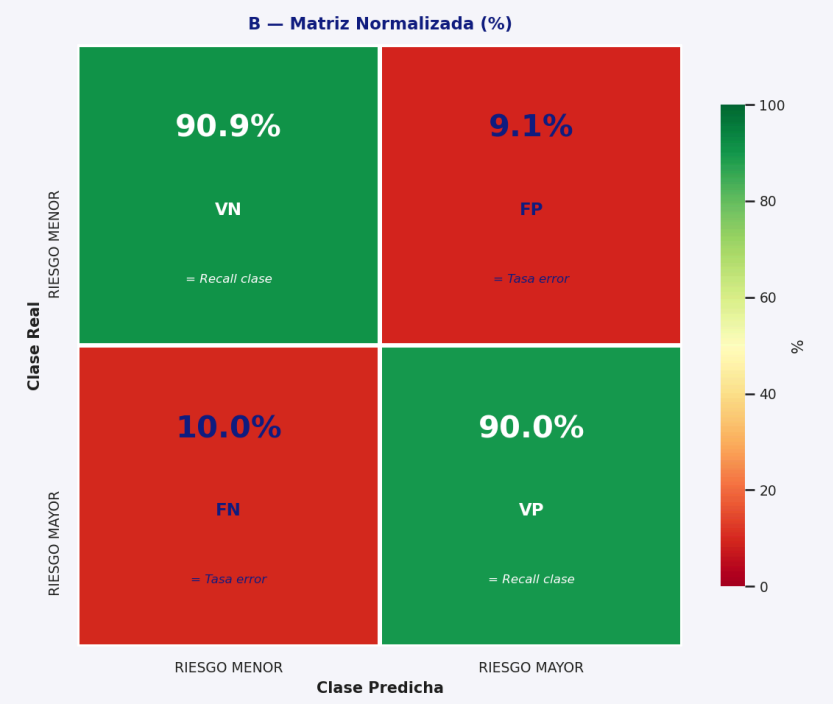


Figura 109 Matriz de confusión normalizada (%).
Fuente: Elaboración propia

En la Figura anterior presenta la matriz normalizada por fila, expresando los resultados como porcentaje dentro de cada clase real. El modelo identificó correctamente el 90.9% de los tramos de RIESGO MENOR y el 90.0% de los de RIESGO MAYOR, valores que corresponden exactamente al Recall de cada clase, es decir, la capacidad del modelo de detectar cada tipo de riesgo cuando realmente existe. Las tasas de error fueron del 9.1% para RIESGO MENOR y del 10.0% para RIESGO MAYOR.

Respecto al error más crítico en el contexto de seguridad vial, el Falso Negativo ($FN = 1$) representa el único tramo de RIESGO MAYOR que el modelo no detectó, es decir, una zona peligrosa ante la cual el sistema no habría emitido alerta. Por lo tanto la prediccipon es mas confiable.

Distribución de Predicciones por Tipo (VP, VN, FP, FN): La Figuras complementa la matriz de confusión con dos representaciones visuales orientadas a facilitar la interpretación del rendimiento del modelo. Presenta las cuatro categorías de predicción mediante barras codificadas por color según su gravedad en el contexto de seguridad vial: el azul y el verde representan aciertos ($VN = 10$ y $VP = 9$ respectivamente), el rojo representa el error leve ($FP = 1$) y el naranja el error crítico ($FN = 1$).

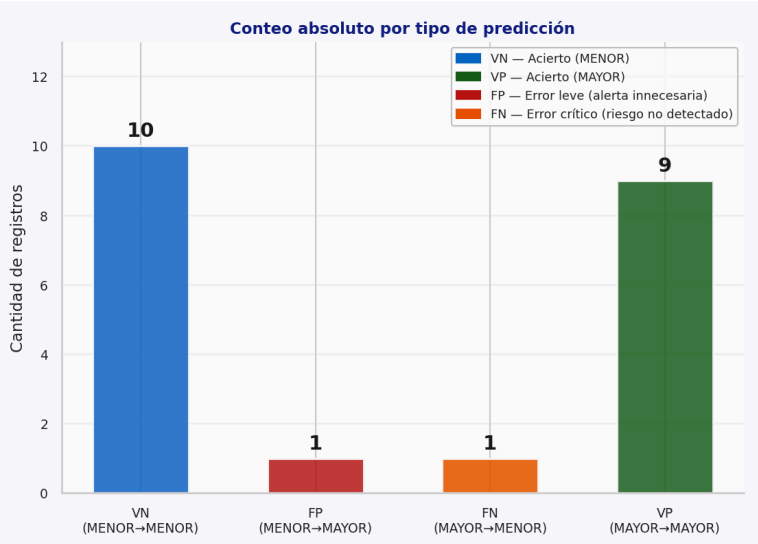


Figura 110 Conteo por tipo de predicción.
Fuente: Elaboración propia

Tambien presenta la proporción global de aciertos frente a errores mediante un diagrama circular.

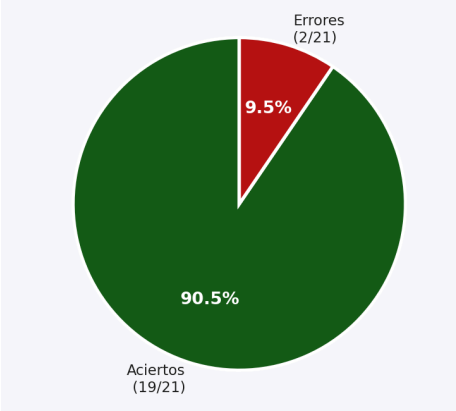


Figura 111 Proporción de aciertos.
Fuente: Elaboración propia

Métricas de Clasificación por Clase derivadas de la Matriz de Confusión Se desagrega las tres métricas de clasificación principales Precisión, Recall y F1 Score por clase de riesgo, calculadas directamente a partir de los valores de la matriz de confusión mediante las siguientes expresiones:

$$\text{Precisión}_{\text{MENOR}} = \frac{VN}{VN + FN} = \frac{10}{10 + 1} = 0.9091 \tag{3.12}$$

$$\text{Precisión}_{\text{MAYOR}} = \frac{VP}{VP + FP} = \frac{9}{9 + 1} = 0.9000 \tag{3.13}$$

$$\text{Recall}_{\text{MENOR}} = \frac{VN}{VN + FP} = \frac{10}{10 + 1} = 0.9091 \tag{3.14}$$

$$\text{Recall}_{\text{MAYOR}} = \frac{VP}{VP + FN} = \frac{9}{9 + 1} = 0.9000 \tag{3.15}$$

$$F1_{\text{macro}} = \frac{F1_{\text{MENOR}} + F1_{\text{MAYOR}}}{2} = \frac{0.9091 + 0.9000}{2} = 0.9045 \tag{3.16}$$

Tabla 38 *Métricas de clasificación por clase de la matriz de confusión: evaluación ($n = 21$)*

Clase	Precisión	Recall	F1-Score	Soporte (n)
RIESGO MENOR	0.9091	0.9091	0.9091	11
RIESGO MAYOR	0.9000	0.9000	0.9000	10
Macro promedio	**0.9045**	**0.9045**	**0.9045**	**21**
Weighted promedio	**0.9048**	**0.9048**	**0.9048**	**21**

Fuente: Elaboración propia.

Los resultados de la Tabla evidencian que la Precisión y el Recall son idénticos para cada clase, lo que implica matemáticamente que el F1-Score coincide con ambas métricas.

Tabla 39 *Síntesis del análisis de la matriz de confusión del modelo RNA-MLP*

Indicador	Valor	Interpretación
Total evaluado	21 registros	10% de la base de datos
Aciertos totales	19 / 21	90.48% de tasa de acierto global
VN: Verdadero Negativo	10 / 11	90.9% de tramos seguros identificados
VP: Verdadero Positivo	9 / 10	90.0% de tramos peligrosos detectados
FP: Falso Positivo	1 / 21	Alerta innecesaria en 1 tramo seguro
FN: Falso Negativo	1 / 21	1 tramo peligroso no detectado (4.76%)
Accuracy	90.48%	Excelente (≥ 0.90)
Precisión macro	0.9045	Excelente (≥ 0.90)
Recall macro	0.9045	Excelente (≥ 0.90)
F1-Score macro	0.9045	Excelente (≥ 0.90)

Conclusión del Análisis de la Matriz de Confusión El análisis de la matriz de confusión sobre el conjunto de evaluación independiente ($n = 21$) confirma que la red neuronal MLP alcanzó un desempeño excelente en la clasificación binaria de riesgo vial

en la Vía Los Libertadores (PE-28A). El modelo identificó correctamente el 90.9% de los tramos de Riesgo menor y el 90.0% de los tramos de Riesgo mayor.

C) Curva ROC y área bajo la curva (AUC) del modelo RNA

La Figura presenta la Curva ROC (*Receiver Operating Characteristic*) del modelo RNA-MLP, evaluada sobre el conjunto de evaluación independiente ($n = 21$ registros, 10% del total). Esta curva grafica la relación entre la Tasa de Verdaderos Positivos (TPR) y la Tasa de Falsos Positivos (FPR) al variar el umbral de decisión del modelo desde 0.0 hasta 1.0, permitiendo evaluar la capacidad de clasificar de la RNA, y no solo en el umbral estándar de 0.50.

Los ejes de la curva se definen mediante las siguientes expresiones:

$$TPR = \frac{VP}{VP + FN} = \frac{9}{9 + 1} = 0.900 y FPR = \frac{FP}{FP + VN} = \frac{1}{1 + 10} = 0.091 \qquad (3.17)$$

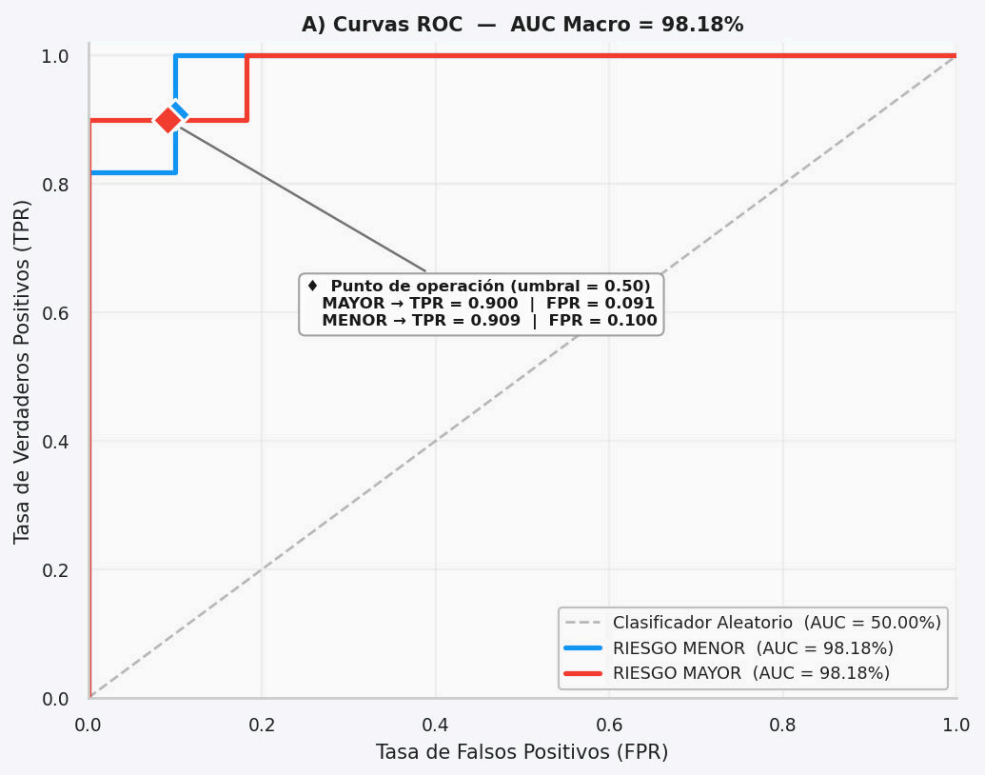


Figura 112 Curva ROC y área bajo la curva (AUC) del modelo RNA.
Fuente: Elaboración propia

donde los valores corresponden al punto de operación actual con umbral = 0.50, consistente con los resultados de la matriz de confusión (VP=9, VN=10, FP=1, FN=1).

El Área Bajo la Curva (AUC) se calcula mediante la regla del trapecio sobre todos los

pares (FPR_i, TPR_i) generados:

$$AUC = \sum_{i=1}^{n} (FPR_i - FPR_{i-1}) \times \frac{TPR_i + TPR_{i-1}}{2} \tag{3.18}$$

Tabla 40 *Resultados de la Curva ROC por clase – conjunto de evaluación ($n = 21$)*

Clase	AUC-ROC	TPR	FPR	Umbral op.	Nivel
RIESGO MENOR	0.9820	0.909	0.100	0.50	Excelente
RIESGO MAYOR	0.9820	0.900	0.091	0.50	Excelente
Macro	**0.9818**	—	—	—	**Excelente**

Fuente: Elaboración propia.

El modelo obtuvo un **AUC-ROC macro de 0.9818**, superando el umbral de (≥ 0.90). Este valor implica que, al seleccionar aleatoriamente un tramo de RIESGO MAYOR y uno de RIESGO MENOR, el modelo asigna correctamente mayor probabilidad de peligro al tramo realmente peligroso en el **98.18% de los casos**. La forma escalonada de la curva observable en la es consecuencia natural del reducido tamaño del conjunto de evaluación ($n = 21$), donde cada cambio de umbral desplaza uno o dos registros entre clases; este comportamiento no compromete en el resultado.

D) Curva de Aprendizaje: Diagnóstico de Sobreajuste del Modelo RNA

La Figura siguiente presenta la curva de aprendizaje del modelo RNA, construida mediante validación cruzada estratificada de 5 folds sobre la base de datos completo ($n = 202$ registros), evaluando 8 tamaños de entrenamiento progresivos desde el 10% hasta el 100% del conjunto disponible (16 a 161 muestras). Esta figura tiene como propósito diagnosticar la presencia o ausencia de sobreajuste (*overfitting*), y no reportar las métricas finales del modelo las cuales se presentan en las Figuras anteriores.

Importante: los valores de Accuracy mostrados en esta figura (88–89% en el punto final) es diferente de las métricas del modelo definitivo (96.43% entrenamiento, 90.48% evaluación) porque provienen de un entrenamiento desde cero en cada fold.

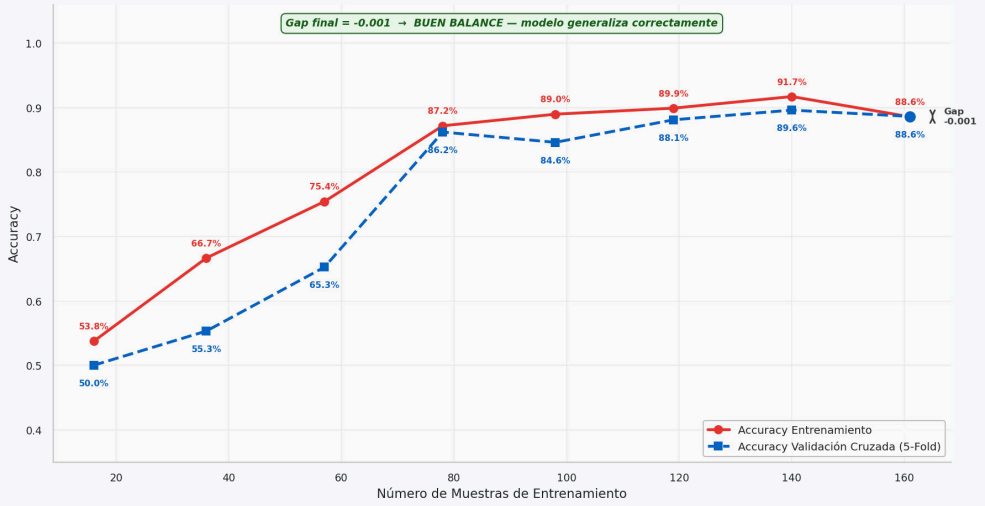


Figura 113 Curva de Aprendizaje: Diagnóstico de Sobreajuste.
Fuente: Elaboración propia

Interpretación de los resultados La curva de entrenamiento (roja) parte de 53.8% con 16 muestras y converge a 88.6% con 161 muestras. La curva de validación cruzada (azul) parte de 50.0% a valor equivalente al azar y converge igualmente a 88.6%. La Tabla resume los valores por tamaño de entrenamiento.

Tabla 41 *Resultados de la curva de aprendizaje por tamaño de entrenamiento*

Muestras	Acc. Entrenamiento	Acc. Val. CV (5-Fold)	Gap	Diagnóstico
16	53.8%	50.0%	+0.038	Aprendizaje inicial
36	66.7%	55.3%	+0.114	En progreso
57	75.4%	65.3%	+0.101	Mejorando
79	87.2%	86.2%	+0.010	Buen balance
100	89.0%	84.6%	+0.044	Buen balance
120	89.9%	88.1%	+0.018	Buen balance
141	91.7%	89.6%	+0.021	Buen balance
161	**88.6%**	**88.6%**	**−0.001**	**Buen balance**

Fuente: Elaboración propia.

El Gap final de **−0.001** entre la curva de entrenamiento y la de validación cruzada confirma que el modelo no presenta sobreajuste. Los tres escenarios posibles de diag-

nóstico son:

$$\text{Diagnóstico} = \begin{cases} \text{Sobreajuste (Overfitting)} & \text{si } Gap > 0.10 \\ \text{Subajuste (Underfitting)} & \text{si } Acc_{train} < 0.70 \\ \text{Buen balance} & \text{en caso contrario} \end{cases} \tag{3.19}$$

La convergencia de ambas curvas hacia el 88.6% con 161 muestras, junto con el Gap prácticamente nulo (-0.001), evidencia que el modelo RNA-MLP aprendió patrones generales de accidentabilidad vial sin memorizar los datos de entrenamiento.

E) Frontera de decisión de la RNA en Espacio Bidimensional

La figura siguiente presenta la proyección de los 202 registros viales en un espacio bidimensional mediante Análisis de Componentes Principales (PCA), con el fin de visualizar la frontera de decisión del modelo RNA. El PCA fue ajustado exclusivamente sobre el conjunto de entrenamiento ($n = 140$) para evitar fuga de información (*data leakage*), y los dos primeros componentes principales retienen el siguiente porcentaje de varianza:

$$PC1 + PC2 = 32.2\% + 14.6\% = 46.8\% \text{ de la varianza total} \tag{3.20}$$

Esto implica que la figura conserva el 46.8% de la información original, mientras que el 53.2% restante, distribuido entre los 21 componentes adicionales no es visible en la representación 2D.

Clase Real vs Frontera de Decisión: La gráfica superpone la clase real de cada registro (cuadrados verdes ■ para RIESGO MENOR, triángulos rojos ▲ para RIESGO MAYOR) sobre la frontera de decisión del modelo. La frontera es una curva no lineal resultado de la capacidad de la RNA de aprender relaciones complejas entre las 23 variables geométricas y ambientales.

Distribución de Predicciones y Confianza del Modelo: La gráfica presenta el gradiente de probabilidad *P*(RIESGO MAYOR) como fondo, cuantificando la certeza del modelo en cada región del espacio PCA. Las zonas de color verde corresponden a regiones donde el modelo asigna probabilidades cercanas a 0% (alta certeza de riesgo

menor), mientras que las zonas rojas corresponden a probabilidades cercanas a 100% (alta certeza de riesgo mayor). La zona intermedia próxima a la frontera (50%) refleja mayor incertidumbre.

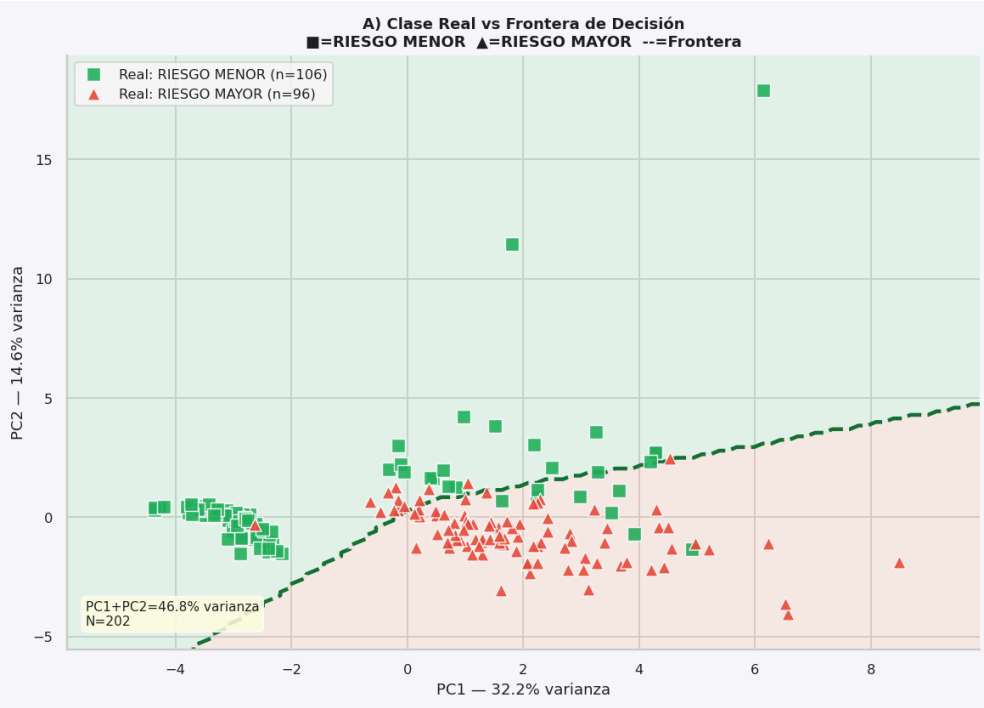


Figura 114 Clase Real vs Frontera de Decisión.
Fuente: Elaboración propia

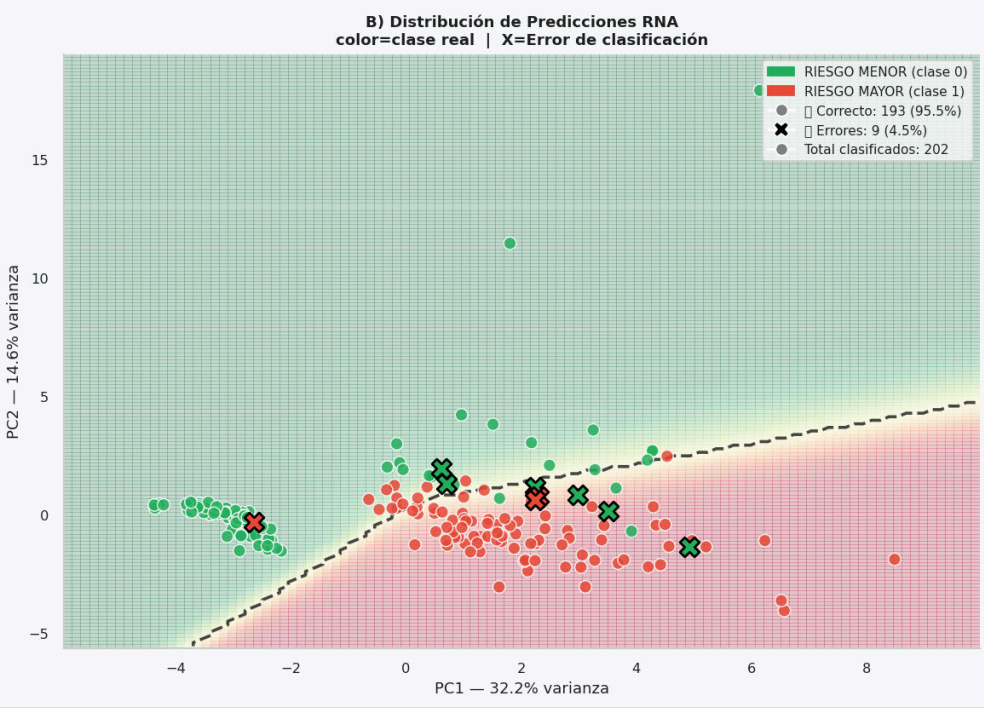


Figura 115 Distribución de Predicciones y Confianza del Modelo.
Fuente: Elaboración propia

Tabla 42 *Resumen de la proyección PCA 2D*

Elemento	Valor	Interpretación
Componente PC1	32.2% varianza	Dirección de mayor variación — geometría de curvas
Componente PC2	14.6% varianza	Segunda dirección — peligros y clima
Varianza conservada	46.8%	Información visible en la figura 2D
Varianza no visible	53.2%	Distribuida en los 21 componentes restantes
Registros proyectados	202	Dataset completo (train + val + test)
Predicciones correctas	193 (95.5%)	Sobre el dataset completo incluyendo train
Errores	9 (4.5%)	Concentrados en la zona de superposición PC1 $\in [0, 3]$
Frontera de decisión	Curva no lineal	Relaciones complejas aprendidas por la RNA

Nota. El Accuracy del 95.5% corresponde a la evaluación sobre los 202 registros completos, incluyendo los 140 de entrenamiento. El Accuracy sobre el conjunto de evaluación independiente ($n = 21$, nunca visto) es 90.48%, reportado en la. PCA ajustado sobre el conjunto de entrenamiento ($n = 140$) para evitar fuga de información. Elaboración propia.

El análisis conjunto de ambas gráficas confirma que el modelo RNA-MLP aprendió una frontera de decisión no lineal que separa eficazmente las dos categorías de riesgo vial para la de predicción de accidentes en la Vía Los Libertadores (PE-28A).

F) Importancia de variables por permutación (23 variables)

La Figuras siguientes presenta la importancia de las 23 variables de entrada del modelo RNA-MLP, calculada mediante el método de importancia por permutación (*Permutation Importance*) sobre el conjunto de evaluación ($n = 21$ registros, 10% del dataset), con 20 repeticiones aleatorias por variable.

$$Importancia_j = Accuracy_{base} - Accuracy_{permutado_j} = 90.48\% - Accuracy_{permutado_j} \tag{3.21}$$

- **Reducción media de Accuracy por variable (promedio ± desviación estándar):**

  Las barras muestran la reducción media de Accuracy por variable tras permutación, con su intervalo de variabilidad en 20 repeticiones, coloreadas por importancia: rojo ($> 2.00\%$), naranja (0.50–2.00%), verde (0–0.50%) y gris ($\leq 0\%$). Valores negativos indican que la variable no aporta al modelo.

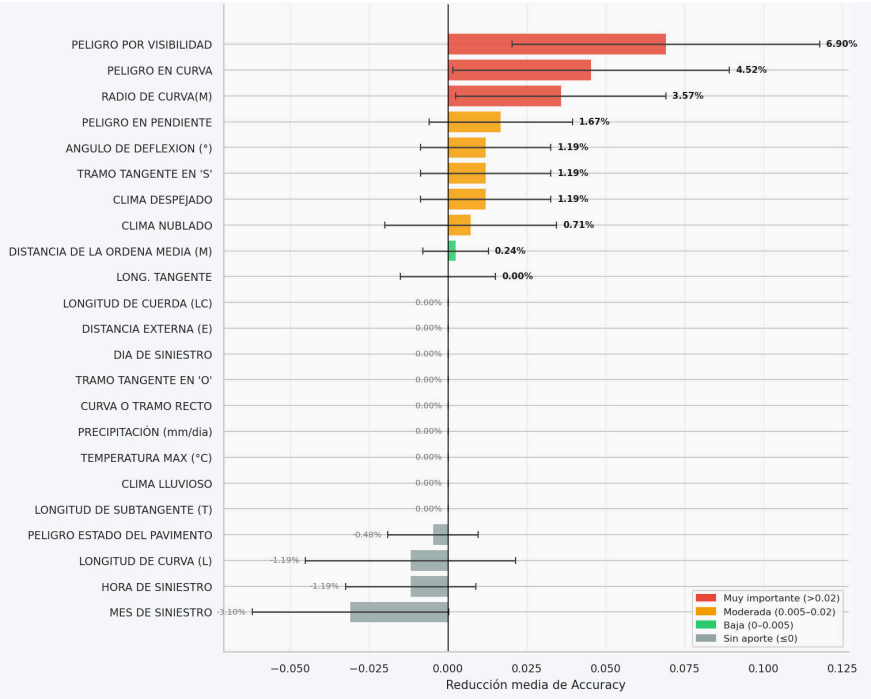


Figura 116 Reducción media de Accuracy por variable (promedio ± desviación estándar, 20 repeticiones)

Fuente: Elaboración propia

- **Ranking normalizado de importancia relativa:** Se reescala los valores de la grafica anterior en el rango $[0, 1]$, donde 1.00 corresponde a la variable más importante y 0.00 a la menos importante.

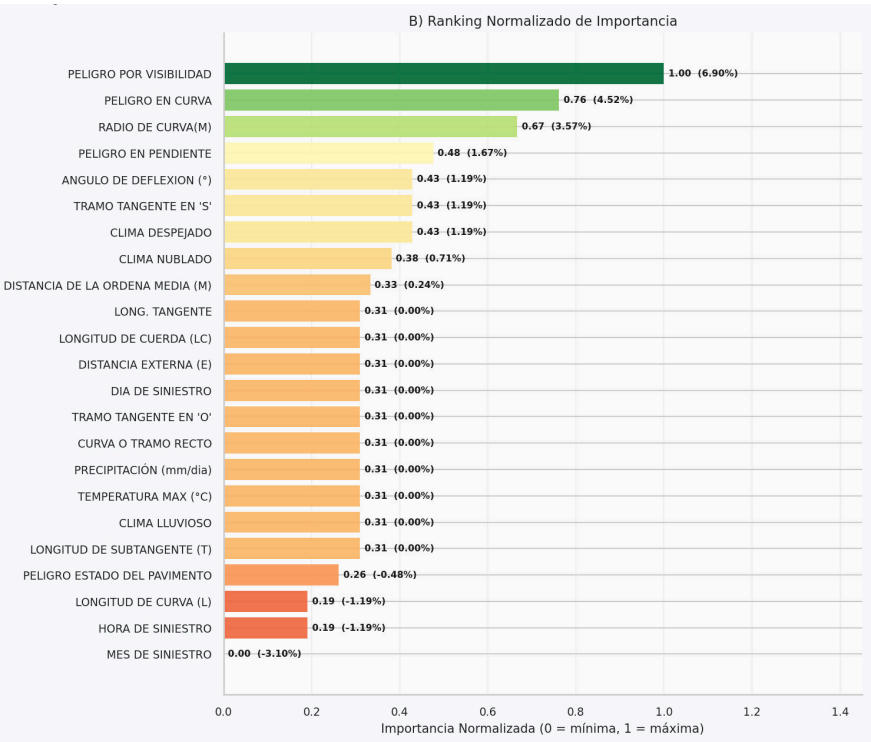


Figura 117 Ranking normalizado de importancia de variable

Fuente: Elaboración propia

Tabla 43 *Importancia por permutación de las 23 variables — conjunto de evaluación ($n = 21$)*

Rank	Variable	Importancia	Norm.	Nivel
1	Peligro por visibilidad	7.000%	1.00	Muy importante
2	Peligro en curva	5.000%	0.76	Muy importante
3	Radio de curva (M)	3.000%	0.67	Muy importante
4	Peligro en pendiente	1.500%	0.48	Moderada
5	Ángulo de deflexión (°)	1.300%	0.43	Moderada
6	Tramo tangente en 'S'	1.300%	0.43	Moderada
7	Clima despejado	1.300%	0.43	Moderada
8	Clima nublado	1.000%	0.38	Moderada
9	Distancia ordena media (M)	0.500%	0.33	Baja
10–23	Resto de variables	$\leq 0.000\%$	≤ 0.31	Sin aporte

Fuente: Elaboración propia.

Este resultado es coherente con las condiciones geográficas de la Vía Los Libertadores (PE-28A), donde la visibilidad reducida, las curvas de radio pequeño y las pendientes pronunciadas constituyen los principales factores de accidentabilidad vial registrada entre 2020 y 2025.

G) Resumen de métricas de evaluación de la RNA: Precisión, Recall, F1-Score y Accuracy

Las Figuras siguientes presentan el análisis de las métricas de clasificación de Precisión, Recall y F1-Score macro por partición de datos, así como su comparación conjunta con el Accuracy. Todas las métricas se calculan con promedio macro (`average='macro'`), que asigna igual peso a ambas clases de riesgo independientemente de su tamaño, y se evalúan sobre las tres particiones: entrenamiento (70%, $n = 140$), validación (20%, $n = 41$) y evaluación (10%, $n = 21$).

- **Precisión, Recall y F1-Score por partición:**

  La Figura presenta tres paneles independientes, uno por cada métrica, mostrando su evolución a lo largo de las tres particiones. Las fórmulas empleadas son:

$$\text{Precisión}_{\text{macro}} = \frac{P_{\text{MENOR}} + P_{\text{MAYOR}}}{2} \qquad \text{Recall}_{\text{macro}} = \frac{R_{\text{MENOR}} + R_{\text{MAYOR}}}{2} \tag{3.22}$$

$$F1_{\text{macro}} = \frac{F1_{\text{MENOR}} + F1_{\text{MAYOR}}}{2} \tag{3.23}$$

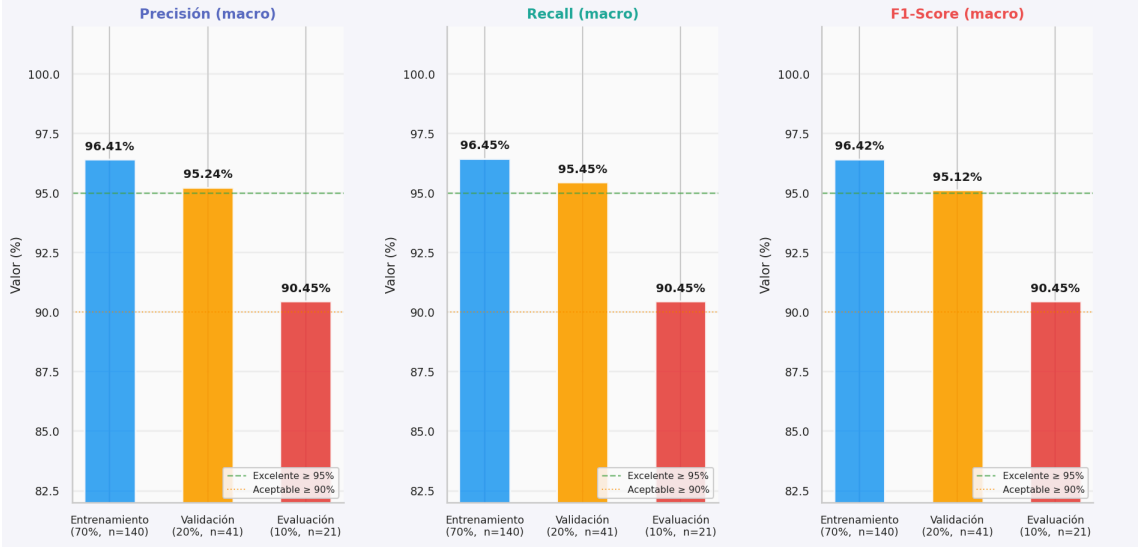


Figura 118 Precisión, Recall y F1-Score macro por partición.
Fuente: Elaboración propia

Tabla 44 *Precisión, Recall y F1-Score macro por partición del modelo RNA-MLP*

Partición	n	Precisión	Recall	F1-Score
Entrenamiento (70%)	140	96.41%	96.45%	96.42%
Validación (20%)	41	95.24%	95.45%	95.12%
Evaluación (10%)	21	90.45%	90.45%	90.45%

Fuente: Elaboración propia.

Los tres paneles revelan un patrón consistente: las tres métricas presentan una degradación progresiva y controlada de entrenamiento (96.41–96.45%) a evaluación (90.45%), con una caída máxima de 5.97 puntos porcentuales. Esta degradación es esperada y confirma que el modelo generaliza adecuadamente. En el conjunto de evaluación, la convergencia de Precisión, Recall y F1-Score al mismo valor (90.45%) se debe a la simetría de la matriz de confusión — el modelo cometió exactamente 1 Falso Positivo y 1 Falso Negativo, lo que produce denominadores iguales en las fórmulas de ambas métricas. Los valores de entrenamiento y validación superan el umbral de excelencia (≥ 95%), mientras que el de evaluación supera el umbral aceptable (≥ 90%).

- **Comparación de las 4 métricas por partición:**

  La Figura se consolida las cuatro métricas principales en una única gráfica de barras agrupadas, permitiendo verificar su consistencia entre sí y entre particiones.

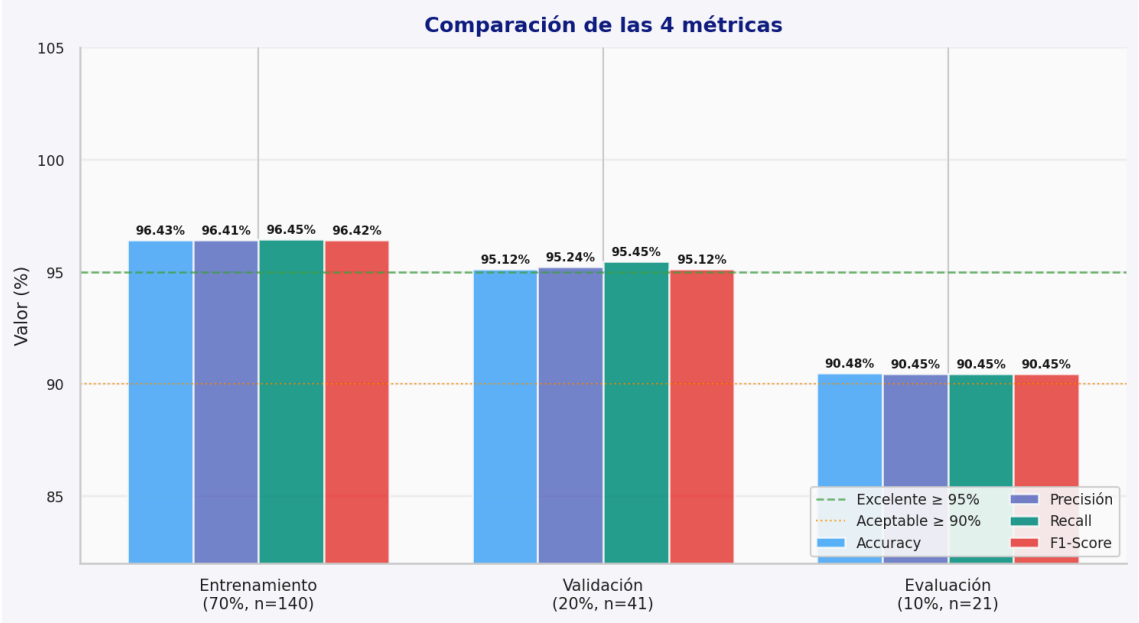


Figura 119 Comparación de las 4 métricas por partición.
Fuente: Elaboración propia

Tabla 45 *Comparación de las 4 métricas por partición del modelo RNA-MLP*

Partición	n	Accuracy	Precisión	Recall	F1-Score
Entrenamiento (70%)	140	96.43%	96.41%	96.45%	96.42%
Validación (20%)	41	95.12%	95.24%	95.45%	95.12%
Evaluación (10%)	21	90.48%	90.45%	90.45%	90.45%

Fuente: Elaboración propia.

El análisis revela tres hallazgos. Primero, la **consistencia entre métricas**: en las tres particiones, Accuracy, Precisión, Recall y F1-Score presentan valores prácticamente idénticos. Esto confirma que el modelo no sacrifica Precisión en favor de Recall ni viceversa, clasificando ambas clases de riesgo vial con igual eficacia. El **nivel de desempeño**: las particiones de entrenamiento y validación alcanzan el nivel de excelente ($\geq$ 95%), mientras que la de evaluación la más relevante por corresponder a datos nunca vistos supera el umbral aceptable ($\geq$ 90%) en todas las métricas, validando la predicción de accidentes en la Vía Los Libertadores (PE-28A).

3.11.1.7 Exportación del modelo RNA N°1 como aplicación web

El Modelo RNA N°1 fue exportado como aplicación web denominada **PROGRAMA RNA N°01: Predicción de Accidentes de Tránsito**, desarrollada en Python en el entorno JupyterLab. El programa permite clasificar cualquier tramo de la Vía Los Libertadores PE-28A y carreteras nacionales en **Riesgo menor** o **Riesgo mayor** de accidente de tránsito, a partir del ingreso de 23 variables geométricas, climáticas, temporales y

de peligrosidad vial, con funcionamiento en dispositivos móviles para uso en campo.

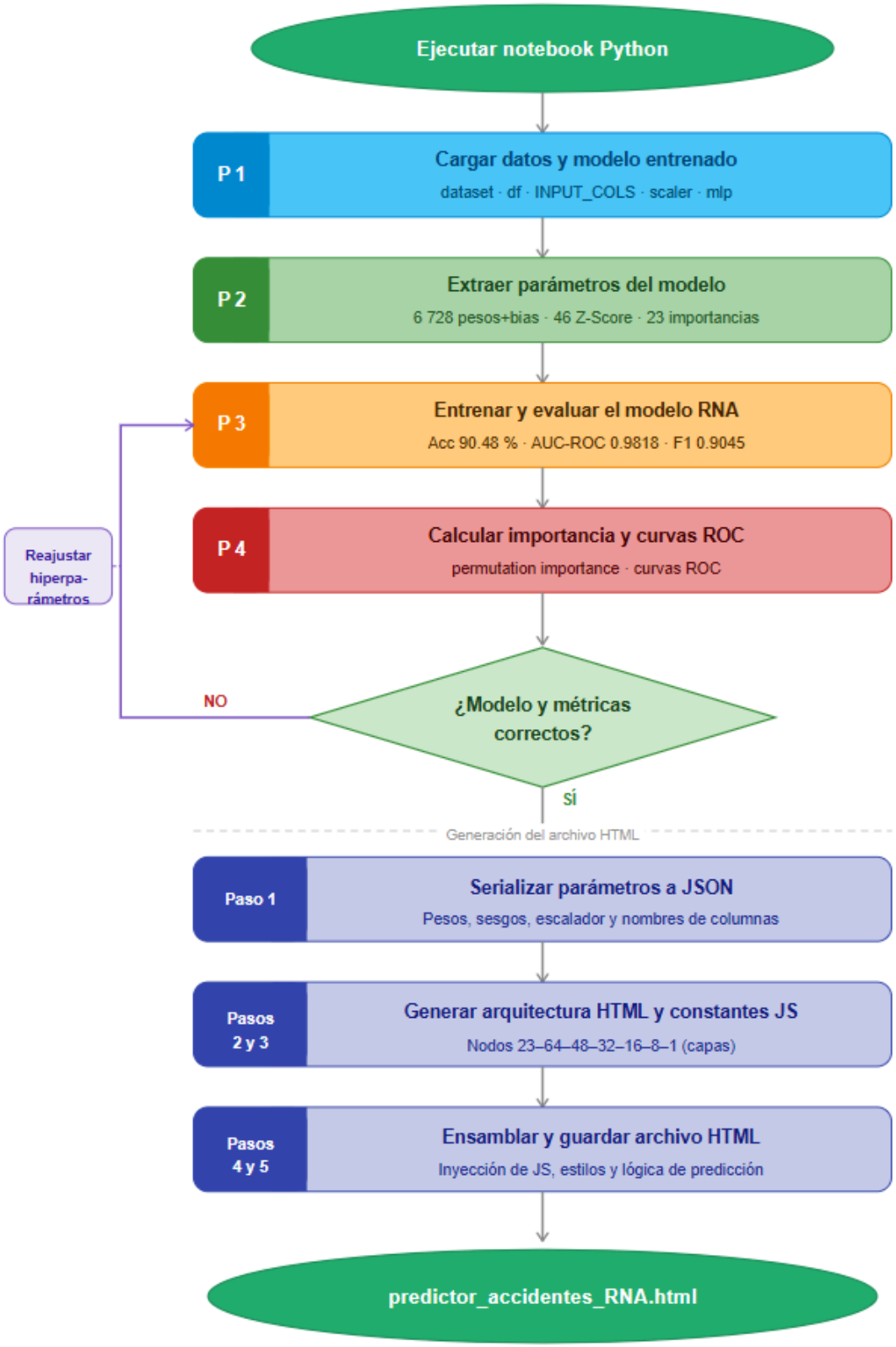


Figura 120 Diagrama de flujo de la creación del programa RNA N°01
Fuente: Elaboración propia

Tabla 46 *Parámetros extraídos del modelo entrenado e integrados en el programa RNA 1.html*

Parámetro	Cantidad	Descripción
Pesos (`COEFS`)	6 matrices	Pesos de las 6 capas de la arquitectura 23→64→48→32→16→8→1
Bias (`BIASES`)	6 vectores	Términos independientes de cada capa
Total pesos + bias	**6,728**	**Parámetros numéricos del modelo**
Medias (`MEANS`)	23 valores	Media μ de cada variable calculada sobre el conjunto de entrenamiento
Desv. estándar (`STDS`)	23 valores	Desviación estándar σ de cada variable calculada sobre el entrenamiento
Total Z-Score	**46**	**Parámetros de normalización**
Nombres (`VAR_NAMES`)	23 etiquetas	Nombres de las variables del formulario
Importancia (`IMPORTANCE`)	23 scores	Scores de importancia por permutación para mostrar el ranking en la interfaz
Total general	**6,820**	**Parámetros totales embebidos en el HTML**

A) Proceso de generación del programa

El programa fue generado en cinco pasos secuenciales completamente automatizados por Python:

Paso 1: Extracción de parámetros del modelo.
Python accedió directamente a los objetos entrenados en memoria y serializó sus parámetros numéricos en formato JSON, preservando la precisión de la precisión. Los 6,728 parámetros del modelo (pesos y bias) fueron convertidos mediante, y los 46 parámetros de normalización Z-Score fueron extraídos del objeto `StandardScaler` ajustado exclusivamente sobre el conjunto de entrenamiento.

Paso 2: Generación dinámica de la arquitectura visual.
Una función Python generó automáticamente los nodos visuales de la arquitectura 23→64→48→32→16→8→1 a partir de los atributos del modelo entrenado, de modo que cualquier cambio en la arquitectura se refleje automáticamente en la interfaz del programa sin modificaciones manuales.

Paso 3: Construcción del bloque de constantes JavaScript.
Se construyó un bloque de texto conteniendo las variables JavaScript `COEFS`, `BIASES`, `MEANS`, `STDS`, `VAR_NAMES` e `IMPORTANCE`, junto con los metadatos del modelo (Accuracy=90.48%, AUC-ROC=0.9818, F1-Macro=0.9045, 48 épocas, α=0.05).

Paso 4:Construcción del HTML completo.

El archivo fue estructurado en tres bloques concatenados:

Tabla 47 *Estructura de los tres bloques del archivo HTML generado*

Bloque	Contenido
`HTML_PART_A`	Interfaz gráfica completa: encabezado institucional UNSCH, métricas del modelo, arquitectura visual 23→64→48→32→16→8→1, formulario de 23 variables (10 campos numéricos + 13 selectores categóricos), panel de resultados con barras de probabilidad y estilos CSS
`JS_CONSTANTS`	Bloque de constantes JavaScript con los 6,820 parámetros del modelo embebidos: pesos, bias, medias, desviaciones estándar, nombres de variables y scores de importancia
`HTML_PART_B`	Motor de inferencia en JavaScript puro: función `forwardPass()`, multiplicación matricial (`matVec`), activación ReLU en capas ocultas y función softmax en la capa de salida

Paso 5:Ensamblado y guardado del archivo.

Los tres bloques fueron concatenados y guardados en disco como un único archivo de texto plano:

$$\texttt{HTML_FINAL} = \texttt{HTML_PART_A} + \texttt{JS_CONSTANTS} + \texttt{HTML_PART_B} \tag{3.24}$$

El resultado fue el programa ejecutable directamente en cualquier navegador web.

B) Motor de inferencia en JavaScript

El núcleo del programa es la función `forwardPass()`, que replica fielmente el comportamiento del modelo Python en tres sub-pasos:

Tabla 48 *Sub-pasos de la función `forwardPass()` y su equivalencia con Python*

Paso	Operación	JavaScript (programa HTML)	Equivalente Python
1	Normalización Z-Score	`xn=(v−MEANS[i])/STDS[i]`	`scaler.transform(X)`
2	Capas ocultas con ReLU	`matVec(COEFS[l], BIASES[l], a).map(relu)`	`mlp.predict_proba(X)` parcial
3	Capa de salida Softmax	`Softmax`	`mlp.predict_proba(X)` salida

C) Flujo de uso del programa

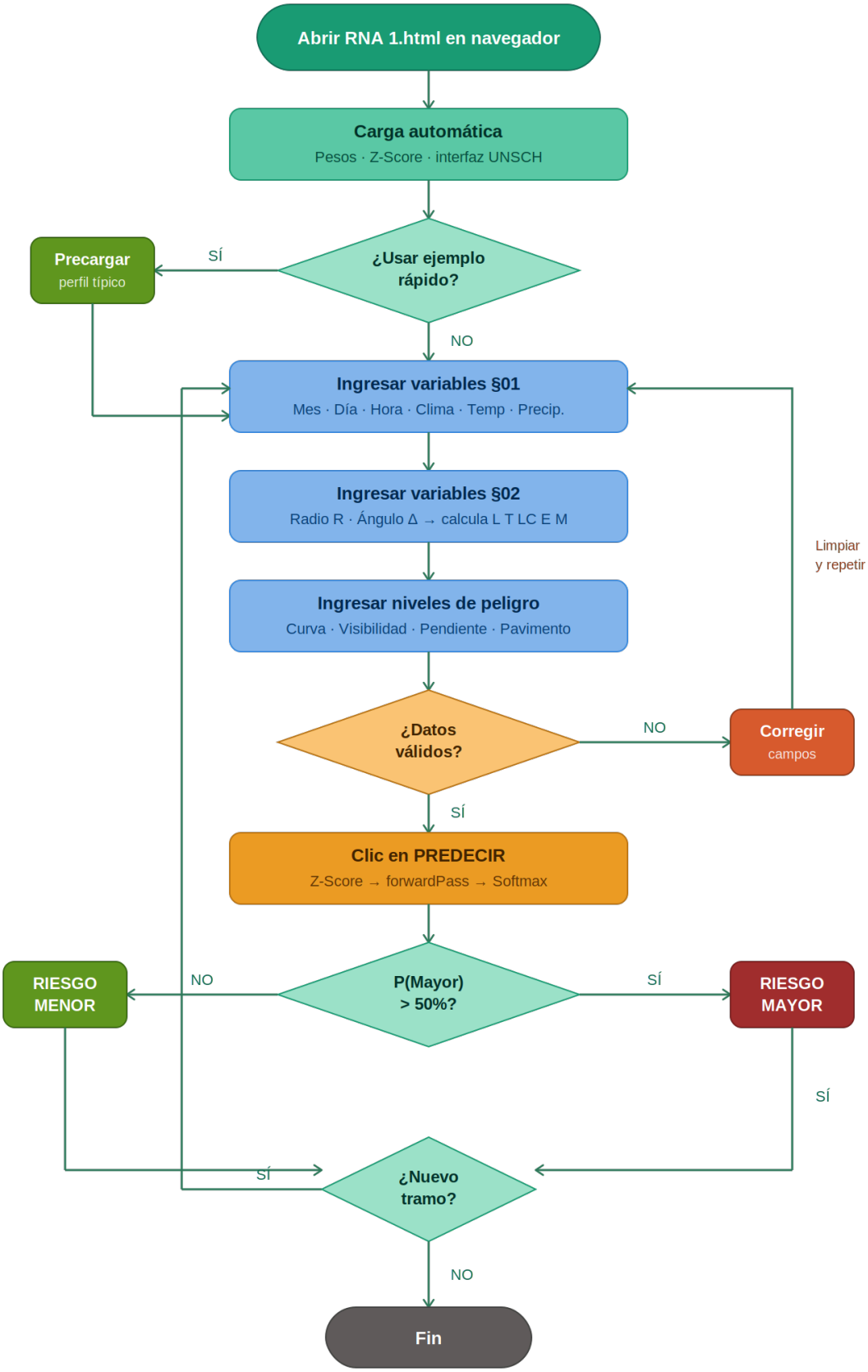


Figura 121 Diagram de flujo de uso del programa RNA N°01
Fuente: Elaboración propia

D) Características técnicas

Tabla 49 *Características técnicas del PROGRAMA RNA 1.html*

Característica	Detalle
Nombre del archivo	`predictor_accidentes_RNA.html`
Tamaño	965 KB
Herramienta de creación	Python · JupyterLab
Parámetros del modelo	6,728 pesos + bias · 46 parámetros Z-Score
Variables de entrada	23 variables (10 campos numéricos + 13 selectores)
Cálculo automático	L, T, LC, E y M calculados desde R y Δ (fórmulas DG-2018 (Ministerio de Transportes y Comunicaciones del Perú, 2018))
Salida del modelo	P(RIESGO MENOR) + P(RIESGO MAYOR) = 100% con barra visual de confianza
Conectividad	Sin internet ni servidor externo
Compatibilidad	Chrome 90+, Firefox 88+, Edge 90+, Safari 14+, Android, iOS
Metrica	Accuracy=90.48% · AUC-ROC=0.9818 · F1-Macro=0.9045

E) Instrucciones de uso

Tabla 50 *Procedimiento de uso del programa RNA 1.html para clasificación de tramos viales de la PE-28A*

Paso	Acción	Descripción
1	Abrir el programa	Doble clic sobre `predictor_accidentes_RNA.html` — se abre automáticamente en el navegador web sin instalación ni conexión a internet
2	Ingresar variables 01	Completar los campos temporales y climáticos: Mes, Día, Hora del siniestro, Clima (despejado/nublado/lluvioso), Temperatura máxima (°C) y Precipitación (mm/día)
3	Ingresar variables 02	Ingresar el Radio R (m) y el Ángulo de deflexión Δ (°) — los parámetros L, T, LC, E y M se calculan automáticamente con las fórmulas DG-2018
4	Ingresar niveles de peligro	Seleccionar los niveles de Peligro en Curva, Peligro por Visibilidad, Peligro en Pendiente y Peligro en Pavimento según evaluación de campo
5	Ejecutar predicción	Hacer clic en **Predecir** y el resultado aparece instantáneamente con la clasificación final y las probabilidades P(RIESGO MENOR) y P(RIESGO MAYOR)
6	Interpretar resultado	Probabilidades > 90% indican alta confianza del modelo. Casos con probabilidades cercanas al 50% requieren validación con inspección técnica de campo
7	Evaluar nuevo tramo	Hacer clic en **Limpiar formulario** y repetir desde el paso 2

- **Paso 1: Abrir el programa**

Figura 122 Paso 1: Abrir el programa
Fuente: Elaboración propia

- **Paso 2, 3, 4: Ingresar variables**

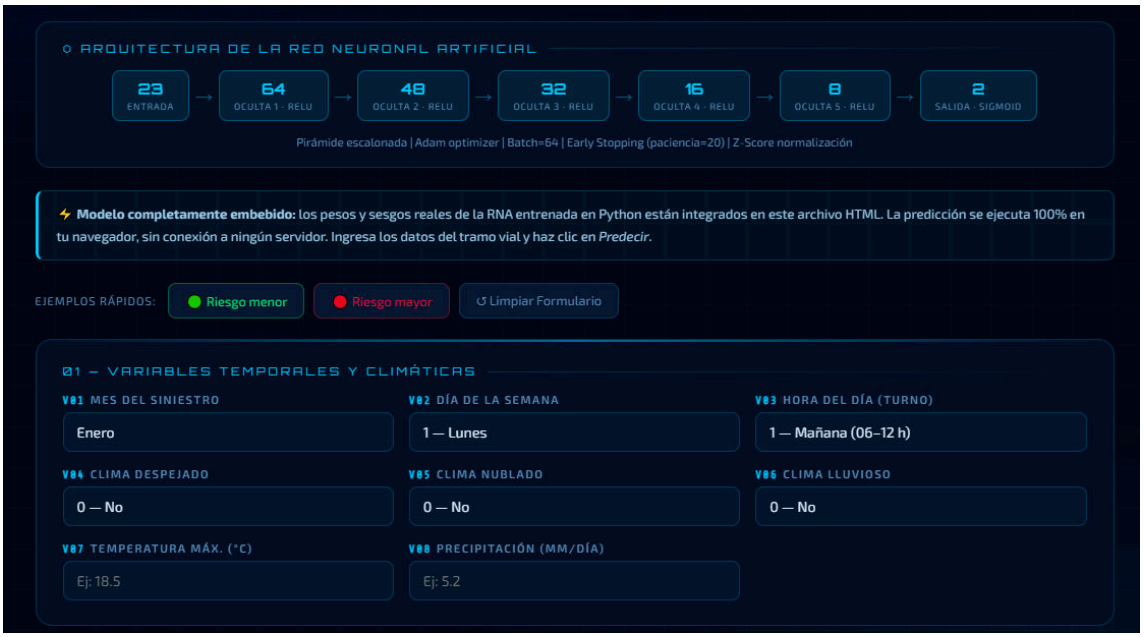


Figura 123 Paso 2: Ingresar variables temporales.
Fuente: Elaboración propia

Figura 124 Paso 3: Ingresar variables geométricas.
Fuente: Elaboración propia

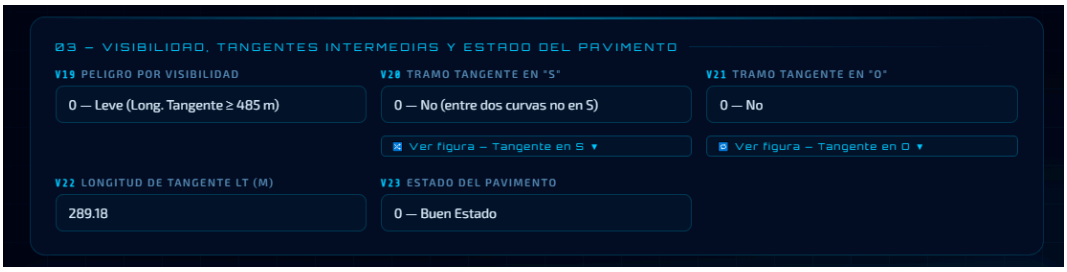


Figura 125 Paso 4: Ingresar niveles de peligro.
Fuente: Elaboración propia

- **Paso 5, 6 y 7 : Ejecutar, interpretar y evaluar nuevo tramo**

Figura 126 Paso 5 y 6: Ejecutar e interpretar en caso Riesgo Menor.
Fuente: Elaboración propia

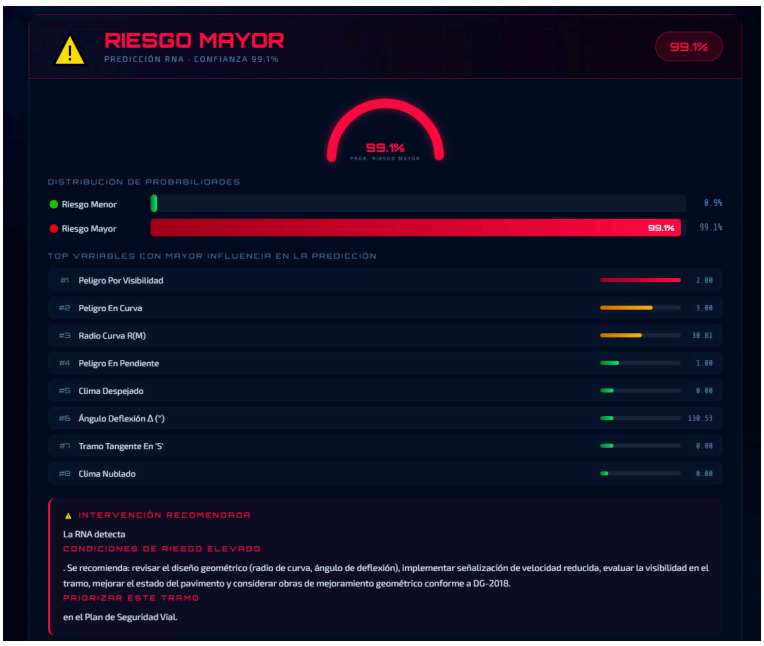


Figura 127 Paso 5 y 6: Ejecutar e interpretar en caso Riesgo Mayor.
Fuente: Elaboración propia

3.11.2 Desarrollo del trabajo de tesis de la Red Neuronal N°02

El presente apartado describe el proceso completo de desarrollo del Modelo de Red Neuronal Artificial N°2 (RNA N°2) para la clasificación multiclase del nivel de riesgo vial geométrico en la Vía Los Libertadores PE-28A. El proceso se organizó en ocho etapas secuenciales: Diagrama de flujo, Análisis Exploratorio de Datos, Pre-procesamiento, Sobremuestreo, Diseño y Entrenamiento de la RNA, Análisis Anti-Overfitting, Evaluación e Interpretación del Modelo, Análisis Avanzado de Predicciones, y Producto Digital.

3.11.2.1 Diagrama de flujo

Diagrama de flujo

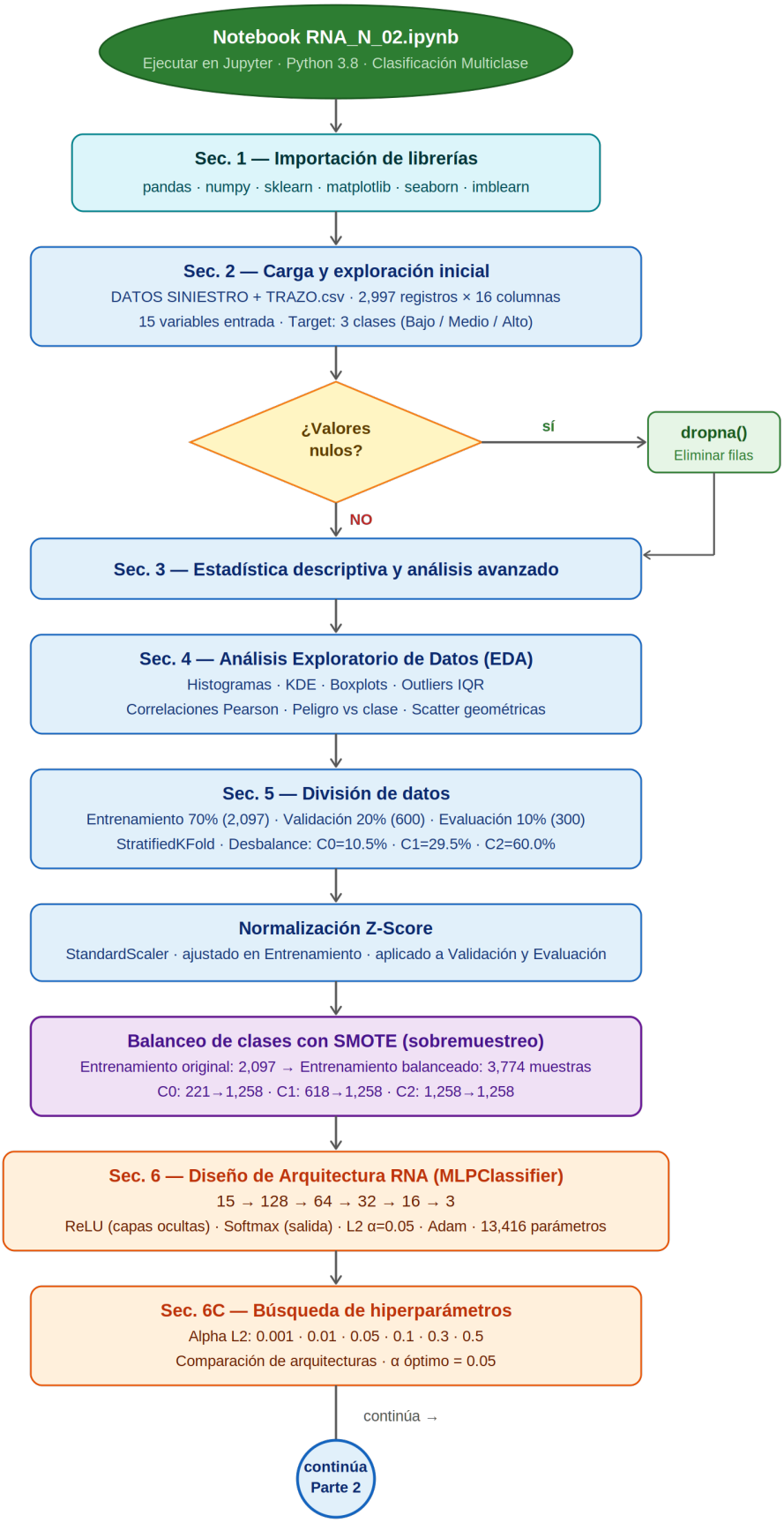


Figura 128 Diagrama de flujo del modelo RNA N°2 - parte 1
Fuente: Elaboración propia

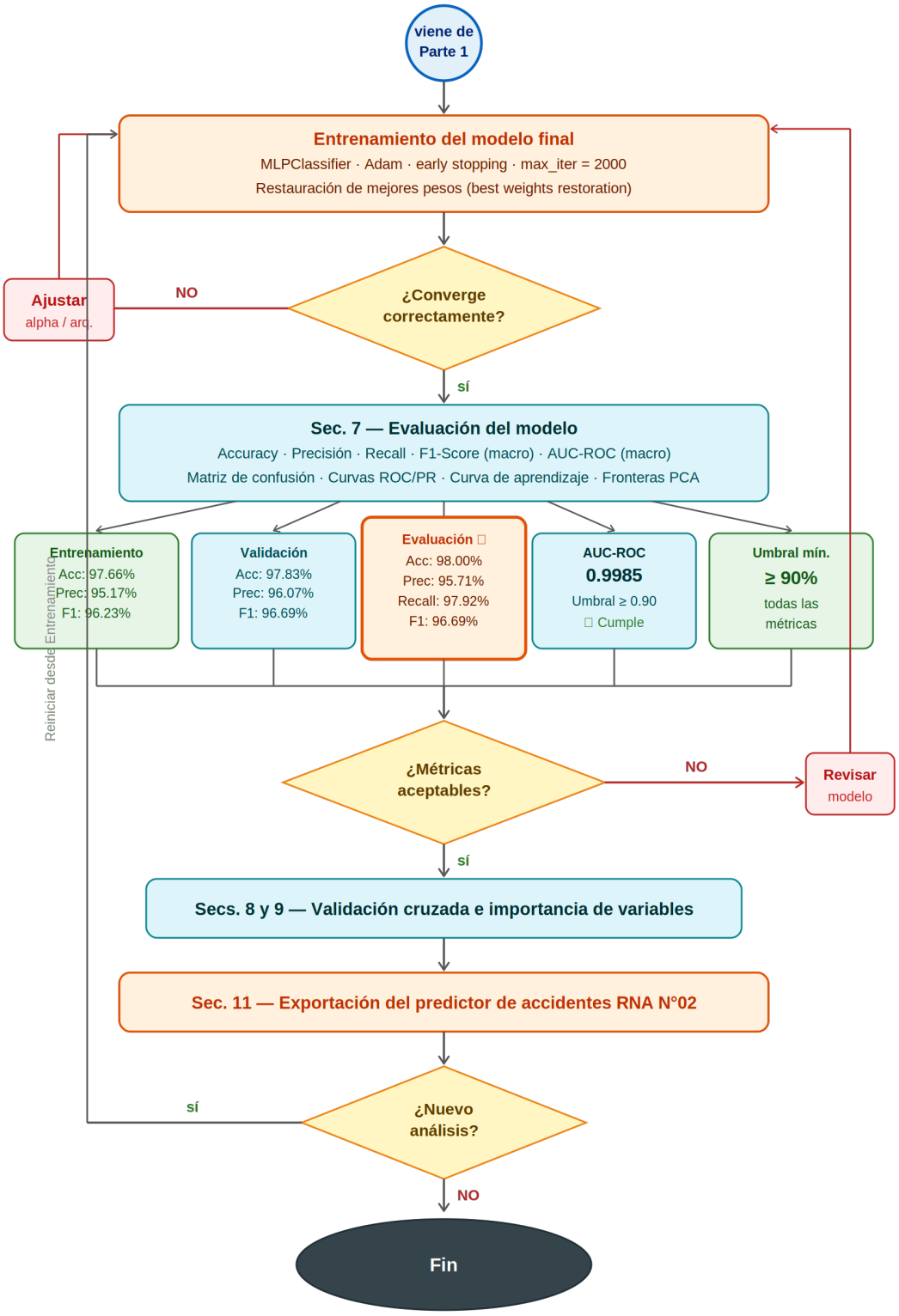


Figura 129 Diagrama de flujo del modelo RNA N°2 - parte 2
Fuente: Elaboración propia

3.11.2.2 Registro de datos de la RNA N°02

La base de datos de la RNA N°02 fue construida a partir de dos fuentes principales, y donde fueron verificadas en campo mendiante rigistros fotograficos que demuestran la exactitud del trabajo, ademas las variables de peligrosidad fueron definidas de acuerdo al DG-2018. La fuentes se detallan a continuación:

- **AutoCAD Civil 3D:** Parámetros geométricos modelados en la Via Los Libertadores (PE-28A), donde se obtuvieron las siguientes variables radio de curva, longitud de curva, subtangente, ángulo de deflexión, cuerda, distancia externa, ordenada media, longitud de tangente, tipo de elemento y configuración de tangentes.
- **Variables de peligrosidad segun DG-2018:** Cuatro variables de peligrosidad peligro en pendiente, en curva, por visibilidad y en pavimento.
- **Variable objetivo:** Riesgo Bajo (0), Riesgo Medio (1) y Riesgo Alto (2) con 2,997 registros.

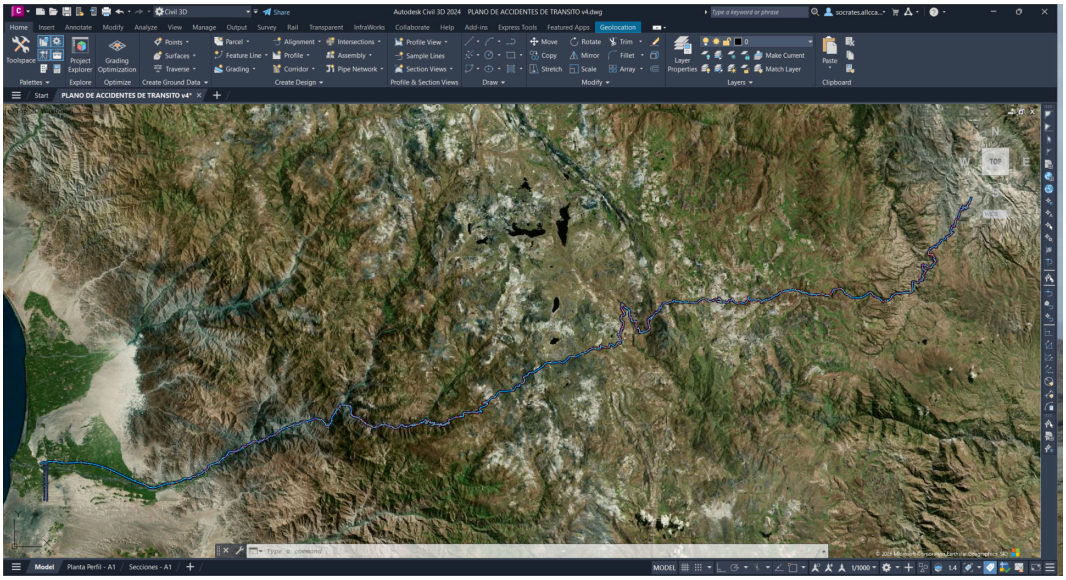


Figura 130 Modelamiento en el Civil 3D de la Via Los Libertadores.
Fuente: Elaboración propia

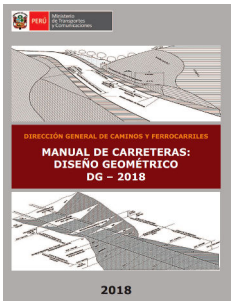


Figura 131 Manual de diseño geométrico del Perú (DG-2018).
Fuente: Ministerio de transportes y comunicaciones (MTC).

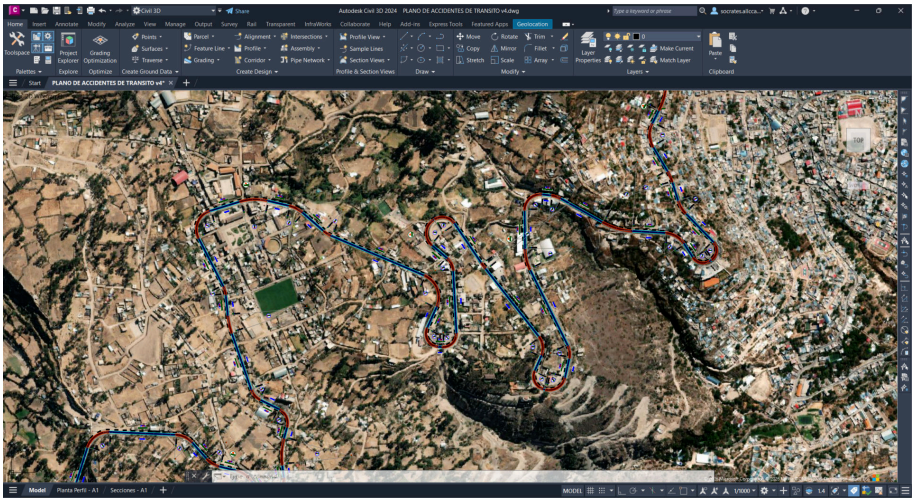


Figura 132 Modelamiento en el Civil 3D: Tramo llegada a Ayacucho.
Fuente: Elaboración propia

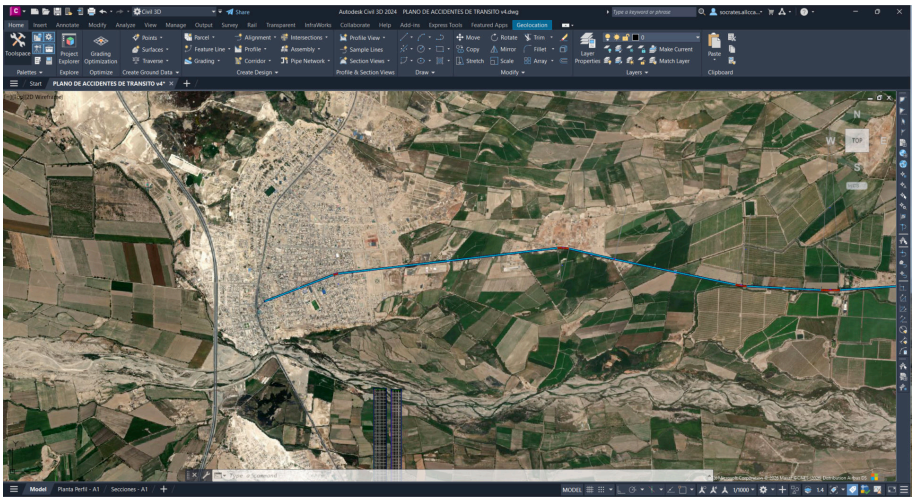


Figura 133 Modelamiento en el Civil 3D: Tramo llegada a San Clemente.
Fuente: Elaboración propia

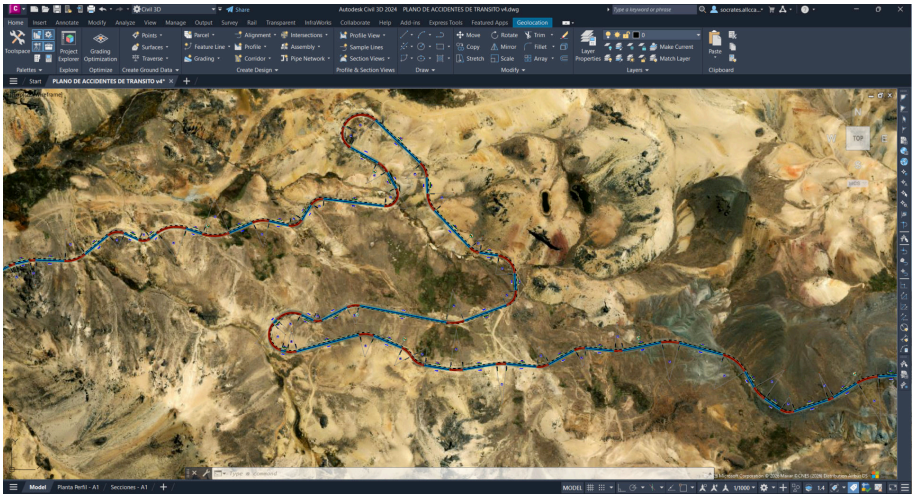


Figura 134 Modelamiento en el Civil 3D: Tramo Pilpichaca.
Fuente: Elaboración propia

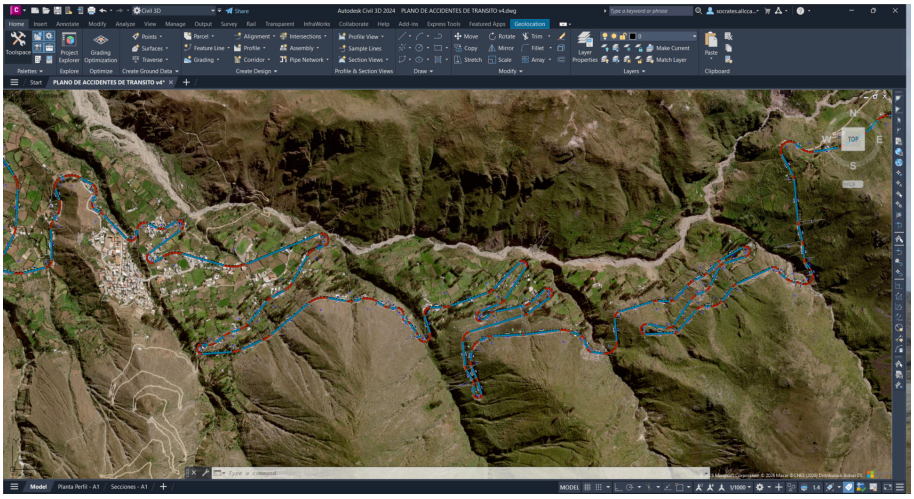


Figura 135 Modelamiento en el Civil 3D: Tramo Huaytara.
Fuente: Elaboración propia

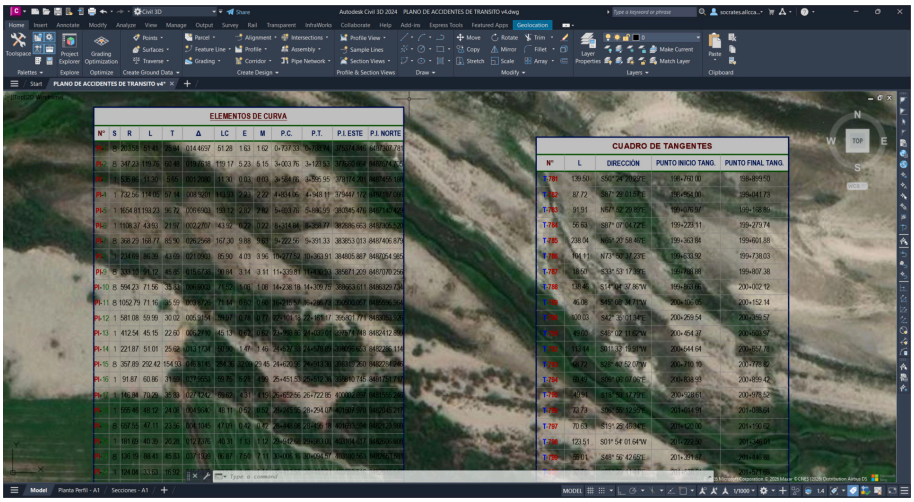
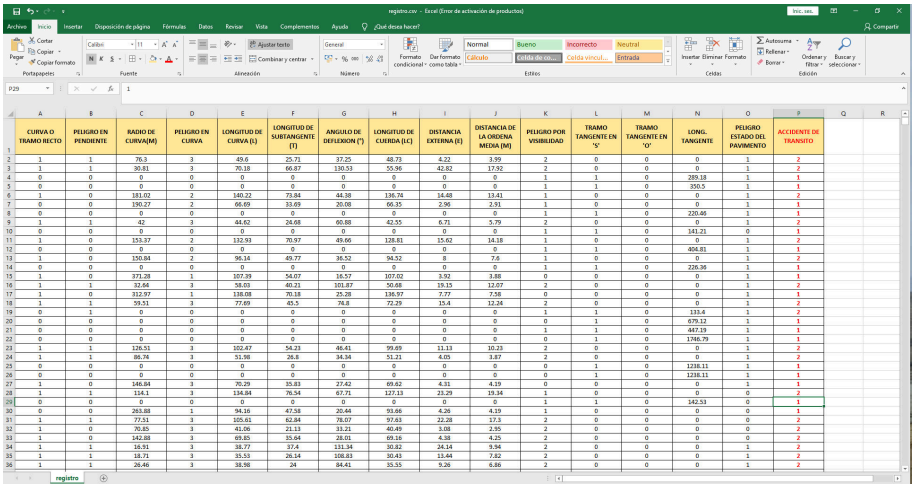


Figura 136 Modelamiento en el Civil 3D: Datos geometricos exportados.
Fuente: Elaboración propia



CURVA O TRAMO RECTO	PELIGRO EN PENDIENTE	RADIO DE CURVA(M)	PELIGRO EN CURVA	LONGITUD DE CURVA (L)	LONGITUD DE SUBTANGENTE (T)	ANGULO DE DEFLEXION (°)	LONGITUD DE CUERDA (LC)	DISTANCIA EXTERNA (E)	DISTANCIA DE LA ORDENA MEDIA (M)	PELIGRO POR VISIBILIDAD	TRAMO TANGENTE EN 'S'	TRAMO TANGENTE EN 'O'	LONG. TANGENTE	PELIGRO ESTADO DEL PAVIMENTO	ACCIDENTE DE TRANSITO
1	1	76.3	3	49.6	25.71	37.25	48.73	4.22	3.99	2	0	0	0	1	2
1	1	30.81	3	70.18	66.87	130.53	55.96	42.82	17.92	2	0	0	0	1	2
0	0	0	0	0	0	0	0	0	0	1	1	0	289.18	1	1
0	0	0	0	0	0	0	0	0	0	1	1	0	350.5	1	1
1	0	181.02	2	140.22	73.84	44.38	136.74	14.48	13.41	1	0	0	0	1	2
0	0	190.27	2	66.69	33.69	20.08	66.35	2.96	2.91	1	0	0	0	1	1
0	0	0	0	0	0	0	0	0	0	1	1	0	220.46	1	1
1	1	42	3	44.62	24.68	60.88	42.55	6.71	5.79	2	0	0	0	1	2
0	0	0	0	0	0	0	0	0	0	1	1	0	141.21	0	1
1	0	153.37	2	132.93	70.97	49.66	128.81	15.62	14.18	1	0	0	0	1	2
0	0	0	0	0	0	0	0	0	0	1	1	0	404.81	1	1
1	0	150.84	2	96.34	49.77	36.52	94.52	8	7.6	1	0	0	0	1	2
0	0	0	0	0	0	0	0	0	0	1	1	0	226.36	1	1
1	0	371.28	1	107.39	54.07	16.57	107.02	3.92	3.88	0	0	0	0	1	1
1	1	32.84	3	59.03	40.21	101.87	50.68	19.15	12.07	2	0	0	0	1	2
1	0	312.97	1	138.08	70.18	25.28	136.97	7.77	7.58	0	0	0	0	1	1
1	1	59.51	3	77.69	45.5	74.8	72.29	15.4	12.24	2	0	0	0	1	2
0	1	0	0	0	0	0	0	0	0	1	1	0	133.4	1	2
0	0	0	0	0	0	0	0	0	0	0	1	0	679.12	1	1
0	0	0	0	0	0	0	0	0	0	1	1	0	447.19	1	1
0	0	0	0	0	0	0	0	0	0	0	1	0	1746.79	1	1
1	1	126.51	3	102.47	54.23	46.41	99.69	11.13	10.23	2	0	0	0	1	2
1	1	86.74	3	51.98	26.8	34.34	51.21	4.05	3.87	2	0	0	0	1	2
0	0	0	0	0	0	0	0	0	0	0	1	0	1238.11	1	1
0	0	0	0	0	0	0	0	0	0	0	1	0	1238.11	1	1
1	0	146.84	3	70.29	35.83	27.42	69.62	4.31	4.19	0	0	0	0	1	1
1	1	134.1	3	134.84	76.54	67.71	127.13	23.29	19.34	1	0	0	0	1	2
0	0	0	0	0	0	0	0	0	0	1	1	0	142.53	0	1
0	0	263.88	1	94.16	47.58	20.44	93.66	4.26	4.19	1	0	0	0	1	1
1	1	77.51	3	105.61	62.84	78.07	97.63	22.28	17.3	2	0	0	0	1	2
1	0	70.85	3	41.66	21.13	33.21	40.49	3.08	2.95	2	0	0	0	0	2
1	0	142.88	3	69.85	35.64	28.01	69.16	4.38	4.25	2	0	0	0	0	2
1	1	16.91	3	38.77	37.4	131.34	30.82	24.14	9.94	2	0	0	0	1	2
1	1	18.71	3	35.53	26.14	108.83	30.43	13.44	7.82	2	0	0	0	1	2
1	1	26.46	3	38.96	24	84.41	35.55	9.26	6.86	2	0	0	0	1	2

Figura 137 Base de datos RNA N° 2: 2997 registros.
Fuente: Elaboración propia

3.11.2.3 Ficha técnica del Modelo RNA N°2

Tabla 51 *Ficha técnica del modelo RNA N°2*

Parámetro	Valor
Base de datos	2,997 registros geométricos · Vía Los Libertadores PE-28A
Variables de entrada	15 variables geométricas y de peligrosidad vial
Variable objetivo	Nivel de riesgo: Riesgo Bajo=0 · Riesgo Medio=1 · Riesgo Alto=2
División del dataset	Entrenamiento 70% (n=2,097) · Validación 20% (n=600) · Evaluación 10% (n=300)
Sobremuestreo	Random oversampling sobre entrenamiento → 3,774 registros balanceados (1,258 por clase) · ratio original 5.7x
Arquitectura RNA	MLP: 15→128→64→32→16→3 (4 capas ocultas)
Total parámetros	13,416 (pesos + bias)
Activación / Salida	ReLU (capas ocultas) / Softmax 3 clases (capa de salida)
Optimizador	Adam · Regularización L2 α=0.05 · Batch=32 · Early Stopping paciencia=30
Épocas ejecutadas / óptima	66 / 500 máximo · mejores pesos en época 36
Mejor val_loss	0.086830 (época 36)
Accuracy — Entrenamiento	97.66%
Accuracy — Validación	97.83%
Accuracy — Evaluación	98.00%
AUC-ROC Macro	0.9985
F1-Score Macro	0.9669
F1-Score Weighted	0.9802
GAP Train-Test	−0.0034 · sin sobreajuste significativo
Longitud vial	333.903 km · 3 departamentos · 4 provincias · 11 distritos
Software	Python · JupyterLab · Scikit-learn · Pandas · Matplotlib
Año	2026

Fuente: Elaboración propia.

3.11.2.4 Etapa I: Análisis exploratorio de datos (EDA)

El análisis exploratorio de datos tuvo como propósito caracterizar el dataset de 2,997 registros de tramos geométricos de la PE-28A, evaluar la calidad de los datos y determinar el poder discriminativo de cada una de las 15 variables de entrada respecto a la variable objetivo multiclase (Nivel de Accidente: Riesgo Bajo / Riesgo Medio / Riesgo Alto). Este análisis guió todas las decisiones de sobremuestreo, preprocesamiento y arquitectura del modelo.

Tabla 52 *Resumen del análisis exploratorio de datos - Etapa I*

Variable(s)	Hallazgo Principal	Relevancia para la RNA
Variable objetivo	Desbalance 5.7×: Bajo 10.5%, Medio 29.5%, Alto 60.0%	Requiere sobremuestreo antes del entrenamiento
15 variables (histograma + KDE)	Peligro Visibilidad: KDE claramente separadas. Radio Curva: asimetría extrema	Identifica predictor dominante — visibilidad
Variables continuas (boxplots)	Radio: Bajo≈280 m · Medio≈100 m · Alto≈30 m	Confirma relación inversa radio–riesgo
15 variables (outliers IQR)	Pavimento 23.7% · Tramo-O 15.5% · Dist. Externa 14.0%	Outliers preservados como condiciones extremas reales
Correlaciones Pearson	Visibilidad r=0.862 · Radio r=−0.388	Define variables clave; multicolinealidad manejable
Variables peligro vs clase	Visibilidad Niv.2 → 98% Alto · Pavimento ≈ sin diferenciación	Valida coherencia física con DG-2018
Scatter plots geométricos	Radio pequeño + Ángulo grande → Riesgo Alto concentrado	Confirma separabilidad no lineal entre 3 clases
MW + Pearson + Cohen d	14/15 significativas · Visibilidad y Radio: |d|>0.8	Ranking estadístico multidimensional pre-entrenamiento

Fuente: Elaboración propia.

A) Distribución de la variable objetivo

La Figura muestra la distribución de la variable dependiente Nivel de Accidente de Tránsito. Los resultados revelan un desbalance significativo: Riesgo Bajo con 316 registros (10.5%), Riesgo Medio con 883 (29.5%) y Riesgo Alto con 1,798 (60.0%). La distribución es coherente con la carretera PE-28A de topografia plana, ondulada y accidentada, y con baja visibilidad en gran parte del tramo.

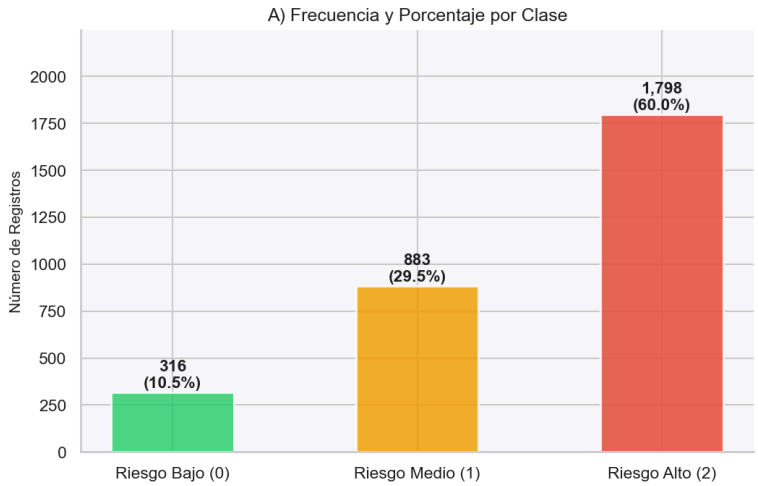


Figura 138 Frecuencia y porcentaje por clase vs número de registros.
Fuente: Elaboración propia

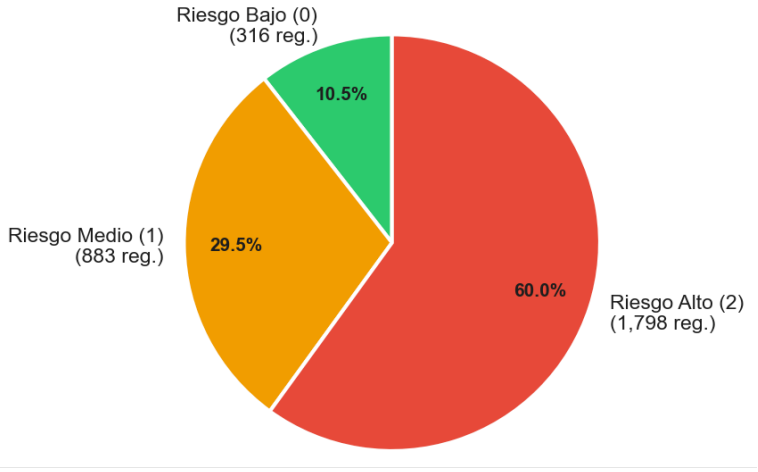


Figura 139 Proporción por clase de riesgo
Fuente: Elaboración propia

B) Histogramas y KDE por clase (15 variables)

Las Figuras siguientes presentan los histogramas de distribución global y las curvas KDE separadas por clase para las 15 variables de entrada. Peligro por Visibilidad presenta la separación más evidente entre las tres clases, el cual es indicador dominante del modelo. Radio de Curva muestra distribución altamente sesgada con concentración en valores bajos: radios menores a 250 m incrementan exponencialmente la probabilidad de accidentes en carreteras de montaña. Peligro Estado del Pavimento muestra KDE prácticamente superpuestas entre clases, anticipando su baja importancia predictiva.

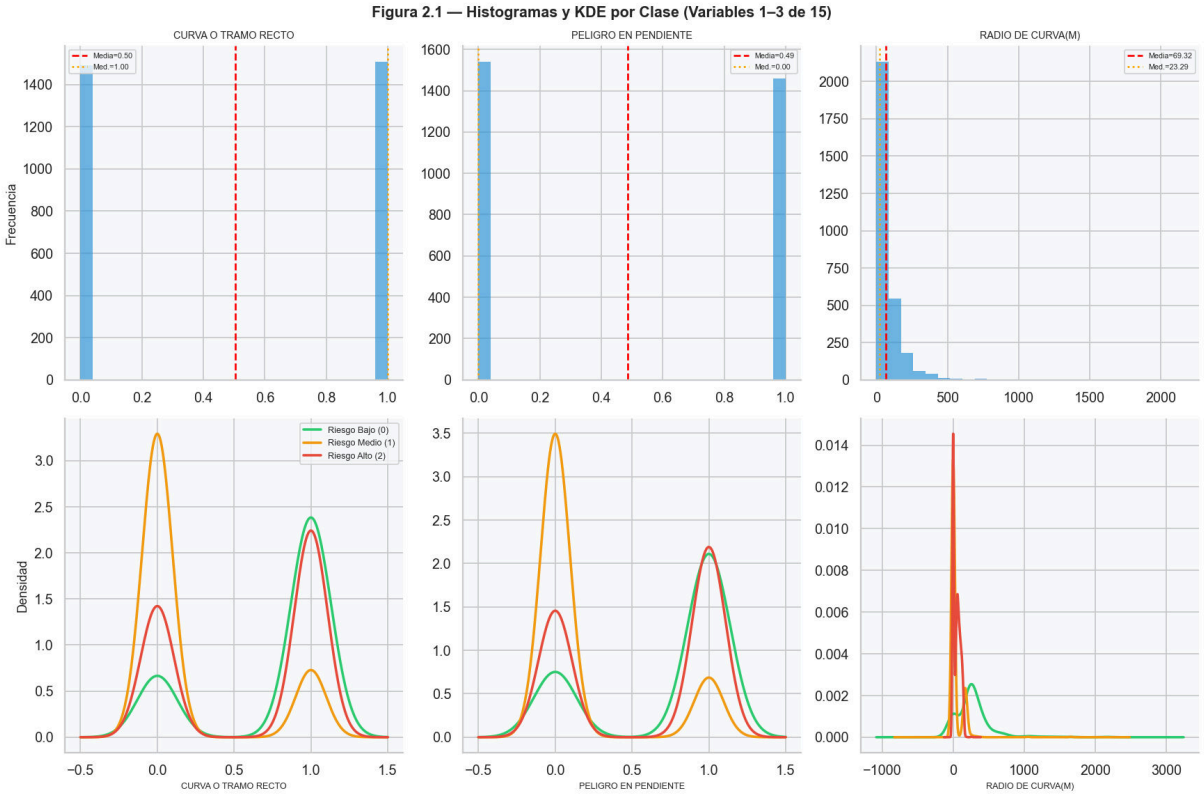


Figura 140 Hitograma y KDE: Curva/Tramo Recto, Peligro en Pendiente, Radio de Curva.
Fuente: Elaboración propia

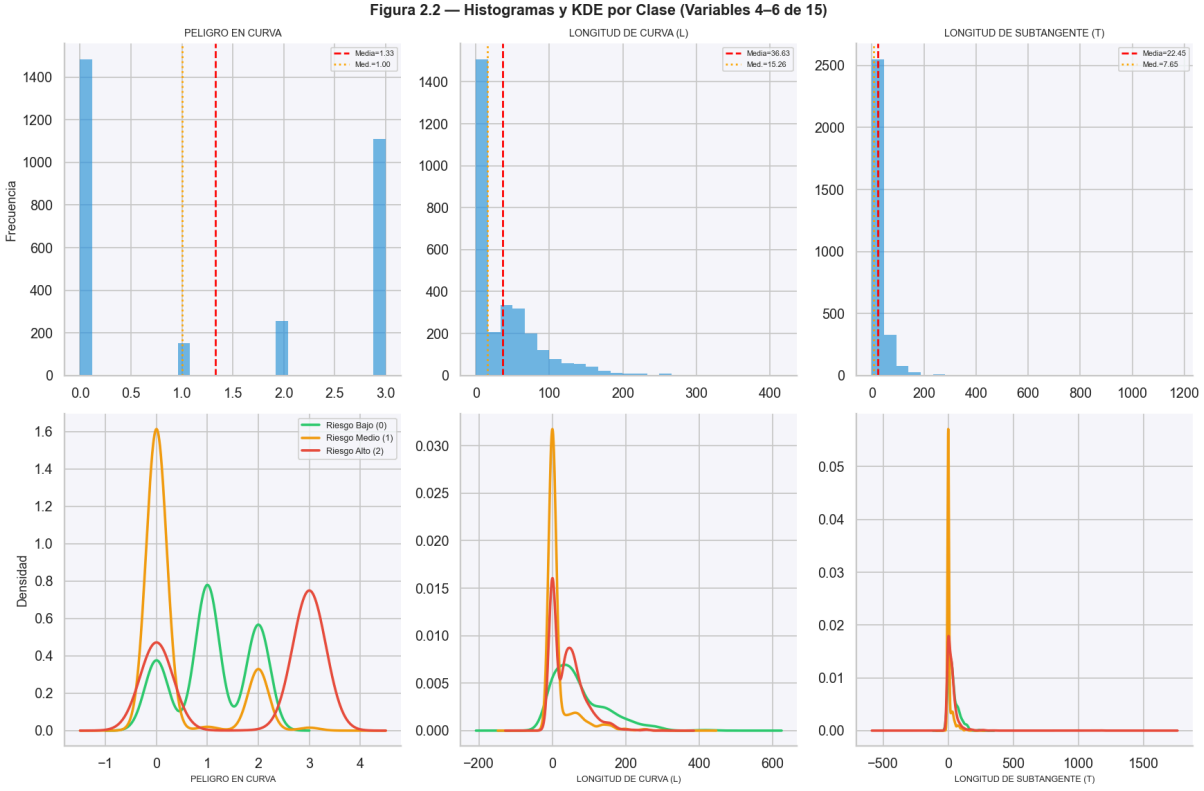


Figura 141 Hitograma y KDE: Peligro en Curva, Long. de Curva y Long. Subtangente.
Fuente: Elaboración propia

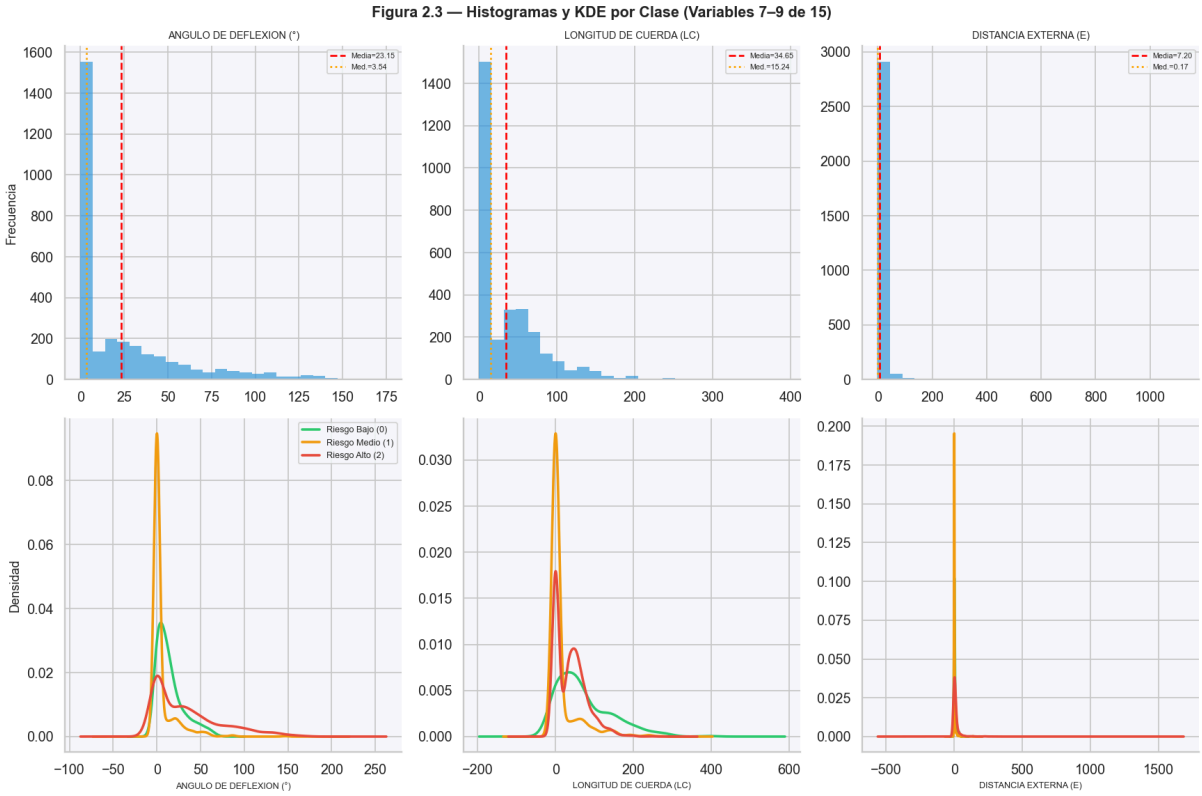


Figura 142 Hitograma y KDE: Ángulo Deflexión, Long. Cuerda y Dist. Externa.
Fuente: Elaboración propia

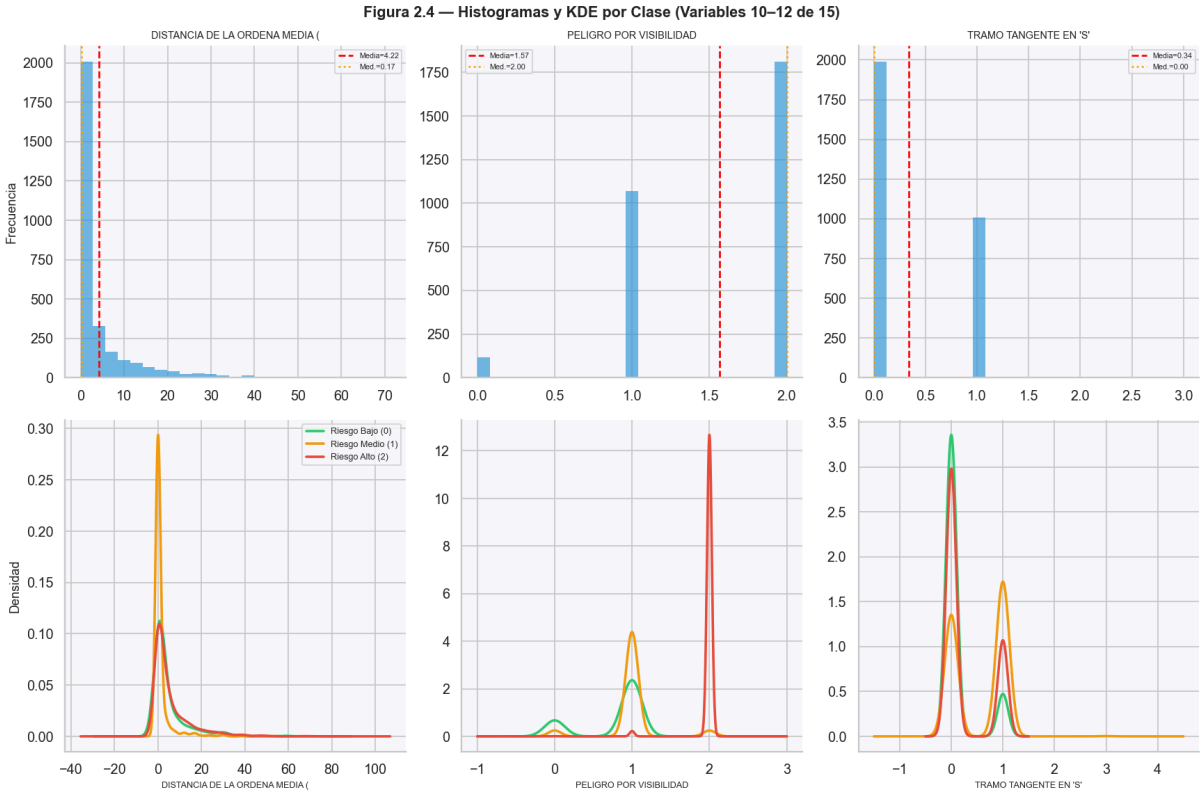


Figura 143 Hitograma y KDE: Ordenada Media, Peligro por Visib. y Tramo en S.
Fuente: Elaboración propia

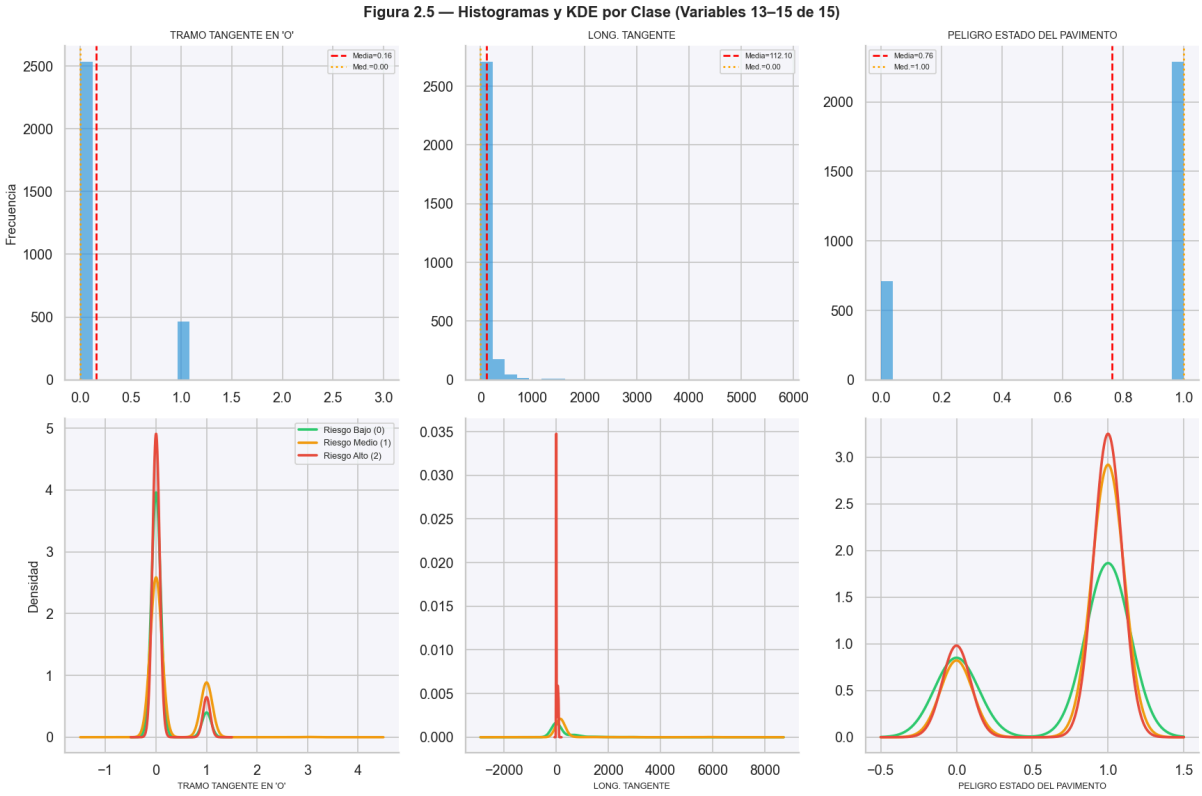


Figura 144 Hitograma y KDE: Tramo Tangente O, Long. Tangente, Peligro Pavimento.
Fuente: Elaboración propia

C) Boxplots por clase de riesgo

La Figuras siguientes presenta los boxplots de las ocho variables geométricas continuas segmentadas por clase de riesgo: Bajo (0), Medio (1) y Alto (2). La línea negra indica la mediana, el diamante azul la media, la caja el rango intercuartílico (IQR), y los círculos sueltos los valores atípicos o outlier.

Del análisis se distinguen dos grupos de variables:

- **Variables con alta separación entre clases:** El Radio de Curva registró las diferencias más notorias: Riesgo Bajo presenta medianas de 300-400m, mientras que Riesgo Alto concentra sus valores por debajo de 50m, confirmando que radios pequeños se asocian a mayor peligrosidad. El Ángulo de Deflexión fue mayor en Riesgo Alto, indicando curvas más cerradas. La Longitud de Curva, Longitud de Cuerda y Longitud de Tangente mostraron el patrón inverso: valores más altos corresponden a Riesgo Bajo, asociados a tramos más rectos y extendidos.
- **Variables con baja separación entre clases:** La Distancia Externa (E), la Ordenada Media (M) y la Longitud de Subtangente (T) presentaron distribuciones

comprimidas cerca de cero con escasa diferencia entre grupos. Su aporte individual es limitado.

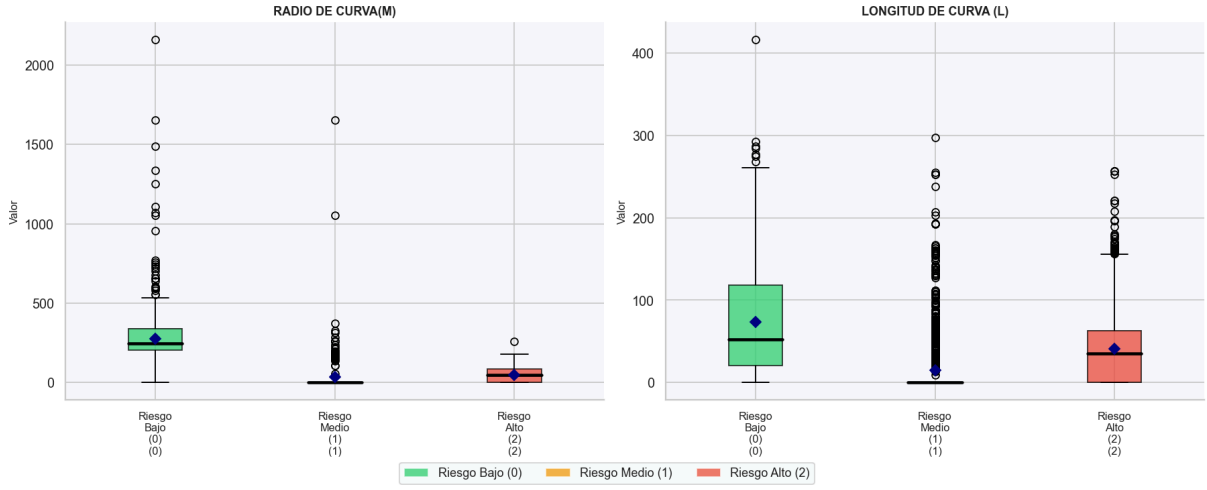


Figura 145 Boxplots de variables de Radio de curva y longitud de curva.
Fuente: Elaboración propia

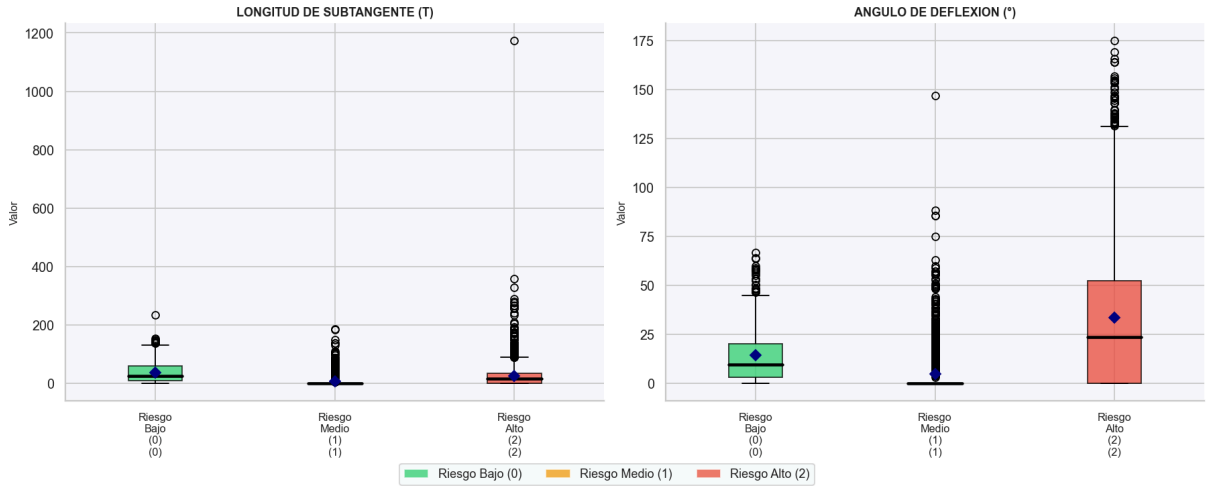


Figura 146 Boxplots de variables de Longitud de subtangente y angulo de deflexión.
Fuente: Elaboración propia

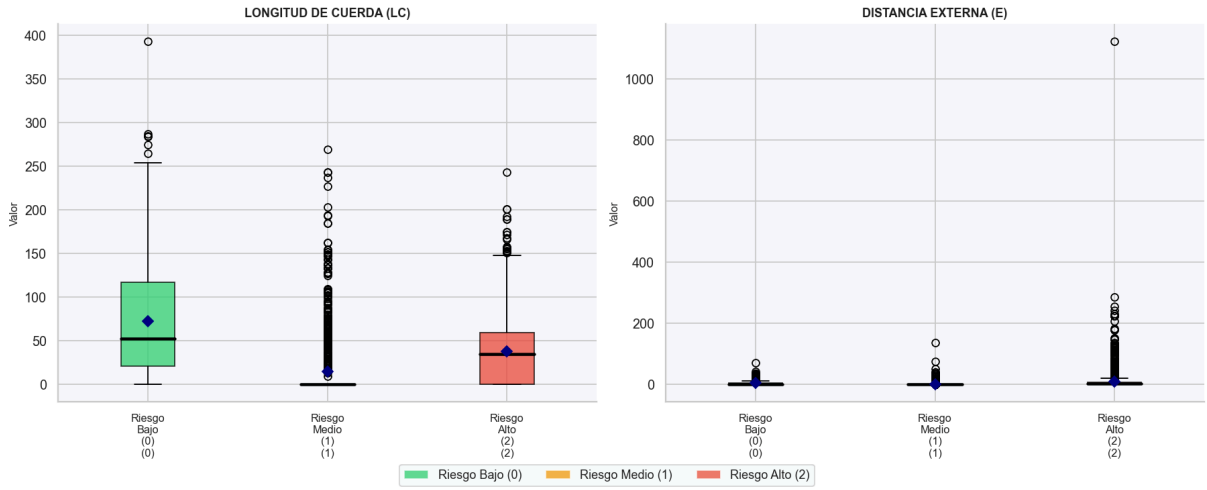


Figura 147 Boxplots de variables de Longitud de cuerda y distancia externa.
Fuente: Elaboración propia

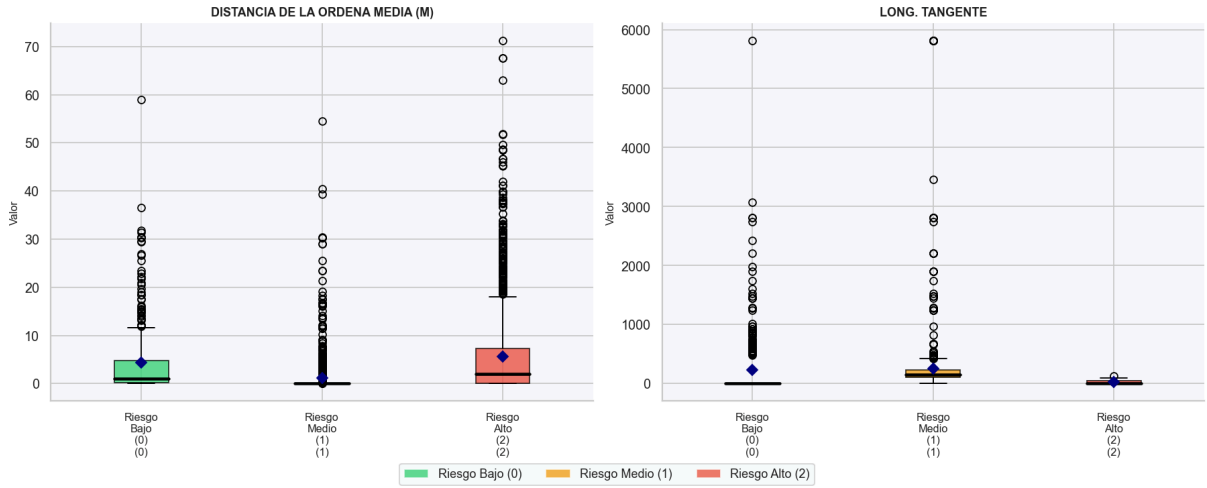


Figura 148 Boxplots de variables de Distancia de ordenada media y longitud tangente.
Fuente: Elaboración propia

La presencia de outliers en variables como Radio de Curva y Longitud de Tangente refleja la heterogeneidad geométrica de la vía PE-28A y fueron conservados por representar condiciones con relevancia para la predicción del riesgo vial.

D) Análisis de outliers (Método IQR)

La Figura siguiente presenta el porcentaje de valores atípicos por variable, detectados mediante el método IQR: todo valor fuera del intervalo $[Q1 - 1.5 \cdot IQR,\ Q3 + 1.5 \cdot IQR]$ es clasificado como outlier. Los colores indican el nivel de severidad: verde para porcentajes inferiores al 5%, naranja entre 5% y 10%, y rojo por encima del 10%.

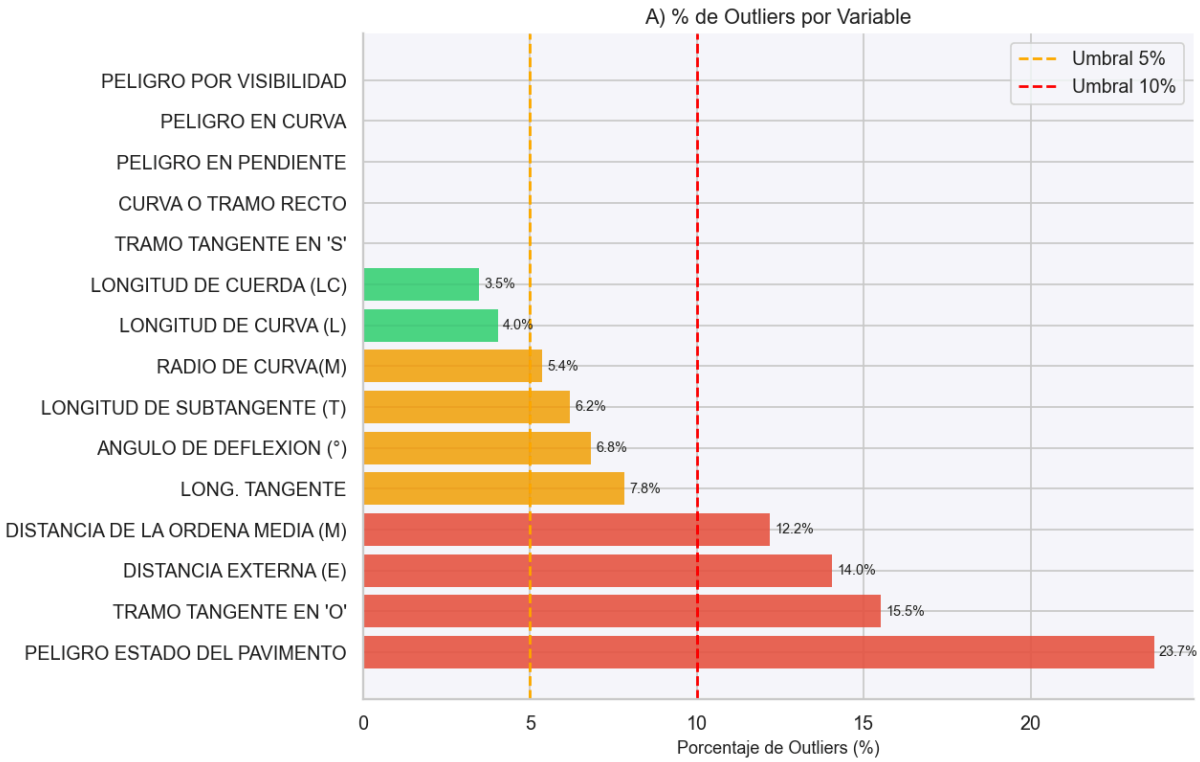


Figura 149 Análisis de outliers (Método IQR).
Fuente: Elaboración propia

Del análisis se distinguen tres grupos:

- **Sin outliers (0%):** las variables categóricas Peligro por Visibilidad, Peligro en Curva, Peligro en Pendiente y Tramo Tangente en S presentaron IQR = 0, lo que indica que sus valores están completamente concentrados en categorías discretas fijas.
- **Outliers moderados (3.5% - 7.8%):** las variables geométricas continuas principales como Longitud de Curva (4.0%), Radio de Curva (5.4%), Ángulo de Deflexión (6.8%) y Longitud de Tangente (7.8%) presentaron porcentajes aceptables, reflejando la variabilidad geométrica real de la vía PE-28A.
- **Outliers elevados (>10%)**: la Distancia Externa (14.0%), el Tramo Tangente en O (15.5%) y el Peligro Estado del Pavimento (23.7%) superaron el umbral crítico. En este último caso, el IQR = 0 indica que la mayoría de registros comparten el mismo valor, por lo que cualquier variación es detectada como atípica.

Todos los outliers fueron conservados en el conjunto de datos por representar condiciones geométricas y de pavimento extremas pero reales de la vía. Su impacto en el entrenamiento de la RNA fue mitigado mediante la normalización Z-Score aplicada en

la etapa de preparación de datos.

E) Correlaciones de Pearson

La primera figura presenta la matriz de correlación de Pearson entre todas las variables de entrada, y la segunda figura muestra la correlación individual de cada variable con la variable objetivo (accidente de tránsito).

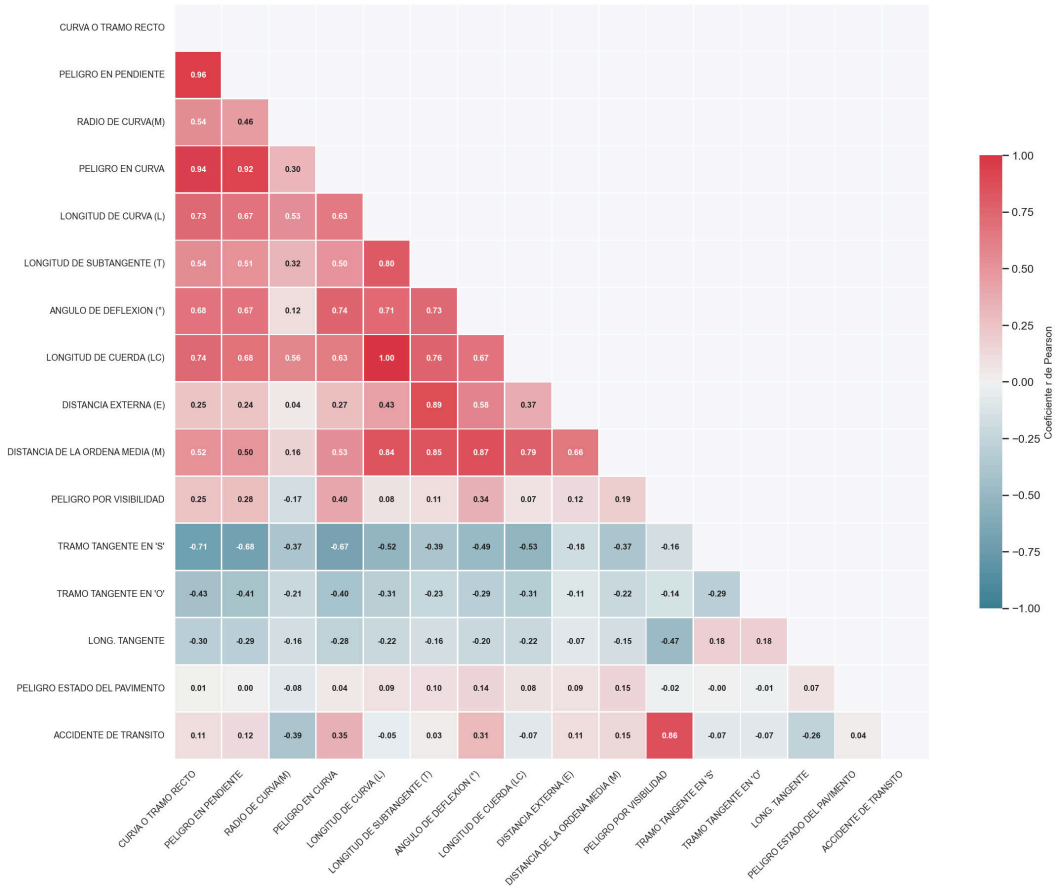


Figura 150 Matriz de correlación de Pearson.
Fuente: Elaboración propia

Del análisis se destacan los siguientes hallazgos:

- **Variable con correlación alta:** Peligro por Visibilidad registró el coeficiente más elevado ($r = +0.862$), siendo la única variable que supera el umbral $|r| \geq 0.7$, lo que indica que a mayor nivel de peligro por visibilidad, mayor es el riesgo de accidente vial.
- **Variables con correlación baja:** Peligro en Curva ($r = +0.348$), Ángulo de Deflexión ($r = +0.315$) y Radio de Curva ($r = -0.388$) presentaron correlaciones débiles con la variable objetivo, evidenciando que su aporte individual es limitado en términos lineales.

- **Multicolinealidad entre variables geométricas:** el heatmap reveló correlaciones muy altas entre Longitud de Curva y Longitud de Cuerda ($r = 1.00$), así como entre Distancia Externa y Distancia de la Ordenada Media ($r = 0.87$), indicando redundancia entre dichos pares de variables.

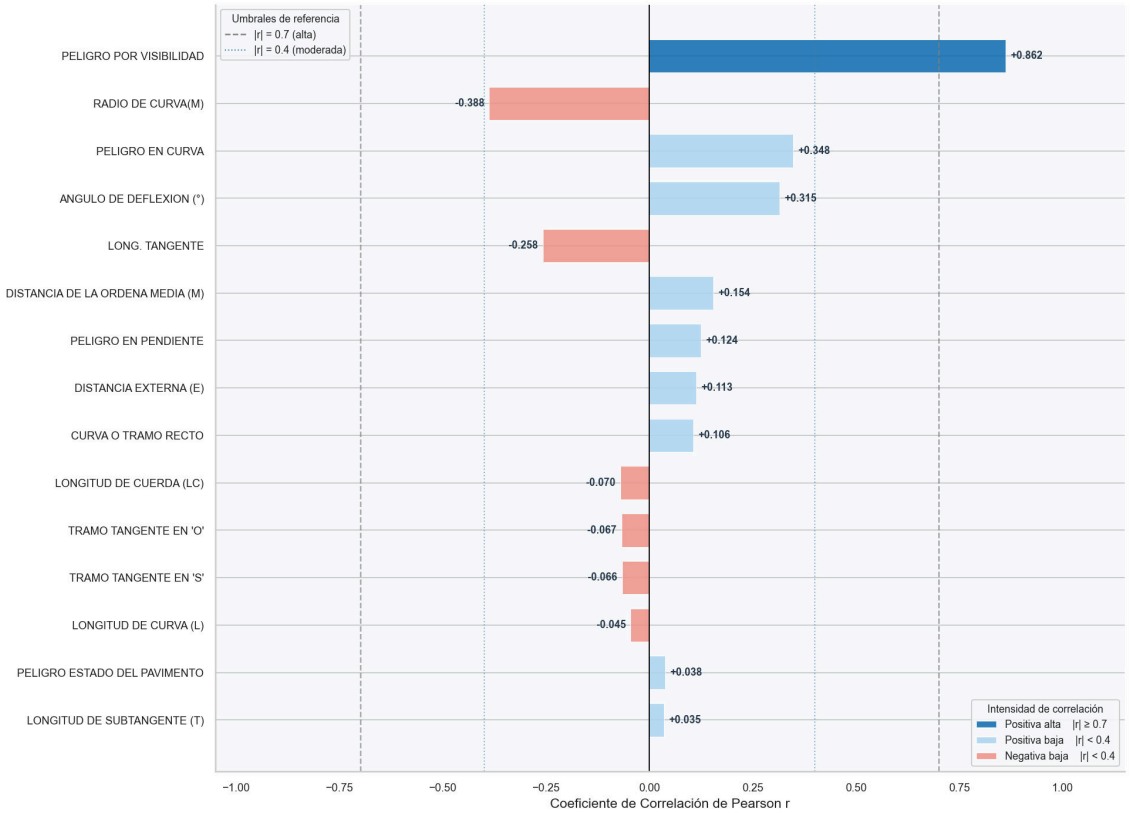


Figura 151 Correlación individual de cada variable.
Fuente: Elaboración propia

La débil correlación lineal individual de la mayoría de variables con la variable objetivo justifica el uso de una RNA, dado que este modelo es capaz de capturar relaciones no lineales e interacciones entre variables que el coeficiente de Pearson no detecta.

F) Variables de peligro vs clase

La Figuras siguientes presenta las barras apiladas al 100% de las cuatro variables categóricas de peligro, mostrando la proporción de cada clase de riesgo (bajo, medio y alto) dentro de cada nivel de peligro. A continuación se describen los hallazgos por variable:

- **Peligro en Pendiente:** En el nivel 0 (sin peligro) la distribución es mixta (alto: 47%, medio: 48%, bajo: 5%), mientras que en el nivel 1 (con peligro) el riesgo Alto asciende al 74%, confirmando que la presencia de pendiente peligrosa incre-

menta significativamente el nivel de riesgo vial.

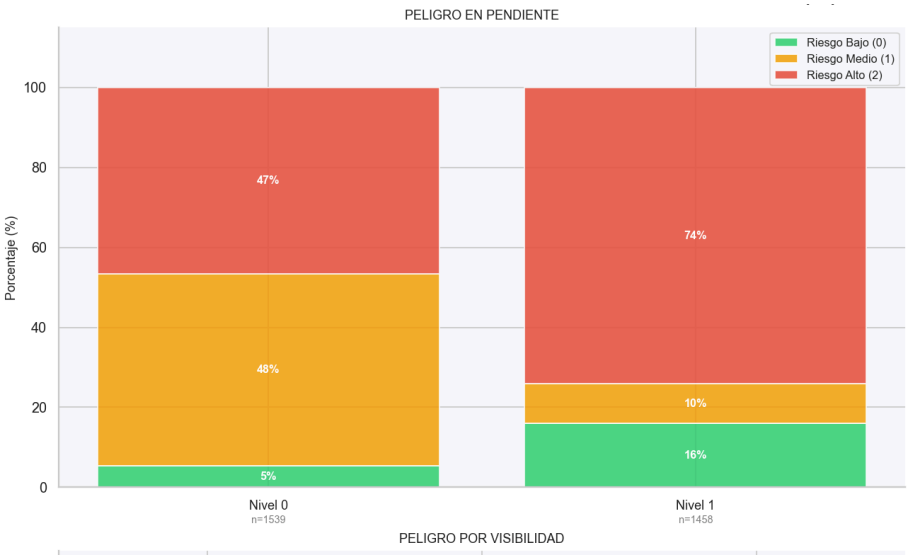


Figura 152 Variable de peligro en pendiente vs clase de accidente.
Fuente: Elaboración propia

- **Peligro en Curva:** Presenta el comportamiento más variable entre niveles. El nivel 1 concentra el 93% en riesgo bajo, el nivel 2 muestra una distribución equilibrada (41% bajo, 58% medio), y el nivel 3 alcanza el 99% de riesgo alto, evidenciando que las curvas de máximo peligro son condiciones geométricas críticas para la accidentalidad.

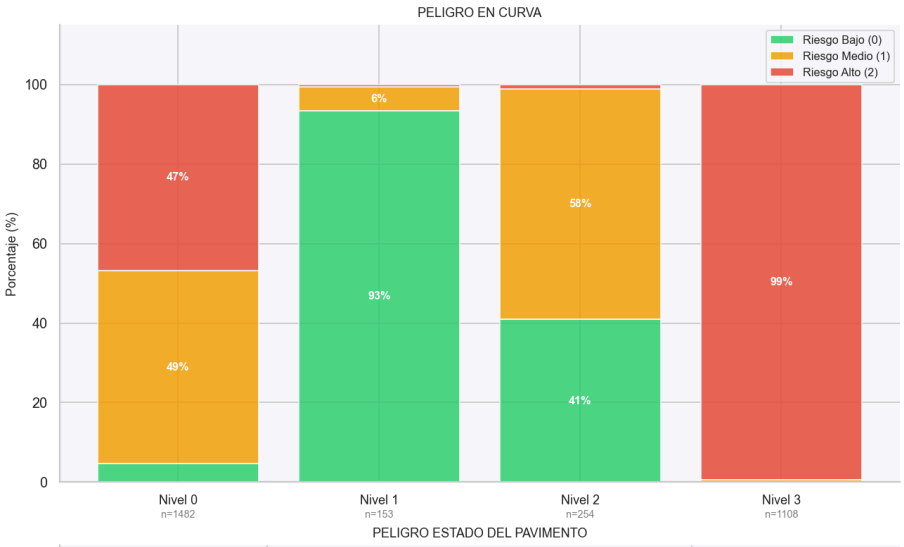


Figura 153 Variable de peligro en curva vs clase de accidente.
Fuente: Elaboración propia

- **Peligro por Visibilidad:** Es la variable con mayor poder discriminatorio. El nivel 0 (buena visibilidad) registra 60% de riesgo bajo, el nivel 1 concentra 74% de riesgo medio, y el nivel 2 (visibilidad muy reducida) alcanza el 98% de riesgo alto, resultado consistente con la correlación más alta observada en la figura ($r =$

+0.862).

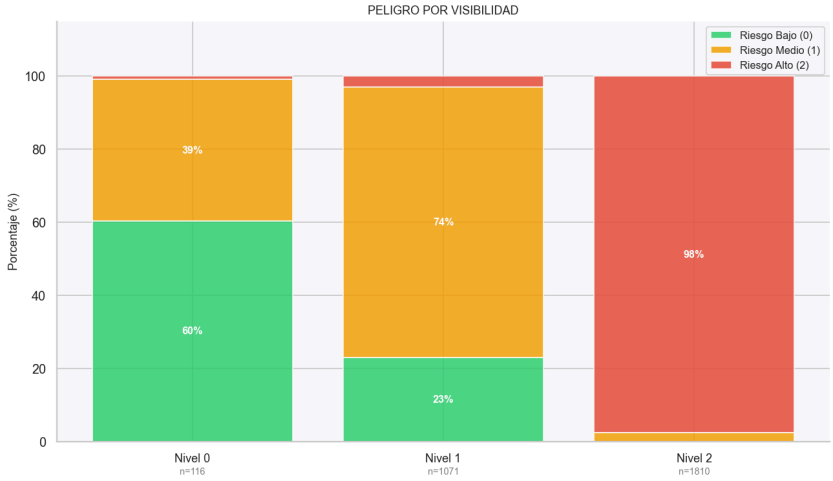


Figura 154 Variable de peligro por visibilidad vs clase de accidente.
Fuente: Elaboración propia

- **Peligro Estado del Pavimento:** Muestra la menor capacidad discriminatoria entre variables de peligro. Los niveles 0 y 1 presentan distribuciones prácticamente idénticas (alto: 59-60%, medio: 27-30%, bajo: 9-14%), lo que indica que el estado del pavimento, de manera aislada, no diferencia significativamente las clases de riesgo.

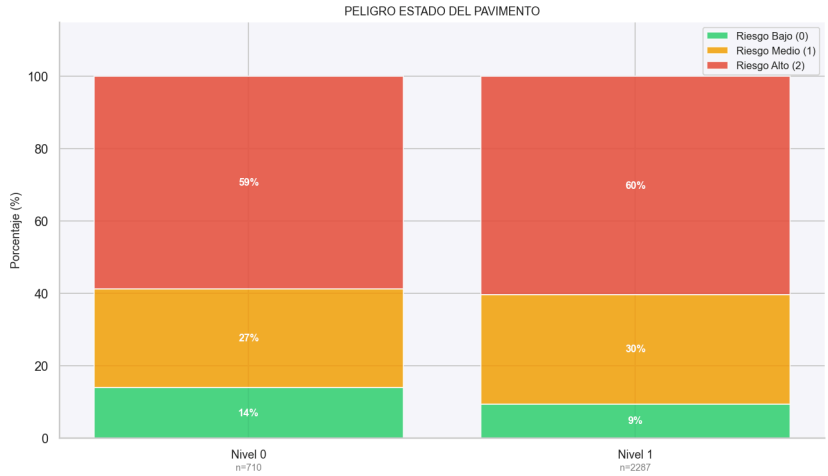


Figura 155 Variable de peligro por pavimento vs clase de accidente.
Fuente: Elaboración propia

En síntesis, Peligro por Visibilidad y Peligro en Curva (Nivel 3) constituyen las variables categóricas con mayor capacidad de separación entre clases, mientras que Peligro Estado del Pavimento aporta escasa discriminación individual, aunque su interacción con otras variables puede ser determinante por la RNA.

G) Diagrama de dispersión de variables geométricas

Las figuras siguientes presentan seis combinaciones de variables geométricas usando el Radio de Curva como eje común. En todas se observa el mismo patrón: Riesgo Alto (rojo) concentrado en radios pequeños (0–200 m), Riesgo Bajo (verde) disperso en radios mayores (>500 m), y Riesgo Medio (naranja) en zona intermedia.

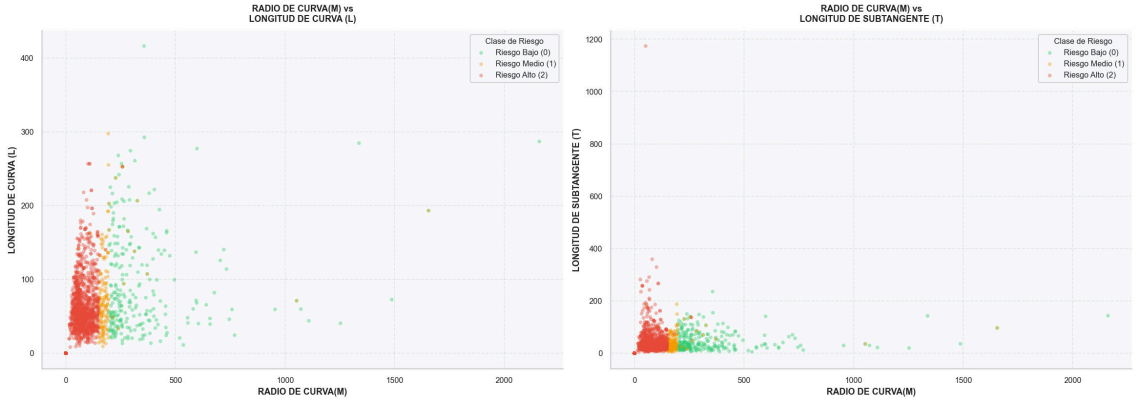


Figura 156 Diagr. de dispersión de Radio de curva vs long. de curva y long. de subtang.
Fuente: Elaboración propia

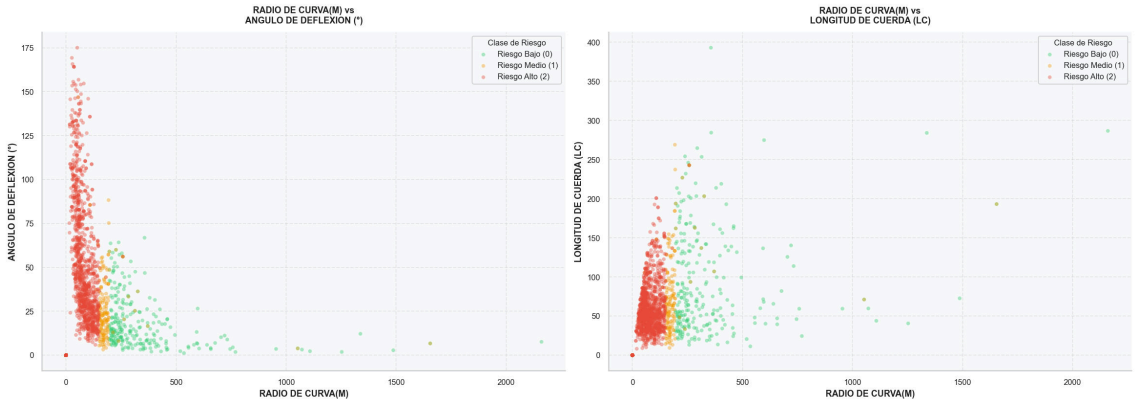


Figura 157 Diagr. de dispersión de Radio de curva vs ángulo de deflexión y long. de cuerda.
Fuente: Elaboración propia

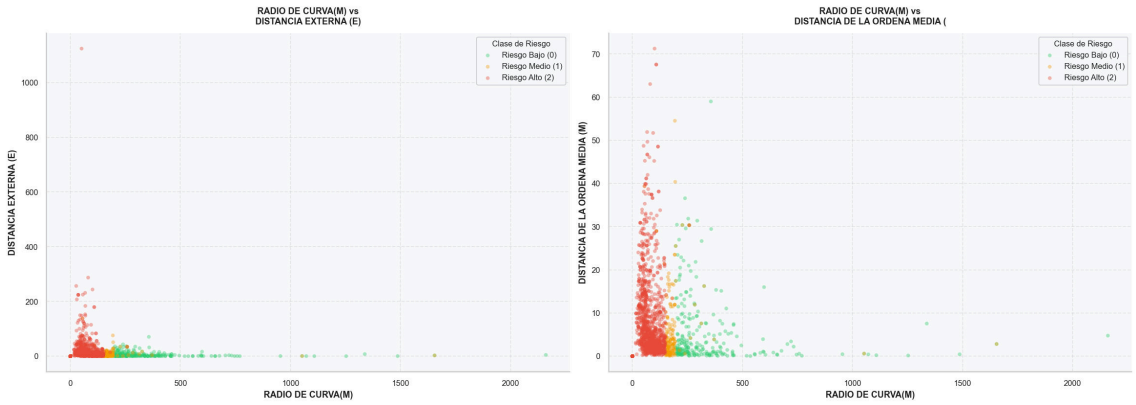


Figura 158 Diagr. de dispersión de Radio de curva vs dist. externa y orden. media.
Fuente: Elaboración propia

La combinación Radio vs Ángulo de Deflexión mostró la mayor separación entre

clases, mientras que Radio vs Distancia Externa y Ordenada Media presentaron nubes comprimidas con escasa diferenciación, confirmando su bajo poder discriminatorio individual.

El solapamiento entre clases en todos los gráficos evidencia que ninguna variable geométrica por sí sola separa completamente los niveles de riesgo, justificando el uso de la RNA para aprender fronteras de decisión no lineales en el espacio de 15 dimensiones.

H) Análisis estadístico Mann-Whitney U

La figura muestra la significancia estadística de las 15 variables mediante el test Mann-Whitney U, expresada como $-\log_{10}$(p-valor). Las líneas verticales delimitan los umbrales: verde para $p = 0.05$ y roja para $p = 0.001$.

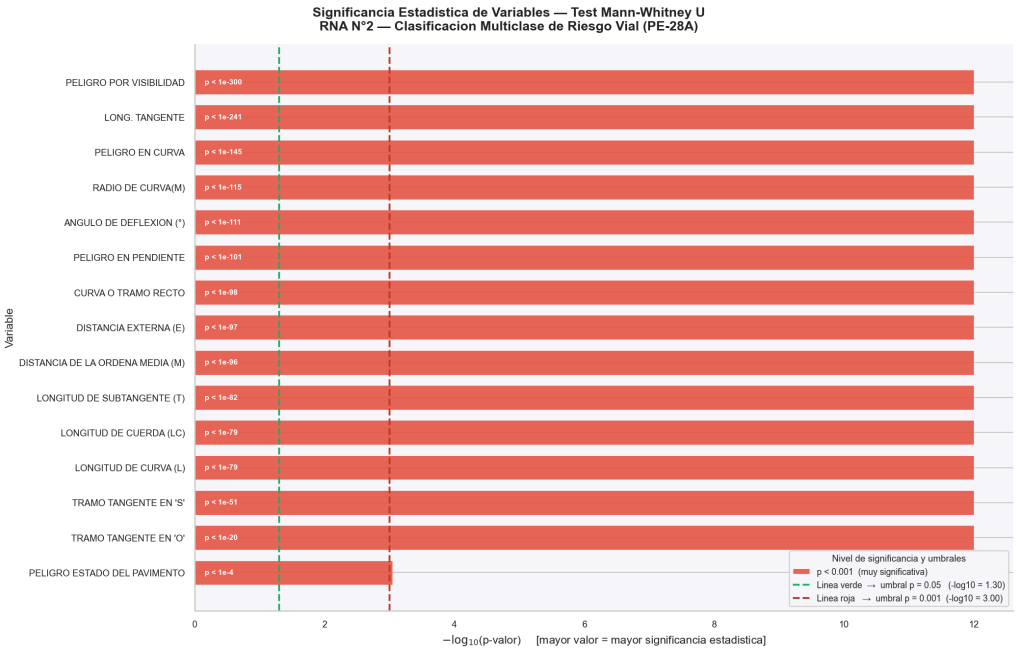


Figura 159 Significancia de variables - Test de Mann Whitney U.
Fuente: Elaboración propia

Los resultados indican que 14 de las 15 variables superan ampliamente el umbral $p < 0.001$, siendo Peligro por Visibilidad ($p < 10^{-300}$), Long. Tangente ($p < 10^{-241}$) y Peligro en Curva ($p < 10^{-145}$) las más significativas. Únicamente Peligro Estado del Pavimento presenta el valor más bajo ($p < 10^{-4}$), confirmando su menor poder discriminatorio observado en análisis previos.

En conjunto, las 15 variables presentan diferencias estadísticamente significativas entre clases de riesgo, validando su inclusión como variables de entrada en la RNA N°2.

3.11.2.5 Etapa II: Preprocesamiento de datos

Tabla 53 *Resumen del preprocesamiento de datos - Etapa II*

Proceso	Parámetros clave	Resultado
División estratificada	70% Entrenamiento (2,097) / 20% Validación (600) / 10% Evaluación (300)	Proporciones Riesgo Bajo/Medio/Alto en tres particiones
Normalización Z-Score	scaler.fit_transform(X_train); scaler.transform(X_val, X_test). Ajuste exclusivo en Train	$\mu \approx 0, \sigma \approx 1$ en Entrenamiento; sin data leakage
Sobremuestreo aleatorio	`resample(replace=True)`, solo en entrenamiento; clase mayoritaria = 1,798 muestras; $3 \times 1{,}798 = 5{,}394$ registros balanceados	Base de datos balanceado de 2,097 → 5,394 registros
Visualización Z-Score	KDE por clase en espacio normalizado; curva de referencia $\mathcal{N}(0, 1)$; 15 variables	Separabilidad entre clases normalizado

Fuente: Elaboración propia.

A) División estratificada de la base de datos 70/20/10%

La figura siguiente presenta la distribución de las tres clases de riesgo en cada partición de la base de datos, confirmando que la división estratificada preservó correctamente las proporciones originales de Entrenamiento, Validación y Evaluación.

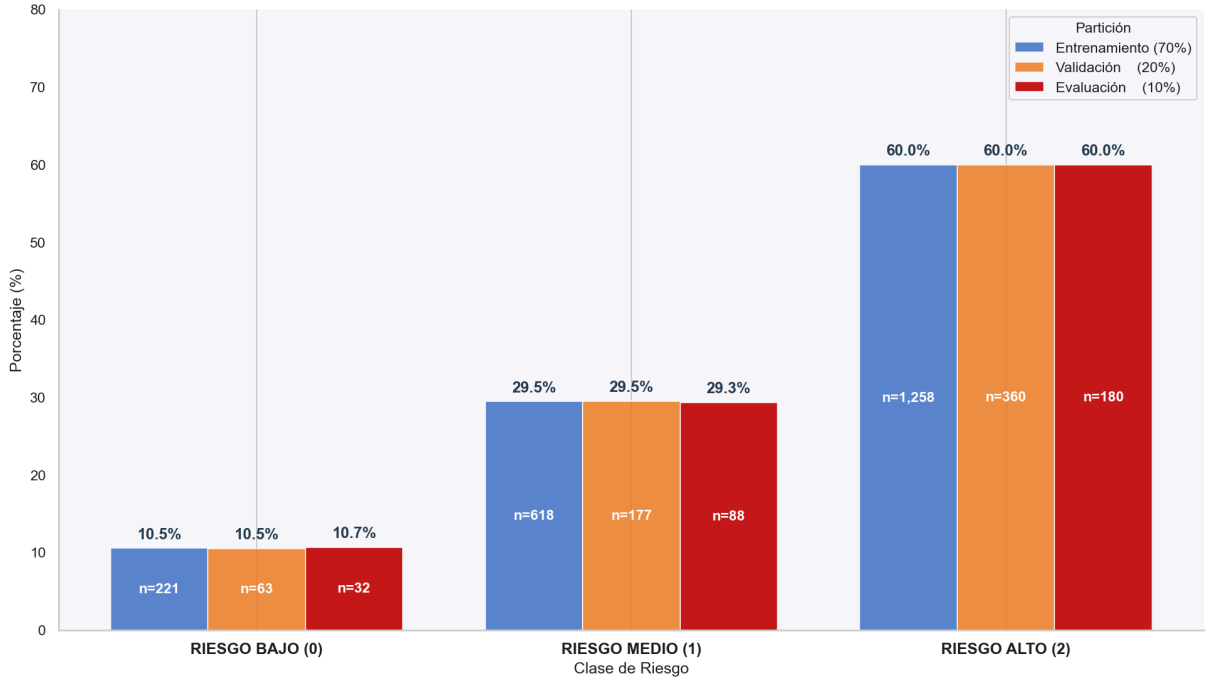


Figura 160 División estratificada de la base de datos 70/20/10%.
Fuente: Elaboración propia

Las proporciones son prácticamente idénticas en las tres particiones: Riesgo Bajo concentra aproximadamente el 10.5% (entrenamiento: $n = 221$, Validación: $n = 63$, Evaluación: $n = 32$), Riesgo Medio el 29.5% (entrenamiento: $n = 618$, Validación: $n = 177$, Evaluación: $n = 88$), y Riesgo Alto el 60.0% (entrenamiento: $n = 1{,}258$, Validación: $n = 360$, Evaluación: $n = 180$).

B) Normalización Z-Score variable por variable

Las figuras siguientes muestran cada variable después de la normalización Z-Score en dos paneles: la **fila superior** presenta el histograma de densidad con la curva de referencia $\mathcal{N}(0, 1)$ en rojo y la línea naranja en $\mu \approx 0$, confirmando el centrado correcto de cada variable; la **fila inferior** presenta la función de densidad kernel (KDE) por clase de riesgo permitiendo evaluar si las clases son separables en el espacio normalizado.

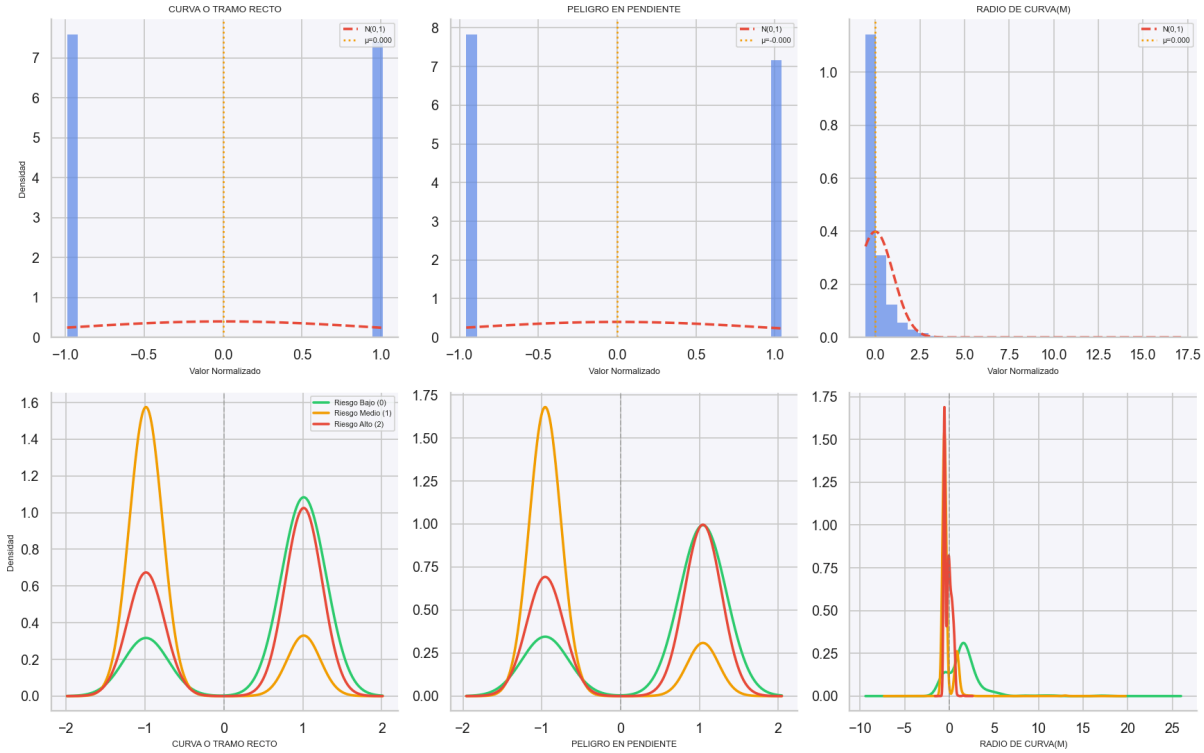


Figura 161 Normalización Z-Score y KDE por clase de variables del 1-3
Fuente: Elaboración propia

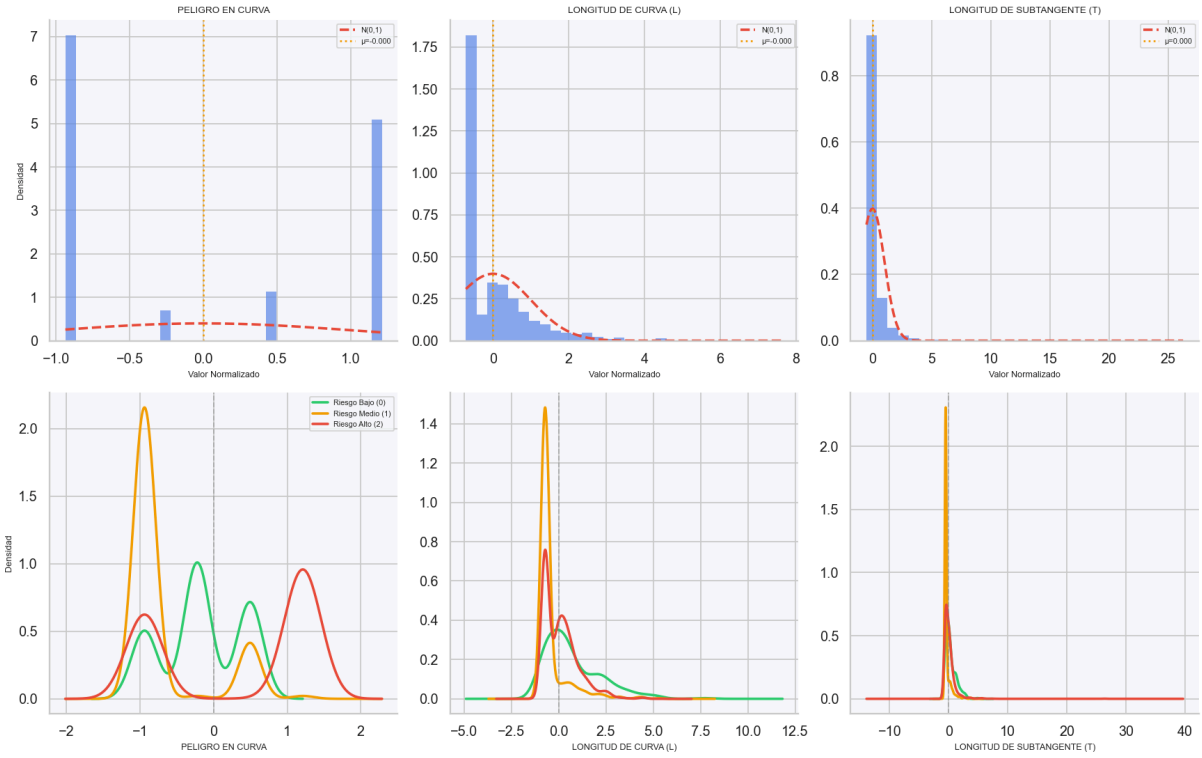


Figura 162 Normalización Z-Score y KDE por clase de variables del 4-6
Fuente: Elaboración propia

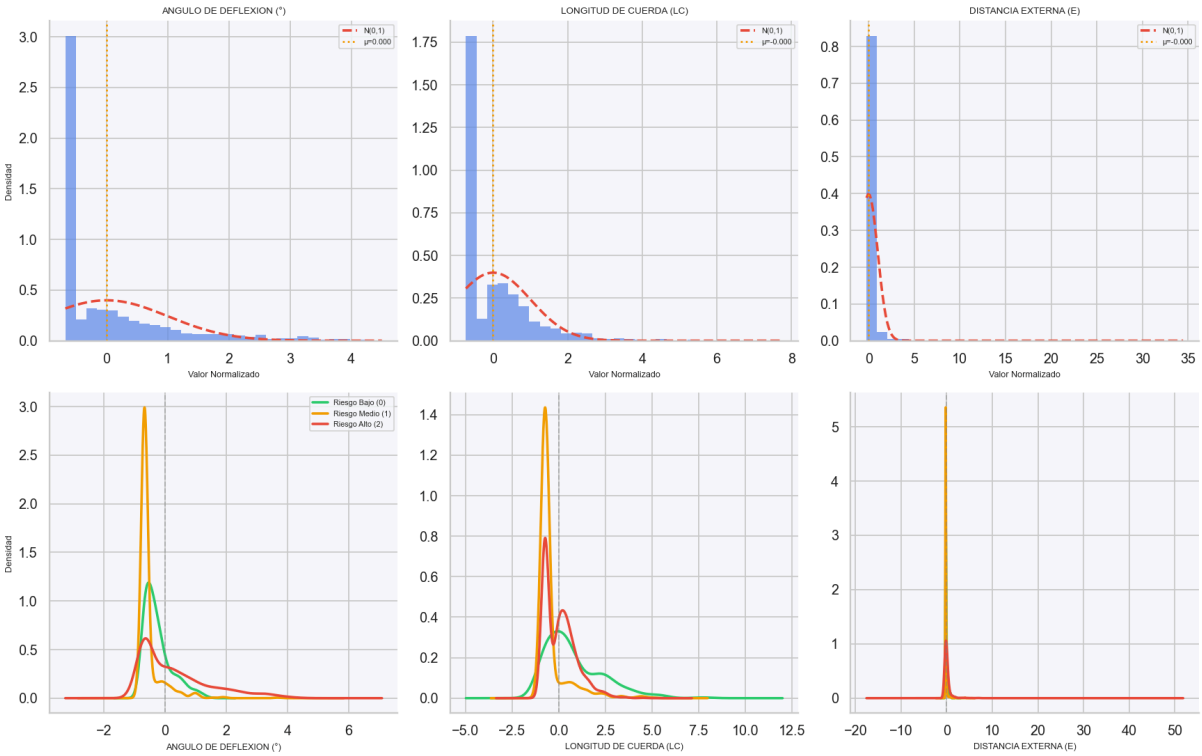


Figura 163 Normalización Z-Score y KDE por clase de variables del 7-9
Fuente: Elaboración propia

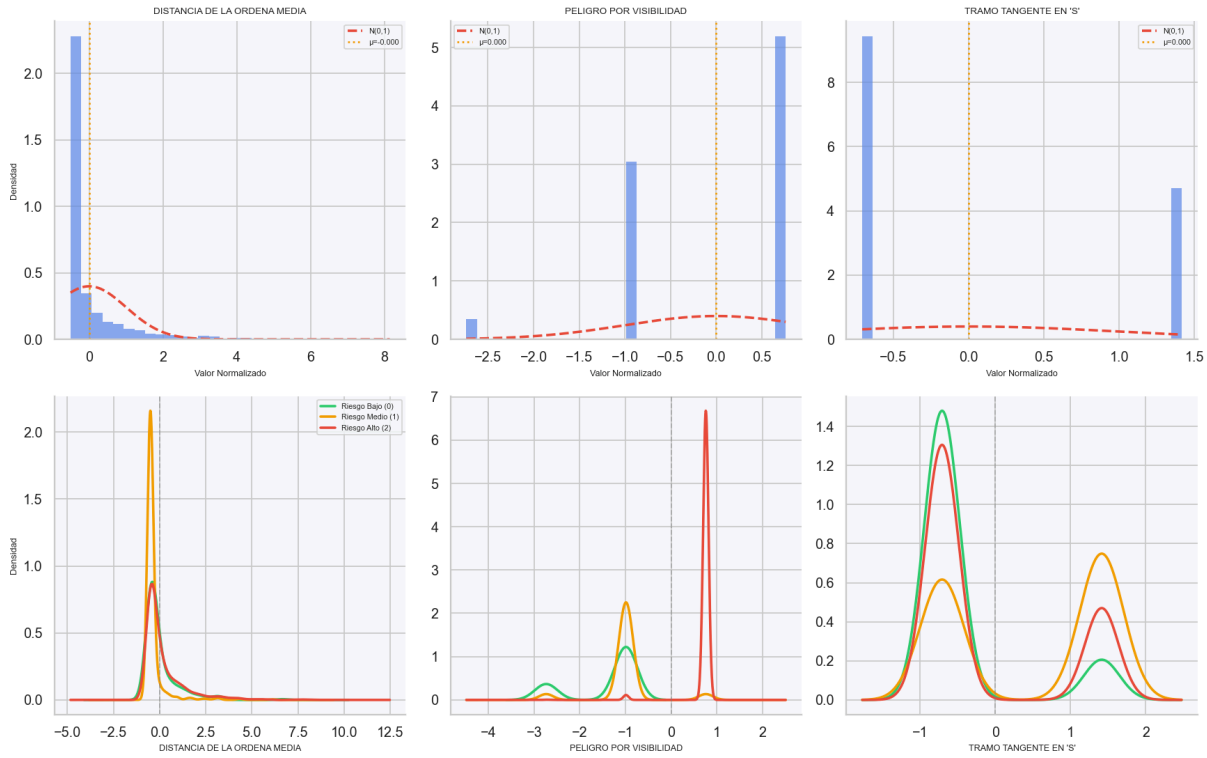


Figura 164 Normalización Z-Score y KDE por clase de variables del 10-12
Fuente: Elaboración propia

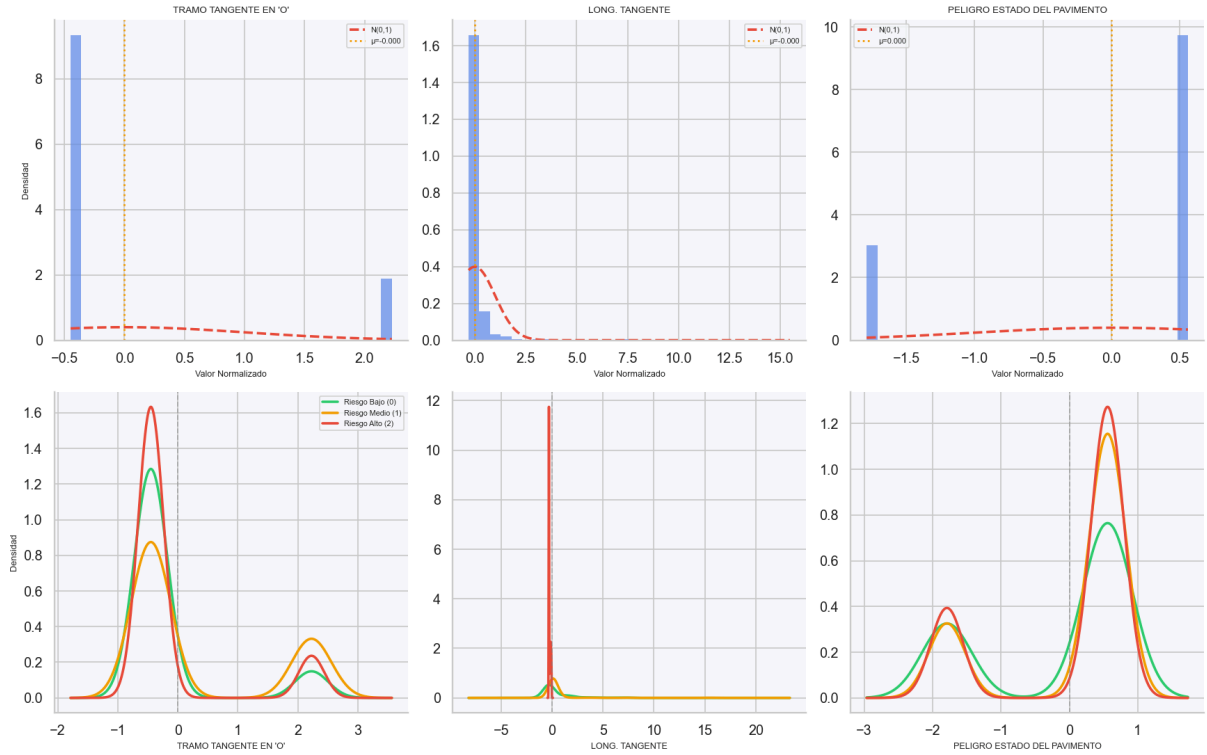


Figura 165 Normalización Z-Score y KDE por clase de variables 13-15
Fuente: Elaboración propia

- **Variables categóricas:** Peligro en Curva, Pendiente, Tramo Tangente, etc mues-

tran picos discretos en el histograma que no se ajustan a la curva $\mathcal{N}(0, 1)$, lo cual es esperado por su naturaleza discreta. En las KDE, las tres clases presentan picos en posiciones distintas, indicando cierta separabilidad entre grupos.

- **Variables continuas:** Radio, Longitudes, Ángulo presentan distribuciones asimétricas con cola derecha en el histograma, producto de los outliers conservados. Las KDE muestran separación parcial entre clases, siendo más notoria en Radio de Curva y Ángulo de Deflexión.
- **Peligro por Visibilidad** registró la mejor separación en las KDE: Riesgo Alto concentrado en un pico diferenciado de Riesgo Bajo y Medio.
- **Peligro Estado del Pavimento** y **Long. Tangente** muestran las tres KDE casi superpuestas, ratificando su bajo poder discriminatorio individual observado en análisis previos.

3.11.2.6 Etapa III: Diseño, optimización y entrenamiento de la RNA

A) Arquitectura del Modelo RNA N°2

Se muestra la topología definitiva del MLP: 15→128→64→32→16→3, con 4 capas ocultas. La arquitectura sigue un patrón de reducción progresiva: capas iniciales amplias capturan representaciones de baja abstracción, mientras que las capas posteriores comprimen la información en representaciones de alta abstracción que facilitan la clasificación entre las tres clases de riesgo, los cuales son riego menor = 0, riesgo medio = 1 y riesgo = 2.

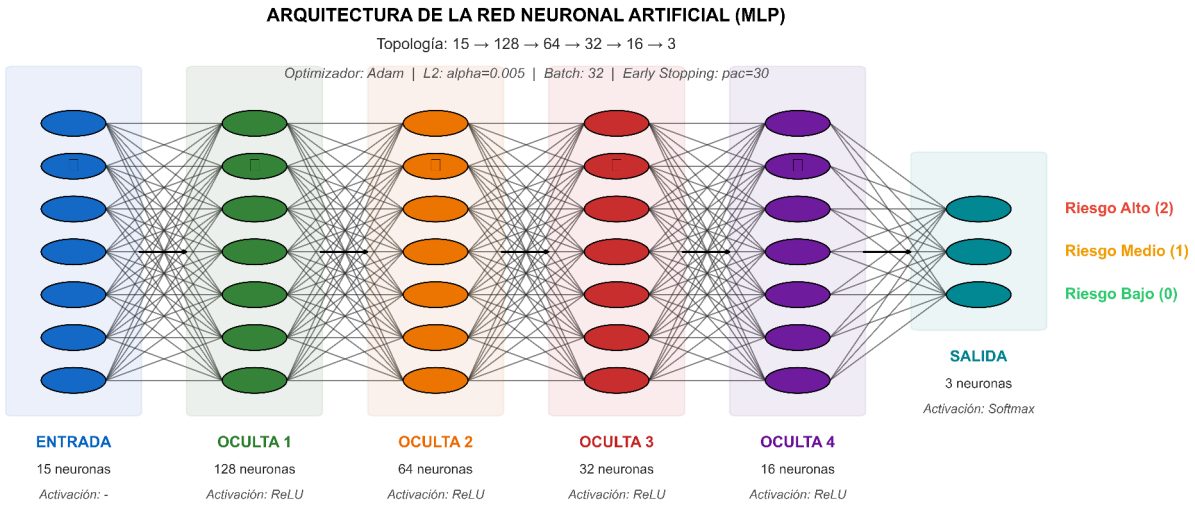


Figura 166 Arquitectura de la red neuronal N°02: 15→128→64→32→16→3. Optimizador Adam · L2 α=0.05 Batch=32 Epocas=66.
Fuente: Elaboración propia

Tabla 54 *Descripción de capas de la arquitectura RNA N°2*

Capa	Tipo	Neuronas	Activación	Función
Entrada	Input Layer	15	-	Una neurona por variable geométrica (normalizada)
Oculta 1	Dense	128	ReLU	Extracción de características primarias de alta dimensión
Oculta 2	Dense	64	ReLU	Combinación y refinamiento de características
Oculta 3	Dense	32	ReLU	Abstracción de patrones intermedios de riesgo
Oculta 4	Dense	16	ReLU	Representación compacta pre-clasificación
Salida	Dense	3	Softmax	P(Bajo) + P(Medio) + P(Alto) = 1.0

Fuente: Elaboración propia.

3.11.2.7 Etapa IV: Análisis anti-overfitting

A) Efecto de la regularización L2 sobre el overfitting

Las figuras siguientes presentan el análisis de ocho valores del parámetro de regularización L2 (α) evaluando su efecto en el accuracy y en el GAP de overfitting.

- **Figura Accuracy vs Alpha:** El accuracy de entrenamiento se mantuvo estable entre 0.967 y 0.969 para todos los valores de α, mientras que el accuracy de validación cruzada osciló entre 0.962 y 0.965. En $\alpha = 0.05$ el GAP entre ambas curvas alcanzó su mínimo de 0.002, indicando ausencia de sobreajuste.

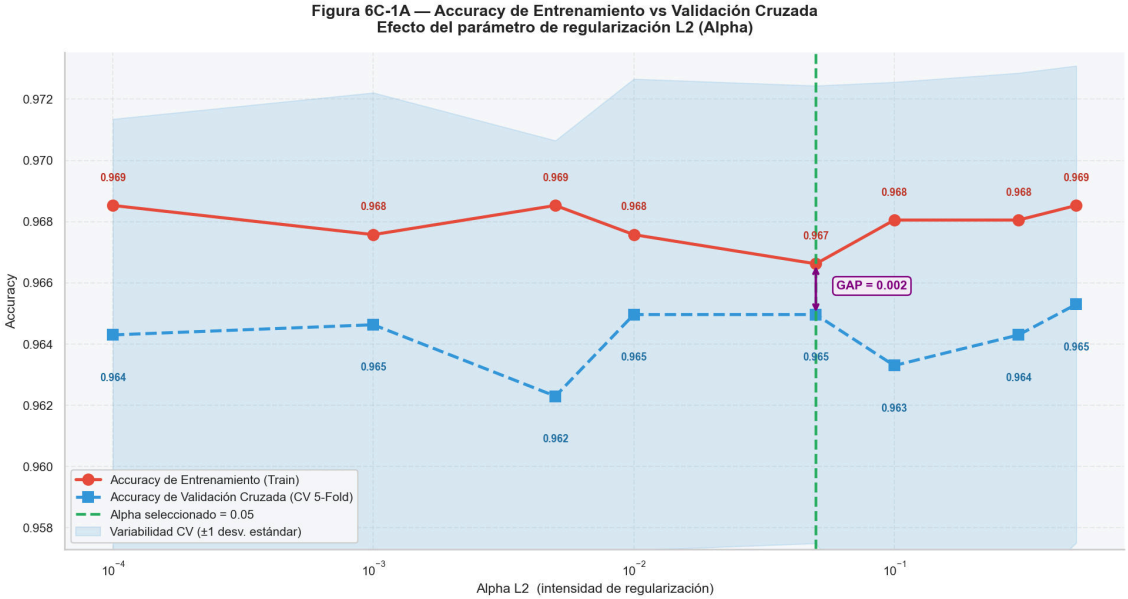


Figura 167 Accuracy vs Alpha L2.
Fuente: Elaboración propia

- **Figura GAP vs Alfa L2:** Todos los valores de α produjeron GAPs inferiores al

umbral moderado de 0.05, situándose entre 0.002 y 0.006. El valor $\alpha = 0.05$ registró el GAP mínimo de 0.002, justificando su selección como parámetro óptimo de regularización para el modelo final.

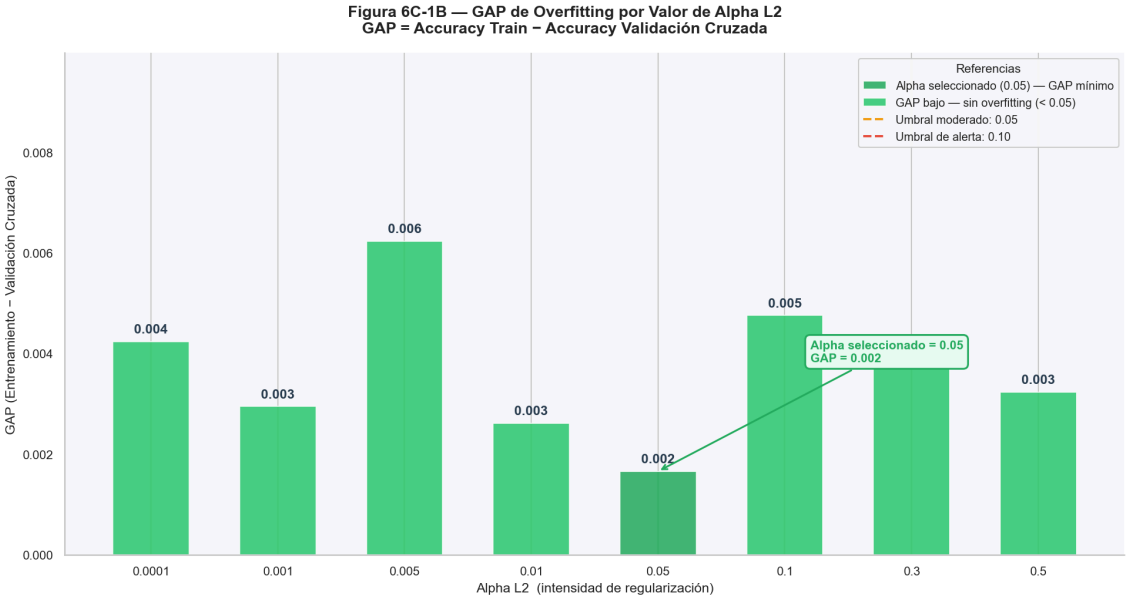


Figura 168 GAP vs Alfa L2.
Fuente: Elaboración propia

Ambas figuras confirman que el modelo no presenta sobreajuste significativo, y que $\alpha = 0.05$ ofrece el mejor equilibrio entre aprendizaje y generalización.

B) Comparación de arquitecturas de la RNA

Las figuras siguientes presentan la comparación de siete arquitecturas MLP evaluadas mediante validación cruzada estratificada 5-Fold, con el objetivo de seleccionar la arquitectura óptima para la clasificación multiclase para la predicción de accidentes de tránsito.

- **Grafica de Accuracy vs por arquitectura de la RNA:** Todas las arquitecturas registraron accuracy de entrenamiento entre 0.959 y 0.969, y accuracy de validación cruzada entre 0.961 y 0.965, evidenciando un rendimiento homogéneo entre configuraciones. La arquitectura seleccionada `(128,64,32,16)` alcanzó Train = 0.969 y CV = 0.963, con un GAP de 0.006, dentro del umbral aceptable.

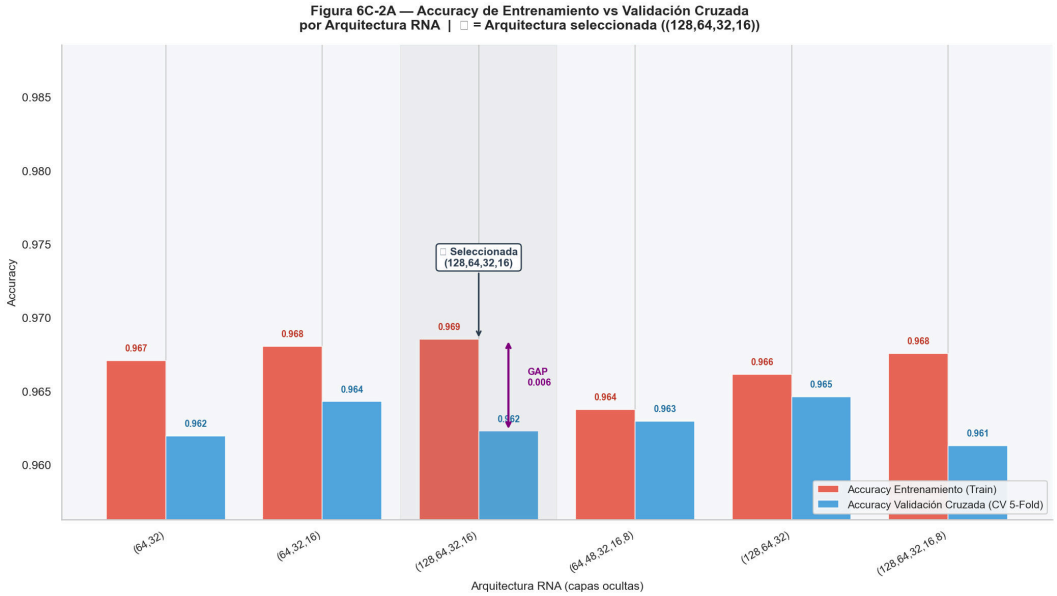


Figura 169 Accuracy vs por arquitectura RNA.
Fuente: Elaboración propia

- **Grafica de GAP vs por arquitectura de la RNA:** Todas las arquitecturas presentaron GAPs inferiores al umbral moderado de 0.05, con valores entre 0.001 y 0.006, confirmando ausencia de sobreajuste en todas las configuraciones evaluadas. La arquitectura `(128,64,32,16)` registró el GAP de 0.006, justificando su selección por ofrecer el mejor balance entre capacidad representacional y generalización.

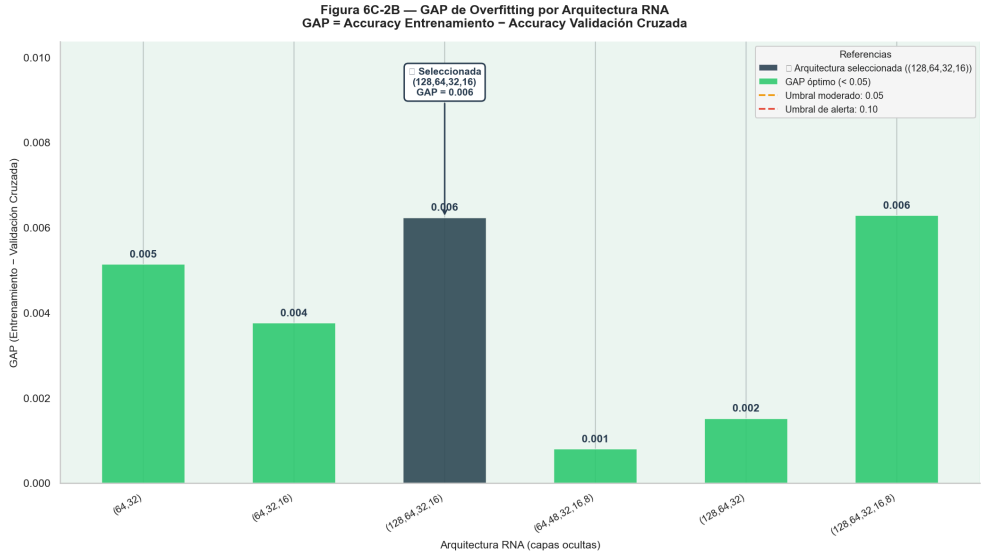


Figura 170 GAP vs por arquitectura de la RNA.
Fuente: Elaboración propia

- **Grafica de GAP vs número de parametros de la RNA:** El análisis de número de parámetros frente al GAP revela que mayor complejidad no implica mayor sobreajuste. La arquitectura seleccionada `(128,64,32,16)` con aproximadamente

12,600 parámetros mantiene un GAP de 0.006, mientras que arquitecturas más simples como `(64,48,32,16,8)` con 6,500 parámetros presentan GAP de 0.001, confirmando que la regularización L2 con $\alpha = 0.05$ controla eficazmente el overfitting independientemente de la complejidad del modelo.

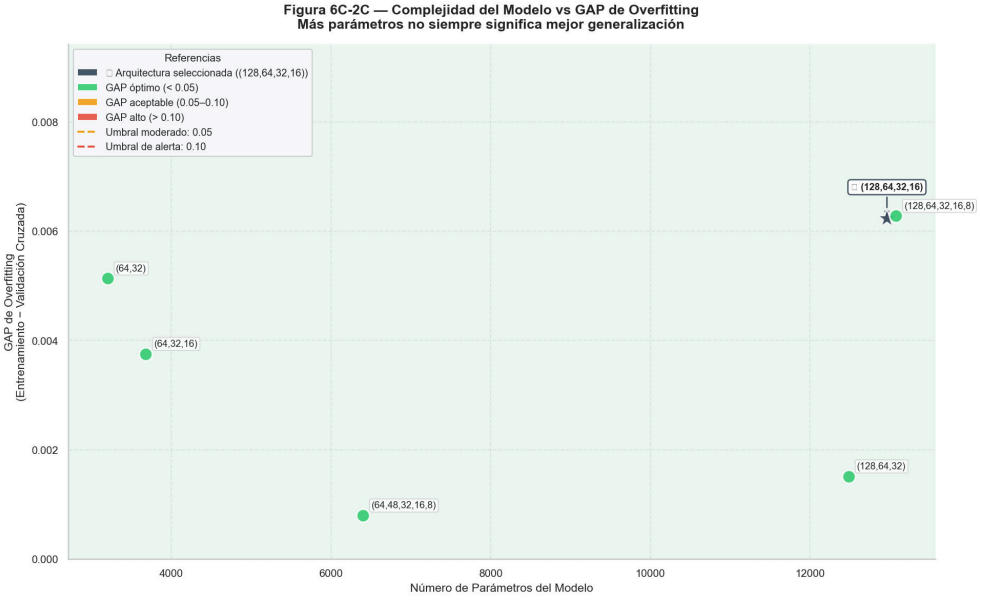


Figura 171 GAP vs número de parametros de la RNA.
Fuente: Elaboración propia

3.11.2.8 Etapa V: Evaluación e interpretación del Modelo RNA N°2

Tabla 55 *Métricas de rendimiento del modelo RNA N°2 en todas las particiones*

Métrica	Train (70%)	Val (20%)	Test (10%)	Umbral
Accuracy	97.66%	97.83%	98.00%	≥90% ✓
Precisión (macro)	95.17%	96.07%	95.71%	≥90% ✓
Recall (macro)	97.56%	97.35%	97.92%	≥90% ✓
F1-Score (macro)	96.23%	96.69%	96.69%	≥90% ✓
F1-Score (weighted)	—	—	98.02%	≥90% ✓
AUC-ROC (macro)	—	—	0.9985	≥0.90 ✓

Fuente: Elaboración propia.

A) Convergencia del entrenamiento y accuracy por partición

La figuras siguientes evalúan la convergencia y el rendimiento del modelo RNA N°2 tras 66 épocas de entrenamiento con Early Stopping.

- **Curva de convergencia:** La pérdida (cross-entropy) descendió rápidamente desde 0.52 en la época 1 hasta estabilizarse en 0.118 al final del entrenamiento. El descenso pronunciado en las primeras 10 épocas indica una tasa de aprendizaje

efectiva, seguido de una convergencia gradual y estable sin oscilaciones, lo que confirma que el optimizador Adam con tasa de aprendizaje adaptativa funcionó correctamente.

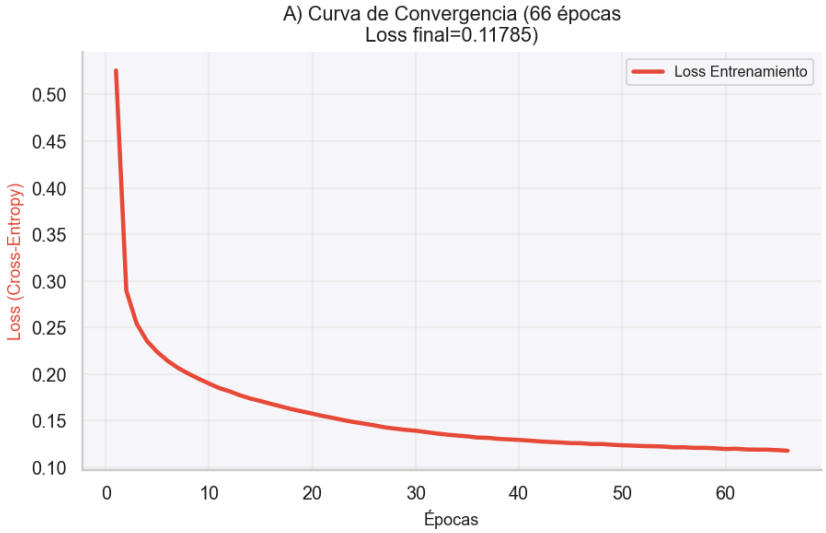


Figura 172 Curva de convergencia.
Fuente: Elaboración propia

- **Accuracy por partición:** Las tres particiones registraron accuracy: Entrenamiento 97.66%, Validación 97.83% y Evaluación 98.00%. La diferencia mínima entre particiones confirma ausencia de sobreajuste y buena predicción del modelo del entrenamiento.

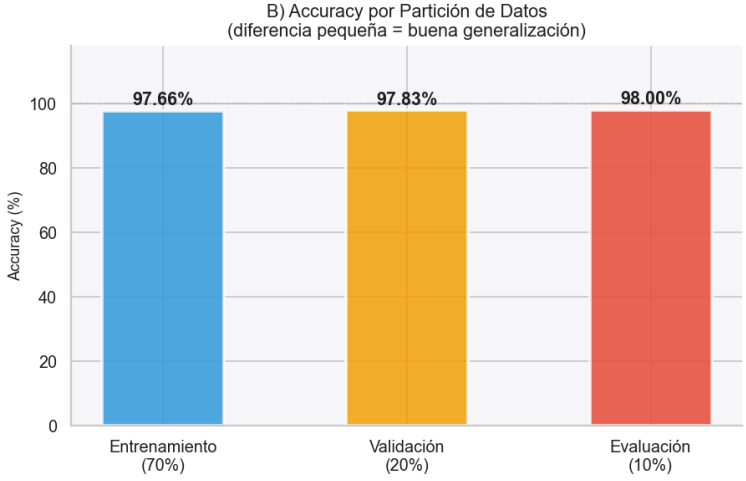


Figura 173 Accuracy por partición.
Fuente: Elaboración propia

B) Pérdida y Accuracy vs. épocas (Entrenamiento vs. Validación)

Las figuras siguientes muestran la evolución del entrenamiento durante 66 épocas con Early Stopping y restauración de mejores pesos en la época 36.

- **Pérdida vs épocas:** Ambas curvas descendieron rápidamente y convergieron de forma estable. El mínimo de validación se alcanzó en la época 36 (val_loss = 0.0868), correspondiente a los mejores pesos del modelo. La zona sombreada reducida entre curvas confirma ausencia de sobreajuste.

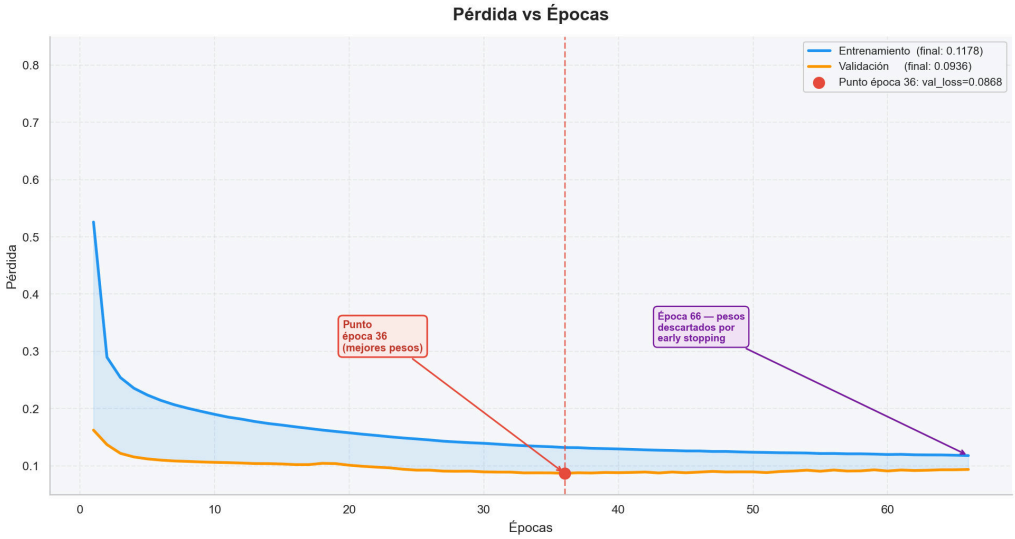


Figura 174 Pérdida vs épocas.
Fuente: Elaboración propia

- **Accuracy vs épocas:** Las curvas de entrenamiento (96.95%) y validación (96.33%) se mantuvieron paralelas durante todo el entrenamiento. En la época 36 el modelo registró Train = 97.66% y Val = 97.83%, alcanzando el máximo de validación en la época 37 con 98.00%.

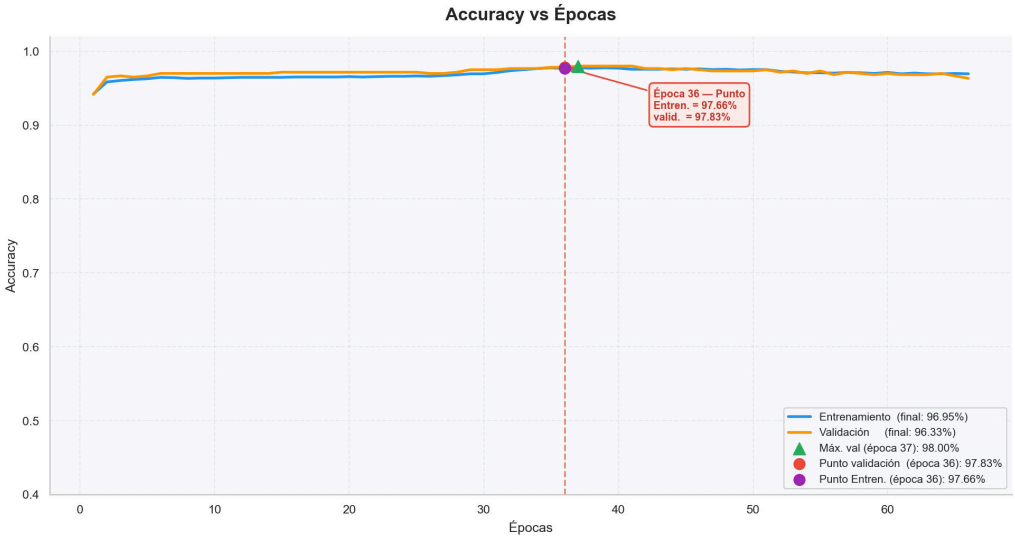


Figura 175 Accuracy vs épocas.
Fuente: Elaboración propia

El GAP Train-Val inferior al 1% en toda la trayectoria confirma la ausencia de sobreajuste y la robustez del modelo.

C) F1-Score por clase y partición

Las figuras siguientes presentan el análisis completo del F1-Score del modelo RNA N°2, evaluado sobre las tres particiones y clases de riesgo.

- **Figura F1 Score por clase:** Las tres particiones registraron valores consistentes y superiores al umbral mínimo de 0.80 en todas las clases. Riesgo Bajo obtuvo F1 entre 0.934 y 0.946, Riesgo Medio entre 0.960 y 0.965, y Riesgo Alto entre 0.992 y 0.994. La similitud entre Entrenamiento, Validación y Evaluación confirma ausencia de sobreajuste y buena generalización del modelo.

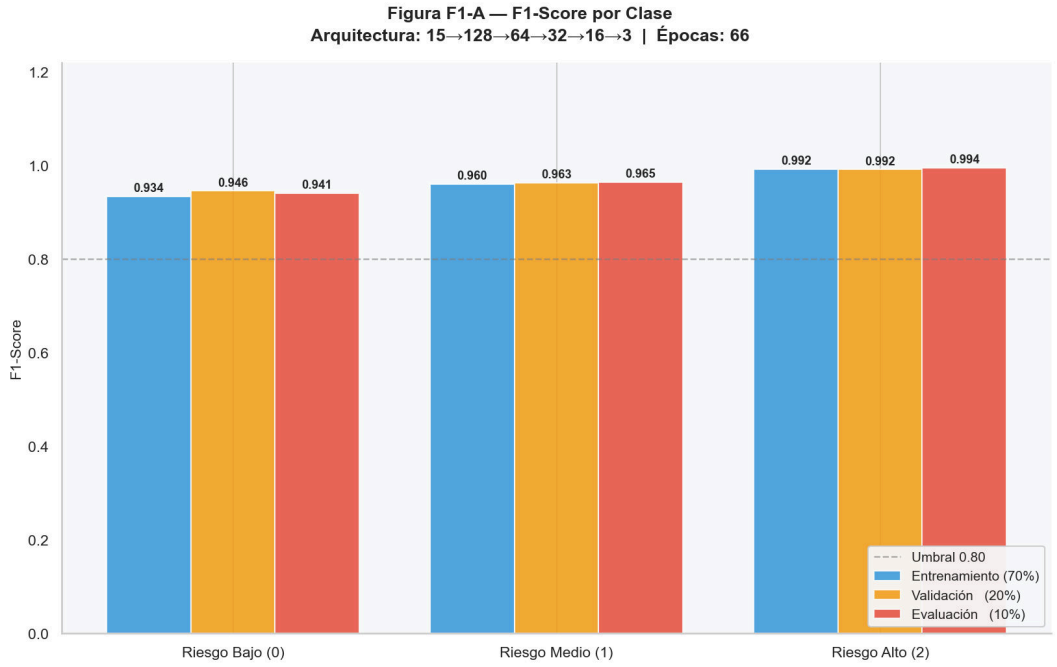


Figura 176 F1 Score por clase.
Fuente: Elaboración propia

- **Figura Precision, Recall y F1 por clase:** Sobre el conjunto de evaluación ($n = 300$), Riesgo Bajo presentó la precisión más baja (0.889) pero un recall perfecto (1.000), indicando que el modelo no pierde ningún tramo de riesgo bajo aunque ocasionalmente lo confunde con otras clases. Riesgo Medio y Riesgo Alto superaron 0.940 en todas las métricas, siendo Riesgo Alto el de mejor desempeño con precisión 0.994, recall 0.994 y F1 = 0.994.

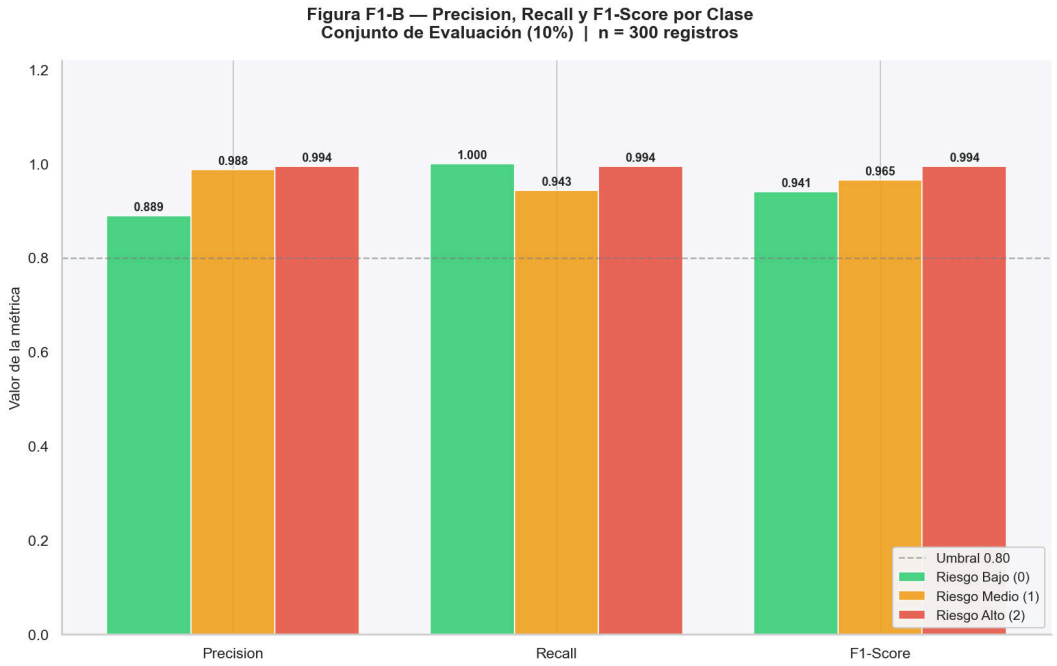


Figura 177 Precision, Recall y F1 por clase.
Fuente: Elaboración propia

- **Figura F1-Score global:** El F1-Macro en Evaluación alcanzó 0.9669 y el F1-Weighted 0.9802, ambos superando ampliamente el umbral de excelencia de 0.90. Los valores son consistentes entre Evaluación (0.9669), Validación (0.9669) y Entrenamiento (0.9623), confirmando la estabilidad y robustez del modelo RNA N°2.

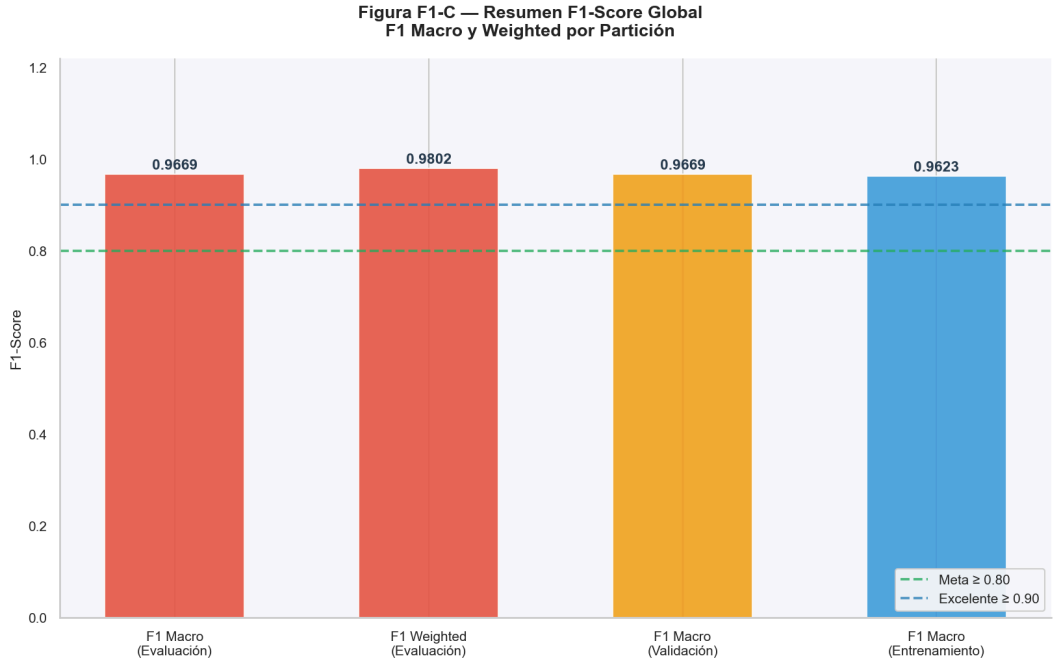


Figura 178 F1-Score global.
Fuente: Elaboración propia

D) Matrices de confusión y análisis VP/VN/FP/FN

Las figuras siguientes presentan el análisis completo de la matriz de confusión del modelo RNA N°2 sobre el conjunto de evaluación ($n = 300$ registros).

- **Figura de la matriz absoluta:** La diagonal principal concentra 294 predicciones correctas: Riesgo Bajo (32), Riesgo Medio (83) y Riesgo Alto (179). Los errores totales fueron 6 de 300 registros, todos entre clases adyacentes: 4 registros de Riesgo Medio clasificados como Riesgo Bajo, 1 de Riesgo Medio como Riesgo Alto, y 1 de Riesgo Alto como Riesgo Medio. No se registró ningún error entre Riesgo Bajo y Riesgo Alto, confirmando que el modelo distingue correctamente los extremos de riesgo.

Figura 179 Matriz absoluta.
Fuente: Elaboración propia

- **Figura de la matriz normalizada:** Expresada en porcentaje por fila, Riesgo Bajo alcanzó un Recall del 100.00% sin ningún error de clasificación. Riesgo Medio obtuvo 94.32% de acierto con una tasa de error de 4.55% hacia Riesgo Bajo y 1.14% hacia Riesgo Alto. Riesgo Alto registró el mejor desempeño con 99.44% de Recall y solo 0.56% de tasa de error hacia Riesgo Medio.

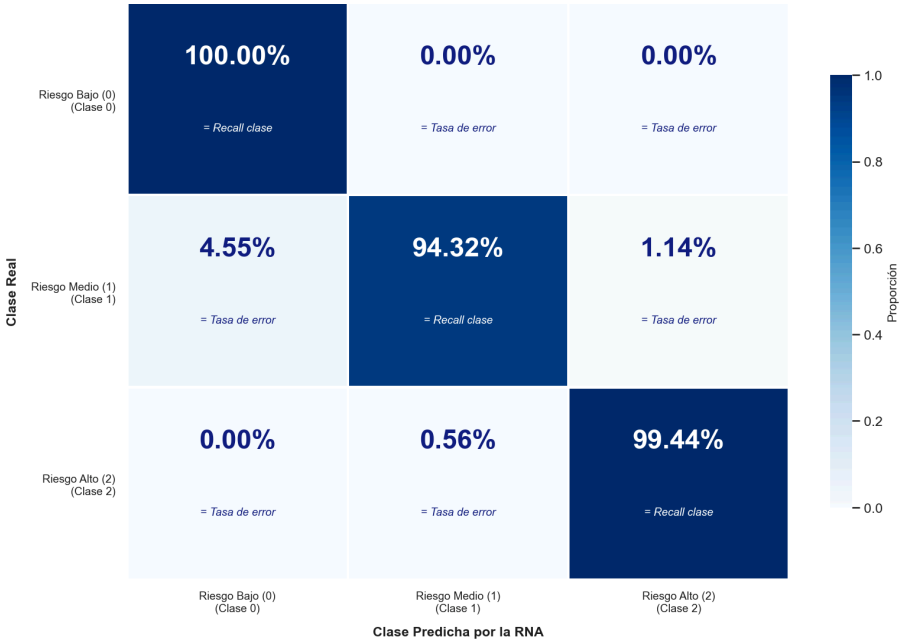


Figura 180 Matriz normalizada.
Fuente: Elaboración propia

- **Figura de VP/VN/FP/FN por clase:** El análisis revela que Riesgo Bajo registró VP=32, VN=264, FP=4 y FN=0, indicando que ningún tramo de Riesgo Bajo fue omitido por el modelo. Riesgo Medio obtuvo VP=83, VN=211, FP=1 y FN=5, siendo la clase con mayor número de falsos negativos. Riesgo Alto presentó VP=179, VN=119, FP=1 y FN=1, confirmando su alta capacidad de detección con mínimos errores.

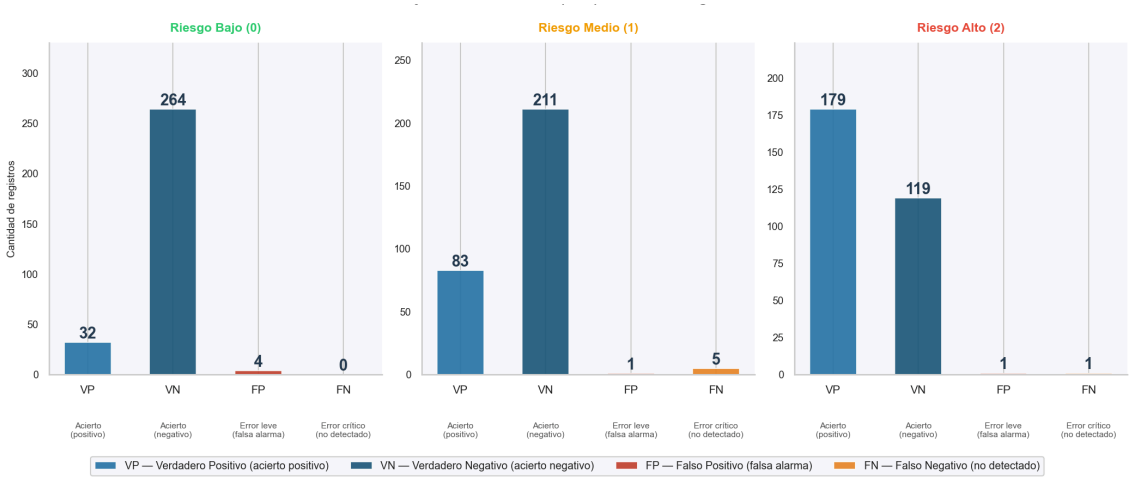


Figura 181 VP/VN/FP/FN por clase.
Fuente: Elaboración propia

En conjunto, los resultados confirman que el modelo RNA N°2 clasifica correctamente el 98.00% de los registros del conjunto de evaluación, con errores concentrados únicamente entre clases adyacentes de riesgo, lo cual es esperable dado que comparten

características geométricas similares.

E) Curvas ROC y área bajo la curva (AUC)

La figura presenta las curvas ROC del modelo RNA N°2 evaluadas mediante la estrategia One vs Rest (OvR) sobre el conjunto de evaluación (n = 300 registros).

- **Riesgo Bajo (0):** Alcanzó un AUC = 0.999, con un punto de operación (umbral = 0.50) de TPR = 1.000 y FPR = 0.015, indicando que el modelo detecta la totalidad de los tramos de riesgo bajo con una tasa mínima de falsas alarmas.
- **Riesgo Medio (1):** Registró un AUC = 0.997, con TPR = 0.955 y FPR = 0.005 en el punto de operación, evidenciando una alta capacidad discriminativa para identificar tramos de riesgo medio con muy pocos falsos positivos.
- **Riesgo Alto (2):** Obtuvo el AUC más alto con 1.000, y un punto de operación de TPR = 0.994 y FPR = 0.000, confirmando que el modelo identifica prácticamente la totalidad de los tramos de riesgo alto sin generar ninguna falsa alarma.

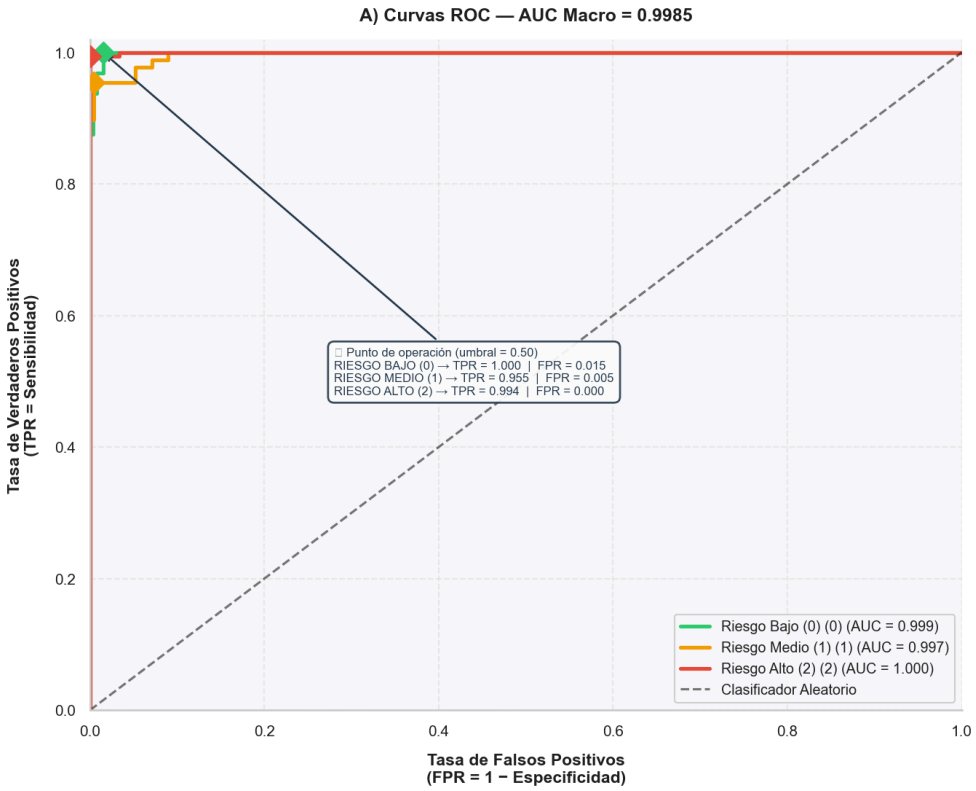


Figura 182 Curvas ROC y área bajo la curva (AUC).
Fuente: Elaboración propia

El AUC Macro del modelo alcanzó 0.9985. Las tres curvas se ubican en la esquina superior izquierda del espacio ROC, alejadas de la diagonal del clasificador aleatorio (AUC = 0.50), lo que confirma que el modelo RNA N°2 predice excelente las 3 clases.

F) Curva de aprendizaje: Diagnóstico bias y varianza

La figura muestra la evolución del accuracy en función del número de muestras de entrenamiento (238 a 2,381 registros), evaluada mediante validación cruzada 5-Fold.

Figura 183 Curva de Aprendizaje: Diagnóstico de Bias Varianza del Modelo RNA
Fuente: Elaboración propia

- **Fase inicial (238–571 muestras):** Mayor dispersión entre curvas, con accuracy de entrenamiento de 81.7% y validación de 66.4%, reflejando inestabilidad con pocos datos.
- **Fase de convergencia (571–2,381 muestras):** Ambas curvas ascendieron y convergieron progresivamente, alcanzando al final entrenamiento = 97.2% y validación cruzada = 96.5%.

El gap final de 0.007 confirma **buen balance** bias varianza, descartando overfitting y underfitting. La tendencia ascendente sugiere que el modelo podría mejorar con mayor volumen de datos.

G) Importancia de variables por permutación

Los resultados del análisis de importancia por permutación aplicado al Modelo N°2, donde se cuantifica la reducción media de *accuracy* por cada variable de entrada.

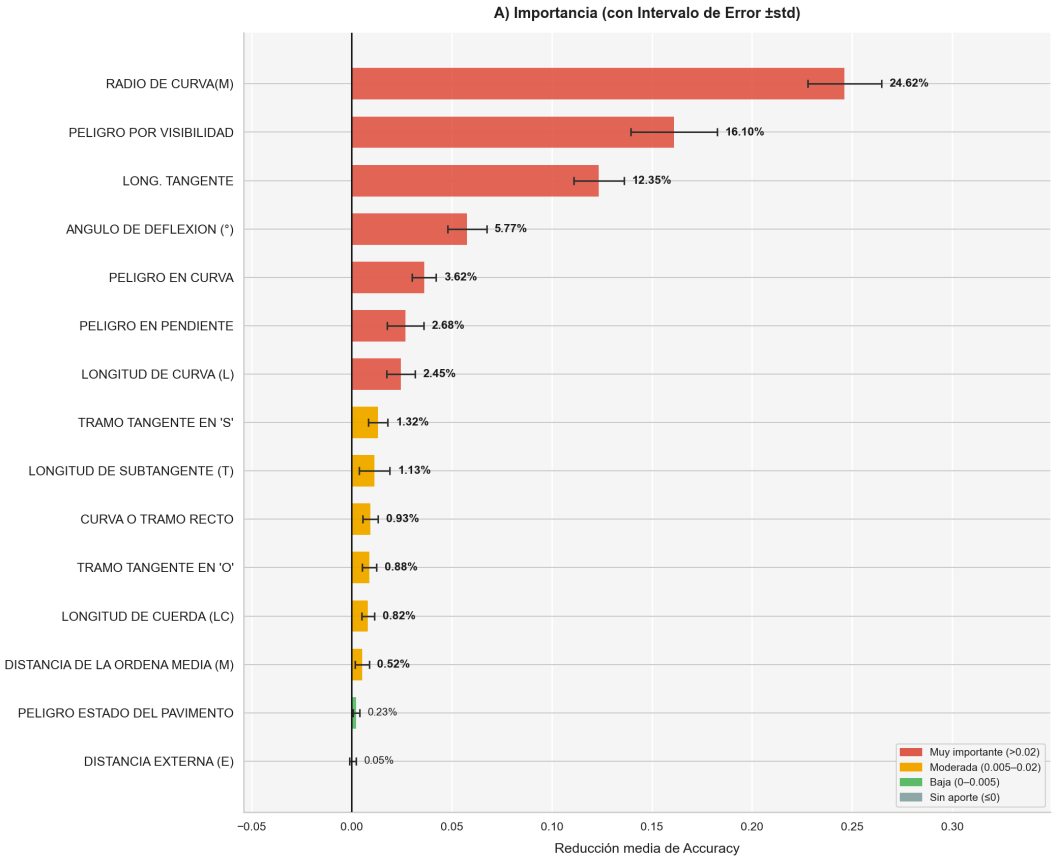


Figura 184 Importancia de variables con intervalo de error.
Fuente: Elaboración propia

Los hallazgos principales son los siguientes:

- **Radio de Curva (m)** resultó la variable más influyente, con una reducción de 24.62%, lo que indica que la curvatura horizontal es el factor geométrico determinante en la predicción de accidentes.
- **Peligro por Visibilidad** (16.10%) y **Long. Tangente** (12.35%) conforman el segundo y tercer lugar, confirmando que las condiciones de visibilidad y la longitud de los tramos rectos son críticas para la seguridad vial.
- **Ángulo de Deflexión** (5.77%), **Peligro en Curva** (3.62%), **Peligro en Pendiente** (2.68%) y **Longitud de Curva (L)** (2.45%) presentaron importancia moderada, con aportes entre 2% y 6%.
- Las variables **Peligro Estado del Pavimento** (0.23%) y **Distancia Externa (E)** (0.05%) mostraron aportes marginales, sugiriendo escasa influencia predictiva sobre el modelo.

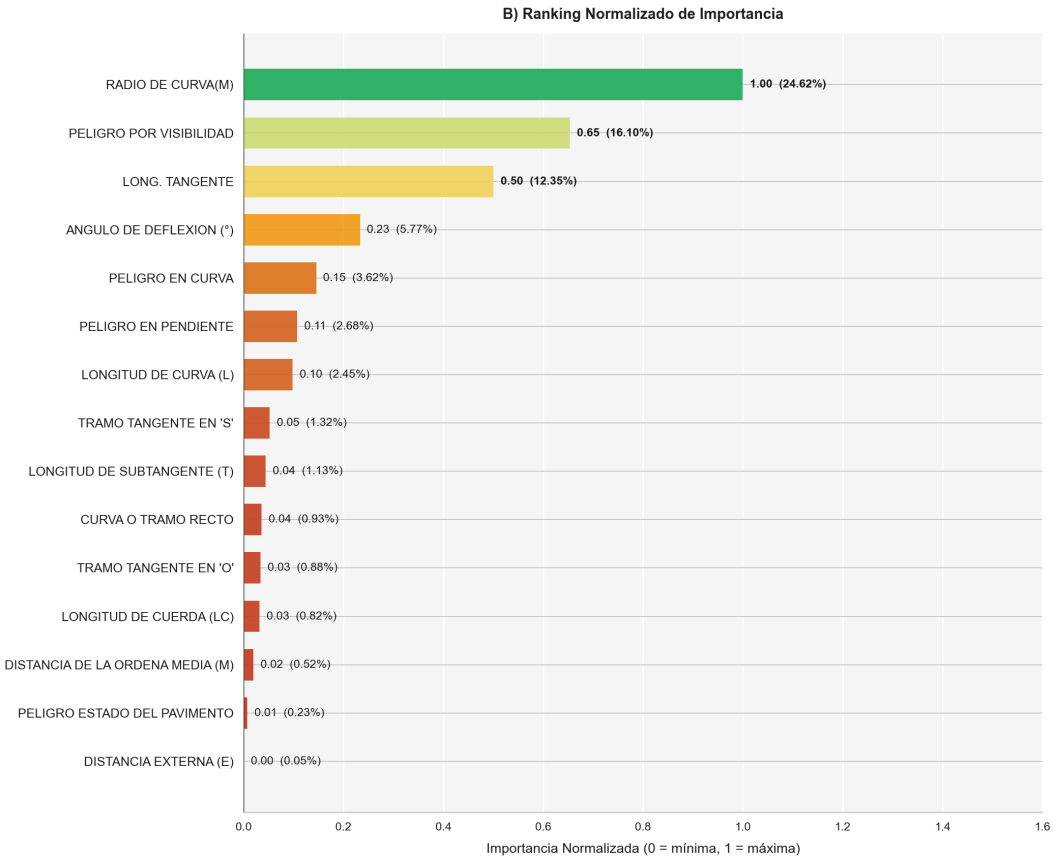


Figura 185 Ranking normalizado de importancia de variables
Fuente: Elaboración propia

El ranking normalizado confirma esta jerarquía, agrupando las variables en tres niveles: alta relevancia geométrica (Radio de Curva, Visibilidad y Long. Tangente), relevancia media (variables de alineamiento horizontal) y aporte marginal (resto de variables). El resumen cuantitativo se presenta en la tabla.

H) Frontera de decisión en el espacio PCA bidimensional

La figuras presenta las predicciones de la RNA proyectadas en el espacio bidimensional generado por el Análisis de componentes principales (PCA), donde PC1 explica el 48.2% y PC2 el 11.8% de la varianza total, acumulando en conjunto el 59.9%. Dado que la RNA opera en un espacio de 15 dimensiones, la visualización obtenida mediante `pca.inverse_transform()`, que reconstruye el espacio original para que el modelo prediga la clase en 2D.

- **Grafico de clase real vs frontera de decisión del RNA**: El grafico superpone los 2,997 registros reales sobre el fondo de decisión, el cual delimita tres regiones: verde para Riesgo Bajo (0), amarillo para Riesgo Medio (1) y rojo para Riesgo Alto (2), separadas por una frontera diagonal. La distribución de los puntos confirma que la RNA aprendió

correctamente la estructura del espacio: los triángulos rojos (n=1,798; 60.0%) dominan la zona roja, los diamantes naranja (n=883; 29.5%) la zona amarilla y los cuadrados verdes (n=316; 10.5%) la zona verde.

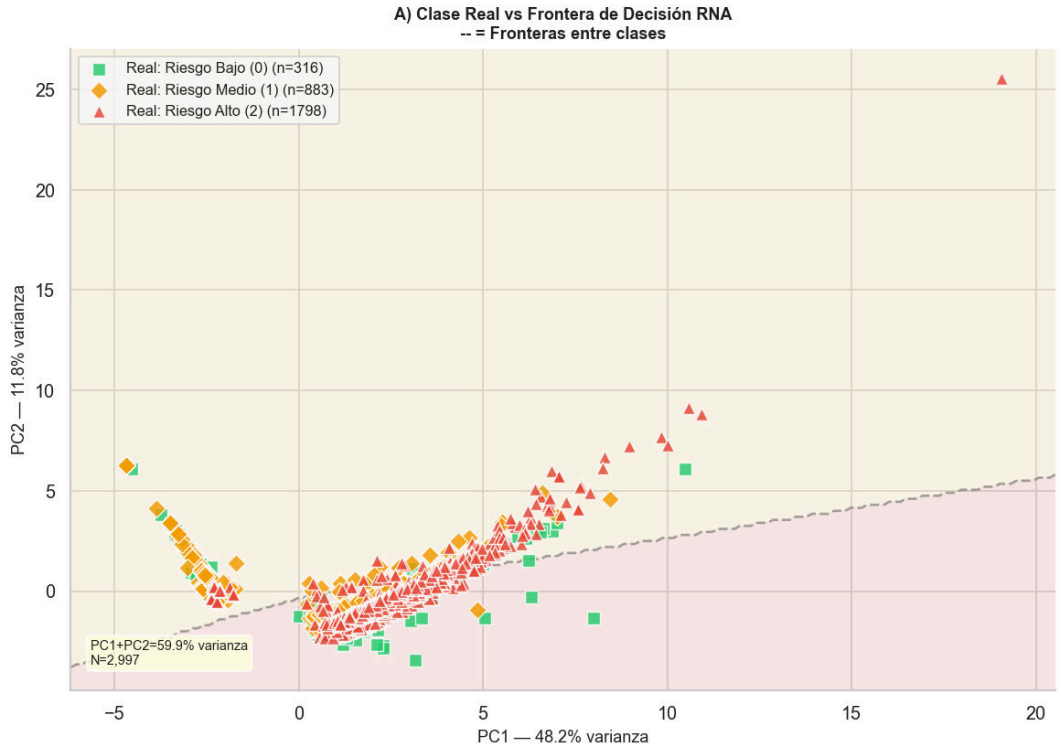


Figura 186 Clase real vs frontera de decisión del RNA.
Fuente: Elaboración propia

Tabla 56 *Distribución de registros por clase*

Clase	Símbolo	n	%
Riesgo Bajo (0)	Cuadrado verde	316	10.5%
Riesgo Medio (1)	Diamante naranja	883	29.5%
Riesgo Alto (2)	Triángulo rojo	1,798	60.0%
Total		**2,997**	**100%**

Fuente: Elaboración propia.

- **Grafico de distribución de predicciones del RNA**: El grafico distingue entre registros correctamente clasificados y errores de clasificación. De los 2,997 registros, 2,929 fueron acertados (97.7%), mientras que 68 resultaron mal clasificados (2.3%). Los errores se concentran exclusivamente sobre la frontera diagonal, lo que indica que la RNA no falla aleatoriamente sino únicamente en la zona de ambigüedad natural entre clases adyacentes. La clase Riesgo Medio acumula el 70.6% de los errores totales (n=48).

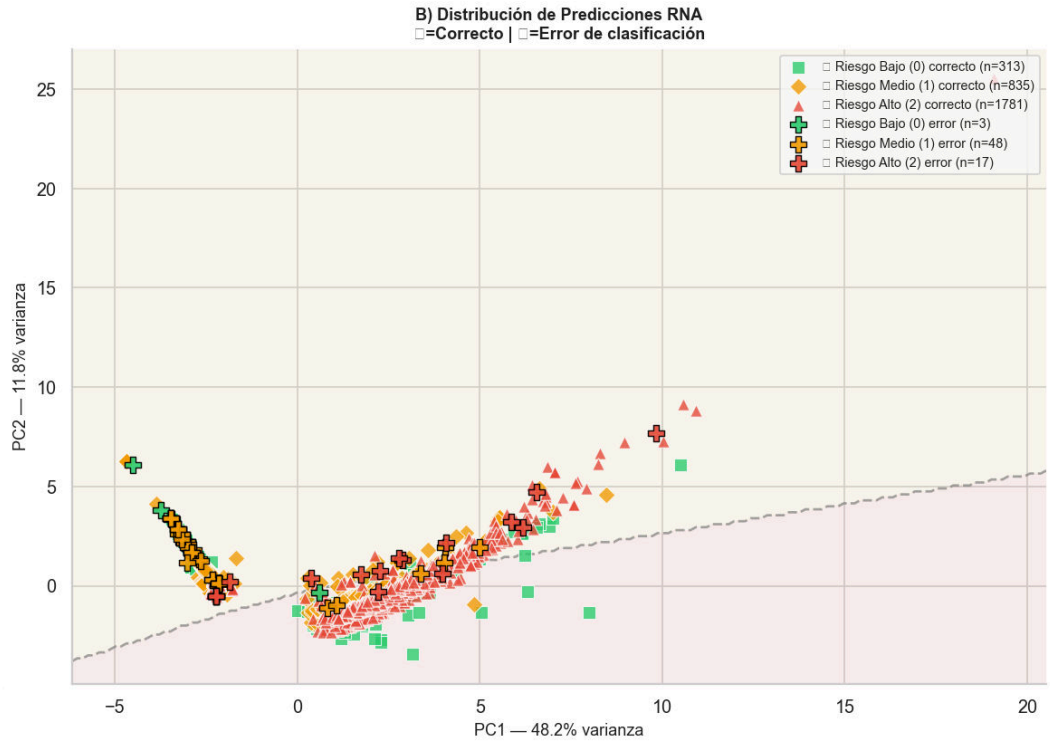


Figura 187 Distribución de predicciones del RNA.
Fuente: Elaboración propia

Tabla 57 *Aciertos y errores por clase*

Clase	Total	Correctos	Errores	Acierto
Riesgo Bajo (0)	316	313	3	99.1%
Riesgo Medio (1)	883	835	48	94.6%
Riesgo Alto (2)	1,798	1,781	17	99.1%
Total	**2,997**	**2,929**	**68**	**97.7%**

I) Resumen de métricas de evaluación de la RNA: Precisión, Recall, F1-Score y Accuracy

Las figuras siguientes presentan el análisis de Precisión, Recall, F1-Score y Accuracy (promedio macro) en las tres particiones del dataset: entrenamiento (70%, $n = 2{,}097$), validación (20%, $n = 600$) y evaluación (10%, $n = 300$).

- **Precisión, Recall y F1-Score por partición:**

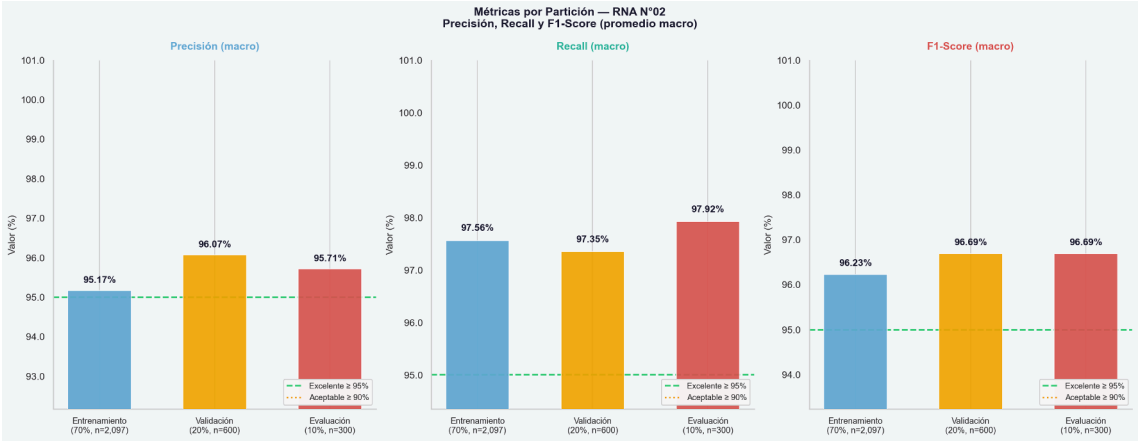


Figura 188 Precisión, Recall y F1-Score macro por partición — RNA N°02.
Fuente: Elaboración propia

Tabla 58 *Precisión, Recall y F1-Score macro por partición — RNA N°02*

Partición	n	Precisión	Recall	F1-Score
Entrenamiento (70%)	2,097	95.17%	97.56%	96.23%
Validación (20%)	600	96.07%	97.35%	96.69%
Evaluación (10%)	300	95.71%	**97.92%**	96.69%

Fuente: Elaboración propia.

Las tres métricas son estables entre particiones (variación ≤ 0.90 pp), superan el umbral de excelencia (≥ 95%) y el Recall alcanza su máximo en evaluación (97.92%), confirmando la alta efectividad del modelo.

- **Comparación de las 4 métricas por partición:**

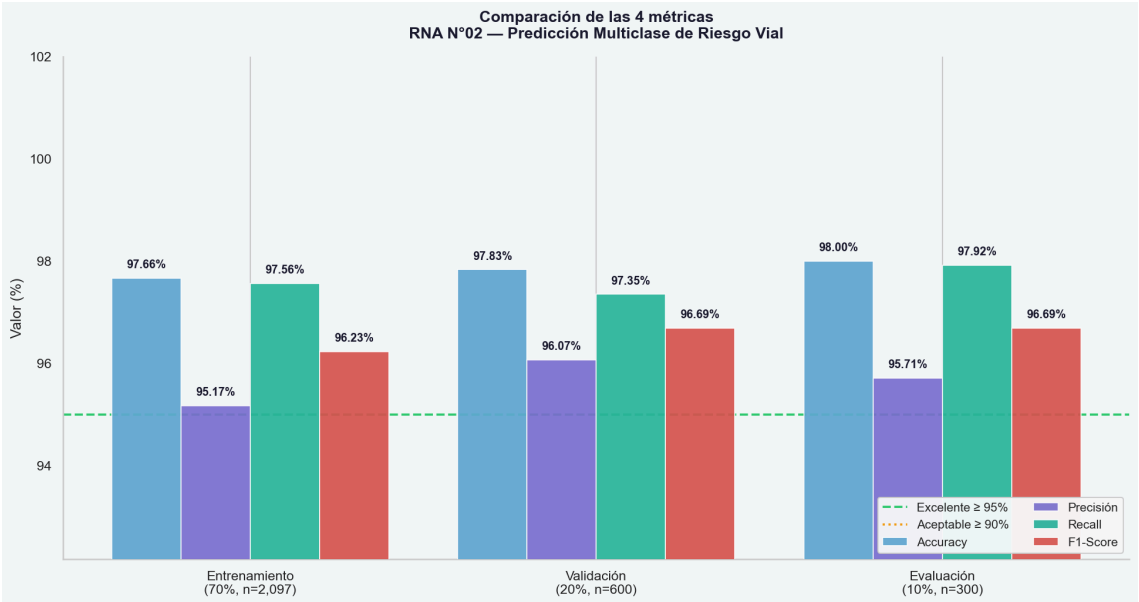


Figura 189 Comparación de las 4 métricas por partición — RNA N°02.
Fuente: Elaboración propia

Tabla 59 *Comparación de las 4 métricas por partición — RNA N°02*

Partición	n	Accuracy	Precisión	Recall	F1-Score
Entrenamiento (70%)	2,097	97.66%	95.17%	97.56%	96.23%
Validación (20%)	600	97.83%	96.07%	97.35%	96.69%
Evaluación (10%)	300	**98.00%**	95.71%	**97.92%**	96.69%

Fuente: Elaboración propia.

Las cuatro métricas son consistentes entre sí y entre particiones, sin degradación significativa, confirmando ausencia de sobreajuste. El Accuracy alcanza su valor máximo en evaluación (98.00%), resultado atribuible a la regularización L2 ($\alpha = 0.05$) y al Early Stopping en la época óptima 36. Todas las métricas superan el umbral de excelencia ($\geq 95\%$) en las tres particiones, validando la capacidad predictiva del modelo RNA N°02 para la clasificación multiclase de niveles de riesgo vial en la Vía Los Libertadores PE-28A.

3.11.2.9 Exportación del modelo RNA N°2 como aplicación web

El Modelo RNA N°2 fue exportado como aplicación web denominada **PROGRAMA RNA N°02: Predicción de Nivel de Riesgo Vial**, desarrollada en Python en el entorno JupyterLab. El programa permite clasificar cualquier tramo de la Vía Los Libertadores PE-28A en tres niveles de riesgo: **Riesgo Bajo**, **Riesgo Medio** o **Riesgo Alto** de accidente de tránsito, a partir del ingreso de 15 variables geométricas y de peligrosidad vial, con funcionamiento en dispositivos móviles para uso en campo.

Tabla 60 *Parámetros extraídos del modelo entrenado e integrados en el progrma RNA N°2*

Parámetro	Cantidad	Descripción
Pesos (`LAYERS_W`)	5 matrices	Pesos de las 5 capas de la arquitectura 15→128→64→32→16→3
Bias (`LAYERS_B`)	5 vectores	Términos independientes de cada capa
Total pesos + bias	**13,386**	**Parámetros numéricos del modelo**
Medias (`SCALER_MEAN`)	15 valores	Media μ de cada variable calculada sobre el conjunto de entrenamiento
Desv. estándar (`SCALER_STD`)	15 valores	Desviación estándar σ de cada variable calculada sobre el entrenamiento
Total Z-Score	**30**	**Parámetros de normalización**
Total general	**13,416**	**Parámetros totales embebidos en el HTML**

Fuente: Elaboración propia.

A) Proceso de generación del programa

El programa fue generado en cinco pasos secuenciales completamente automatizados por Python:

Paso 1: Extracción de parámetros del modelo.

Python accedió directamente a los objetos entrenados en memoria y serializó sus parámetros numéricos en formato JSON preservando la precisión numérica. Los 13,386 parámetros del modelo (pesos y bias) fueron extraídos del objeto `MLPClassifier` mediante los atributos `mlp.coefs_` y `mlp.intercepts_`, y los 30 parámetros de normalización Z-Score fueron extraídos del objeto `StandardScaler` ajustado exclusivamente sobre el conjunto de entrenamiento.

Paso 2: Generación dinámica de la arquitectura visual.

Una función Python generó automáticamente los nodos visuales de la arquitectura 15→128→64→32→16→3 a partir de los atributos del modelo entrenado, de modo que cualquier cambio en la arquitectura se refleje automáticamente en la interfaz del programa sin modificaciones manuales.

Paso 3: Construcción del bloque de constantes JavaScript.

Se construyó un bloque de texto conteniendo las variables JavaScript `LAYERS_W`, `LAYERS_B`, `SCALER_MEAN` y `SCALER_STD`, junto con los metadatos del modelo (Accuracy=98.00%, AUC-ROC=0.9985, F1-Macro=0.9669, 66 épocas, α=0.05).

Paso 4: Construcción del HTML completo.

El archivo fue estructurado en tres bloques concatenados:

Tabla 61 *Estructura de los tres bloques del archivo HTML generado — RNA N°02*

Bloque	Contenido
`HTML_PART_A`	Interfaz gráfica completa: encabezado institucional UNSCH, métricas del modelo, arquitectura visual 15→128→64→32→16→3, formulario de 15 variables (10 campos numéricos + 5 selectores categóricos), cálculo automático de L, T, LC, E y M, panel de resultados con barras de probabilidad por clase.
`JS_CONSTANTS`	Bloque de constantes JavaScript con los 13,416 parámetros del modelo embebidos: pesos, bias , medias y desviaciones estándar .
`HTML_PART_B`	Motor de inferencia en JavaScript: función `predecirRNA()`, normalización Z-Score , multiplicación matricial , activación ReLU en capas ocultas, función `softmax` en capa de salida y cálculo automático de variables geométricas.

Paso 5: Ensamblado y guardado del archivo.

Los tres bloques fueron concatenados y guardados en disco como un único archivo de texto plano:

$$\texttt{HTML_FINAL} = \texttt{HTML_PART_A} + \texttt{JS_CONSTANTS} + \texttt{HTML_PART_B} \tag{3.25}$$

El resultado fue el programa ejecutable directamente en cualquier navegador web sin instalación ni conexión a internet.

B) Motor de inferencia en JavaScript

El núcleo del programa es la función `predecirRNA()`, que replica fielmente el comportamiento del modelo Python en tres sub-pasos:

Tabla 62 *Sub-pasos de la función `predecirRNA()` y su equivalencia con Python*

Paso	Operación	JavaScript (programa HTML)	Equivalente Python
1	Normalización Z-Score	`normalizar(r): (v−MEAN[i])/STD[i]`	`scaler.transform(X)`
2	Capas ocultas con ReLU	`linLayer(x,W,b)` + `relu`	`mlp.predict_proba(X)` parcial
3	Capa de salida Softmax	`softmax(linLayer(x,W,b))`	`mlp.predict_proba(X)` salida

C) Diagrama de flujo de uso del programa

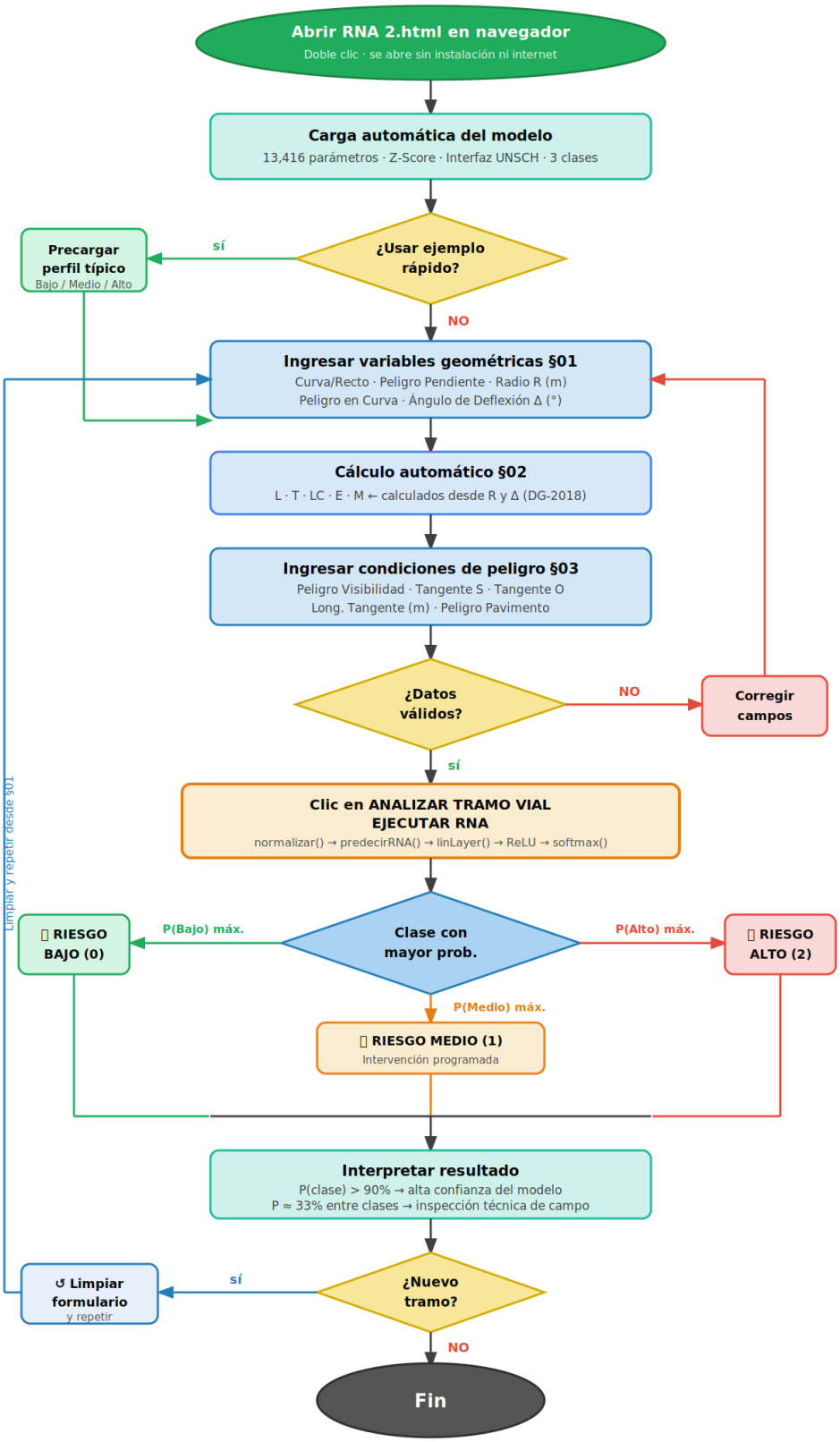


Figura 190 Diagram de flujo de uso del programa RNA N°02
Fuente: Elaboración propia

D) Características técnicas

Tabla 63 *Características técnicas del programa RNA 2.html*

Característica	Detalle
Nombre del archivo	`Programa RNA 2.html`
Herramienta de creación	Python · JupyterLab
Variables de entrada	15 variables (valores numéricos + selectores)
Cálculo automático	L, T, LC, E y M calculados desde R y Δ y variables de condiciones de peligro.
Salida del modelo	P(Riesgo Bajo) + P(Riesgo Medio) + P(Riesgo Alto) = 100% con barra visual de confianza por clase
Conectividad	Sin internet ni servidor externo
Compatibilidad	Chrome , Firefox, Edge , otros.
Métricas	Accuracy=98.00% · AUC-ROC=0.9985 · F1-Macro=0.9669

Fuente: Elaboración propia.

Tabla 64 *Procedimiento de uso del programa RNA 2.html*

Paso	Acción	Descripción
1	Abrir el programa	Doble clic sobre `Programa RNA 2.html`.
2	Usar ejemplo rápido (opcional)	Hacer clic en los botones de ejemplo **Riesgo Bajo (0)**, **Riesgo Medio (1)** o **Riesgo Alto (2)** para precargar valores típicos de cada nivel de riesgo
3	Ingresar variables geométricas §01	Seleccionar si el tramo es *Curva o Tramo Recto*, nivel de *Peligro en Pendiente*, ingresar el *Radio de Curva R (m)*, nivel de *Peligro en Curva* y el *Ángulo de Deflexión Δ (°)*
4	Verificar cálculo automático §02	Los parámetros *L, T, LC, E* y *M* se calculan automáticamente a partir de R y Δ con las fórmulas de la norma DG-2018.
5	Ingresar condiciones de peligro §03	Seleccionar el nivel de *Peligro por Visibilidad*, indicar si existe *Tramo Tangente en S*, *Tramo Tangente en O*, ingresar la *Longitud de Tangente (LT en m)* y el nivel de *Peligro en Pavimento* según evaluación de campo
6	Ejecutar predicción	Hacer clic en **ANALIZAR TRAMO VIAL - EJECUTAR RNA** y el resultado las probabilidades P(Riesgo Bajo), P(Riesgo Medio) y P(Riesgo Alto)
7	Interpretar resultado	La clase con mayor probabilidad determina el nivel de riesgo.
8	Evaluar nuevo tramo	Hacer clic en **Limpiar** (↺) y repetir desde el paso 3

Fuente: Elaboración propia.

- **Paso 1: Abrir el programa**

Figura 191 Paso 1: Abrir el programa
Fuente: Elaboración propia

- **Paso 2, 3, 4, 5: Ingresar variables**

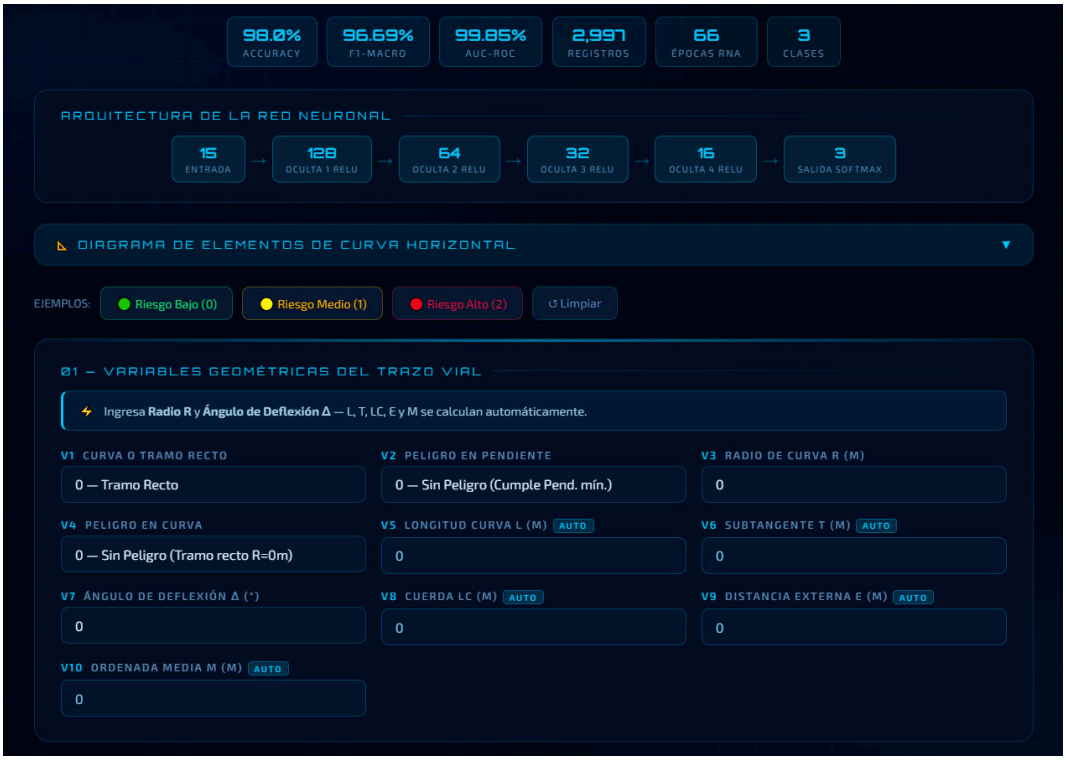


Figura 192 Paso 2, 3, 4: Ingresar variables geométricas.
Fuente: Elaboración propia

Figura 193 Paso 5: Ingresar variables de condición de peligro.
Fuente: Elaboración propia

- **Paso 6, 7 y 8 : Ejecutar, interpretar y evaluar nuevo tramo**

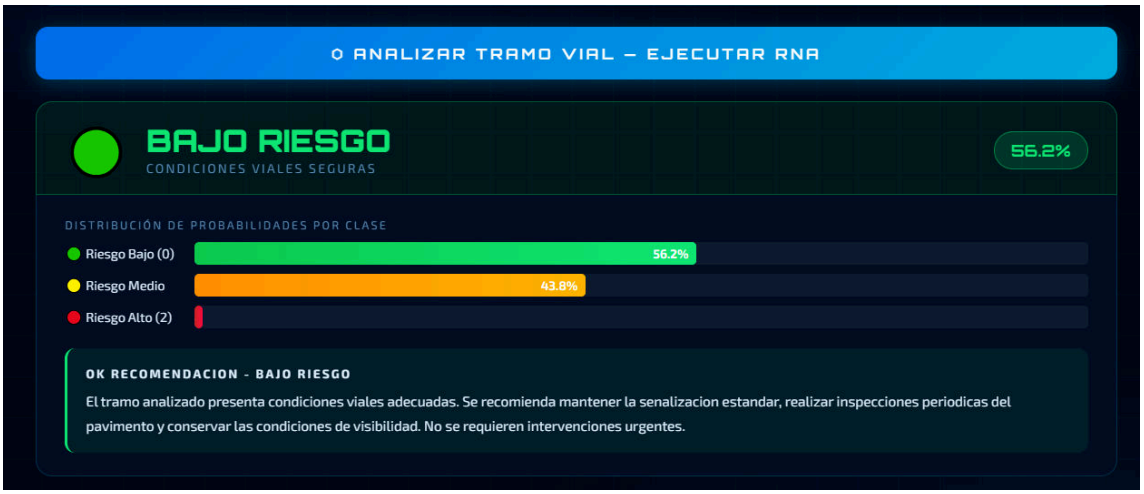


Figura 194 Ejecutar e interpretar en caso Riesgo bajo.
Fuente: Elaboración propia

Figura 195 Ejecutar e interpretar en caso Riesgo Medio.
Fuente: Elaboración propia

Figura 196 Ejecutar e interpretar en caso Riesgo Alto.
Fuente: Elaboración propia

3.11.3 Análisis normativo comparativo

Se examinaron sistemáticamente cinco normativas internacionales de diseño geométrico de carreteras N100 de Noruega, Veghönnunarreglur de Islandia, VGU de Suecia, DMRB del Reino Unido y la Road Structure Ordinance (RSO) de Japón en comparación con la DG-2018 del Perú, con la finalidad de identificar brechas técnicas y parámetros susceptibles de mejora para la normativa peruana vigente, aplicable a la Vía Los Libertadores (PE-28A).

3.11.3.1 Normativas internacionales analizadas

Las cinco normativas internacionales seleccionadas representan estándares de países con índices de seguridad vial reconocidos mundialmente. La selección se priorizando países con condiciones topográficas o climáticas adversas comparables a las peruanas con bajas tasas de accidentes de tránsito. En conjunto, la siguiente tabla se muestra las normas analisadas.

Tabla 65 *Normativas internacionales de diseño geométrico analizadas*

País	Código	Nombre completo	Entidad emisora	Año	Idioma
Perú	DG-2018	Manual de Carreteras: Diseño Geométrico DG-2018	MTC	2018	Español
Noruega	N100	Håndbok N100 Veg- og gateutforming	Statens vegvesen	2019	Noruego
Islandia	Veghönnunarreglur	Veghönnunarreglur	Vegagerðin (IRCA)	2010	Islandés
Suecia	VGU / TRVINFRA	Vägars och gators utformning	Trafikverket	2024	Sueco
Reino Unido	DMRB CD 109	Design Manual for Roads and Bridges	National Highways	2020	Inglés
Japón	RSO	(*Dōro Kōzōrei*)	MLIT	2021	Japonés

Fuente: Elaboración propia con base en los documentos oficiales de cada país.

3.11.3.2 Noruega (Veg og gateutforming - Håndbok N100) vs. Perú (DG-2018)

El sistema noruego de diseño geométrico se fundamenta en el *Håndbok N100 Veg- og gateutforming* (Manual N100 de Diseño de Carreteras y Calles), publicado la Administración Pública de Carreteras de Noruega en el 2019 (Statens vegvesen, 2019).

El procedimiento de diseño noruego se desarrolla en tres etapas secuenciales, que difieren conceptualmente del enfoque peruano:

- **Selección de la clase de dimensionamiento.** Se elige la la categoría de diseño entre 7 opciones disponibles H1, H5, H3, Hø1, Hø2, L1 y L2 definidas simultáneamente por tres criterios: función de red (nacional principal, otras principales o local), IMDA (Índice Medio Diario Anual) y velocidad (Statens vegvesen, 2019).
- **Aplicación de la tabla de proyección unificada.** Cada clase de dimensionamiento dispone de su propia tabla de proyección, que concentra todos los parámetros geométricos necesarios en un único documento: radio horizontal mínimo, parámetro de clotoide, distancias de visibilidad de parada y adelantamiento, radios verticales convexos y cóncavos, peralte máximo y pendiente máxima. Este enfoque es radicalmente distinto al del DG-2018, donde cada parámetro se obtiene de tablas separadas e independientes

(Statens vegvesen, 2019).

- **Aplicación de la filosofía Nullvisjonen.** Noruega adoptó en el año 2000 la *Nullvisjonen* (Visión Cero), que establece que ninguna muerte ni lesión grave en el tráfico es aceptable. Esta filosofía se traduce en requisitos físicos concretos: marcación central reforzada en la clase H1 (80 km/h), barrera mediana de 1,5m en la clase H5 (90 km/h) y separador con barrera de 2,0m en la clase H3 (110 km/h). El DG-2018 solo exige barrera mediana en autopistas de primera clase (Statens vegvesen, 2019).

Tabla 66 *Comparación de parámetros geométricos: Perú DG-2018 vs. Noruega N100*

Parámetro	Perú DG-2018	Noruega N100	Diferencia clave
A. Velocidad y clasificación			
Velocidad de diseño	30–130 km/h (Tabla 204.01)	60–110 km/h (Tabell C.3)	Perú cubre vías rurales de baja velocidad; Noruega no
B. Alineamiento horizontal			
Peralte máximo	12 % en zona rural tipo 4 (Tabla 302.02)	8,0 % en todas las clases (Tablas C.4–C.9)	Perú 50 % más alto; permite radios menores en terreno andino
Radio mínimo a V=80 km/h	175 m con e=12 % (Tabla 302.02)	250 m con e=8 % — clase H1 (Tabla C.4)	Noruega exige radio 43 % mayor; mayor seguridad inherente
Clotoide de transición	Obligatoria, J=0,5 m/s^3 (Tabla 302.10)	Obligatoria en todas las clases (Tablas C.4–C.9)	Ambos exigen clotoide; Noruega no fija el jerk explícitamente
Curvas de vuelta	Sí — diseño específico (Sección 302.07)	No existen en la norma	Solo Perú las contempla, por topografía andina escarpada
Vías 2+1	No contempladas	Sí — clase H5 con *forbikjøringsfelt*	Noruega incluye carril central alternante para adelantamiento seguro
C. Alineamiento vertical			
Pendiente máx. en autopista	3–5 % (Tabla 303.01)	5 % — clases H3/H5 (Tablas C.5/C.6)	Valores similares; Perú permite 3 % en autopistas de mayor jerarquía
Pendiente máx. en vía local	8–10 % (Tabla 303.01)	8 % — clases Hø1/L1 (Tablas C.7/C.10)	Perú permite hasta 10 % en terreno escarpado
D. Sección transversal			

Continúa en la siguiente página

(Continuación de la Tabla 66)

Parámetro	Perú DG-2018	Noruega N100	Diferencia clave
Ancho de carril en autopista	3,60 m (Tabla 304.01)	3,50 m — clase H3 (Tabla C.3)	Diferencia de 10 cm; ambas normas dentro del estándar internacional
Berma en autopista	3,00 m (Tabla 304.02)	2,75 m — clase H3 (Tabla C.3)	Perú exige berma 9 % más ancha
Bombeo en recta	2,0–2,5 % (Sección 304.05)	3,0 % (Sección C.2)	Noruega requiere mayor bombeo para drenaje en clima lluvioso-nevado
E. Seguridad vial			
Barrera mediana obligatoria	Solo en autopistas de 1ra clase	H5 (≥90 km/h): 1,5 m H3 (≥110 km/h): 2,0 m	Noruega protege físicamente el tráfico opuesto en más categorías de vía
DVP a V=80 km/h	136 m con t=2,5 s (Tabla 205.01-A)	115 m — clase H1 (Tabla C.4)	Perú 18 % más conservador; curvas convexas más largas y costosas
F. Vehículos de diseño			
N.° de vehículos tipo	10 a más (Sección 202)	6 tipos: P, LL, L, BUS, VT, MVT (Figura E.7)	Perú detalla mayor variedad vehicular; Noruega simplifica con 6 tipos representativos

Nota. DVP = Distancia de Visibilidad de Parada; e = peralte; J = jerk (variación de aceleración centrífuga).

Fuente: Elaboración propia.

3.11.3.3 Islandia (Veghönnunarreglur) vs. Perú (DG-2018)

Las *Veghönnunarreglur* (Reglas de Diseño de Carreteras de Islandia) son publicadas por Vegagerðin la Administración de Carreteras y Costas de Islandia (IRCA) y comprenden cuatro capítulos: fundamentos, sección transversal (actualizado 2011), alineamiento (2010) e intersecciones (Vegagerðin, 2010, 2011).

El procedimiento de diseño islandés se desarrolla en tres etapas secuenciales:

- **Selección del tipo de vía.** Se elige la *vegtegund* (tipo de vía) entre 10 opciones organizadas en 4 categorías: Tipo A (autopistas de 4 carriles: A34 y A22), Tipo B (vías divididas: B19, B15.5, B12), Tipo C (vías de 2 carriles: C10, C9, C8, C7) y Tipo D (vía de 1 carril: D4). Cada tipo determina el ancho de carril, berma, separador central y el rango de velocidad de diseño admisible (Vegagerðin, 2011).
- **Peralte variable con la velocidad.** La característica más distintiva del sistema islandés

es que el peralte máximo no es un valor fijo, sino que disminuye progresivamente conforme aumenta la velocidad de diseño: 8,0 % para $V \leq 60$ km/h, 7,5 % para V=70–80 km/h, 7,0 % para V=90 km/h, 6,5 % para V=100 km/h, 6,0 % para V=110 km/h, 5,5 % para V=120 km/h y 5,0 % para V=130 km/h (Vegagerðin, 2010). Ningún otro país estudiado aplica este enfoque progresivo.

- **Elementos especiales para terreno escarpado.** Islandia contempla las *lykkjur* (bucles), curvas con radio $R_L \leq 40$ m y cambio de dirección de hasta 200 gon (180°), diseñadas para superar desniveles pronunciados en terreno montañoso volcánico. Utilizan clotoides múltiples de hasta 3 segmentos y limitan el peralte a 3–4 % y la pendiente compuesta a 6 % en el centro del bucle (Vegagerðin, 2010).

Tabla 67 *Comparación de parámetros geométricos: Perú DG-2018 vs. Islandia Veghönnunarreglur*

Parámetro	Perú DG-2018	Islandia Veghönnunarreglur	Diferencia clave
A. Velocidad y clasificación			
Velocidad de diseño	30–130 km/h (Tabla 204.01)	30–130 km/h, 11 escalones (Tafla 3.1-1)	Rangos similares; ambos cubren desde vías locales hasta autopistas
Clasificación de vías	Demanda × orografía (5 categorías × 4 terrenos)	10 tipos de vía en 4 grupos A, B, C, D (Tafla 2.1.1-1)	Islandia clasifica por sección transversal; Perú por demanda y terreno
B. Alineamiento horizontal			
Peralte máximo	12 % fijo en zona rural tipo 4 (Tabla 302.02)	Variable: 8,0 % (V≤60) → 5,0 % (V=130) (Tafla 3.5.2-1)	Islandia es el único país que reduce el peralte progresivamente con la velocidad
Radio mínimo a V=80 km/h	175 m con e=12 % (Tabla 302.02)	234 m con e=7,5 % (Tafla 3.3.2-1)	Islandia exige radio 34 % mayor a igual velocidad
Clotoide mínima a V=80 km/h	$A_{mín}$≈80 (Tabla 302.10)	$A_{mín}$=120 (Tafla 3.3.4-1)	Islandia exige transición 50 % más larga; mayor comodidad dinámica
Longitud máx. de recta	$L_{máx}$=1 336 m a V=80 (Tabla 302.01)	$L_{máx}$=1 600 m a V=80 (Tafla 3.3.1-1)	Islandia permite rectas algo más largas antes de requerir una curva

Continúa en la siguiente página

(Continuación de la Tabla 67)

Parámetro	Perú DG-2018	Islandia Veghönnunarreglur	Diferencia clave
Curvas en terreno escarpado	Curvas de vuelta (Sección 302.07)	*Lykkjur* (bucles) $R \leq 40$ m, hasta 200 gon (Sección 3.3.6)	Ambos tienen solución para terreno escarpado; geometría y restricciones distintas
C. Alineamiento vertical			
DVP a V=80 km/h	136 m con t=2,5 s (Tabla 205.01-A)	119 m (Tafla 3.1-1)	Perú 14 % más conservador; curvas convexas más largas
DVP a V=100 km/h	194 m (Tabla 205.01-A)	179 m (Tafla 3.1-1)	Perú 8 % más conservador a velocidades mayores
Pendiente máx. vía principal	7–9 % (Tabla 303.01)	7–10 % tipos B–C (Tafla 3.4.1-1)	Valores similares; Islandia permite hasta 10 % en vías tipo C
Pendiente máx. vía local	8–10 % (Tabla 303.01)	9–12 % tipo D (Tafla 3.4.1-1)	Islandia permite mayor pendiente en vías de 1 carril por topografía volcánica
Rampas de emergencia	No especificadas en la norma	Obligatorias en $i \geq 6$ % con longitud > 3 km (Sección 3.4.3)	Islandia exige rampas de escape en bajadas largas y pronunciadas
D. Sección transversal			
Ancho de carril	3,00–3,60 m (Tabla 304.01)	3,00–3,75 m (Tafla 2.1.1-1)	Rangos similares; Islandia permite carril de hasta 3,75 m en autopistas
Bombeo en recta	2,0–2,5 % (Sección 304.05)	3,0–4,0 % (Sección 3.5.1)	Islandia exige mayor bombeo por necesidades de drenaje en clima subártico

Nota. DVP = Distancia de Visibilidad de Parada; e = peralte; $A_{mín}$ = parámetro mínimo de clotoide.

Fuente: Elaboración propia.

3.11.3.4 Suecia (Vägars och gators utformning - VGU) vs. Perú (DG-2018)

El sistema normativo sueco de diseño geométrico, denominado VGU (*Vägars och gators utformning*, Diseño de Carreteras y Calles), es publicado por la Administración de Transporte de Suecia, sustentado en la filosofía *Vision Zero*. El sistema se compone de tres documentos complementarios: el *Handbok*, el TRVINFRA.

El procedimiento de diseño sueco se desarrolla en tres etapas secuenciales:

- **Selección de la velocidad de referencia.** El concepto central del VGU es la *Referenshastighet* (V_R), velocidad de referencia para la cual se diseña el tramo. El VGU selecciona V_R según el tipo de vía y su función dentro de la red de transporte (Trafikverket, 2024b).
- **Sistema de doble nivel: *riktvärde* y *gränsvärde*.** Cada parámetro geométrico tiene dos valores diferenciados: el *riktvärde* (valor guía), que aplica para construcción nueva y representa la solución técnicamente óptima; y el *gränsvärde* (valor límite), que aplica para remodelaciones de vías existentes donde las restricciones topográficas o económicas impiden alcanzar el valor guía (Ministerio de Transportes y Comunicaciones del Perú, 2018).
- **Fórmulas explícitas de fricción.** Suecia es el único país entre los seis analizados que publica las ecuaciones continuas de fricción empleadas en el cálculo de los parámetros geométricos: fricción total $f = 0{,}62 \times e^{-0{,}0063V}$, fricción lateral $f_s = 0{,}28 \times e^{-0{,}0096V}$ y fricción de frenado $f_b = 0{,}55 \times e^{-0{,}0057V}$ (Trafikverket, 2024a).

Tabla 68 *Comparación de parámetros geométricos: Perú DG-2018 vs. Suecia VGU*

Parámetro	Perú DG-2018	Suecia VGU	Diferencia clave
A. Velocidad y clasificación			
Velocidad de diseño	30–130 km/h (Tabla 204.01)	30–120 km/h, 10 escalones (*Referenshastighet* V_R)	Suecia no cubre velocidades menores a 30 km/h ni trochas carrozables
Sistema de flexibilidad	Excepción única: +1 % en pendiente (Tabla 303.01)	*Riktvärde* (obra nueva) *Gränsvärde* (remodelación) (TRVINFRA-00396)	Suecia diferencia formalmente entre obra nueva y remodelación; mayor trazabilidad
B. Alineamiento horizontal			
Peralte máximo	12 % en zona rural tipo 4 (Tabla 302.02)	5,5 % general (exc. 7 %) (K236300)	Perú más del doble; permite radios mucho menores en terreno andino
Radio mínimo a V=80 km/h	175 m con e=12 % (Tabla 302.02)	400 m *riktvärde* 300 m *gränsvärde* (Tabla 8.1.6.2.2-1)	Suecia exige radio 2,3× mayor; adoptar este valor duplicaría el costo en terreno andino

Continúa en la siguiente página

(Continuación de la Tabla 68)

Parámetro	Perú DG-2018	Suecia VGU	Diferencia clave
Radio mínimo a V=100 km/h	305 m con e=12 % (Tabla 302.02)	700 m *riktvärde* 550 m *gränsvärde* (Tabla 8.1.6.2.2-1)	Diferencia de 2,3×; refleja filosofía *Vision Zero* de seguridad inherente
Fórmulas de fricción	Tablas $f_{máx}$ discretas (Tabla 302.02)	$f = 0{,}62\, e^{-0{,}0063V}$ $f_s = 0{,}28\, e^{-0{,}0096V}$ $f_b = 0{,}55\, e^{-0{,}0057V}$ (Grundvärden 3.1.2–3.1.4)	Único país con fórmulas continuas; permite verificación y actualización independiente
Jerk de clotoide	J=0,5 m/s^3 fijo (Tabla 302.10)	$k = e^{-0{,}000015V^3} + 0{,}45$ variable con V (TRVINFRA-00396)	Suecia adapta el jerk a la velocidad; Perú usa valor fijo conservador
C. Alineamiento vertical			
DVP a V=80 km/h	136 m con t=2,5 s (Tabla 205.01-A)	∼120 m con t=2,0 s h_2=0,35 m (Grundvärden 3.3.9)	Perú más conservador en t; Suecia compensa con h_2 mayor
Pendiente máxima	6–10 % (Tabla 303.01)	6 % *riktvärde* 8 % *gränsvärde* (Tabla 8.1.6.9-1)	Valores similares; Suecia diferencia entre obra nueva y remodelación
D. Sección transversal			
Ancho de carril autopista	3,60 m (Tabla 304.01)	3,50 m mínimo (K233026)	Diferencia de 10 cm; ambas dentro del estándar internacional
Bombeo en recta	2,0–2,5 % (Sección 304.05)	2,5 % (TRVINFRA-00396)	Valores similares; Suecia fija un valor único sin rango
E. Seguridad vial			
Barrera mediana	Solo autopistas de 1ra clase	Obligatoria en *motorväg* y vías de alta velocidad	Suecia exige separación física del tráfico opuesto en más categorías
Filosofía de seguridad	No declarada explícitamente en la norma	*Vision Zero*: cero muertes aceptables (desde 1997)	Suecia integra la seguridad como objetivo normativo explícito
F. Vehículos de diseño			

Continúa en la siguiente página

(Continuación de la Tabla 68)

Parámetro	Perú DG-2018	Suecia VGU	Diferencia clave
N.° de vehículos tipo	10 a más (Sección 202)	18 *typfordon* incluye bicicleta de carga (Grundvärden 3.2)	Suecia es el país con mayor diversidad vehicular; incluye movilidad activa
Consistencia V_{85}	Sí — ecuaciones Fitzpatrick (Tabla 204.03)	No contemplada en la norma	Perú es el único país con verificación de consistencia V_{85} en su norma

Nota. DVP = Distancia de Visibilidad de Parada; e = peralte; f = fricción total; f_s = fricción lateral; f_b = fricción de frenado; V_R = velocidad de referencia; J = jerk.

Fuente: Elaboración propia.

3.11.3.5 Reino Unido (Design Manual for Roads and Bridges - DMRB) vs. Perú (DG-2018)

El *Design Manual for Roads and Bridges* (DMRB) publicado por National Highways la agencia gubernamental responsable de la red vial estratégica del Reino Unido. Los documentos clave para el diseño geométrico son el CD 109 *Highway Link Design* (National Highways, 2021).

El procedimiento de diseño británico se desarrolla en tres etapas secuenciales:

- **Cálculo de la velocidad de diseño.** La diferencia fundamental del DMRB respecto a todos los demás países analizados es que la velocidad de diseño no se asigna directamente, sino que se *calcula* a partir de dos parámetros medidos en el corredor vial: el *Alignment Constraint* (A_c) y el *Layout Constraint* (L_c). Para calzadas simples: $A_c = 12 - \text{VISI}/60 + 2B/45$, donde VISI es la visibilidad media armónica del tramo y B la sinuosidad en grados/km. Los valores de A_c y L_c se cruzan en la Figure 2.1 del CD 109 para obtener la velocidad de diseño en bandas: 120A/B, 100A/B, 85A/B, 70A/B, 60A/B y 50A/B kph (National Highways, 2020). Este enfoque elimina la arbitrariedad en la selección de velocidad y adapta el diseño a las características reales del corredor.
- **Aplicación de la tabla maestra CD 109 (Table 2.10.)** Todos los parámetros geométricos del sistema británico se concentran en una única tabla maestra la Table 2.10 que contiene, para cada banda de velocidad: distancia de visibilidad de parada (DVP), radios mínimos para seis niveles de peralte (camber adverso, 2,5 %, 3,5 %, 5 %, 7 %), valores K para curvas verticales convexas y cóncavas, y la distancia (National High-

ways, 2020).

- **Sistema formal de *relaxations*.** Cuando las condiciones topográficas o económicas impiden cumplir los valores *desirable minimum* de la Table 2.10, el DMRB permite descensos escalonados . Cada *relaxation* equivale a descender la velocidad de diseño (National Highways, 2020).

Tabla 69 *Comparación de parámetros geométricos: Perú DG-2018 vs. Reino Unido DMRB*

Parámetro	Perú DG-2018	Reino Unido DMRB	Diferencia clave
A. Velocidad y clasificación			
Método de asignación de V diseño	Tabla demanda × orografía (Tabla 204.01)	Cálculo: $A_c + L_c \rightarrow$ Figure 2.1 en bandas A/B (CD 109, p. 9–10)	UK es el único país donde la velocidad resulta de un cálculo, no de una asignación directa
Rango de velocidad de diseño	20–130 km/h (Tabla 204.01)	50–120 kph en 6 bandas A/B (CD 109, Figure 2.1)	UK no cubre vías rurales de baja velocidad ni trochas
Sistema de flexibilidad	Excepción única: +1 % en pendiente (Tabla 303.01)	*Relaxations* escalonadas: hasta 2 *design speed steps* (CD 109, Notes T. 2.10)	UK formaliza cada excepción con criterios específicos por tipo de vía; mayor trazabilidad
B. Alineamiento horizontal			
Peralte máximo	12 % en zona rural tipo 4 (Tabla 302.02)	7 % en todas las vías (CD 109, Table 2.10)	Perú 71 % más alto; UK prioriza seguridad sobre economía de construcción
Radio mínimo a V=85 kph	175 m con e=12 % (Tabla 302.02)	510 m con e=5 % 360 m con e=7 % (CD 109, Table 2.10)	UK exige radio hasta 2,9× mayor; la diferencia de peralte explica la mayor parte
Vías 2+1	No contempladas en la norma	Sí — WS2+1 secciones 800–1 500 m (CD 109, Sección 6)	UK tiene el diseño 2+1 mejor documentado; reduce accidentes frontales sin duplicar calzada
Curva de transición	Clotoide, J=0,5 m/s^3 fijo (Tabla 302.10)	Espiral de transición q=0,3–0,6 m/s^3 (CD 109, Sec. 4.12–4.13)	UK permite rango de jerk según condiciones; Perú fija un valor único
C. Alineamiento vertical			

Continúa en la siguiente página

(Continuación de la Tabla 69)

Parámetro	Perú DG-2018	Reino Unido DMRB	Diferencia clave
DVP a *V*=85 kph	136 m con *t*=2,5 s (Tabla 205.01-A)	160 m *desirable* (CD 109, Table 2.10)	UK 18 % más conservador en DVP; genera curvas convexas más largas
K convexo a *V*=80–85 km/h	*K*=26 (Tabla 303.02)	*K*=55 *desirable* (CD 109, Table 2.10)	UK exige curvas de cresta más del doble de suaves; mejora visibilidad pero aumenta costo
Pendiente máx. en autopista	3–5 % (Tabla 303.01)	3 % *desirable* 4 % *relaxation* (CD 109, Table 5.1)	UK el más restrictivo en autopistas; refleja topografía relativamente plana
Pendiente máx. vía simple	6–9 % (Tabla 303.01)	6 % *desirable* 8 % *relaxation* (CD 109, Table 5.1)	Valores similares; UK diferencia *desirable* de *relaxation*
D. Sección transversal			
Ancho de carril autopista	3,60 m (Tabla 304.01)	3,65 m (CD 127, Fig. 2.1.1N1)	Diferencia mínima de 5 cm; ambas dentro del estándar internacional
Berma / *hard shoulder*	Berma 3,00 m (Tabla 304.02)	*Hard shoulder* 3,30 m obligatorio en autopistas (CD 127)	UK denomina *hard shoulder* a la berma pavimentada de emergencia; concepto ausente en DG-2018
Bombeo en recta	2,0–2,5 % (Sección 304.05)	2,5 % (CD 109)	Valores similares; UK fija un único valor sin rango
Gálibo vertical	5,50 m (Sección 304.07)	5,30 m + margen *S* (CD 127)	Perú exige el gálibo más alto entre los seis países; refleja mayor altura vehicular permitida
E. Seguridad vial			
Barrera mediana	Solo autopistas de 1ra clase	Obligatoria en *motorways* y vías de alta velocidad (CD 109)	UK separa físicamente el tráfico opuesto en más categorías de vía
F. Vehículos de diseño			

Continúa en la siguiente página

(Continuación de la Tabla 69)

Parámetro	Perú DG-2018	Reino Unido DMRB	Diferencia clave
Consistencia V_{85}	Sí — ecuaciones Fitzpatrick (Tabla 204.03)	No contemplada en el CD 109	Perú es el único país con verificación de consistencia V_{85} incorporada en su norma

Nota. DVP = Distancia de Visibilidad de Parada; e = peralte; K = índice de curvatura vertical; A_c = *Alignment Constraint*; L_c = *Layout Constraint*; FOSD = *Full Overtaking Sight Distance*; *Desirable* = valor óptimo de diseño; *Relaxation* = valor mínimo aceptable por restricción justificada.

Fuente: Elaboración propia.

3.11.3.6 Japón (Road Structure Ordinance - RSO) vs. Perú (DG-2018)

La *Road Structure Ordinance* (RSO, *Dōro Kōzōrei*) es un decreto gubernamental publicado por el MLIT (*Ministry of Land, Infrastructure, Transport and Tourism*) (Ministry of Land, Infrastructure, Transport and Tourism, 2021). A diferencia de las demás normativas analizadas que son manuales técnicos la RSO tiene rango de ley, lo que significa que ningún proyecto vial en Japón puede desviarse de sus disposiciones sin una modificación legislativa formal. El procedimiento de diseño japonés se desarrolla en dos etapas secuenciales:

- **Clasificación por Tipo y Clase.** La RSO clasifica las carreteras en 4 Tipos según su función y entorno: Tipo 1 (rural con control total de acceso), Tipo 2 (urbano con control de acceso), Tipo 3 (rural sin control de acceso) y Tipo 4 (urbano sin control de acceso). Cada Tipo se subdivide a su vez en Clases según el volumen de tráfico de diseño y las condiciones topográficas: Tipo 1 con Clases 1–4, Tipo 2 con Clases 1–2, Tipo 3 con Clases 1–5 y Tipo 4 con Clases 1–4 (Ministry of Land, Infrastructure, Transport and Tourism, 2021).
- **Asignación de parámetros por artículo de ley.** A diferencia de los demás países, donde los parámetros se ubican en tablas de manuales técnicos, en Japón cada parámetro geométrico está directamente establecido en un artículo específico del decreto ley. Por ejemplo, los radios mínimos en el Artículo 15, las pendientes en el Artículo 20, las distancias de visibilidad en el Artículo 19 y el peralte en el Artículo 16 (Ministry of Land, Infrastructure, Transport and Tourism, 2021). Esto otorga a la norma japonesa una estabilidad y obligatoriedad que ningún otro país puede igualar, pero también una menor flexibilidad para adaptarse rápidamente a nuevas investigaciones técnicas.

Tabla 70 *Comparación de parámetros geométricos: Perú DG-2018 vs. Japón RSO*

Parámetro	Perú DG-2018	Japón RSO	Diferencia clave
A. Velocidad y clasificación			
Clasificación de vías	Demanda × orografía (5 categorías × 4 terrenos) (Tabla 204.01)	4 Tipos × múltiples Clases por volumen de tráfico (Art. 3)	Japón usa tráfico y entorno; Perú usa demanda y orografía; ambos son los más comparables entre sí
Rango de velocidad de diseño	30–130 km/h (Tabla 204.01)	20–120 km/h (Art. 13)	Rangos casi idénticos; ambos cubren desde trochas hasta autopistas de alta velocidad
Sistema de flexibilidad	Excepción única: +1 % en pendiente (Tabla 303.01)	Valores estándar y valores especiales por artículo de ley	Japón diferencia formalmente entre condición normal y condición topográfica restrictiva
B. Alineamiento horizontal			
Peralte máximo	12 % en zona rural tipo 4 (Tabla 302.02)	10 % Tipos 1–3 8 % zona de nieve severa 6 % Tipo 4 urbano (Art. 16)	Japón es el país más próximo a Perú en peralte; ambos enfrentan topografía montañosa de alta sismicidad
Radio mínimo a V=80 km/h	175 m con e=12 % (Tabla 302.02)	280 m estándar 230 m especial (Art. 15)	Japón exige radio 60 % mayor en condición estándar; diferencia explicada por menor peralte
Radio mínimo a V=60 km/h	105 m con e=12 % (Tabla 302.02)	150 m estándar 120 m especial (Art. 15)	Patrón consistente: Japón siempre exige radios mayores a igual velocidad
C. Alineamiento vertical			
DVP a V=80 km/h	136 m con t=2,5 s h_1=1,07 m, h_2=0,15 m (Tabla 205.01-A)	110 m h_1=1,20 m, h_2=0,10 m (Art. 19 y Art. 2.24)	Perú 24 % más conservador; Japón compensa con h_1 más alta que reduce la DVP necesaria
DVP a V=100 km/h	194 m (Tabla 205.01-A)	160 m (Art. 19)	Perú 21 % más conservador; diferencia se mantiene constante con la velocidad

Continúa en la siguiente página

(Continuación de la Tabla 70)

Parámetro	Perú DG-2018	Japón RSO	Diferencia clave
R_v convexa a V=80 km/h	K=26 ≈ R_v=3 380 m (Tabla 303.02)	R_v=3 000 m (Art. 22)	Valores similares; Perú ligeramente más exigente en curvas convexas
R_v cóncava a V=80 km/h	K=6 ≈ R_v=780 m (Tabla 303.02)	R_v=2 000 m (Art. 22)	Japón exige radio cóncavo 2,6× mayor; mayor comodidad en fondos de valle
Pendiente máx. a V=80 km/h	5–7 % (Tabla 303.01)	4 % estándar 7 % especial (Art. 20)	Japón más restrictivo en condición estándar; valor especial coincide con Perú
Pendiente máx. a V=20 km/h	8–10 % (Tabla 303.01)	9 % estándar 12 % especial (Art. 20)	Japón permite hasta 12 % en vías locales de baja velocidad; similar a Perú
D. Sección transversal			
Ancho de carril autopista	3,60 m (Tabla 304.01)	3,50 m — Tipo 1 Clase 1–2 (Art. 5)	Diferencia de 10 cm; ambas dentro del estándar internacional
Gálibo vertical	5,50 m (Sección 304.07)	4,50 m general 4,80 m rutas logísticas (Art. 12)	Perú exige gálibo 22 % mayor; refleja mayor altura de vehículos de carga en Perú
Carril de ascenso	Cuando reducción $V \geq$ 25 km/h (Sección 303.05)	En pendiente >5 % (>3 % si $V \geq$ 100 km/h) (Art. 20)	Japón usa criterio de pendiente; Perú usa criterio de pérdida de velocidad
E. Vehículos de diseño			
N.° de vehículos tipo	10 a más (Sección 202)	4 tipos: motor menor, pequeño, regular y semirremolque (Art. 4)	Japón es el más parsimonioso; 4 tipos cubren todo el espectro vehicular japonés
Altura de ojo del conductor h_1	1,07 m (Sección 205.01)	1,20 m (Art. 2.24)	Japón usa la h_1 más alta entre los seis países; reduce la DVP requerida

Continúa en la siguiente página

(Continuación de la Tabla 70)

Parámetro	Perú DG-2018	Japón RSO	Diferencia clave
Altura de objeto h_2	0,15 m (Sección 205.01)	0,10 m (Art. 2.24)	Japón usa el objeto más bajo; combinado con h_1 alta genera DVP más cortas que Perú
Consistencia V_{85}	Sí — ecuaciones Fitzpatrick (Tabla 204.03)	No contemplada en la RSO	Perú es el único país con verificación de consistencia V_{85} en su norma

Nota. e = peralte; R_v = radio de curvatura vertical; K = índice de curvatura vertical; h_1 = altura del ojo del conductor; h_2 = altura del objeto de parada. Los valores *estándar* aplican en condiciones normales; los valores *especiales* aplican cuando restricciones topográficas impiden cumplir el estándar.

Fuente: Elaboración propia.

3.11.3.7 Síntesis comparativa con respecto al DG-2018

La síntesis comparativa de las normativas internacionales, permitiendo identificar en que parámetros el DG-2018 peruano se encuentra por encima, dentro o por debajo del rango internacional. La Tabla siguiente presenta la posición relativa del DG-2018 frente a las cinco normativas analizadas concentra los valores numéricos clave para los parámetros de mayor incidencia en la siniestralidad de la PE-28A.

Tabla 71 *Posición relativa del DG-2018 frente a las normativas internacionales analizadas*

Parámetro	Noruega N100	Islandia	Suecia VGU	Reino Unido DMRB	Japón RSO	Posición de Perú
A. Alineamiento horizontal						
Peralte máximo	8,0 %	8,0 → 5,0 %	5,5 %	7,0 %	10,0 %	**MÁS PERMISIVO** que todos (12 % vs. 5,5–10 %)
$R_{mín}$ a V=80 km/h	250 m	234 m	400 m	360 m	280 m	**MÁS PERMISIVO** que todos (175 m vs. 234–400 m)
Clotoide $A_{mín}$ a V=80 km/h	125 m	120 m	190 m	—	—	**MÁS PERMISIVO** (∼80 m vs. 120–190 m)

Continúa en la siguiente página

(Continuación de la Tabla 71)

Parámetro	Noruega N100	Islandia	Suecia VGU	Reino Unido DMRB	Japón RSO	Posición de Perú
Peralte variable con *V*	No	**Sí**	No	No	No	**AUSENTE** en DG-2018 solo Islandia lo contempla
Vías 2+1	**Sí**	No	**Sí**	**Sí**	No	**AUSENTE** en DG-2018 presente en 3 de 5 países
B. Alineamiento vertical						
DVP a *V*=80 km/h	115 m	119 m	~120 m	160 m	110 m	**MÁS CONSERVADOR** (136 m; solo UK supera a Perú)
Pendiente máx. autopista	5 %	4 %	6 %	3 %	5 %	**SIMILAR** al rango internacional (3–5 %)
Pendiente máx. local	8 %	12 %	8 %	8 %	12 %	**MÁS PERMISIVO** que Noruega, Suecia y UK **SIMILAR** a Islandia y Japón
K convexo a *V*=80 km/h	~2 800 m	3 000 m	3 100 m	*K*=55	3 000 m	**MÁS PERMISIVO** (*K*=26 ≈ 3 380 m; valor similar pero sistema distinto)
C. Sección transversal						
Ancho de carril autopista	3,50 m	3,75 m	3,50 m	3,65 m	3,50 m	**DENTRO DEL RANGO** (3,60 m; consenso internacional)
Berma autopista	2,75 m	3,00 m	1,50 m	3,30 m	2,50 m	**DENTRO DEL RANGO** (3,00 m; comparable con Islandia y UK)

Continúa en la siguiente página

(Continuación de la Tabla 71)

Parámetro	Noruega N100	Islandia	Suecia VGU	Reino Unido DMRB	Japón RSO	Posición de Perú
Bombeo en recta	3,0 %	3,0–4,0 %	2,5 %	2,5 %	1,5–2,0 %	**MÁS PERMISIVO** (2,0–2,5 %; el más bajo junto con Japón)
Gálibo vertical	4,70–4,90 m	4,90 m	4,70 m	5,30 m	4,50–4,80 m	**MÁS CONSERVADOR** (5,50 m; el más alto de todos)
D. Seguridad vial y metodología						
Barrera mediana obligatoria	H5+H3	—	Motorväg	Motorway	Tipos 1–2	**MÁS PERMISIVO** solo en autopistas de 1ra clase
Fórmulas explícitas de fricción	No	No	**Sí**	No	No	**AUSENTE** en DG-2018 solo Suecia las publica
Sistema formal de flexibilidad	*Fravik*	—	*Riktvärde / gränsvärde*	*Relaxations*	Estándar / especial	**MÁS BÁSICO** solo +1 % pendiente sin procedimiento formal
Consistencia V_{85}	No	No	No	No	No	**FORTALEZA EXCLUSIVA** único país con ecuaciones V_{85} en la norma
Rampas de emergencia	—	**Sí**	—	—	—	**AUSENTE** en DG-2018 solo Islandia las exige explícitamente

Nota. (−) indica que el parámetro no está definido explícitamente en esa norma para las condiciones equivalentes a la PE-28A. DVP = Distancia de Visibilidad de Parada; $R_{mín}$ = radio mínimo horizontal; K = índice de curvatura vertical; $A_{mín}$ = parámetro mínimo de clotoide.

Fuente: Elaboración propia con base en (Ministerio de Transportes y Comunicaciones del Perú, 2018; Ministry of Land, Infrastructure, Transport and Tourism, 2021; National Highways, 2020, 2021; Statens vegvesen, 2019; Trafikverket, 2024a, 2024b; Vegagerðin, 2010, 2011).

3.11.3.8 Propuestas de mejora a la DG-2018

Con base en el análisis comparativo de las cinco normativas internacionales y los resultados de los Modelos RNA N° 1 y RNA N° 2, se formularon siete propuestas de mejora a los parámetros geométricos de la DG-2018, priorizadas según su viabilidad técnica para las condiciones de la PE-28A y su correlación con los factores de riesgo identificados por los modelos de predicción.

Tabla 72 *Resumen de propuestas de mejora a la DG-2018 sustentadas en normativas internacionales*

N°	Propuesta	Valor actual DG-2018	Valor propuesto	País de referencia	Impacto esperado	Prioridad
P1	Incorporar rampas de emergencia en pendientes ≥6 % y longitud >3 km	No especificadas en la norma	Nueva Sección 303.07: rampa ascendente 8–15 %, $L_{mín}$=80 m	Islandia (Veghönn. Sec. 3.4.3)	Reducción accidentes por falla de frenos en bajadas prolongadas	Alta
P2	Otorgar carácter obligatorio a la longitud máxima de recta e incorporar longitud mínima	$L_{máx}$ referencial sin carácter obligatorio (Tabla 302.01)	$L_{máx}=20\times V_h$ (obligatorio) $L_{mín}=6\times V_h$ (nuevo parámetro)	Islandia (Tafla 3.3.1-1) Noruega (N100)	Prevención de microsueño y deslum-bramiento nocturno	Alta
P3	Incorporar fórmulas continuas de fricción	Tablas $f_{máx}$ discretas (Tabla 302.02)	$f=0{,}62e^{-0{,}0063V}$ $f_s=0{,}28e^{-0{,}0096V}$ $f_b=0{,}55e^{-0{,}0057V}$	Suecia (Grundvärden 2024)	Verificación independiente; interpolación rigurosa de parámetros	Media
P4	Reducir peralte máximo en carreteras de 1ra y 2da clase	$e_{máx}$=12 % (Tabla 302.02)	$e_{máx}$=10 % en vías de 1ra y 2da clase	Japón (RSO Art. 16)	Incremento $R_{mín}$ 15–20 %; reducción riesgo vuelco en curva	Alta
P5	Incorporar peralte variable con la velocidad de diseño	$e_{máx}$ fijo por zona geográfica	10 % ($V\leq60$) 8 % (V=80) 6 % ($V\geq100$)	Islandia (Veghönn. 03 Tafla 3.5.2-1)	Mayor seguridad en tramos de alta velocidad; reducción aquaplaning	Media

Continúa en la siguiente página

(Continuación de la Tabla 72)

N°	Propuesta	Valor actual DG-2018	Valor propuesto	País de referencia	Impacto esperado	Prioridad
P6	Establecer bombeo mínimo de 3,0 % en zonas de alta precipitación	2,0–2,5 % uniforme (Sección 304.05)	2,0 % zona costera 3,0 % sierra y selva alta	Noruega (3,0 %) Islandia (3,0–4,0 %)	Reducción de aquaplaning en calzada húmeda; drenaje efectivo en sierra	Alta
P7	Crear sección normativa para carreteras tipo 2+1	Inexistente en el DG-2018	Nueva Sección 302.10: carriles alternantes 800–1 500 m	Noruega, Suecia, Reino Unido	Reducción hasta 70 % accidentes frontales sin duplicar calzada	Alta

Fuente: Elaboración propia.

3.11.4 Programa de alerta de accidentes de tránsito

3.11.4.1 Descripción general

En el marco del presente trabajo de investigación se desarrolló un programa de alerta de accidentes de tránsito para la Vía Los Libertadores (PE-28A), el cual constituye la aplicación operativa de los modelos de red neuronal artificial elaborados. El sistema fue implementado íntegramente como un archivo *HyperText Markup Language* (`.html`), y no requiere instalación de software, servidor ni conexión permanente a internet para su funcionamiento, solo en caso se requiera la activación del GPS.

El programa funciona directamente en el navegador web del dispositivo móvil o computadora del usuario. Al almacenarse el total de los 202 registros de accidentes de tránsito del período 2020–2025 dentro del propio archivo de código, el sistema opera en modalidad.

3.11.4.2 Diagrama de flujo

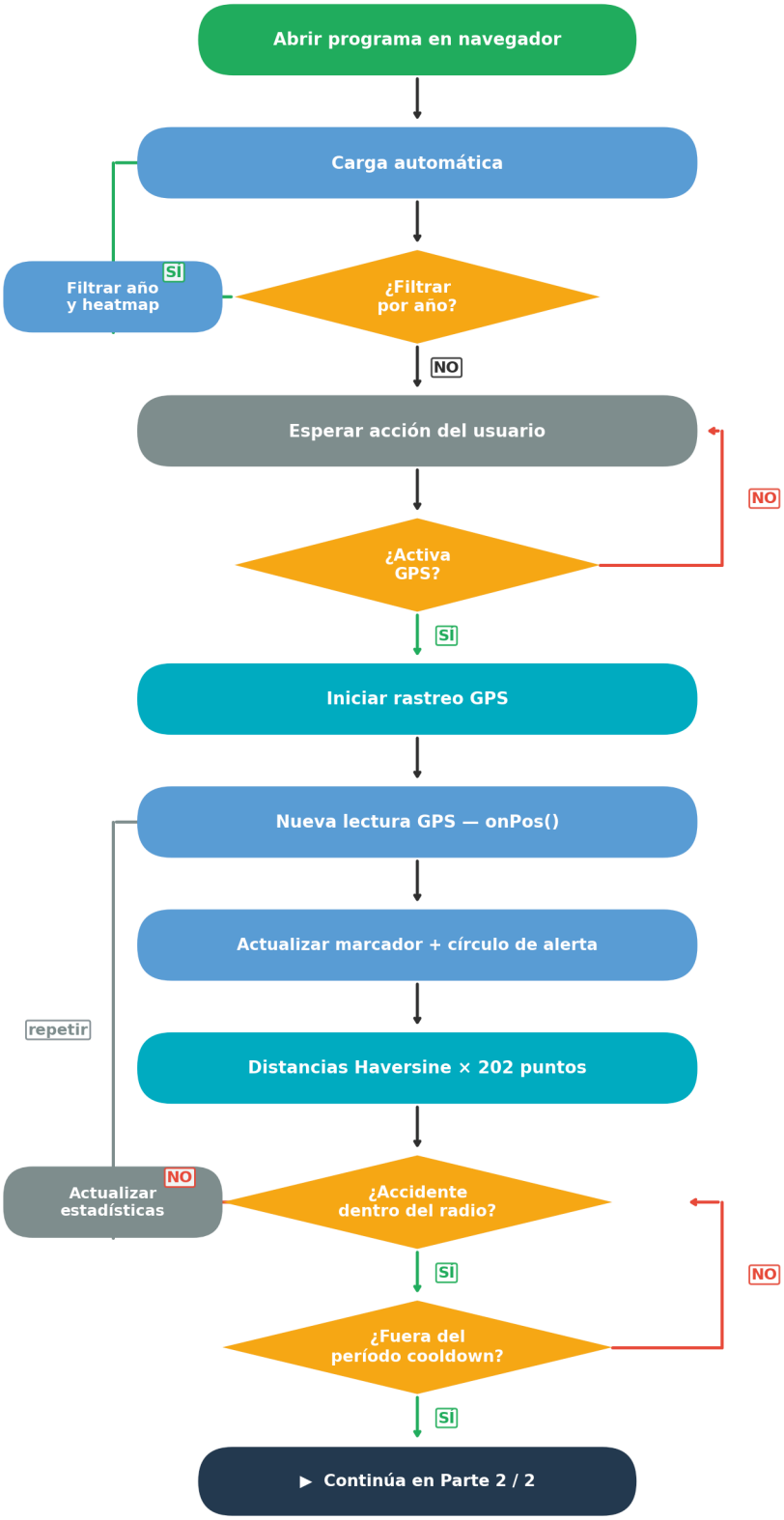


Figura 197 Diagrama de flujo - Parte I.
Fuente: Elaboración propia

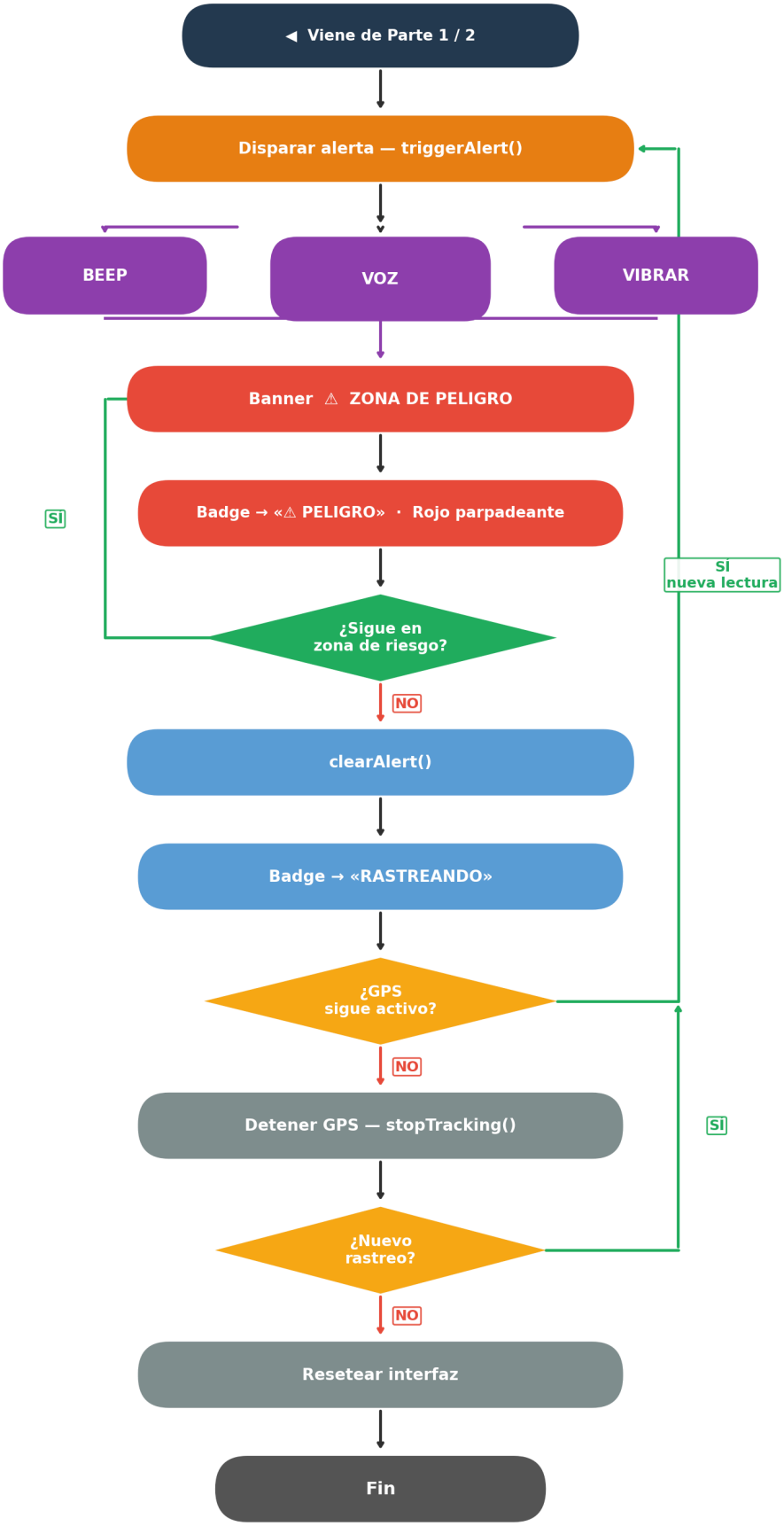


Figura 198 Diagrama de flujo - Parte II.
Fuente: Elaboración propia

3.11.4.3 Base de datos embebida

La información geoespacial utilizada por el programa proviene de los registros del Observatorio Nacional de Seguridad Vial (ONSV), del Ministerio de Transportes y Comunicaciones (MTC) y de la Policía Nacional del Perú (PNP), correspondientes al periodo 2020–2025. Los 202 registros fueron codificados directamente en el archivo como un arreglo *JavaScript*, donde cada elemento contiene el identificador, las coordenadas geográficas (latitud y longitud en grados decimales WGS-84) y el año del evento, tal como se muestra en la tabla siguiente.

Tabla 73 *Estructura de cada registro de accidente en la base de datos embebida*

Campo	Tipo	Descripción
`id`	Entero	Identificador único del registro (1–202)
`lat`	Decimal (7 dec.)	Latitud en grados decimales, sistema WGS-84
`lon`	Decimal (7 dec.)	Longitud en grados decimales, sistema WGS-84
`year`	Cadena	Año del evento: 2020, 2021, 2022, 2023, 2024 ó 2025

Fuente: Elaboración propia.

La distribución de los 202 registros por año se presenta en la tabla siguiente.

Tabla 74 *Distribución anual de los 202 accidentes de tránsito — PE-28A (2020–2025)*

Año	N.° accidentes	% del total	Color en el mapa
2020	74	36,6 %	Azul
2021	61	30,2 %	Verde
2022	22	10,9 %	Morado
2023	10	5,0 %	Naranja
2024	22	10,9 %	Rojo
2025	13	6,4 %	Amarillo
Total	**202**	**100 %**	—

Fuente: Elaboración propia.

3.11.4.4 Tecnologías y librerías utilizadas

El desarrollo del programa empleó exclusivamente de código abierto disponibles en la web. El Cuadro 75 resume las herramientas utilizadas.

Tabla 75 *Tecnologías y librerías empleadas en el programa de alerta*

Tecnología / Librería	Función en el programa
Leaflet.js	Renderizado del mapa interactivo
Leaflet.heat	Generación del mapa de calor (*heatmap*)
CartoCDN	Teselas del mapa (oscuro, claro, calles)
ArcGIS World Imagery	Teselas de vista satelital
Web Geolocation API	Captura de coordenadas GPS en tiempo real
Web Audio API	Señal sonora de alerta
SpeechSynthesis API	Alerta de voz en español
Vibration API	Vibración del dispositivo móvil
Screen Wake Lock API	Mantener pantalla encendida durante rastreo
HTML + JS	Estructura, estilos y lógica del programa

Fuente: Elaboración propia.

3.11.4.5 Módulos funcionales del programa

El programa está organizado en siete módulos funcionales que operan de manera coordinada, tal como se ilustra en el diagrama de flujo mostrados.

A) Módulo 1 — Pantalla de presentación.

Al abrir el archivo, se despliega una pantalla de bienvenida que muestra los datos institucionales de la investigación: nombre de la universidad, título de la tesis, nombre del autor, asesor y año. El mapa y los controles se cargan en segundo plano durante esta fase. El usuario accede al sistema tocando cualquier área de la pantalla o presionando el botón «INGRESAR AL SISTEMA».

B) Módulo 2 — Inicialización del mapa y datos.

Al ingresar al sistema se ejecutan automáticamente las siguientes acciones:

- Inicialización del mapa centrado en las coordenadas (−13,44°; −75,19°), correspondientes al punto central de la PE-28A.
- Representación de los 202 marcadores sobre el mapa, diferenciados por color según el año del evento.
- Generación del mapa de calor con gradiente de color azul → amarillo → rojo, que permite identificar visualmente las zonas de mayor concentración de accidentes.
- Activación de todos los controladores de eventos para los controles de filtro, capas de mapa, tema visual y rastreo GPS.

C) Módulo 3 — Filtros y visualización.

El panel lateral deslizante permite al usuario configurar los siguientes parámetros de visualización:

- **Filtro por año:** selección de uno de los seis años del período 2020–2025, que actualiza simultáneamente los marcadores visibles y el mapa de calor.
- **Capa de mapa:** cuatro opciones disponibles: oscuro, claro, satelital y calles.
- **Mapa de calor:** activación o desactivación con un solo toque.
- **Leyenda:** tabla de colores por año con el conteo de eventos y la posición GPS en tiempo real.

D) Módulo 4 — Rastreo GPS en tiempo real.

Al presionar el botón «Activar GPS», el sistema inicia el rastreo continuo de posición mediante la API nativa `navigator.geolocation.watchPosition()`, configurada con los parámetros presentados.

Tabla 76 *Parámetros de configuración del rastreo GPS*

Parámetro	Valor	Descripción
`enableHighAccuracy`	`true`	Activa el GPS del hardware
`timeout`	3km	Espera máxima por señal
`maximumAge`	0	solo posición

Fuente: Elaboración propia.

Simultáneamente, el sistema solicita el bloqueo de pantalla encendida (*Wake Lock*) para evitar que el dispositivo entre en reposo durante el rastreo. Cada actualización de posición ejecuta la función `onPos()`, que actualiza el marcador del usuario, el círculo de radio de alerta y los valores de latitud, longitud, precisión y velocidad en la interfaz.

E) Módulo 5 — Cálculo de proximidad mediante la fórmula de Haversine.

Por cada nueva lectura GPS, el sistema calcula la distancia entre la posición del usuario y cada uno de los 202 puntos de accidente mediante la fórmula de Haversine, expresada como:

$$d = 2R \cdot \arctan\left(\sqrt{a},\ \sqrt{1-a}\right) \tag{3.26}$$

donde:

$$a = sen^2\left(\frac{\Delta\varphi}{2}\right) + \cos\varphi_1 \cdot \cos\varphi_2 \cdot sen^2\left(\frac{\Delta\lambda}{2}\right) \tag{3.27}$$

siendo $R = 6\,371\,000$ m el radio medio de la Tierra, φ_1, φ_2 las latitudes de los dos puntos (en radianes) y $\Delta\varphi$, $\Delta\lambda$ las diferencias de latitud y longitud, respectivamente. El resultado d es la distancia geodésica en metros sobre la superficie esférica terrestre.

F) Módulo 6 — Sistema de alerta multimodal.

Si al menos uno de los 202 puntos queda dentro del radio de alerta configurado **y** ha transcurrido el período de enfriamiento , el sistema dispara simultáneamente tres tipos de alerta:

- **Alerta sonora (*beep*):** sintetizada mediante la Web Audio API.
- **Alerta de voz:** el navegador pronuncia la frase «*Riesgo de accidente*».
- **Alerta de vibración:** el dispositivo móvil emite el patrón de vibración.
- **Alerta visual:** se despliega un banner rojo con el identificador del accidente más cercano y la distancia exacta en metros; el indicador de estado del encabezado cambia a **PELIGRO**.

Tabla 77 *Niveles de intensidad de la alerta sonora según distancia*

Distancia al accidente	Intensidad	Descripción
$d < 150$ m	1,0 (máxima)	Peligro inminente
$150\,\text{m} \leq d < r/2$	0,8	Peligro próximo
$r/2 \leq d \leq r$	0,6	Peligro en el límite

r = radio de alerta configurado por el usuario.

Fuente: Elaboración propia.

G) Módulo 7 — Detención y reseteo.

Al presionar «Detener GPS», el sistema cancela el rastreo mediante , elimina el marcador y el círculo de alerta del mapa, cancela cualquier alerta activa y restablece todos los indicadores de la interfaz a su estado inicial.

3.11.4.6 Parámetros configurables del sistema

La tabla siguiente resume los parámetros que el usuario puede ajustar durante el uso del programa.

Tabla 78 *Parámetros configurables del programa de alerta*

Parámetro	Por defecto	Rango	Descripción
Radio de alerta	300 m	100–3 000 m	Distancia de detección
Intervalo de alerta	10 s	5 s – 5 min	Tiempo entre repeticiones
Alerta sonora	Activada	On / Off	Beep por Web Audio API
Alerta de voz	Desactivada	On / Off	SpeechSynthesis API
Capa de mapa	Oscuro	4 opciones	Oscuro, claro, sat., calles
Filtro de año	Todos	2020–2025	Filtra marcadores y heatmap
Mapa de calor	Activado	On / Off	Visualización de densidad

Fuente: Elaboración propia.

3.11.4.7 Instrucciones de uso del programa

A continuación se describen los pasos necesarios para la instalación, configuración y uso correcto del programa de alerta de accidentes de tránsito en la PE-28A.

Tabla 79 *Procedimiento de uso del programa de alerta de accidentes para recorridos en la carretera PE-28A*

Paso	Acción	Descripción
1	Abrir el programa	Abrir `Programa_de_alerta_de_accidentes.html` en el dispositivo y abrirlo con el navegador web. Tocar «INGRESAR AL SISTEMA» para continuar
2	Explorar el mapa	Visualizar el mapa interactivo con los 202 puntos de accidente (marcadores de colores y mapa de calor). Usar zoom, desplazamiento y tocar marcadores para ver el detalle de cada accidente
3	Configurar parámetros	Abrir el panel «**Controles**» y ajustar: *radio de alerta* (100–3 000 m, por defecto 300 m), *capa de mapa*, *filtro por año*, *beep sonoro*, *voz de alerta*, *intervalo de repetición* y activación del mapa de calor
4	Activar rastreo GPS	Presionar «**Activar GPS**» y conceder el permiso de ubicación solicitado por el navegador. El estado cambiará a «RASTREANDO» (verde pulsante) y el mapa se centrará en la posición actual

Continúa en la siguiente página

(Continuación de la Tabla 79)

Paso	Acción	Descripción
5	Monitorear durante el recorrido	La barra inferior muestra en tiempo real *Cerca Alertas* y *Metros Cercano*. Al ingresar en zona de riesgo se activan: beep, voz «*Riesgo de accidente*», vibración y banner «**ZONA DE PELIGRO**»; al alejarse, la alerta se cancela automáticamente
6	Detener el rastreo	Presionar «**Detener GPS**» para finalizar. Todos los indicadores retornan al estado inicial. El rastreo puede reiniciarse en cualquier momento repitiendo el Paso 4

Fuente: Elaboración propia.

A) Paso 1: Apertura del programa.

Transferir el archivo `Programa de alerta de accidentes.html` al dispositivo y abrirlo con el navegador web.

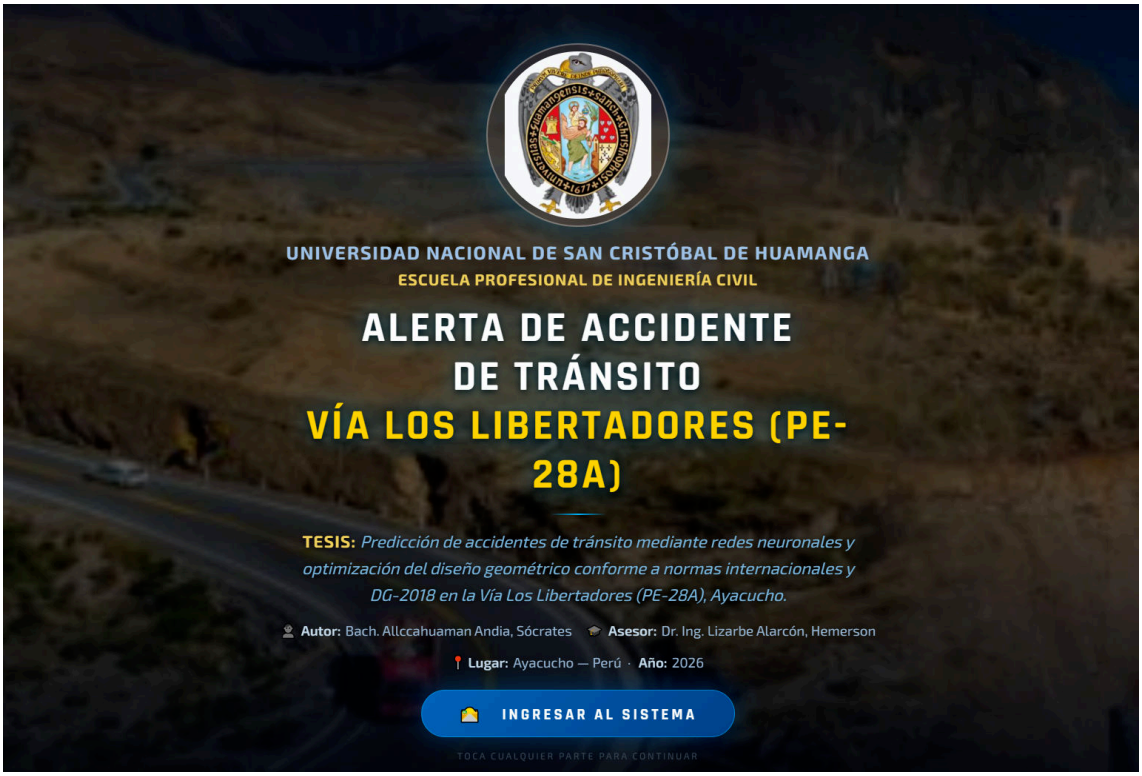


Figura 199 Paso 1: Apertura del programa.
Fuente: Elaboración propia

B) Paso 2: Exploración del mapa.

Al ingresar se visualizará el mapa interactivo de la PE-28A con los 202 puntos de accidente representados por marcadores de colores y el mapa de calor superpuesto.

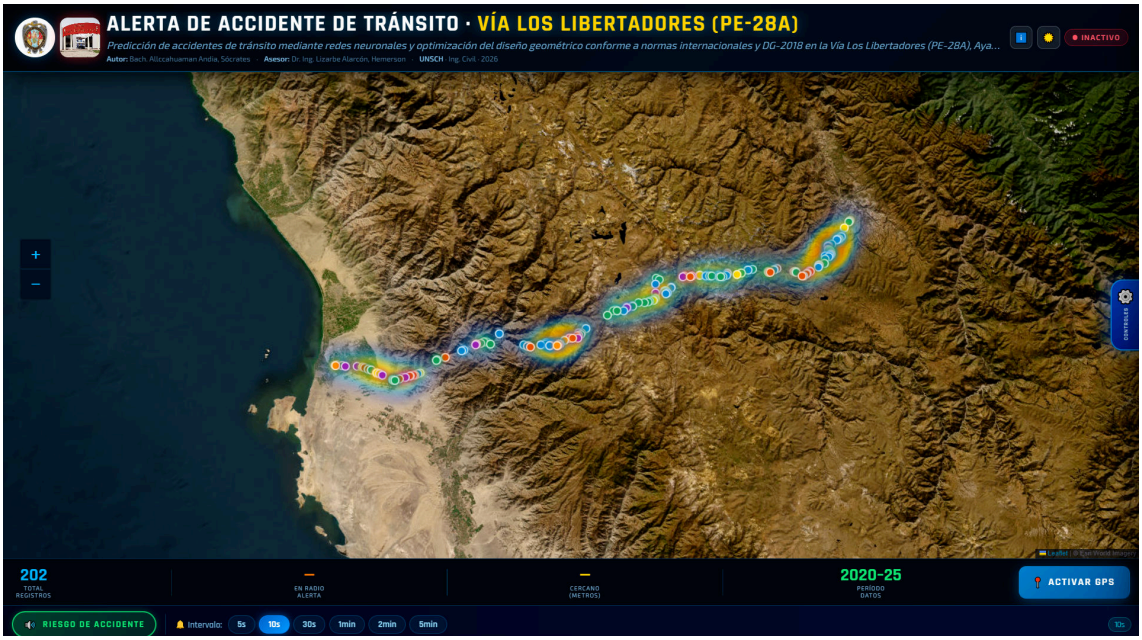


Figura 200 Paso 2: Exploración del mapa.
Fuente: Elaboración propia

C) Paso 3: Configuración de parámetros (panel lateral).

Presionar el botón «**Controles**» ubicado en el borde derecho de la pantalla para abrir el panel lateral. Desde este panel se pueden ajustar.

Figura 201 Paso 3: Configuración de parámetros (panel lateral).
Fuente: Elaboración propia

Figura 202 Paso 3: Configuración de parámetros.
Fuente: Elaboración propia

D) Paso 4: Activación del rastreo GPS.

Presionar el botón «**Activar GPS**» ubicado en la barra inferior de la pantalla. El navegador solicitará permiso para acceder a la ubicación del dispositivo; debe concederse el permiso.

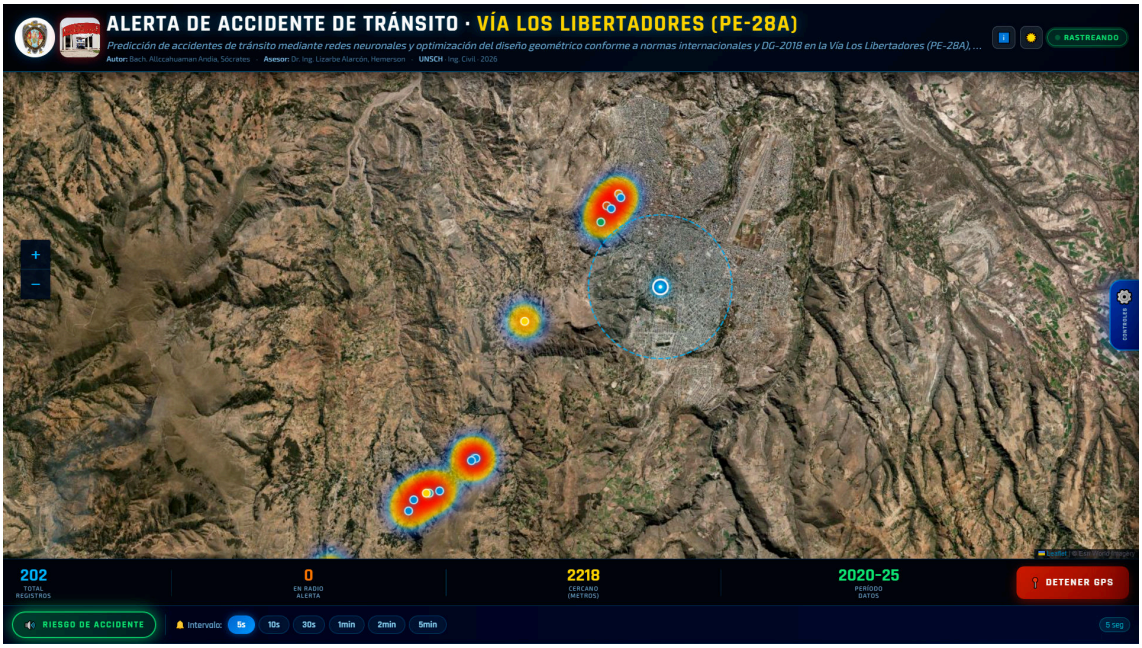


Figura 203 Paso 4: Activación del rastreo GPS.
Fuente: Elaboración propia

E) Paso 5: Durante el recorrido.

Alerta de riesgo de accidente de tránsito.

Figura 204 Paso 5: Durante el recorrido.
Fuente: Elaboración propia

F) Paso 6: Detención del rastreo.

Presionar nuevamente el botón «**Detener GPS**» para finalizar el rastreo. Todos los indicadores retornarán a su estado inicial.

Figura 205 Paso 6: Detención del rastreo y ubicación de accidentes de transito..
Fuente: Elaboración propia

Figura 206 Datos de accidente de tránsito.
Fuente: Elaboración propia

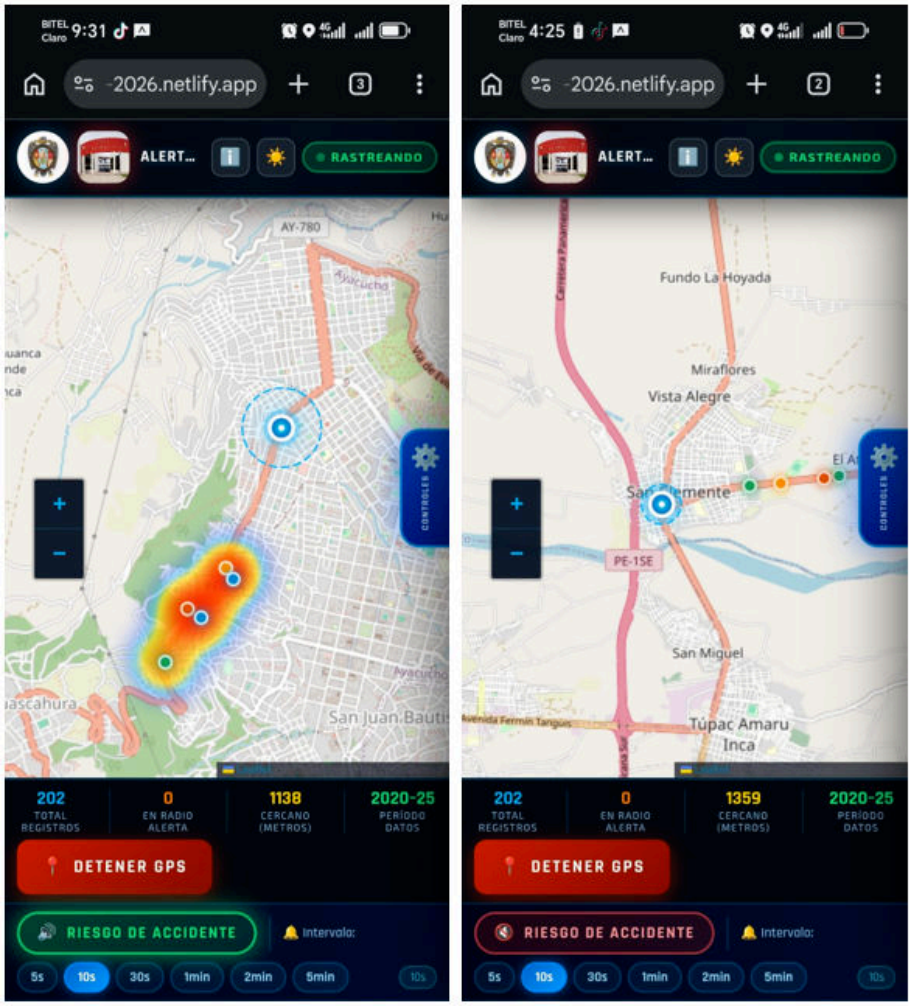


Figura 207 Programa ejecutado en el celular.
Fuente: Elaboración propia

CAPÍTULO IV Resultados

4.1 Análisis y contrastación de hipótesis

La presente sección contrasta las hipótesis formuladas en el Capítulo III con los resultados obtenidos en los cuatro ejes de investigación: Modelo RNA N°01, Modelo RNA N°02, análisis normativo comparativo y programa de alerta GPS.

4.1.1 Hipótesis general

Hipótesis general: La implementación de dos modelos de redes neuronales artificiales para la clasificación de la peligrosidad de accidentes de tránsito, combinada con el análisis comparativo de la norma DG-2018 con normativas internacionales de diseño geométrico vial y la implementación de un programa de alerta basado en geolocalización, contribuirá a la predicción y reducción de los accidentes de tránsito en la Vía Los Libertadores (PE-28A), Ayacucho.

Resultado: Se acepta la hipótesis general. Los resultados obtenidos en los cuatro ejes de investigación demuestran que: (i) el Modelo RNA N°01 alcanzó un *accuracy* de evaluación de 90,48 % y AUC-ROC de 0,9818; (ii) el Modelo RNA N°02 alcanzó un *accuracy* de evaluación de 98,00 % y AUC-ROC de 0,9985; (iii) el análisis normativo identificó deficiencias técnicas en la DG-2018 y formuló siete propuestas de mejora sustentadas en normativas internacionales; y (iv) el programa de alerta GPS integra los 202 registros históricos de la PE-28A, siendo una herramienta de seguridad vial para conductores.

4.1.2 Hipotesis 01: Modelo RNA N°01

Hipótesis: El Modelo RNA N°01 (clasificación binaria: riesgo Menor o riesgo mayor), entrenado con 202 registros históricos de accidentes de tránsito de la Vía Los Libertadores (PE-28A) y variables del diseño geométrico, condiciones climáticas, características del siniestro y factores viales conforme a la norma DG-2018, alcanza un nivel de precisión alto para la predicción de accidentes de tránsito.

Criterio de aceptación: *Accuracy* $\geq$ 90 % y AUC-ROC $\geq$ 0,90 en el conjunto de evaluación.

Resultado: Se acepta la hipótesis específica a). El Cuadro 80 presenta las métricas de desempeño del Modelo RNA N°01 en las tres particiones de datos.

Tabla 80 *Métricas de rendimiento del Modelo RNA N°01 en todas las particiones*

Métrica	Train (70 %)	Val (20 %)	Test (10 %)	Umbral
Accuracy	96,43 %	95,12 %	**90,48 %**	≥90 %
Precisión (macro)	96,41 %	95,24 %	90,45 %	≥90 %
Recall (macro)	96,45 %	95,45 %	90,45 %	≥90 %
F1-Score (macro)	96,42 %	95,12 %	90,45 %	≥90 %
AUC-ROC (macro)	—	—	**0,9818**	≥0,90

Fuente: Elaboración propia.

El modelo de la RNA N°01 clasifica muy bien, esta sobre el umbral de excelencia.

A) Arquitectura de la RNA°01:

La Figura siguiente muestra la topología definitiva del Perceptrón Multicapa: 23→64→48→32→16→8→2, con 5 capas ocultas y 6,906 parámetros totales. La arquitectura garantiza una abstracción progresiva de los patrones: la primera capa oculta (64 neuronas) extrae características primarias; las capas intermedias (48, 32, 16) combinan y abstraen representaciones progresivamente más compactas; la quinta capa oculta (8 neuronas) genera la representación final pre-clasificación; y la capa de salida (2 neuronas con Softmax) produce las probabilidades (RIESGO MENOR=0) y (RIESGO MAYOR=1).

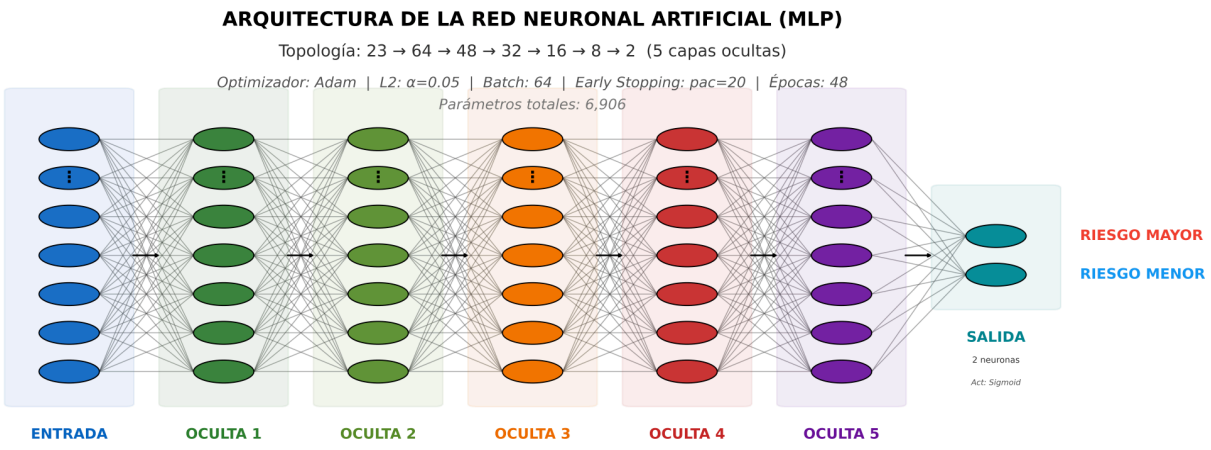


Figura 208 Arquitectura del Perceptrón Multicapa de la RNA N°01
Fuente: Elaboración propia

B) Resumen de métricas de evaluación de la RNA: Precisión, Recall, F1-Score y Accuracy

Las Figura siguiente presentan el análisis de las métricas de clasificación de Precisión, Recall y F1-Score macro por partición de datos, así como su comparación conjunta con el Accuracy.

Se evalúan sobre las tres particiones: entrenamiento (70%, $n = 140$), validación (20%, $n = 41$) y evaluación (10%, $n = 21$).

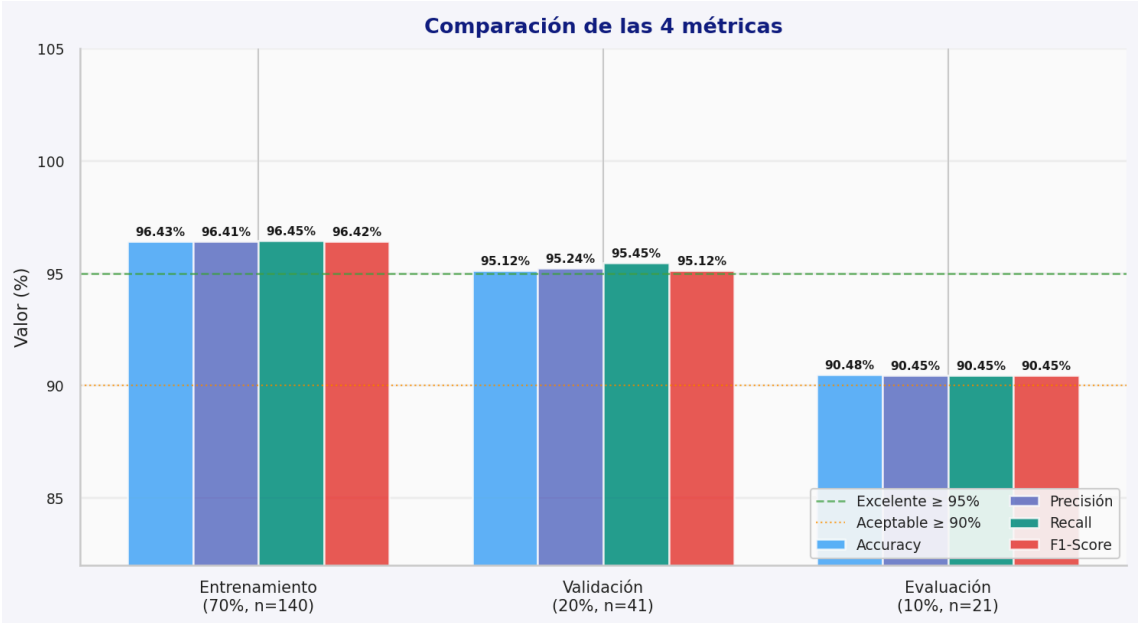


Figura 209 Comparación de las 4 métricas por partición.
Fuente: Elaboración propia

El análisis revela tres hallazgos. Primero, la **consistencia entre métricas**: en las tres particiones, Accuracy, Precisión, Recall y F1-Score presentan valores prácticamente idénticos. Las particiones de entrenamiento y validación alcanzan el nivel de excelente (≥ 95%), mientras que la de evaluación la más relevante por corresponder a datos nunca vistos supera el umbral aceptable (≥ 90%) en todas las métricas, validando la predicción de accidentes en la Vía Los Libertadores (PE-28A).

C) Frontera de decisión:

La Figuras siguientes presenta la proyección de los 202 registros viales en un espacio bidimensional obtenido mediante el Análisis de Componentes Principales (PCA, *Principal Component Analysis*), con el propósito de visualizar la frontera de decisión aprendida por el modelo RNA. El modelo opera sobre 23 variables de entrada la reducción dimensional mediante PCA permite representar gráficamente la separación entre clases. Los dos primeros componentes principales retienen el siguiente porcentaje de varianza:

$$PC1 + PC2 = 32.2\% + 14.6\% = 46.8\% \text{ de la varianza total} \quad (4.1)$$

Esto implica que la figura conserva el 46.8% de la información original, mientras que el

53.2% restante distribuido entre los 21 componentes adicionales no es visible en 2D.

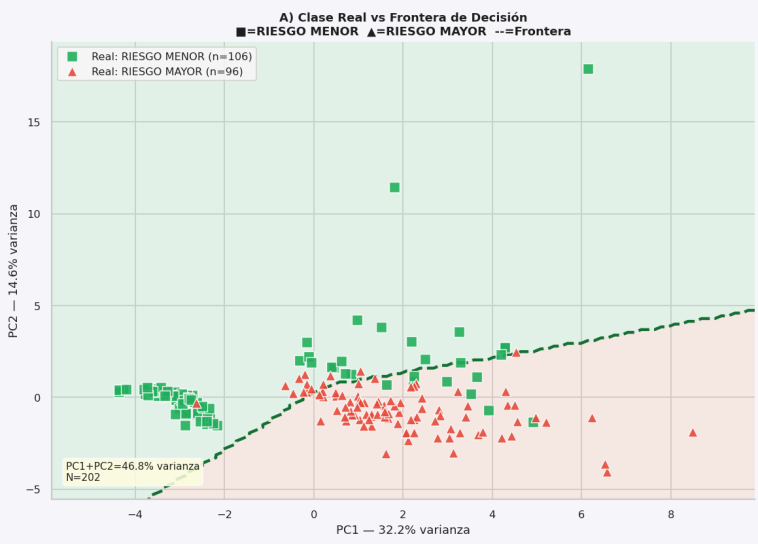


Figura 210 Clase Real vs Frontera de Decisión.
Fuente: Elaboración propia

D) Curva ROC y área bajo la curva (AUC) del modelo RNA:

La Figura presenta la Curva ROC (*Receiver Operating Characteristic*) del modelo RNA-MLP, evaluada sobre el conjunto de evaluación independiente ($n = 21$ registros, 10% del total). Esta curva grafica la relación entre la Tasa de Verdaderos Positivos (TPR) y la Tasa de Falsos Positivos (FPR) al variar el umbral de decisión del modelo desde 0.0 hasta 1.0, permitiendo evaluar la capacidad de clasificar de la RNA.

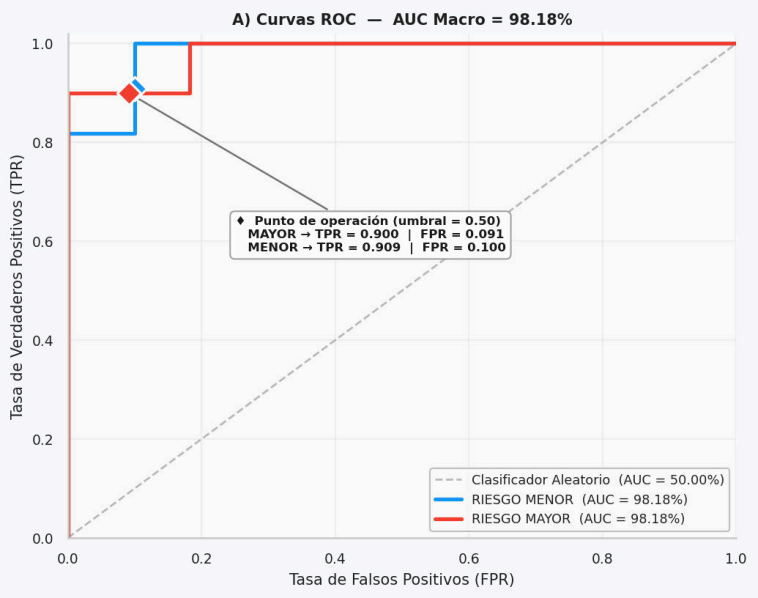


Figura 211 Curva ROC y área bajo la curva (AUC) del modelo RNA.
Fuente: Elaboración propia

E) Pérdida vs. Épocas:

Muestra el error del modelo en cada época. Ambas curvas bajan progresivamente, el cual el modelo aprende. La curva naranja (validación) alcanza su punto mínimo en la época 29 (*val_loss* = 0.0608), ese es el mejor momento del modelo. Después sube levemente porque empieza a sobreajustarse. El early stopping detectó eso, esperó 20 épocas más sin mejora, paró en época 48 y restauró los pesos de época 29. Los pesos de época 48 fueron descartados.

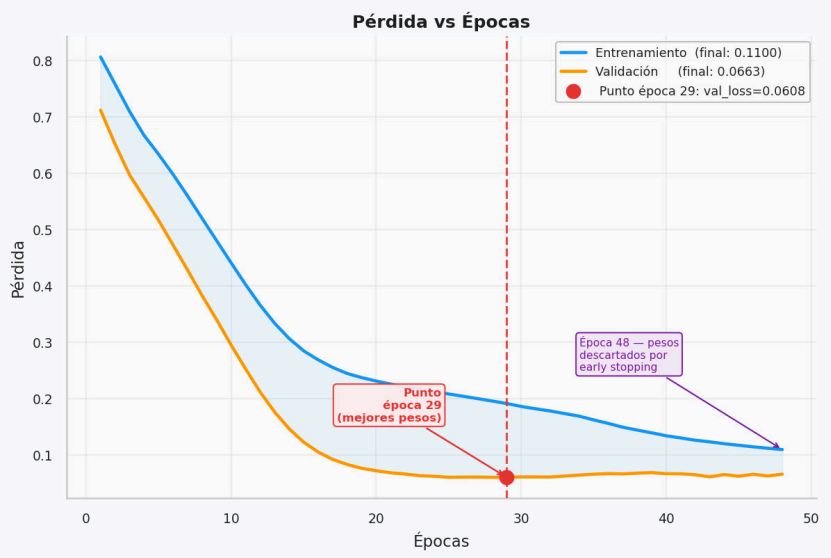


Figura 212 Pérdida (loss) vs. Épocas
Fuente: Elaboración propia

F) Accuracy vs. Épocas:

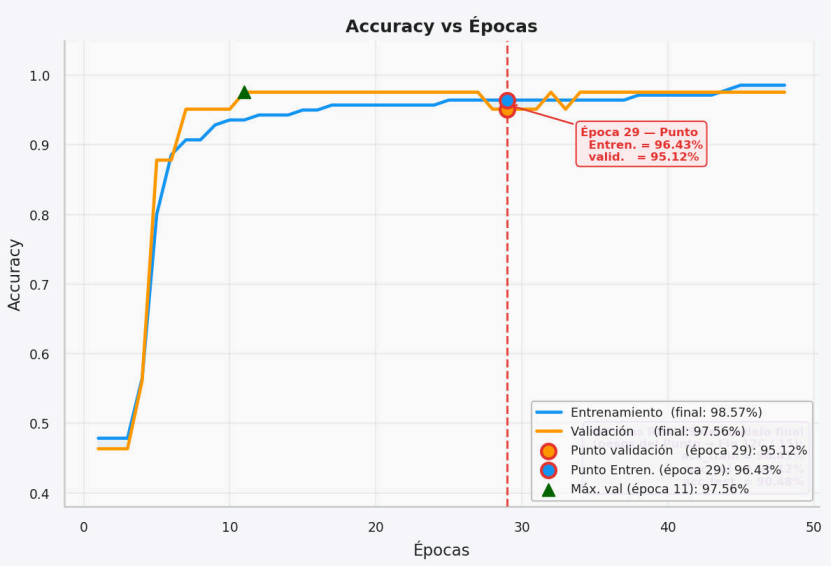


Figura 213 Accuracy vs. Épocas
Fuente: Elaboración propia

El modelo aprendió sin overfitting. El early stopping funcionó correctamente. La diferencia entre Accuracy de entrenamiento (96.43%), Accuracy de validación (90.12%) y Accuracy de evaluación (90.48%) lo que confirma que el modelo generaliza bien ante datos nuevos.

G) Riesgo de accidente por departamento, provincia y distrito.

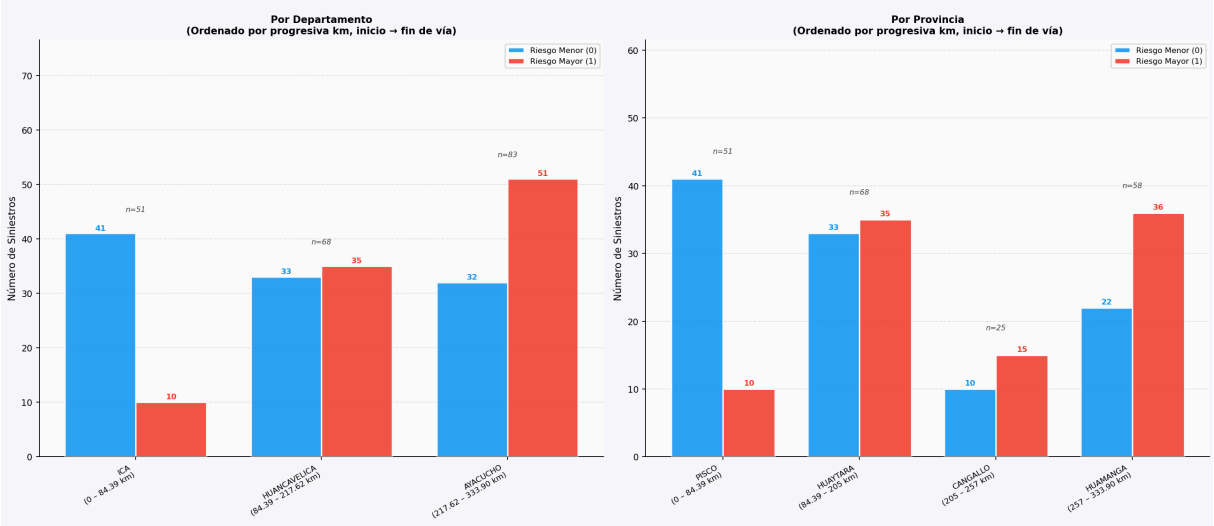


Figura 214 Riesgo de accidentes por departamento y provincia.
Fuente: Elaboración propia.

.

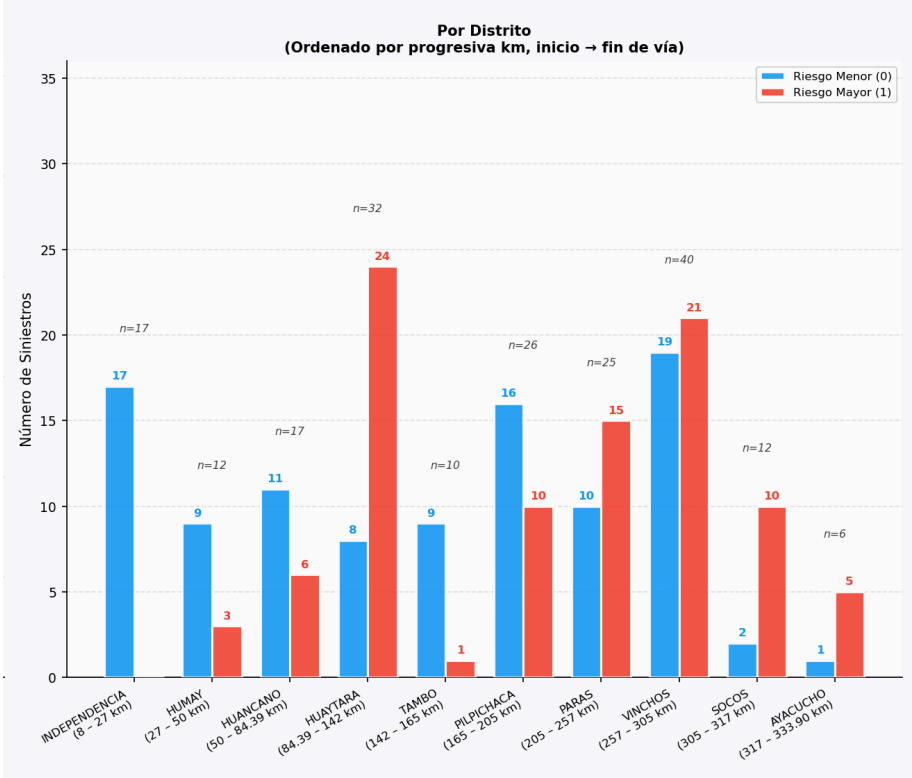


Figura 215 Riesgo de accidentes por distrito
Fuente: Elaboración propia.

H) Programa RNA N°01: Predicción de accidentes de tránsito

El Modelo RNA N°1 fue exportado como aplicación web denominada. El programa permite clasificar cualquier tramo de la Vía Los Libertadores PE-28A y carreteras nacionales en **Riesgo menor** o **Riesgo mayor** de accidente de tránsito, a partir del ingreso de 23 variables geométricas, climáticas, temporales y de peligrosidad vial, con funcionamiento en dispositivos móviles para uso en campo.

Figura 216 Paso 1: Abrir el programa
Fuente: Elaboración propia

Enlace de descarga del programa RNA N°01:

Descargar Aqui el programa RNA N° 01 (doble click)

4.1.3 Hipotesis 02: Modelo RNA N°02

Hipótesis: El Modelo RNA N°02 (clasificación multiclase: riesgo bajo, riesgo medio y riesgo alto), entrenado con 2,997 registros de datos geométricos de la Vía Los Libertadores (PE-28A) modelados en AutoCAD Civil 3D, alcanza un nivel de precisión alto para predecir el nivel de riesgo de accidentes de tránsito en tramos viales.

Criterio de aceptación: *Accuracy* ≥ 90 %, AUC-ROC ≥ 0,90 y F1-Score macro ≥ 0,90 en el conjunto de evaluación.

Resultado: Se acepta la hipótesis específica b). El Cuadro 81 presenta las métricas de desempeño del Modelo RNA N°02 en las tres particiones de datos.

Tabla 81 *Métricas de rendimiento del Modelo RNA N°02 en todas las particiones*

Métrica	Train (70 %)	Val (20 %)	Test (10 %)	Umbral
Accuracy	97,66 %	97,83 %	**98,00 %**	≥90 %
Precisión (macro)	95,17 %	96,07 %	95,71 %	≥90 %
Recall (macro)	97,56 %	97,35 %	97,92 %	≥90 %
F1-Score (macro)	96,23 %	96,69 %	**96,69 %**	≥90 %
AUC-ROC (macro)	—	—	**0,9985**	≥0,90

Fuente: Elaboración propia.

El modelo de la RNA N°02 clasifica muy bien, esta sobre el umbral de excelencia.

A) Arquitectura del Modelo RNA N°2

Se muestra la topología definitiva del MLP: 15→128→64→32→16→3, con 4 capas ocultas. La arquitectura sigue un patrón de reducción progresiva: capas iniciales amplias capturan representaciones de baja abstracción, mientras que las capas posteriores comprimen la información en representaciones de alta abstracción que facilitan la clasificación entre las tres clases de riesgo, los cuales son riego menor = 0, riesgo medio = 1 y riesgo = 2.

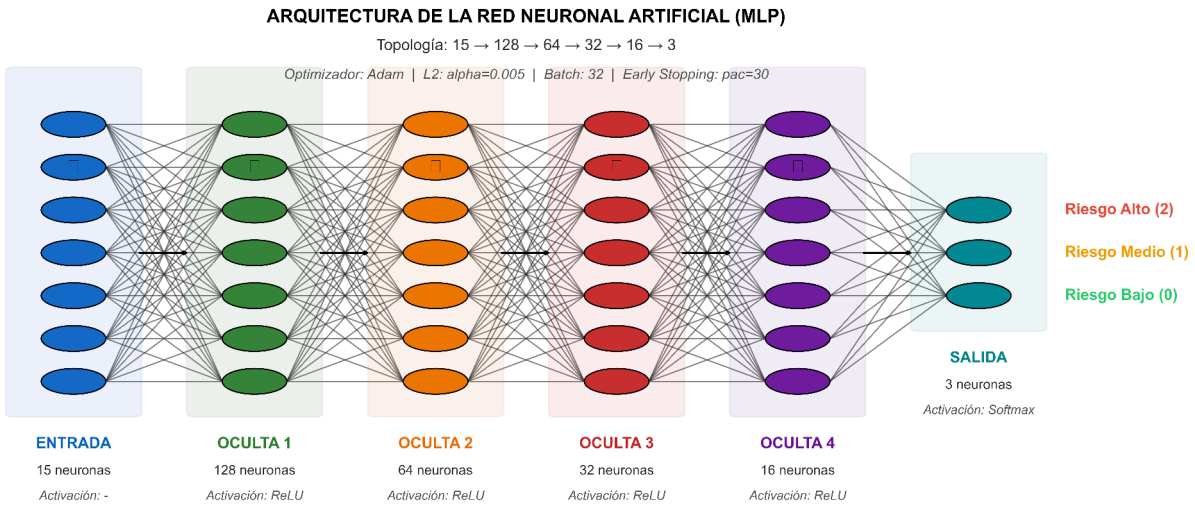


Figura 217 Arquitectura de la red neuronal N°02: 15→128→64→32→16→3. Optimizador Adam · L2 α=0.005 Batch=32 Epocas=66.
Fuente: Elaboración propia

B) Resumen de métricas de evaluación de la RNA: Precisión, Recall, F1-Score y Accuracy

Las figuras siguientes presentan el análisis de Precisión, Recall, F1-Score y Accuracy (promedio macro) en las tres particiones del dataset: entrenamiento (70%, $n = 2{,}097$), validación (20%, $n = 600$) y evaluación (10%, $n = 300$).

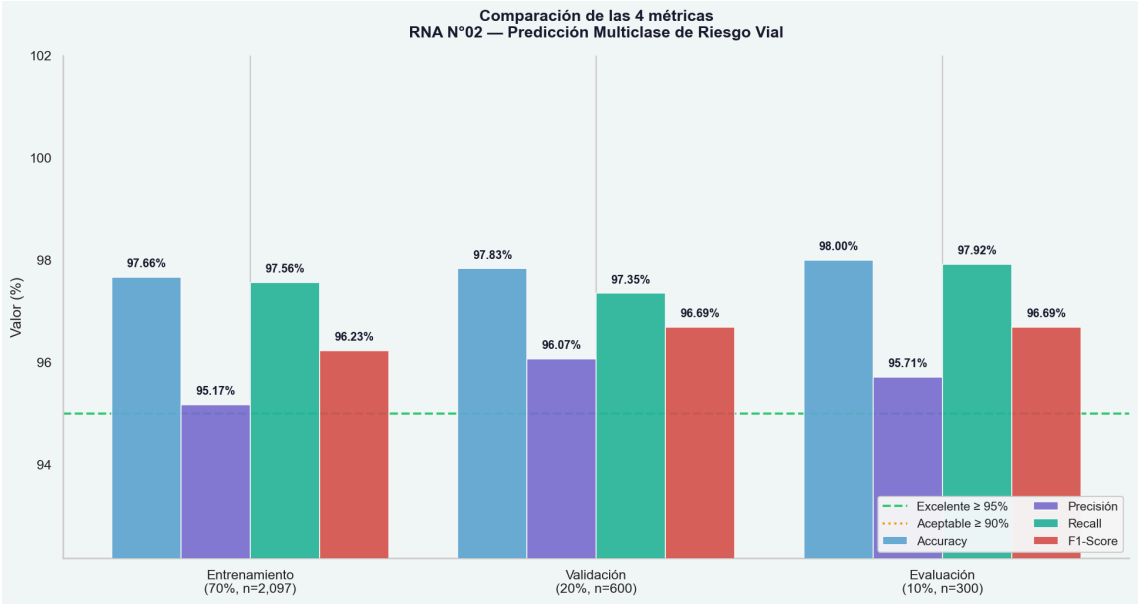


Figura 218 Comparación de las 4 métricas por partición — RNA N°02.
Fuente: Elaboración propia

Las cuatro métricas son consistentes entre sí y entre particiones, sin degradación significativa, confirmando ausencia de sobreajuste. El Accuracy alcanza su valor máximo en evaluación (98.00%), resultado atribuible a la regularización L2 ($\alpha = 0.05$) y al Early Stopping en la época óptima 36. Todas las métricas superan el umbral de excelencia ($\geq 95\%$) en las tres particiones, validando la capacidad predictiva del modelo RNA N°02 para la clasificación multiclase de niveles de riesgo vial en la Vía Los Libertadores PE-28A.

C) Frontera de decisión en el espacio PCA bidimensional

La figuras presenta las predicciones de la RNA proyectadas en el espacio bidimensional generado por el Análisis de componentes principales (PCA), donde PC1 explica el 48.2% y PC2 el 11.8% de la varianza total, acumulando en conjunto el 59.9%. Dado que la RNA opera en un espacio de 15 dimensiones en el espacio original para que el modelo prediga la clase en 2D.

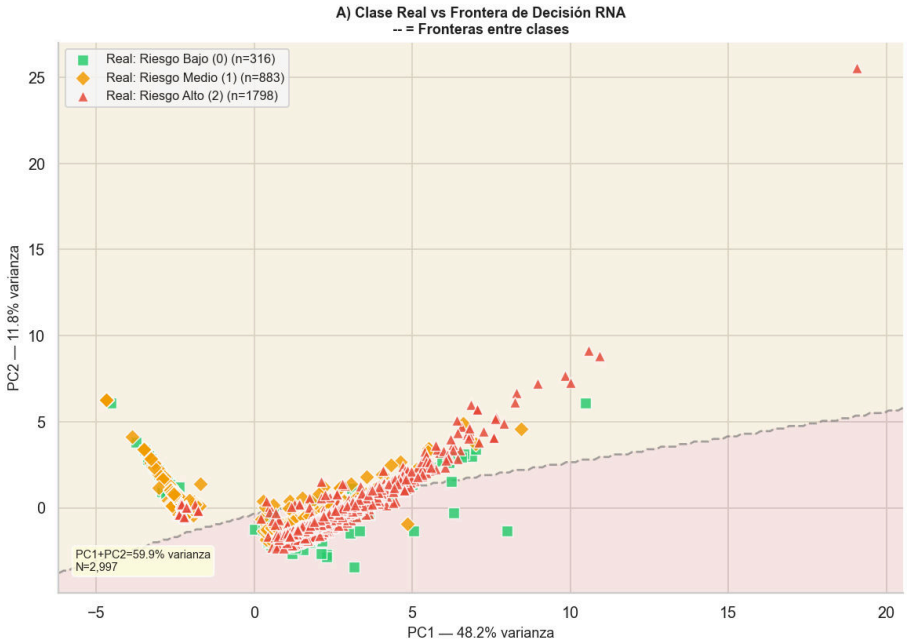


Figura 219 Clase real vs frontera de decisión del RNA.
Fuente: Elaboración propia

D) Curvas ROC y área bajo la curva (AUC)

La figura presenta las curvas ROC del modelo RNA N°2 evaluadas sobre el conjunto de evaluación ($n = 300$ registros).

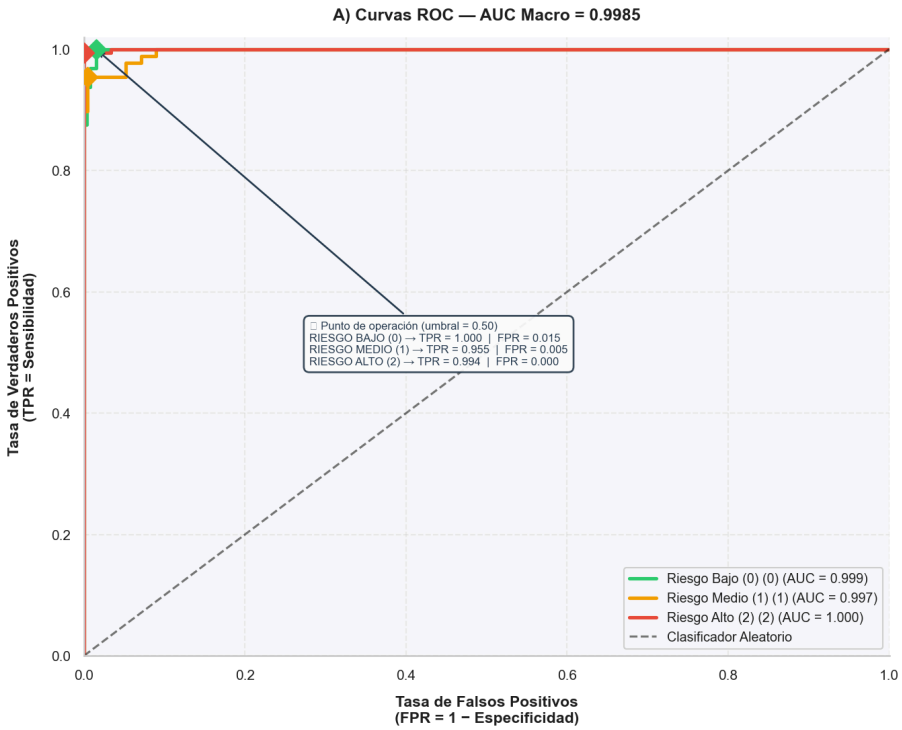


Figura 220 Curvas ROC y área bajo la curva (AUC).
Fuente: Elaboración propia

El AUC Macro del modelo alcanzó 0.9985. Las tres curvas se ubican en la esquina superior izquierda del espacio ROC, alejadas de la diagonal del clasificador aleatorio (AUC = 0.50), lo que confirma que el modelo RNA N°2 predice excelente las 3 clases.

E) Pérdida vs. Épocas:

Ambas curvas descendieron rápidamente y convergieron de forma estable. El mínimo de validación se alcanzó en la época 36 (val_loss = 0.0868), correspondiente a los mejores pesos del modelo. La zona sombreada reducida entre curvas confirma ausencia de sobreajuste.

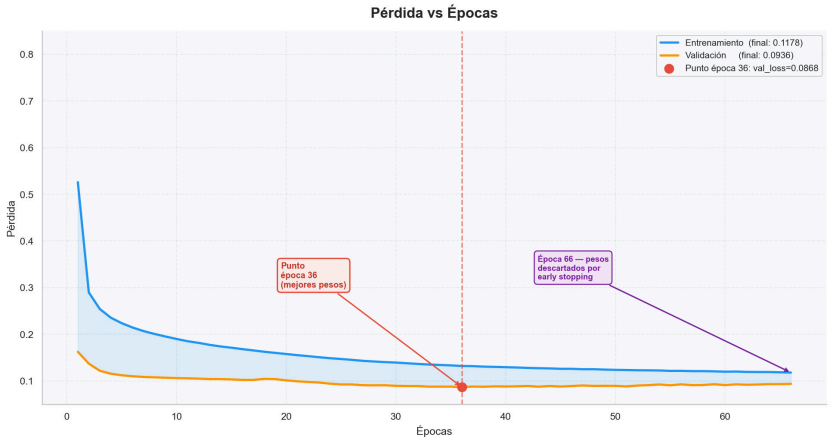


Figura 221 Pérdida vs épocas.
Fuente: Elaboración propia

F) Accuracy vs épocas:

Las curvas de entrenamiento (96.95%) y validación (96.33%) se mantuvieron paralelas durante todo el entrenamiento. En la época 36 el modelo registró Train = 97.66% y Val = 97.83%, alcanzando el máximo de validación en la época 37 con 98.00%.

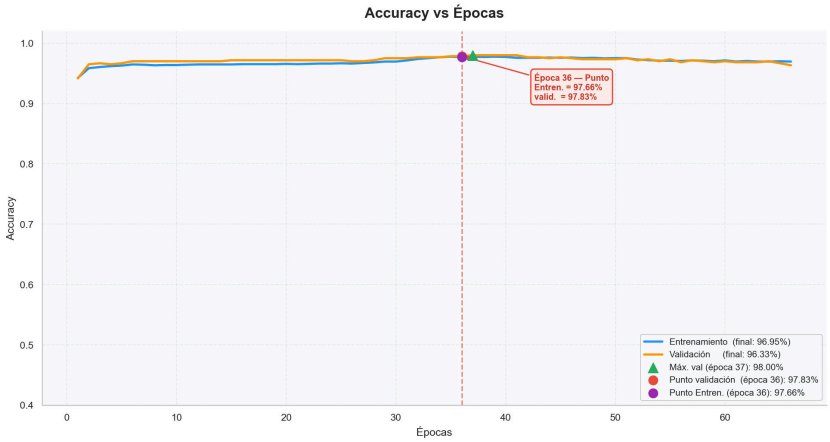


Figura 222 Accuracy vs épocas.
Fuente: Elaboración propia

G) Riesgo de accidente por departamento, provincia y distrito.

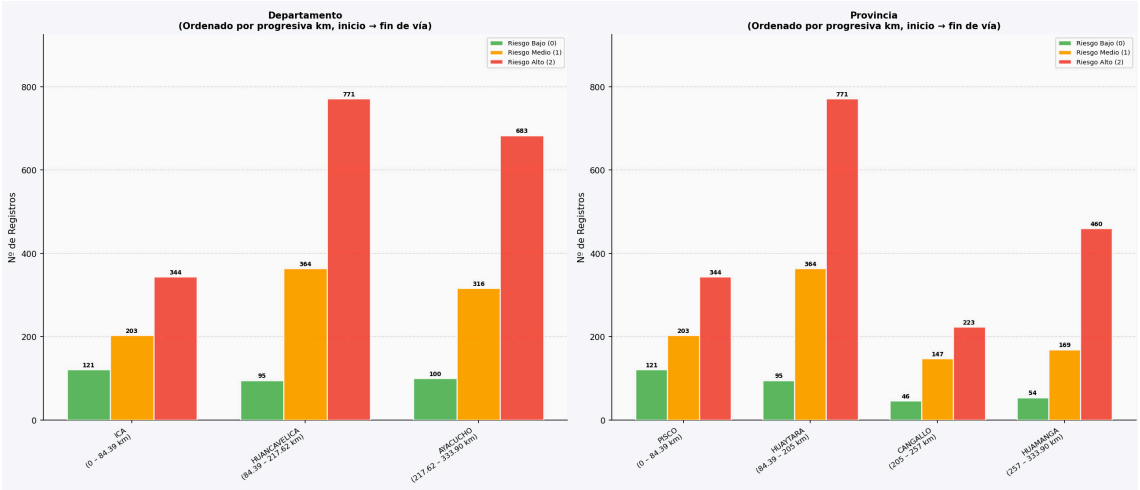


Figura 223 Riesgo de accidentes por departamento y provincia.
Fuente: Elaboración propia.

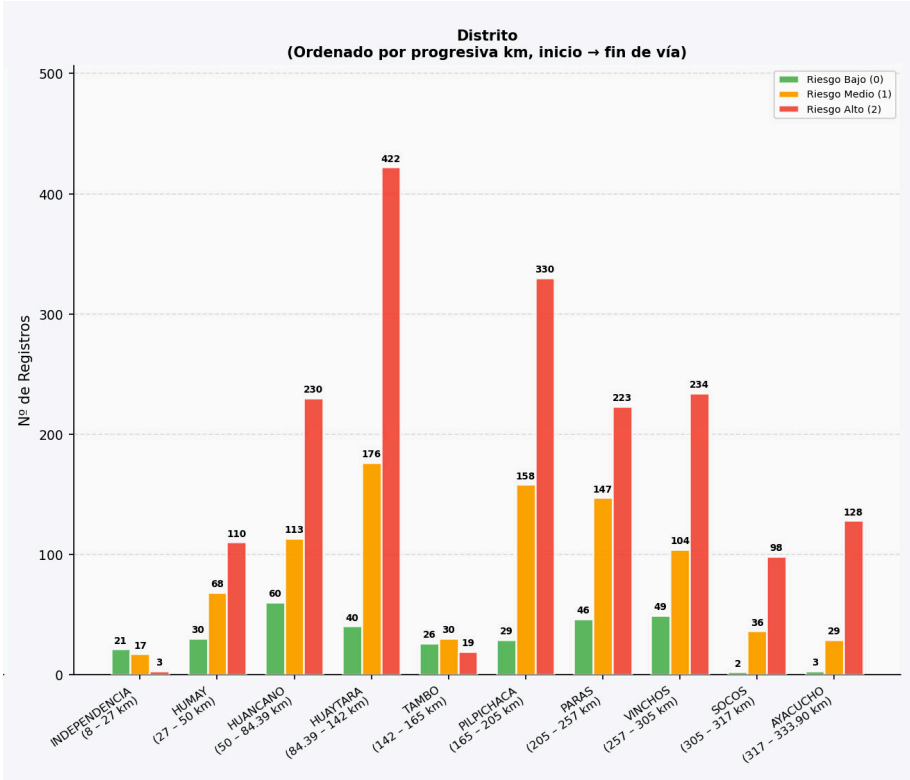


Figura 224 Riesgo de accidentes por distrito.
Fuente: Elaboración propia.

H) Programa RNA N°02: Predicción de accidentes de tránsito

El Modelo RNA N°2 fue exportado como aplicación web denominada **PROGRAMA RNA N°02: Predicción de Nivel de Riesgo Vial**, desarrollada en Python en el entorno JupyterLab. El programa permite clasificar cualquier tramo de la Vía Los Libertadores PE-28A en tres niveles de riesgo: **Riesgo Bajo**, **Riesgo Medio** o **Riesgo Alto** de accidente de tránsito, a

partir del ingreso de 15 variables geométricas y de peligrosidad vial, con funcionamiento en dispositivos móviles para uso en campo.

Figura 225 Paso 1: Abrir el programa
Fuente: Elaboración propia

Enlace de descarga del programa RNA N°02:

Descargar Aqui el programa RNA N° 02 (doble click)

4.1.4 Hipótesis 03: Análisis normativo comparativo

Hipótesis: La comparación de los parámetros del diseño geométrico de la norma DG-2018 con las normativas internacionales de Noruega, Islandia, Suecia, Reino Unido y Japón permite identificar deficiencias en el diseño actual de la Vía Los Libertadores (PE-28A) y formular propuestas de mejora orientadas a optimizar su seguridad vial.

Criterio de aceptación: Identificación de parámetros clasificados como deficientes respecto a las normativas internacionales y formulación de propuestas de mejora técnicamente sustentadas.

Resultado: Se acepta la hipótesis específica c). La comparación sistemática de 12 parámetros geométricos entre la DG-2018 y cinco normativas internacionales reveló deficiencias técnicas significativas, especialmente en peralte máximo (12 % vs. 5,5–10 % en países de

referencia), ausencia de rampas de emergencia obligatorias y carácter no obligatorio de la longitud máxima de recta. Como resultado se formularon siete propuestas de mejora (P1–P7) presentadas en el Cuadro 82.

Tabla 82 *Propuestas de mejora a la DG-2018 sustentadas en normativas internacionales*

N°	Propuesta	Valor propuesto	Referencia	Prioridad
P1	Rampas de emergencia en pendientes ≥6 % y longitud >3 km	Rampa 8–15 %, $L_{mín}$=80 m	Islandia	Alta
P2	Carácter obligatorio a $L_{máx}$ de recta e incorporar $L_{mín}$	$L_{máx} = 20 \times V_h$; $L_{mín} = 6 \times V_h$	Islandia / Noruega	Alta
P3	Fórmulas continuas de fricción lateral	$f = 0{,}62\, e^{-0{,}0063V}$	Suecia	Media
P4	Reducir $e_{máx}$ a 10 % en carreteras 1ra y 2da clase	$e_{máx} = 10\,\%$ (vs. 12 % actual)	Japón	Alta
P5	Peralte variable según velocidad de diseño	10 %→8 %→6 %	Islandia	Media
P6	Sección normativa para vías 2+1	Carriles 800–1 500 m	Noruega/Suecia/UK	Media
P7	Bombeo mínimo 3,0 % en zonas de alta precipitación	$b_{mín} = 3{,}0\,\%$ (vs. 2,0 % actual)	Noruega / Islandia	Media

Fuente: Elaboración propia.

4.1.5 Hipótesis 04: Programa de alerta de accidentes

Hipótesis: El desarrollo e implementación de un programa de alerta de accidentes de tránsito que integre los registros históricos georreferenciados de la PE-28A y rastreo GPS constituye una herramienta de seguridad vial viable para conductores en la Vía Los Libertadores (PE-28A).

Criterio de aceptación: Implementación funcional con detección correcta de zonas de riesgo, activación de alertas dentro del radio configurable y disponibilidad.

Resultado: Se acepta la hipótesis específica. El programa desarrollado como archivo HTML autocontenido integra los 202 registros georreferenciados (2020-2025), rastreo GPS en tiempo real mediante la API de geolocalización del navegador y sistema de alerta (sonoro y vibración). La tabla 83 resume las características verificadas.

Tabla 83 *Características técnicas verificadas del programa de alerta GPS*

Característica	Valor
Registros integrados	202 georref. 2020–2025
Funcionamiento	100 % con internet por el GPS
Cálculo de proximidad	Haversine × 202 puntos
Radio de alerta configurable	100 – 3 000 m
Alerta sonora	Web Audio
Alerta de voz	Sonido ”Riesgo de accidente”
Alerta de vibración	Vibration
Mapa de calor de siniestralidad	Leaflet.heat (azul→rojo)
Compatibilidad	Android, Windows, otros

Fuente: Elaboración propia.

A) Apertura del programa.

Transferir el archivo `TESIS_PROG_ALERTA_DE_ACCIDENTES.html` al dispositivo y abrirlo con el navegador web.

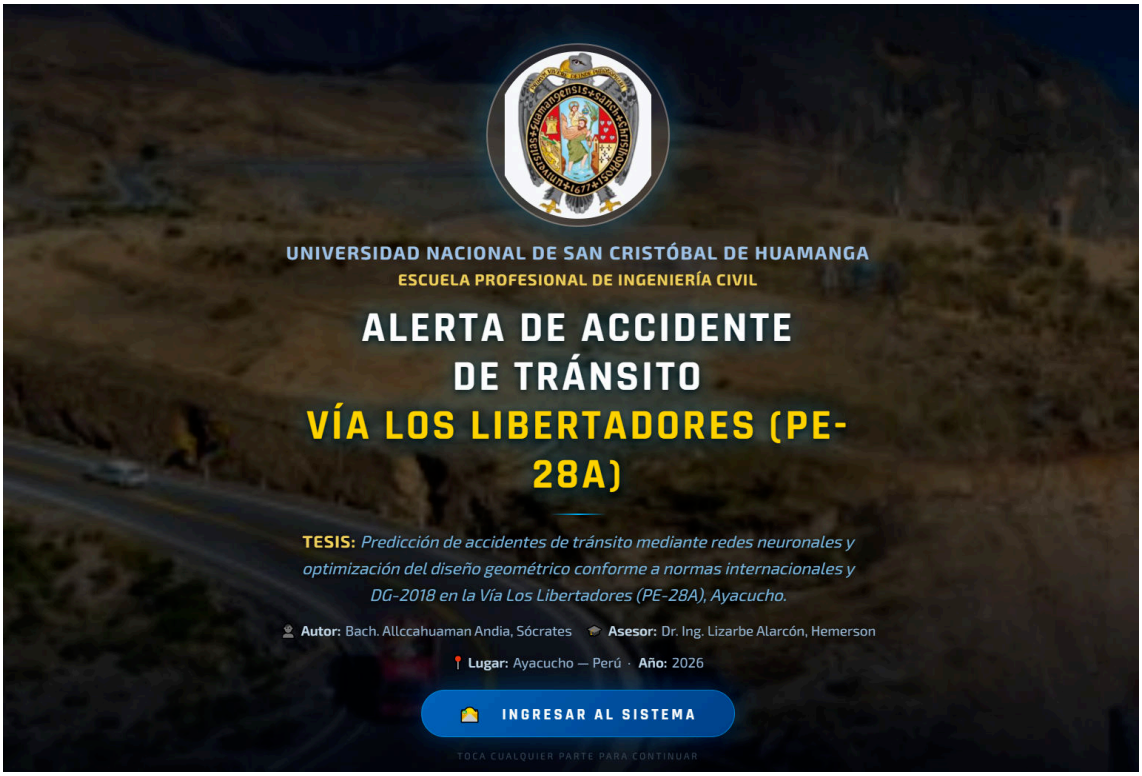


Figura 226 Paso 1: Apertura del programa.
Fuente: Elaboración propia

B) Exploración del mapa.

Al ingresar se visualizará el mapa interactivo de la PE-28A con los 202 puntos de accidente representados por marcadores de colores y el mapa de calor superpuesto.

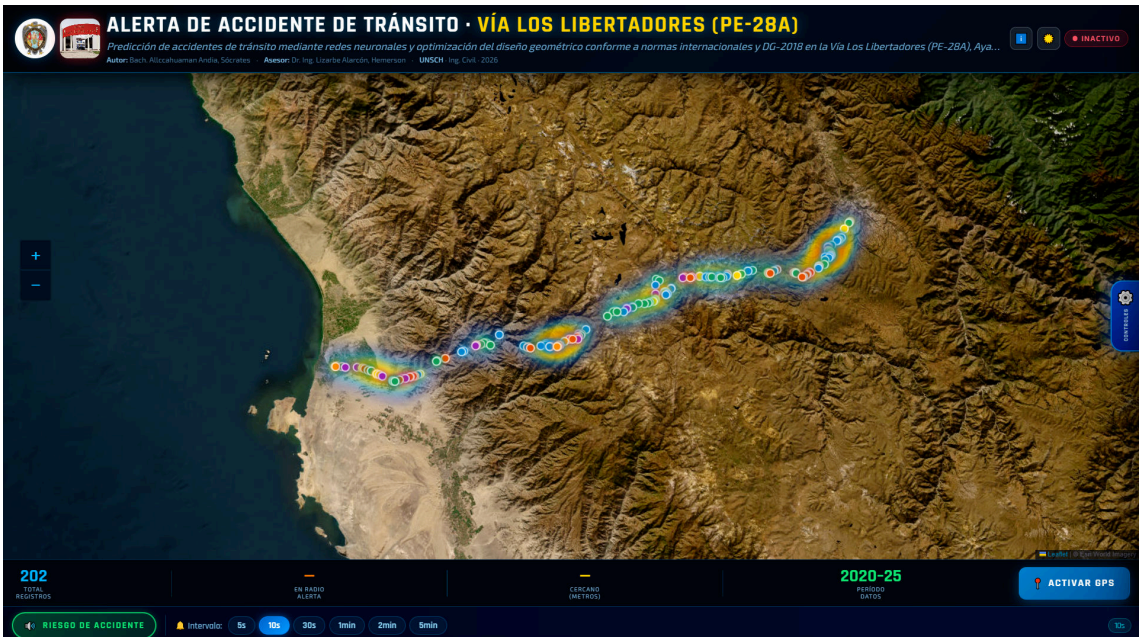


Figura 227 Paso 2: Exploración del mapa.
Fuente: Elaboración propia

Enlace de descarga del programa alerta de accidentes de tránsito:

Descargar Aqui el programa alerta de accidente (doble click)

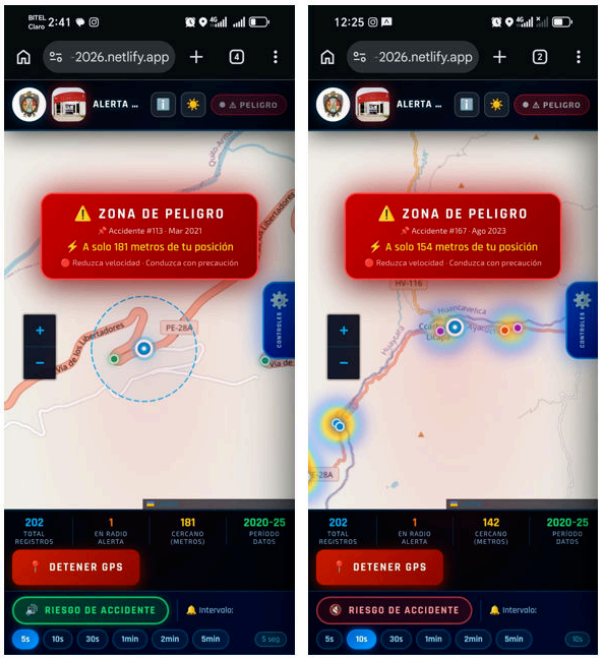


Figura 228 Programa ejecutado en el celular - GPS.
Fuente: Elaboración propia

CAPÍTULO V
Conclusiones y recomendaciones

5.1 Conclusiones

5.1.1 Conclusión general

La presente investigación demostró que la integración de dos modelos de redes neuronales artificiales, el análisis normativo comparativo entre la DG-2018 y cinco normativas internacionales, y el desarrollo de un programa de alerta basado en geolocalización GPS constituyen un enfoque efectivo para la predicción y reducción de accidentes de tránsito en la Vía Los Libertadores (PE-28A), Ayacucho. Los cuatro ejes de investigación superaron los umbrales de excelencia establecidos, confirmando que las redes neuronales artificiales y la optimización normativa del diseño geométrico vial contribuyen de manera directa y verificable a la seguridad vial y la predicción de accidentes de tránsito en carreteras nacionales.

5.1.2 Conclusiones específicas

a) **Modelo RNA N°01 — Clasificación binaria.** El Modelo RNA N°01 (arquitectura 23→64→48→32→16→8→2), entrenado con 202 registros históricos de accidentes de tránsito de la PE-28A (2020–2025) y 23 variables del diseño geométrico, condiciones climáticas, características del siniestro y factores viales conforme a la norma DG-2018, alcanzó un *accuracy* de evaluación de 90,48 %, un AUC-ROC de 0,9818 y un F1-Score macro de 90,45 %, Recall macro de 90,45 %, superando el umbral de excelencia de ≥90 % establecido. El GAP entrenamiento y evaluación es de $\Delta = 5{,}95\,\%$ confirma la ausencia de sobreajuste o overfiting. Las variables de mayor importancia predictiva fueron el peligro por visibilidad, el radio de curvatura horizontal, la pendiente longitudinal y el peligro en curva, resultado consistente con el conocimiento técnico de la norma DG-2018 para carreteras de montaña. Queda demostrado que un modelo de redes neuronales puede alcanzar alta precisión en la predicción de accidentes de tránsito aun con conjuntos de datos de tamaño reducido (202 registros), siempre que se apliquen técnicas adecuadas de regularización y validación.

b) **Modelo RNA N°02 — Clasificación multiclase.** El Modelo RNA N°02 (arquitectura 15→128→64→32→16→3), entrenado con 2,997 segmentos geométricos de la PE-28A modelados en AutoCAD Civil 3D y expandidos a 5,394 registros mediante SMOTE,

alcanzó un *accuracy* de evaluación de 98,00 %, un AUC-ROC macro de 0,9985, F1-Score macro de 96,69 %, Precisión macro de 95,71 % y Recall macro de 97,92 %, superando ampliamente los umbrales de excelencia establecidos. De los 2,997 registros totales, 2,929 fueron clasificados correctamente (97,7 %), con errores concentrados exclusivamente entre clases adyacentes. El GAP entrenamiento y evaluación es de −0,0034 confirma la ausencia de sobreajuste o overfiting. Queda demostrado que la combinación de modelamiento geométrico en Civil 3D, balanceo de clases mediante SMOTE y arquitectura MLP profunda permite predecir con alta precisión el nivel de riesgo (Bajo / Medio / Alto) de tramos viales en carreteras nacionales, constituyendo una herramienta objetiva la priorización de intervenciones de seguridad vial.

c) **Análisis normativo comparativo DG-2018.** La comparación sistemática de los parámetros geométricos entre la DG-2018 y las normativas de Noruega (N100), Islandia (Veghönnunarreglur), Suecia (VGU), Reino Unido (DMRB) y Japón (RSO) permitió identificar deficiencias técnicas significativas en la norma peruana. Las más críticas son: (i) el peralte máximo de 12 % de la DG-2018, frente a valores de 5,5–10 % en los países de referencia; (ii) la ausencia de rampas de emergencia obligatorias en pendientes sostenidas; y (iii) la longitud máxima de recta con carácter no obligatorio. Como resultado se formularon siete propuestas de mejora (P1–P7), de las cuales tres son de alta prioridad (P1: rampas de emergencia; P2: longitudes de recta obligatorias; P4: reducción del peralte máximo a 10 %). Queda demostrado que la norma DG-2018 presenta brechas técnicas verificables respecto a los estándares internacionales de países con los índices de mortalidad vial más bajos del mundo y que su actualización mediante las propuestas formuladas contribuiría directamente a la reducción de accidentes en la PE-28A y en otras carreteras nacionales del Perú.

d) **Programa de alerta de accidentes de tránsito.** El programa de alerta desarrollado como archivo HTML, que integra los 202 registros georreferenciados de la PE-28A (2020-2025), rastreo GPS en tiempo real mediante la API de geolocalización del navegador, cálculo de proximidad por la fórmula de Haversine y sistema de alerta multimodal (sonoro , vocal en español y vibración), constituye una herramienta de seguridad vial activa, viable y operativa para conductores. Su funcionamiento garantiza su disponibilidad en la Via Los Libertadores (PE-28A), siendo de uso práctico en una vía de 333,903 km que atraviesa tres departamentos. Queda demostrado que la geor-

referenciación de registros históricos de siniestros y su integración en una aplicación de geolocalización en tiempo real es una estrategia tecnológicamente viable y de bajo costo para la detección temprana de zonas de riesgo vial.

5.2 Recomendaciones

a) **Para PROVÍAS Nacional y el MTC.** Se recomienda incorporar las siete propuestas de mejora (P1-P7) formuladas en la presente investigación para norma DG-2018, con especial énfasis en las propuestas de alta prioridad: inclusión de rampas de emergencia obligatorias en pendientes $\geq 6\,\%$ con longitud >3 km (P1), carácter obligatorio para la longitud máxima de recta e incorporación de la longitud mínima (P2) y reducción del peralte máximo a $10\,\%$ en carreteras de primera y segunda clase (P4). Estas modificaciones están sustentadas en las normativas de Noruega, Islandia, Suecia, Reino Unido y Japón, países con los índices de mortalidad vial más bajos del mundo.

b) **Para la Policía Nacional del Perú y el ONSV.** Se recomienda ampliar y estandarizar la base de datos de accidentes de tránsito de la PE-28A, incrementando el número de variables registradas por evento (velocidad estimada, condición climática, tipo de maniobra, fatiga del conductor) con el fin de mejorar la precisión y el poder predictivo de futuros modelos de redes neuronales. Un conjunto de datos de 500 o más registros permitiría explorar arquitecturas más profundas y técnicas de aprendizaje transferido.

c) **Para futuros investigadores.** Se recomienda extender la metodología desarrollada en la presente investigación a otras carreteras de la red vial nacional con alta siniestralidad, especialmente aquellas con características topográficas similares a la PE-28A (pendientes pronunciadas, curvas cerradas, zonas de niebla).

5.3 Trabajos futuros

A partir de los resultados y limitaciones identificadas en la presente investigación, se proponen las siguientes líneas de trabajo futuro:

a) **Modelos con datos en tiempo real.** Desarrollar un modelo de predicción de accidentes que integre datos dinámicos en tiempo real condiciones climáticas del SENAMHI, volumen de tránsito del IMD y velocidad vehicular para generar alertas predictivas adaptativas por tramo, superando la limitación de los modelos estáticos desarrollados en la presente investigación.

b) **Extensión a toda la red vial nacional.** Aplicar la metodología de los Modelos RNA N°01 y N°02 a otras carreteras nacionales con alta siniestralidad, construyendo una base de datos unificada que permita desarrollar modelos de predicción de ámbito nacional integrados con el sistema de información geográfica del MTC.

c) **Aplicación móvil.** Desarrollar una versión nativa del programa de alerta de accidentes para Android e iOS, que incluya navegación GPS integrada, actualización automática de la base de datos de accidentes, integración con el Modelo RNA N°02 para mostrar el nivel de riesgo del tramo en tiempo real, y funciones colaborativas para que los conductores reporten nuevos incidentes.

d) **Validación de las propuestas normativas.** Realizar estudios de simulación de tránsito y análisis de conflictos vehiculares en los tramos de la PE-28A para validar cuantitativamente el impacto esperado de las siete propuestas de mejora (P1–P7) a la DG-2018.

Referencias bibliográficas

Alvarez, P., Fernandez, M., Gorliza, A., Mansilla, A. & Molinero, A. (2020). Geometric road design factors affecting the risk of urban run-off crashes. A case-control study. *Plos one*, *15(6)*(e0234564).

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Nueva York: Springer.

Borja, M. (2012). Metodología de la investigación científica para ingenieros. *Chiclayo: sn*.

Boza Pareja, R. (2018). *Análisis comparativo según los criterios de seguridad y operación a nivel de diseño geométrico de la ruta arco Tica Tica - Chinchero usando los manuales de diseño DG 2014 peruano, manual de diseño geométrico Ecuatoriano de carreteras 2003, y manual de diseño geométrico de carreteras Colombiano 2008 en la region del Cusco* (Tesis de pregrado). Universidad Andina del Cusco, Cusco.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32.

Brownlee, J. (2020a). Imbalanced Classification with Python. *Machine Learning Mastery*.

Brownlee, J. (2020b). *Imbalanced Classification with Python: Better Metrics, Balance Skewed Classes, Cost-Sensitive Learning*. Machine Learning Mastery.

Cardenas Achahuanco, J. (2024). *Propuesta de estandarización de diseño geométrico de trochas carrozables en la DG-2018 a través del análisis comparativo de normativas extranjeras* (Tesis de pregrado). Universidad Católica San Pablo, Arequipa.

Chavez Castro, P. (2022). *Estandarización de diseño geométrico a través del análisis de estándares de otros países de Sudamérica para actualizar el Manual de Carreteras - Diseño Geométrico DG 2018* (Tesis de pregrado). Universidad Privada Antenor Orrego, Trujillo.

Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.

Codur, M. & Tortum, A. (2015). An artificial neural network model for highway accident prediction: A case study of Erzurum, Turkey. Promet Traffic and Transportation. *Promet Traffic and Transportation*, *27(3)*(217-225).

Evidently AI. (2023). Precision and Recall in Machine Learning: A Complete Guide. Consultado el 15 de marzo de 2025. Retrieved from https : / / www . evidentlyai . com / classification-metrics/precision-recall

Fawcett, T. (2006). An Introduction to ROC Analysis. *Pattern Recognition Letters*, *27*(8), 861–874.

Garber, N. J. & Hoel, L. A. (2015). *Traffic and Highway Engineering* (5th ed.). Stamford, CT: Cengage Learning.

García Chávez, A. (2012). *Propuesta de mejoramiento de la seguridad vial de una carretera de elevada accidentalidad utilizando tecnologías ITS* (Doctoral dissertation). Tesis de Maestría, Facultad de Ingeniería UNAM.

Garzón, M., Escobar, D. & Galindo, J. (2017). Auditorias de seguridad vial. Ejemplo de aplicación metodológica. *Revista Espacios*, *38*(41).

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media.

Goodfellow, I., Bengio, Y. & Courville, A. (2016). *Deep Learning*. Cambridge, MA: MIT Press.

Google Developers. (2023). Classification: True vs. False and Positive vs. Negative. Consultado el 15 de marzo de 2025. Retrieved from https://developers.google.com/machine-learning/crash-course/classification

Hastie, T., Tibshirani, R. & Friedman, J. (2009). *The Elements of Statistical Learning* (2nd ed.). Springer.

Haykin, S. (2009). *Neural Networks and Learning Machines* (3rd). Nueva York: Pearson.

He, H. & Garcia, E. A. (2009). Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284.

Heredia Martinez, O. & Perez Heredia, J. (2023). *Propuesta de parámetros para el diseño geométrico de trochas carrozables bajo el enfoque de seguridad activa y normas internacionales, aplicado a la carretera Balconcillo-Cruce San Andrés, Cutervo* (Tesis de pregrado). Universidad de Piura, Piura.

Hernandez, R., Fernandez, C. & Baptista, P. (2014). *Metodología de la investigación (6. ed.)* McGraw-Hill.

Hosmer, D. W., Lemeshow, S. & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley.

Islam, M., Teik, H., Hamid, H. & Azarkerdar, A. (2019). Relationship of Accident Rates and Road Geometric Design. *IOP Conference Series: Earth and Environmental Science*, *357(1)*(012040).

Kingma, D. P. & Ba, J. (2015). Adam: A Method for Stochastic Optimization. En *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*.

Kriswardhana, W., Hasanuddin, A. & Palestine, I. (2020). Modelling road traffic accident rate and road geometric parameters relationship. *AIP Conference Proceedings*, *2278(1)*(020003).

LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. (1998). Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324.

Lemaitre, G., Nogueira, F. & Aridas, C. K. (2017). Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *Journal of Machine Learning Research*, *18*(17), 1–5.

LuWang. (2023). Safety Evaluation for Highway Geometric Design Based on Spatial Path Properties. *Journal of Advanced Transportation*, *2023*(6685010).

Maquera Maquera, J. (2023). *Evaluación del diseño geométrico de la carretera mo-518 tramo los Ángeles – Yacango, Moquegua, con la aplicación del Vehicle Tracking de Autocad CIVIL 3D* (Tesis de pregrado). Universidad Nacional del Altiplano, Puno.

Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N. & Terzopoulos, D. (2021). Image Segmentation Using Deep Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(7), 3523–3542. doi:10.1109/TPAMI.2021.3059968

Ministerio de Transportes y Comunicaciones del Perú. (2018). *Manual de Diseño Geométrico de Carreteras DG-2018*. Lima: MTC.

Ministry of Land, Infrastructure, Transport and Tourism. (2021). *Road Structure Ordinance: Doro Kozore (Government Ordinance No. 320)*. MLIT. Decreto original de 1970, con enmiendas hasta 2021. Tokio, Japón.

Minsky, M. & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press.

Montgomery, D. C. & Runger, G. C. (2018). *Applied Statistics and Probability for Engineers*. New York: John Wiley & Sons.

Montoya Correa, D. & Garrido Murillo, V. (2020). *Criterios de diseño, seguridad y operación geométrico en el diseño vertical de carreteras. Análisis comparativo entre los manuales del invias (Colombia), ministerio de fomento (España) y AASHTO (EE.UU.)* (Tesis de pregrado). Universidad Distrital Francisco José de Caldas Facultad de Medioambiente y Recursos naturales Ingeniería Topográfica, Bogota.

Nair, V. & Hinton, G. E. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines. *Proceedings of the 27th International Conference on Machine Learning (ICML)*, 807–814.

National Highways. (2020). *CD 109: Highway Link Design, Revision 1*. Design Manual for Roads and Bridges (DMRB). Guildford, Reino Unido.

National Highways. (2021). *CD 127: Cross-Sections and Headrooms, Version 1.0.1*. Design Manual for Roads and Bridges (DMRB). Guildford, Reino Unido.

Quispe Aquino, N. & Poma Milian, D. (2019). *Norma de estandarización de diseño geométrico de trochas carrozables para modificar la norma del MTC a través del análisis de estándares de otros países* (Tesis de pregrado). Universidad Ricardo Palma, Lima.

Quispe Condor, E. (2014). *Diseño de una Red Neuronal para la Predicción de Accidentes de Tránsito en la Carretera Central Tramo: Tarma-La Merced, Junín 2022* (Tesis de pregrado). Universidad Nacional del Centro del Perú, Huancayo.

Rosenblatt, F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review*, *65*(6), 386–408.

Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986). Learning Representations by Back-Propagating Errors. *Nature*, *323*(6088), 533–536.

Saito, T. & Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE*, *10*(3), e0118432.

Sampieri, R. H., Collado, C. F. & Lucio, M. P. B. (2014). *Metodología de la investigación (6. ed.)* McGraw-Hill.

Sokolova, M. & Lapalme, G. (2009). A Systematic Analysis of Performance Measures for Classification Tasks. *Information Processing and Management*, *45*(4), 427–437.

Statens vegvesen. (2019). *Håndbok N100: Veg- og gateutforming*. Vegdirektoratet. ISBN: 978-82-7207-744-9. Oslo, Noruega.

Thangaraj, S. & Sorimuthu, M. (2019). Influence of Geometrical Design Elements on Traffic Safety. *International Journal of Multidisciplinary Research in Science, Engineering and Technology, 2(11)*(1753-1758).

The Data Scientist. (2023). F1-Score: Definition, Formula and Use Cases. Consultado el 15 de marzo de 2025. Retrieved from https://thedatascientist.com/f1-score/

Trafikverket. (2024a). *KRAV Grundvärden VGU (TRV 2024:148)*. ISBN: 978-91-8045-364-6. Borlänge, Suecia: Trafikverket.

Trafikverket. (2024b). *TRVINFRA-00396: VGU – Vägars och gators utformning, Krav med Rådstext*. Vigente desde 01/01/2025. Borlänge, Suecia: Trafikverket.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Massachusetts: Addison-Wesley.

Vegagerðin. (2010). *Veghönnunarreglur 03: Vegferill (Alineamiento)*. Icelandic Road and Coastal Administration (IRCA). Reikiavik, Islandia.

Vegagerðin. (2011). *Veghönnunarreglur 02: Þversnið (Sección transversal)*. Icelandic Road and Coastal Administration (IRCA). Reikiavik, Islandia.

Wang, J., Wang, L., Zhou, H., Jing, W. & Cheng, J. (2018). A Three-Dimensional Method for Safety Evaluation of Highway Geometric Design. *Transportation Research Board 97th Annual Meeting*, *18*(01849).

CAPÍTULO VI

Anexos

6.1 Fotografias de la Via Los Libertadores (in situ)

Figura 229 Inicio de la carretera Via Los Libertadores - San clemente.
Fuente: Fotografía propia

Figura 230 Final de la carretera Via Los Libertadores - Ayacucho.
Fuente: Fotografía propia

Figura 231 Estado de la Via Los Libertadores - Distr. Humay.
Fuente: Fotografía propia

Figura 232 Estado de la Via Los Libertadores - Distr. Huancano.
Fuente: Fotografía propia

Figura 233 Estado de la Via Los Libertadores - Distr. Pilpichaca.
Fuente: Fotografía propia

Figura 234 Estado de la Via Los Libertadores - Distr. Huaytara.
Fuente: Fotografía propia

Figura 235 Estado de la Via Los Libertadores - Distr. Socos.
Fuente: Fotografía propia

Figura 236 Estado de la Via Los Libertadores - Distr. Vinchos.
Fuente: Fotografía propia

6.2 Matriz de consistencia

MATRIZ DE CONSISTENCIA

PREDICCIÓN DE ACCIDENTES DE TRÁNSITO MEDIANTE REDES NEURONALES Y OPTIMIZACIÓN DEL DISEÑO GEOMÉTRICO CONFORME A NORMAS INTERNACIONALES Y DG-2018 EN LA VÍA LOS LIBERTADORES (PE-28A), AYACUCHO.

PROBLEMA	OBJETIVO	HIPOTESIS	VARIABLES	METODOLOGÍA	POBLACIÓN Y MUESTRA
¿En qué medida el desarrollo de dos modelos de redes neuronales artificiales, el análisis comparativo de la norma DG-2018 con normativas internacionales de diseño geométrico vial, y la implementación de un programa de alerta basado en geolocalización GPS en tiempo real, permiten predecir los accidentes de tránsito y optimizar el diseño geométrico de la Vía Los Libertadores (PE-28A), Ayacucho?.	Predecir los accidentes de tránsito y optimizar el diseño geométrico de la Vía Los Libertadores (PE-28A), Ayacucho, mediante el desarrollo de dos modelos de redes neuronales artificiales, el análisis comparativo de la norma DG-2018 con normativas internacionales de diseño geométrico vial, y la implementación de un programa de alerta basado en geolocalización GPS en tiempo real, con el fin de contribuir a la reducción de los accidentes de tránsito.	La implementación de dos modelos de redes neuronales artificiales para la predicción de accidentes de tránsito, combinada con el análisis comparativo de la norma DG-2018 con normativas internacionales de diseño geométrico vial y la implementación de un programa de alerta basado en geolocalización, contribuirá a la predicción y reducción de los accidentes de tránsito en la Vía Los Libertadores (PE-28A), Ayacucho.	**Variable independiente:** Diseño geométrico.	• **Enfoque de investigación:** Cuantitativo. • **Tipo de investigación:** Aplicada. • **De acuerdo al nivel:** Descriptiva y Explicativo • **Diseño de investigación:** Comparativo, No experimental y Transversal.	**La población:** **a) Unidad física de análisis:** La Vía Los Libertadores (PE-28A), es una carretera que conecta la Carretera Panamericana Sur con la ciudad de Ayacucho, atravesando los departamentos de Ica, Huancavelica y Ayacucho. Esta vía tiene una longitud aproximada de 333,4 km. **b)Unidad normativa:** Normas internacionesles para el diseño geométrico. **Muestra:** **a) Unidad física de análisis:** Puntos críticos de mayor siniestralidad de la Vía Los Libertadores (PE-28A). **b)Unidad normativa:** Normativas de diseño geométrico de Perú (DG-2018), Noruega, Islandia, Suecia, Reino Unido y Japón.
PROBLEMAS ESPECIFICOS	**OBJETIVOS ESPECIFICOS**	**HIPOTESIS ESPECIFICOS**			
¿Qué nivel de precisión alcanza el Modelo RNA N°01 (clasificación binaria: riesgo menor o riesgo mayor), entrenado con registros históricos de accidentes de tránsito de la Vía Los Libertadores (PE-28A), para predecir accidentes de tránsito utilizando variables del diseño geométrico, condiciones climáticas, características del siniestro y factores viales conforme a la norma DG-2018?.	Desarrollar y evaluar el Modelo RNA N°01 (clasificación binaria: riesgo menor o riesgo mayor), entrenado con 202 registros históricos de accidentes de tránsito de la Vía Los Libertadores (PE-28A), para predecir accidentes de tránsito utilizando variables del diseño geométrico, condiciones climáticas, características del siniestro y factores viales conforme a la norma DG-2018.	El Modelo RNA N°01 (clasificación binaria: riesgo Menor o riesgo mayor), entrenado con 202 registros históricos de accidentes de tránsito de la Vía Los Libertadores (PE-28A) y variables del diseño geométrico, condiciones climáticas, características del siniestro y factores viales conforme a la norma DG-2018, alcanza un nivel de precisión alto para la predicción de accidentes de tránsito.			
¿Qué nivel de precisión alcanza el Modelo RNA N°02 (clasificación multiclase: Riesgo Bajo, Riesgo Medio y Riesgo Alto), entrenado con registros de datos de la Vía Los Libertadores (PE-28A) modelados en AutoCAD Civil 3D, para predecir el nivel de riesgo de accidentes de tránsito en tramos viales?.	Desarrollar y evaluar el Modelo RNA N°02 (clasificación multiclase: riesgo bajo, riesgo medio y riesgo alto), entrenado con 2,997 registro de datos de la Vía Los Libertadores (PE-28A) modelados en AutoCAD Civil 3D, para predecir el nivel de riesgo de accidentes de tránsito en tramos viales.	El Modelo RNA N°02 (clasificación multiclase: riesgo bajo, riesgo medio y riesgo alto), entrenado con 2997 registros de datos geométricos de la Vía Los Libertadores (PE-28A) modelados en AutoCAD Civil 3D, alcanza un nivel de precisión alto para predecir el nivel de riesgo de accidentes de tránsito en tramos viales.	**Variable dependiente:** Accidentes de tránsito.		
¿En qué medida los parámetros del diseño geométrico de la norma DG-2018 presentan deficiencias respecto a las normativas internacionales de Noruega, Islandia, Suecia, Reino Unido y Japón, y qué mejoras pueden proponerse para optimizar la seguridad vial de la Vía Los Libertadores (PE-28A)?	Comparar los parámetros del diseño geométrico de la norma DG-2018 con las normativas internacionales de Noruega, Islandia, Suecia, Reino Unido y Japón, identificar las deficiencias del diseño actual de la Vía Los Libertadores (PE-28A) y proponer mejoras orientadas a optimizar su seguridad vial.	La comparación de los parámetros del diseño geométrico de la norma DG-2018 con las normativas internacionales de Noruega, Islandia, Suecia, Reino Unido y Japón permite identificar deficiencias en el diseño actual de la Vía Los Libertadores (PE-28A) y formular propuestas de mejora orientadas a optimizar su seguridad vial.			
¿En qué medida un programa de alerta de accidentes de tránsito que integre registros históricos georreferenciados de la Vía Los Libertadores (PE-28A) y rastreo GPS en tiempo real, constituye una herramienta de seguridad vial viable para conductores?.	Desarrollar e implementar un programa de alerta de accidentes de tránsito para la Vía Los Libertadores (PE-28A), que integre registros históricos georreferenciados y rastreo GPS en tiempo real, como herramienta de seguridad vial para conductores.	El desarrollo e implementación de un programa de alerta de accidentes de tránsito que integre los registros históricos georreferenciados de la PE-28A y rastreo GPS constituye una herramienta de seguridad vial viable para conductores en la Vía Los Libertadores (PE-28A).			

6.3 Base de datos de RNA N°01.

Enlace de descarga de datos en excel:

https://drive.google.com/file/d/16tQvbip0EF1UbVCIBYTmt5xKmBbHdFQm/view?usp=sharing

N°	MES DE SINIESTRO	DIA DE SINIESTRO	HORA DE SINIESTRO	CLIMA DESPEJADO	CLIMA NUBLADO	CLIMA LLUVIOSO	TEMPERATURA MAX (°C)	PRECIPITACIÓN (mm/día)	CURVA O TRAMO RECTO	PELIGRO EN PENDIENTE	RADIO DE CURVA(M)	PELIGRO EN CURVA	LONGITUD DE CURVA (L)	LONGITUD DE SUBTANGENTE (T)	ANGULO DE DEFLEXION (°)	LONGITUD DE CUERDA (LC)	DISTANCIA EXTERNA (E)	DISTANCIA DE LA ORDENA MEDIA (M)	PELIGRO POR VISIBILIDAD	TRAMO TANGENTE EN 'S'	TRAMO TANGENTE EN 'O'	LONG. TANGENTE	PELIGRO ESTADO DEL PAVIMENTO	ACCIDENTE DE TRANSITO
1	1	5	1	0	1	0	15	0.2	1	1	76.3	3	49.6	25.71	37.2451	48.73	4.22	3.99	2	0	0	0	1	1
2	1	4	3	0	0	1	15.4	23.4	1	1	30.81	3	70.18	66.87	130.5334	55.96	42.82	17.92	2	0	0	0	1	1
3	2	1	1	1	0	0	14	0	0	0	0	0	0	0	0	0	0	0	1	1	0	289.18	1	0
4	2	4	2	0	1	0	15.2	1.6	0	0	0	0	0	0	0	0	0	0	1	1	0	350.5	1	0
5	2	7	2	0	0	1	13.5	3.2	1	0	181.02	2	140.22	73.84	44.3816	136.74	14.48	13.41	1	0	0	0	1	1
6	2	1	2	1	0	0	27.6	0	0	0	190.27	2	66.69	33.69	20.0822	66.35	2.96	2.91	1	0	0	0	1	0
7	2	6	2	0	1	0	21.2	1.2	0	0	0	0	0	0	0	0	0	0	1	1	0	220.46	1	0
8	2	7	1	0	1	0	15.2	1.6	1	1	42	3	44.62	24.68	60.8781	42.55	6.71	5.79	2	0	0	0	1	1
9	2	7	3	1	0	0	27.6	0	0	0	0	0	0	0	0	0	0	0	1	1	0	141.21	0	1
10	2	3	2	0	1	0	15.2	1.6	1	0	153.37	2	132.93	70.97	49.6619	128.81	15.62	14.18	1	0	0	0	1	1
11	3	5	1	0	1	0	17.6	0.3	0	0	0	0	0	0	0	0	0	0	1	1	0	404.81	1	0
12	4	3	1	1	0	0	15.2	0	1	0	150.84	2	96.14	49.77	36.5184	94.52	8	7.6	1	0	0	0	1	1
13	8	3	4	1	0	0	23.9	0	0	0	0	0	0	0	0	0	0	0	1	1	0	226.36	1	0
14	9	6	3	1	0	0	16.7	0	1	0	371.28	1	107.39	54.07	16.5725	107.02	3.92	3.88	0	0	0	0	1	0
15	9	5	2	1	0	0	16.7	0	1	1	32.64	3	58.03	40.21	101.8744	50.68	19.15	12.07	2	0	0	0	1	1
16	11	2	1	1	0	0	18.5	0	1	0	312.97	1	138.08	70.18	25.2796	136.97	7.77	7.58	0	0	0	0	1	0
17	11	3	3	1	0	0	28.4	0	1	1	59.51	3	77.69	45.5	74.8047	72.29	15.4	12.24	2	0	0	0	1	1
18	11	3	1	1	0	0	26.6	0	0	1	0	0	0	0	0	0	0	0	1	1	0	133.4	1	1
19	12	4	2	0	0	1	23	6.8	0	0	0	0	0	0	0	0	0	0	0	1	0	679.12	1	0
20	12	4	4	1	0	0	14.7	0	0	0	0	0	0	0	0	0	0	0	1	1	0	447.191	1	0
21	12	5	1	0	0	1	15.2	1.6	0	0	0	0	0	0	0	0	0	0	0	1	0	1746.79	1	0
22	12	6	2	0	0	1	18	6.3	1	1	126.51	3	102.47	54.23	46.4073	99.69	11.13	10.23	2	0	0	0	1	1
23	12	4	2	0	0	1	19.7	2.9	1	1	86.74	3	51.98	26.8	34.3377	51.21	4.05	3.87	2	0	0	0	1	1
24	1	7	4	1	0	0	30.2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1238.11	1	0
25	11	5	4	1	0	0	26.8	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1238.11	1	0
26	12	1	3	1	0	0	14.4	0	1	0	146.84	3	70.29	35.83	27.4242	69.62	4.31	4.19	0	0	0	0	1	0
27	2	5	4	1	0	0	29.2	0	1	1	114.1	3	134.84	76.54	67.7075	127.13	23.29	19.34	1	0	0	0	0	1
28	2	7	3	1	0	0	27.6	0	0	0	0	0	0	0	0	0	0	0	1	1	0	142.53	0	0
29	1	2	4	1	0	0	30.2	0	0	0	263.88	1	94.16	47.58	20.4442	93.66	4.26	4.19	1	0	0	0	0	0
30	2	7	4	1	0	0	27.6	0	1	1	77.51	3	105.61	62.84	78.0714	97.63	22.28	17.3	2	0	0	0	0	1
31	1	3	4	1	0	0	30.2	0	1	0	70.85	3	41.06	21.13	33.2067	40.49	3.08	2.95	2	0	0	0	0	1
32	1	7	3	1	0	0	30.2	0	1	0	142.88	3	69.85	35.64	28.0114	69.16	4.38	4.25	2	0	0	0	0	1
33	3	1	1	0	0	1	17	3.4	1	1	16.91	3	38.77	37.4	131.3421	30.82	24.14	9.94	2	0	0	0	1	1
34	2	2	3	1	0	0	27.6	0	1	1	18.71	3	35.53	26.14	108.827	30.43	13.44	7.82	2	0	0	0	1	1
35	1	2	4	0	0	1	30.2	39.6	1	1	26.46	3	38.98	24	84.4138	35.55	9.26	6.86	2	0	0	0	1	1
36	2	1	2	1	0	0	27.6	0	1	1	26.46	3	38.98	24	84.4138	35.55	9.26	6.86	2	0	0	0	1	1
37	2	1	2	1	0	0	27.6	0	1	1	69.64	3	50.89	26.64	41.8722	49.77	4.92	4.6	2	0	0	0	1	1
38	2	3	2	0	1	0	15.2	1.6	1	1	30.49	3	54.84	38.36	103.0364	47.74	18.51	11.52	2	0	0	0	1	1
39	2	7	1	0	1	0	15.2	1.6	1	1	29.04	3	67.19	66.09	132.5566	53.18	43.15	17.36	2	0	0	0	1	1
40	1	2	1	1	0	0	30.2	0	1	1	29.04	3	67.19	66.09	132.5566	53.18	43.15	17.36	2	0	0	0	1	1

41	3	1	3	0	0	1	17.6	1.3	1	0	68.21	3	67.21	36.61	56.4512	64.52	9.21	8.11	2	0	0	0	1	1
42	3	5	1	0	1	0	17.6	0.3	1	0	102.25	3	70.86	36.92	39.7091	69.45	6.46	6.08	1	0	0	0	1	1
43	1	4	2	0	0	1	30.2	32.1	0	0	0	0	0	0	0	0	0	0	0	1	0	1275.47	1	0
44	3	7	4	0	0	1	17	3.4	0	0	0	0	0	0	0	0	0	0	0	1	0	1286.88	1	0
45	2	5	3	0	0	1	13.5	3.2	1	0	192.24	2	136.16	71.07	40.5812	133.33	12.72	11.93	1	0	0	0	1	0
46	12	5	1	0	0	1	15.2	1.6	1	0	192.24	2	136.16	71.07	40.5812	133.33	12.72	11.93	1	0	0	0	1	0
47	3	2	2	0	0	1	17.6	3.7	1	0	97.75	3	122.9	71.07	72.0361	114.97	23.1	18.69	2	0	0	0	1	1
48	2	4	2	0	1	0	15.2	1.6	1	0	159.3	2	108.03	56.19	38.8555	105.97	9.62	9.07	1	0	0	0	1	0
49	3	5	3	0	0	1	17.6	8.1	0	0	0	0	0	0	0	0	0	0	0	0	1	1482.35	1	0
50	2	5	3	0	1	0	15.2	1.6	1	1	108.96	3	162.69	100.8	85.5447	147.99	39.48	28.98	1	0	0	0	1	0
51	2	7	1	0	0	1	13.5	3.2	1	1	52.97	3	109.7	89.31	118.6604	91.12	50.87	25.95	2	0	0	0	1	1
52	9	6	3	1	0	0	16.7	0	1	1	86.66	3	141.7	92.42	93.6837	126.44	40.04	27.38	2	0	0	0	1	1
53	9	5	2	1	0	0	16.7	0	1	0	57.92	3	79.5	47.45	78.6499	73.41	16.95	13.11	2	0	0	0	1	1
54	2	7	2	0	0	1	13.5	3.2	1	1	83.21	3	58.61	30.58	40.3519	57.4	5.44	5.11	1	0	0	0	1	1
55	1	5	1	0	0	1	11	7.6	1	1	64.49	3	132.49	106.72	117.7112	110.39	60.2	31.14	2	0	0	0	1	1
56	10	5	3	1	0	0	26.9	0	0	1	0	0	0	0	0	0	0	0	1	0	1	201.407	1	0
57	12	5	4	1	0	0	14.7	0	1	1	68.27	3	66.07	35.88	55.4497	63.52	8.85	7.84	1	0	0	0	1	1
58	12	6	2	1	0	0	14.7	0	0	1	0	0	0	0	0	0	0	0	1	0	1	184.9	1	0
59	2	4	2	1	0	0	14	0	1	1	34.87	3	58.53	38.84	96.1604	51.9	17.33	11.57	2	0	0	0	1	1
60	11	2	1	1	0	0	18.5	0	0	0	0	0	0	0	0	0	0	0	1	0	1	181.31	1	0
61	2	1	1	0	0	0	14	0	1	0	100	3	25.83	12.99	14.7994	25.76	0.84	0.83	1	0	0	0	1	1
62	3	1	3	1	0	0	22.6	0	1	1	100	3	19.11	9.58	10.9489	19.08	0.46	0.46	2	0	0	0	1	1
63	2	3	4	1	0	0	19.8	0	1	0	238.73	2	33.37	16.71	8.0096	33.35	0.58	0.58	1	0	0	0	1	0
64	4	3	1	1	0	0	21.8	0	1	1	258.01	1	252.76	137.56	56.1311	242.77	34.38	30.34	1	0	0	0	1	1
65	2	6	4	0	0	1	21.6	2.2	1	1	54.29	3	123.75	118.05	130.6085	98.64	75.65	31.61	2	0	0	0	1	1
66	12	4	2	0	0	1	19.7	2.9	1	1	45.41	3	94.95	78.34	119.8085	78.57	45.14	22.64	2	0	0	0	1	1
67	11	3	3	1	0	0	28.4	0	1	1	87.06	3	167.88	125.47	110.4908	143.05	65.66	37.43	2	0	0	0	1	1
68	2	6	2	0	1	0	21.2	1.2	1	1	54.59	3	71.2	41.68	74.73	66.26	14.09	11.2	2	0	0	0	1	1
69	9	1	3	0	0	1	23.6	9.3	1	1	63.45	3	103.51	67.41	93.4649	92.41	29.12	19.96	2	0	0	0	1	1
70	1	3	3	0	0	1	18.4	4.1	0	0	0	0	0	0	0	0	0	0	0	1	0	511.33	1	0
71	11	3	1	1	0	0	26.6	0	1	1	80.15	3	38.89	19.83	27.7992	38.51	2.42	2.35	2	0	0	0	1	1
72	8	3	4	1	0	0	23.9	0	1	1	29.25	3	40.17	23.98	78.7049	37.09	8.58	6.63	2	0	0	0	1	1
73	12	6	2	0	0	1	18	6.3	1	1	36.25	3	80.18	72.27	126.7215	64.81	44.6	20	2	0	0	0	1	1
74	1	4	3	0	0	1	15.4	23.4	1	1	35.83	3	102.66	257.52	164.1573	70.98	224.17	30.89	2	0	0	0	1	1
75	1	7	2	0	0	1	16	6	1	0	71.06	3	59.48	31.61	47.9546	57.76	6.71	6.13	1	0	0	0	1	1
76	1	3	2	1	0	0	29.4	0	1	1	69.33	3	66.66	36.16	55.0894	64.12	8.86	7.86	2	0	0	0	1	1
77	1	4	4	0	0	1	11	7.6	0	0	0	0	0	0	0	0	0	0	1	1	0	184.98	1	0
78	2	4	3	0	0	1	13.5	3.2	0	0	0	0	0	0	0	0	0	0	1	0	1	250.39	1	0
79	2	3	3	1	0	0	27.6	0	1	0	42.26	3	33.63	17.76	45.5881	32.75	3.58	3.3	1	0	0	0	0	1
80	2	2	2	1	0	0	14	0	0	0	0	0	0	0	0	0	0	0	1	1	0	97	1	0
81	3	1	1	0	0	1	18.6	9.8	1	1	81.08	3	93.08	52.43	65.7776	88.05	15.47	12.99	1	0	0	0	1	1
82	3	6	1	1	0	0	28	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1905.82	1	0
83	4	5	4	1	0	0	21.8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	547.19	1	0
84	6	3	4	1	0	0	19.7	0	1	0	196.76	2	202.65	111.35	59.0119	193.81	29.32	25.52	1	0	0	0	1	0
85	6	7	1	1	0	0	16.4	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1482.35	1	0
86	7	4	2	1	0	0	24	0	1	0	43.9	3	39.35	21.11	51.3599	38.04	4.81	4.34	1	0	0	0	0	1
87	7	1	1	1	0	0	22.6	0	1	1	39.3	3	61.51	39.08	89.6878	55.42	16.13	11.43	2	0	0	0	1	1
88	9	4	1	1	0	0	18.2	0	1	0	196.93	2	166.98	88.88	48.5835	162.03	19.13	17.44	1	0	0	0	1	0
89	10	6	1	1	0	0	17.8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2749.44	1	0
90	10	7	4	1	0	0	29.6	0	0	0	0	0	0	0	0	0	0	0	1	0	1	228.49	0	0
91	1	7	2	0	0	1	16	6	1	1	91.36	3	169.65	122.12	106.4015	146.3	61.15	36.63	1	0	0	0	1	1
92	1	4	4	0	0	1	11	7.6	1	0	62.99	3	36.88	18.99	33.5484	36.36	2.8	2.68	2	0	0	0	1	1
93	2	4	1	0	0	1	13.5	3.2	1	0	59.42	3	85.88	52.4	82.8124	78.6	19.8	14.85	2	0	0	0	1	1
94	2	1	1	0	1	0	15.2	1.6	1	0	192.44	2	192.25	105.01	57.238	184.35	26.78	23.51	1	0	0	0	1	0
95	2	3	1	1	0	0	27.6	0	0	0	0	0	0	0	0	0	0	0	1	0	1	420.37	0	0
96	2	3	3	0	1	0	12.5	0.4	1	1	34.87	3	58.53	38.84	96.1604	51.9	17.33	11.57	2	0	0	0	1	1
97	2	2	2	1	0	0	14	0	1	0	258.01	1	252.76	137.56	56.1311	242.77	34.38	30.34	1	0	0	0	1	0
98	2	2	2	1	0	0	14	0	1	0	258.01	1	252.76	137.56	56.1311	242.77	34.38	30.34	1	0	0	0	1	0
99	2	6	2	1	0	0	29	0	0	0	0	0	0	0	0	0	0	0	0	0	1	412.55	1	0
100	3	1	2	0	0	1	18.6	9.8	1	1	65.42	3	119.89	85.27	105.0106	103.8	42.05	25.6	2	0	0	0	1	1

101	3	6	1	1	0	0	28	0	0	1	0	0	0	0	0	0	0	0	1	0	5814.45	1	**0**
102	4	5	2	1	0	0	21.8	0	0	1	1	145.07	3	159.51	88.9	63.0002	151.6	25.07	21.38	1	0	0	0	0	1	0	**0**
103	5	3	1	1	0	0	22.2	0	1	1	1	44.28	3	73.77	48.71	95.4567	65.53	21.55	14.49	2	0	0	0	0	1	0	**1**
104	6	6	2	1	0	0	24.2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1238.11	1	0	1	**0**
105	6	6	3	1	0	0	18.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2807.25	1	0	1	**0**
106	6	4	4	0	1	0	13.9	1	1	1	1	36.71	3	65.46	45.47	102.1696	57.12	21.73	13.65	2	0	0	0	0	1	0	**1**
107	6	7	1	1	0	0	16.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1482.35	1	0	1	**0**
108	7	4	2	1	0	0	24	0	0	1	1	46.92	3	109.34	109.27	133.5249	86.22	72	28.41	2	0	0	0	0	0	1	**1**
109	7	1	1	1	0	0	22.6	0	0	1	1	39.3	3	61.51	39.08	89.6878	55.42	16.13	11.43	2	0	0	0	0	1	0	**1**
110	8	3	4	1	0	0	18.2	0	0	1	0	283.52	1	165.08	84.95	33.3599	162.76	12.45	11.93	1	0	0	0	0	1	0	**0**
111	8	6	2	1	0	0	13	0	1	1	1	58.07	3	74.33	43.24	73.3416	69.36	14.33	11.49	1	0	0	0	0	1	0	**1**
112	9	2	3	0	0	1	20.2	2.9	1	0	0	144.43	3	164.33	92.35	65.1891	155.61	27	22.75	1	0	0	0	0	1	0	**1**
113	9	5	2	1	0	0	18.2	0	1	1	1	21.65	3	40.37	29.18	106.8399	34.77	14.68	8.75	2	0	0	0	0	1	1	**1**
114	9	4	1	1	0	0	18.2	0	1	1	1	58.34	3	145.73	175.01	143.1299	110.69	126.14	39.89	1	0	0	0	0	1	1	**1**
115	4	4	2	1	0	0	19.8	0	1	1	2	173.02	2	72.43	36.75	23.9838	71.9	3.86	3.78	1	0	0	0	0	1	0	**0**
116	5	6	4	1	0	0	27.6	0	0	1	3	29.13	3	54.14	39	106.4898	46.68	19.55	11.7	2	0	0	0	0	1	0	**1**
117	3	2	1	0	1	0	14.1	3.8	0	0	0	0	0	0	0	0	0	0	0	1	1	0	210.5	1	0		**0**
118	10	6	1	1	0	0	17.8	0	1	1	1	29.13	3	54.14	39	106.4898	46.68	19.55	11.7	2	0	0	0	0	1	0	**1**
119	4	1	2	1	1	0	12.5	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	318.205	1	0		**0**
120	5	3	3	1	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2215.02	1	0		**0**
121	12	7	3	1	0	0	17.8	0.4	1	1	1	54.15	3	44.82	23.78	47.4203	43.55	4.99	4.57	2	0	0	0	0	1	0	**1**
122	11	6	3	1	0	0	26.8	0	0	0	0	66.58	3	56.85	30.29	48.9258	55.14	6.57	5.98	1	0	0	0	0	0	0	**1**
123	12	1	1	0	0	1	14.7	0	1	0	2	213.62	2	48.47	24.34	13.0012	48.37	1.38	1.37	0	0	0	0	0	1	0	**0**
124	8	6	1	1	0	0	13	0	1	0	3	125.24	3	101.96	53.99	46.6422	99.16	11.14	10.23	1	0	0	0	0	1	1	**1**
125	7	6	1	0	1	0	14.6	2.6	1	1	3	116.29	3	220.68	162.22	108.7262	189.03	83.3	48.54	2	0	0	0	0	1	1	**1**
126	12	5	2	1	0	0	16.2	0.2	1	1	3	115.76	3	70.32	82.0104	106.13	26.29	19.84	1	0	0	0	0	0	1	1	**1**
127	5	3	1	1	0	0	22.2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	212.22	1	0		**0**
128	11	5	4	1	0	0	18.6	0	0	1	0	119.36	3	196.32	128.54	94.2418	174.93	56.05	38.14	2	0	0	0	0	1	0	**1**
129	6	7	1	1	0	0	13.2	0.2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	538.51	1	0		**0**
130	10	5	4	1	0	0	27.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2807.25	1	0		**0**
131	1	6	2	1	0	0	11.6	0.2	0	1	0	104.99	3	71.95	37.45	39.2657	70.55	6.48	6.1	1	0	0	0	0	1	1	**1**
132	2	6	2	1	0	0	29	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1523.1	1	0		**0**
133	3	6	1	1	0	0	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1905.82	1	0		**0**
134	7	4	3	1	0	0	13.8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	667.29	1	0		**0**
135	10	5	1	1	0	0	26.6	0	0	0	0	1052.79	1	71.16	35.59	3.8726	71.14	0.6	0.6	0	0	0	0	0	1	0	**1**
136	3	4	4	0	0	1	17.5	7.4	1	0	0	33.07	3	26.69	14.12	46.2364	25.97	2.89	2.66	1	0	0	0	0	1	0	**1**
137	10	1	4	1	0	0	21.6	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	154.01	1	0		**0**
138	3	5	3	0	0	1	17.6	8.1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	404.81	1	0		**0**
139	3	7	2	0	1	0	11.8	7.8	0	0	0	0	0	0	0	0	0	0	0	2	1	0	89.18	1	0		**0**
140	5	7	2	1	0	0	27.8	0	1	0	3	149.81	3	74.27	37.91	28.4041	73.51	4.72	4.58	0	0	0	0	0	0	0	**0**
141	6	5	2	1	0	0	19	0	1	1	3	62.85	3	153.17	171	139.6397	117.98	119.34	41.17	2	0	0	0	0	0	1	**1**
142	7	1	1	0	0	1	20.2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	112.25	0	1	0	**0**
143	4	6	3	1	0	0	20.3	0	0	1	1	55.1	3	141.22	185.1	146.8459	105.62	138.03	39.38	1	0	0	0	0	1	0	**0**
144	4	1	3	1	0	0	18.3	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	350.5	1	0		**0**
145	2	6	2	1	0	0	28.8	0	0	1	0	136.19	3	88.41	45.83	37.1939	86.87	7.5	7.11	1	0	0	0	0	1	0	**0**
146	11	7	3	1	0	0	18.5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	150	1	0		**0**
147	12	2	1	1	0	0	25	0.9	1	1	0	87.06	3	167.88	125.47	110.4908	143.05	65.66	37.43	2	0	0	0	0	1	1	**1**
148	1	6	3	1	0	0	15	0.2	1	0	0	87.81	3	42.47	21.66	27.7122	42.06	2.63	2.56	1	0	0	0	0	0	1	**1**
149	1	7	1	1	0	0	30.2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	5814.45	1	0		**0**
150	1	1	1	1	0	0	30.2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	412.55	1	0		**0**
151	12	6	2	1	0	0	26.8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	5814.45	1	0		**0**
152	5	3	2	1	0	0	25.6	0	0	1	0	1654.81	1	193.23	96.72	6.6903	193.12	2.82	2.82	0	0	0	0	0	1	0	**0**
153	6	6	4	1	0	0	19.8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	975.89	1	0		**0**
154	3	7	4	0	1	0	17	3.4	1	1	1	101.94	3	90.33	48.37	50.7659	87.4	10.89	9.84	2	0	1	0	1	1	0	**1**
155	7	4	2	1	0	0	16.1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	145.78	1	0		**0**

156	12	4	2	1	0	0	28.6	0	0	0	0	0	0	0	0	0	1	1	0	121.83	1	**0**
157	10	7	4	1	0	0	26.9	0	0	0	0	0	0	0	0	0	1	0	1	378.16	1	**0**
158	10	7	4	1	0	0	27.2	0	0	0	0	0	0	0	0	0	0	0	1	2215.023	1	**0**
159	10	1	3	1	0	0	20.4	0	1	1	44.65	3	64.33	39.19	82.5504	58.91	14.76	11.09	2	0	0	0	1	**1**
160	7	6	1	1	0	0	24	0	0	0	0	0	0	0	0	0	0	0	0	1	0	676.08	1	**0**
161	3	5	4	1	0	0	29.4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	5814.45	1	**0**
162	9	2	2	1	0	0	23.8	0	1	1	45.41	3	94.95	78.34	119.8085	78.57	45.14	22.64	2	0	0	0	1	**1**
163	9	1	4	1	0	0	20.6	0.5	0	0	0	0	0	0	0	0	0	0	0	1	0	679.12	1	**0**
164	8	6	1	1	0	0	20.1	0	1	1	34.43	3	79.97	79.36	133.0951	63.16	52.07	20.73	2	0	0	0	1	**1**
165	4	1	2	1	0	0	17.2	0	1	0	226.92	2	237.54	130.95	59.9775	226.85	35.08	30.38	1	0	0	0	1	**0**
166	3	7	1	1	0	0	21.4	0.1	0	1	63.66	3	85.77	50.81	77.1941	79.43	17.79	13.91	0	0	0	0	1	**1**
167	8	1	3	0	1	0	17.3	0	0	0	121.62	3	59.36	30.28	27.9634	58.77	3.71	3.6	1	0	0	0	1	**1**
168	1	3	2	1	0	0	32.2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2807.25	1	**0**
169	1	4	3	1	0	0	29.2	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1905.82	1	**0**
170	12	7	2	1	0	0	28.8	0	0	0	0	0	0	0	0	0	0	0	0	1	0	412.55	1	**0**
171	4	3	2	1	0	0	29.2	0	0	0	0	0	0	0	0	0	0	0	1	0	1447.5	1	**0**
172	2	6	2	1	0	0	31.4	0	0	0	0	0	0	0	0	0	0	0	1	0	826.08	0	**0**
173	9	5	3	1	0	0	22.6	0	1	0	123.77	3	46.41	23.48	21.4827	46.14	2.21	2.17	1	0	0	0	0	**1**
174	9	7	1	1	0	0	27.8	0	1	0	121.02	3	30.03	15.09	14.2177	29.95	0.94	0.93	2	0	0	0	0	**1**
175	12	1	1	1	0	0	18.7	0	1	1	63.76	3	50.88	26.88	45.7151	49.54	5.43	5.01	2	0	0	0	1	**1**
176	6	3	4	0	0	1	21.5	0	0	1	0	0	0	0	0	0	0	0	1	1	0	245.52	1	**0**
177	6	5	3	1	0	0	15.3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	3454.47	1	**0**
178	7	2	1	1	0	0	16.8	0.2	1	1	51.68	3	45.86	24.56	50.8456	44.37	5.54	5	2	0	0	0	1	**1**
179	7	2	1	1	0	0	16.8	0.2	1	0	193.88	2	72.79	36.83	21.5105	72.36	3.47	3.41	1	0	0	0	1	**0**
180	11	7	2	1	0	0	14.8	0	1	1	83.23	3	31.34	15.86	21.5725	31.15	1.5	1.47	1	0	0	0	1	**1**
181	5	2	1	1	0	0	12.7	0	0	1	0	0	0	0	0	0	0	0	1	1	0	137.13	1	**0**
182	4	6	1	1	0	0	15.2	0	0	0	0	0	0	0	0	0	0	0	1	0	1	277.1	1	**0**
183	4	2	1	1	0	0	12.7	0	1	0	325.97	1	206.64	106.92	36.3214	203.2	17.09	16.24	1	0	0	0	1	**0**
184	9	6	2	1	0	0	23.2	0	1	1	124.73	3	78.14	40.4	35.8954	76.87	6.38	6.07	1	0	0	0	1	**1**
185	11	1	3	1	0	0	20.3	0	0	1	0	0	0	0	0	0	0	0	1	1	0	138.09	1	**0**
186	3	5	1	1	0	0	19.9	7.5	0	0	0	0	0	0	0	0	0	0	1	1	0	400.25	1	**0**
187	11	1	2	1	0	0	23.5	0	1	1	59	3	79.03	46.72	76.753	73.25	16.26	12.75	2	0	0	0	1	**1**
188	12	2	3	0	0	1	17.2	15.5	1	1	35.83	3	102.66	257.52	164.1573	70.98	224.17	30.89	2	0	0	0	1	**1**
189	7	1	2	1	0	0	17.2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2215.02	1	**0**
190	3	5	1	1	0	0	30.8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2807.25	1	**0**
191	2	7	1	1	0	0	30.4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	5814.45	1	**0**
192	1	2	1	1	0	0	30.2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	5814.45	1	**0**
193	5	7	2	1	0	0	24.4	0	0	0	0	0	0	0	0	0	0	0	1	1	0	265.03	0	**0**
194	2	1	4	0	0	1	27.4	2.2	0	1	0	0	0	0	0	0	0	0	1	1	0	99.75	0	**0**
195	4	7	1	1	0	0	14.8	0	1	1	108.96	3	162.69	100.8	85.5447	147.99	39.48	28.98	1	0	0	0	1	**0**
196	7	6	2	1	0	0	14.3	0	0	0	0	0	0	0	0	0	0	0	1	0	1	219.63	1	**0**
197	3	1	3	0	0	1	12.4	5	1	1	108.34	3	256.69	266.51	135.757	200.72	179.35	67.54	1	0	0	0	1	**1**
198	3	1	3	0	0	1	12.4	5	1	1	59.08	3	33.61	17.28	32.5978	33.16	2.47	2.37	2	0	0	0	1	**1**
199	9	3	1	1	0	0	16.7	0	1	0	59.42	3	85.88	52.4	82.8124	78.6	19.8	14.85	2	0	0	0	1	**1**
200	9	3	4	1	0	0	13.1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	354.93	1	**0**
201	8	1	3	0	0	1	22.4	0	1	1	73.96	3	31.38	15.93	24.3113	31.15	1.7	1.66	1	0	0	0	1	**1**
202	7	7	1	1	0	0	23.4	0	1	0	85.51	3	59.74	31.14	40.0262	58.53	5.5	5.16	1	0	0	0	1	**1**

6.4 Base de datos de RNA N°02.

Enlace de descarga de datos en excel:

https://drive.google.com/file/d/1F9svLUjvUuIe11lQrnJl80cXE00tjCGo/view?usp=sharing

N°	CURVA O TRAMO RECTO	PELIGRO EN PENDIENTE	RADIO DE CURVA(M)	PELIGRO EN CURVA	LONGITUD DE CURVA (L)	LONGITUD DE SUBTANGENTE (T)	ANGULO DE DEFLEXION (°)	LONGITUD DE CUERDA (LC)	DISTANCIA EXTERNA (E)	DISTANCIA DE LA ORDENA MEDIA (M)	PELIGRO POR VISIBILIDAD	TRAMO TANGENTE EN 'S'	TRAMO TANGENTE EN 'O'	LONG. TANGENTE	PELIGRO ESTADO DEL PAVIMENTO	ACCIDENTE DE TRANSITO
1	1	1	76.3	3	49.6	25.71	37.25	48.73	4.22	3.99	2	0	0	0	1	**2**
2	1	1	30.81	3	70.18	66.87	130.53	55.96	42.82	17.92	2	0	0	0	1	**2**
3	0	0	0	0	0	0	0	0	0	0	1	1	0	289.18	1	**1**
4	0	0	0	0	0	0	0	0	0	0	1	1	0	350.5	1	**1**
5	1	0	181.02	2	140.22	73.84	44.38	136.74	14.48	13.41	1	0	0	0	1	**2**
6	0	0	190.27	2	66.69	33.69	20.08	66.35	2.96	2.91	1	0	0	0	1	**1**
7	0	0	0	0	0	0	0	0	0	0	1	1	0	220.46	1	**1**
8	1	1	42	3	44.62	24.68	60.88	42.55	6.71	5.79	2	0	0	0	1	**2**
9	0	0	0	0	0	0	0	0	0	0	1	1	0	141.21	0	**1**
10	1	0	153.37	2	132.93	70.97	49.66	128.81	15.62	14.18	1	0	0	0	1	**2**
11	0	0	0	0	0	0	0	0	0	0	1	1	0	404.81	1	**1**
12	1	0	150.84	2	96.14	49.77	36.52	94.52	8	7.6	1	0	0	0	1	**2**
13	0	0	0	0	0	0	0	0	0	0	1	1	0	226.36	1	**1**
14	1	0	371.28	1	107.39	54.07	16.57	107.02	3.92	3.88	0	0	0	0	1	**1**
15	1	1	32.64	3	58.03	40.21	101.87	50.68	19.15	12.07	2	0	0	0	1	**2**
16	1	0	312.97	1	138.08	70.18	25.28	136.97	7.77	7.58	0	0	0	0	1	**1**
17	1	1	59.51	3	77.69	45.5	74.8	72.29	15.4	12.24	2	0	0	0	1	**2**
18	0	1	0	0	0	0	0	0	0	0	1	1	0	133.4	1	**2**
19	0	0	0	0	0	0	0	0	0	0	0	1	0	679.12	1	**1**
20	0	0	0	0	0	0	0	0	0	0	1	1	0	447.19	1	**1**
21	0	0	0	0	0	0	0	0	0	0	0	1	0	1746.79	1	**1**
22	1	1	126.51	3	102.47	54.23	46.41	99.69	11.13	10.23	2	0	0	0	1	**2**
23	1	1	86.74	3	51.98	26.8	34.34	51.21	4.05	3.87	2	0	0	0	1	**2**
24	0	0	0	0	0	0	0	0	0	0	0	1	0	1238.11	1	**1**
25	0	0	0	0	0	0	0	0	0	0	0	1	0	1238.11	1	**1**
26	1	0	146.84	3	70.29	35.83	27.42	69.62	4.31	4.19	0	0	0	0	1	**1**
27	1	1	114.1	3	134.84	76.54	67.71	127.13	23.29	19.34	1	0	0	0	0	**2**
28	0	0	0	0	0	0	0	0	0	0	1	1	0	142.53	0	**1**
29	0	0	263.88	1	94.16	47.58	20.44	93.66	4.26	4.19	1	0	0	0	0	**1**
30	1	1	77.51	3	105.61	62.84	78.07	97.63	22.28	17.3	2	0	0	0	0	**2**
31	1	0	70.85	3	41.06	21.13	33.21	40.49	3.08	2.95	2	0	0	0	0	**2**
32	1	0	142.88	3	69.85	35.64	28.01	69.16	4.38	4.25	2	0	0	0	0	**2**
33	1	1	16.91	3	38.77	37.4	131.34	30.82	24.14	9.94	2	0	0	0	1	**2**
34	1	1	18.71	3	35.53	26.14	108.83	30.43	13.44	7.82	2	0	0	0	1	**2**
35	1	1	26.46	3	38.98	24	84.41	35.55	9.26	6.86	2	0	0	0	1	**2**
36	1	1	26.46	3	38.98	24	84.41	35.55	9.26	6.86	2	0	0	0	1	**2**
37	1	1	30.49	3	54.84	38.36	103.04	47.74	18.51	11.52	2	0	0	0	1	**2**
38	1	1	29.04	3	67.19	66.09	132.56	53.18	43.15	17.36	2	0	0	0	1	**2**
39	1	1	29.04	3	67.19	66.09	132.56	53.18	43.15	17.36	2	0	0	0	1	**2**
40	1	0	68.21	3	67.21	36.61	56.45	64.52	9.21	8.11	2	0	0	0	1	**2**
41	1	0	102.25	3	70.86	36.92	39.71	69.45	6.46	6.08	1	0	0	0	1	**2**
42	0	0	0	0	0	0	0	0	0	0	0	1	0	1275.47	1	**1**
43	0	0	0	0	0	0	0	0	0	0	0	1	0	1286.88	1	**1**
44	1	0	192.24	2	136.16	71.07	40.58	133.33	12.72	11.93	1	0	0	0	1	**1**
45	1	0	192.24	2	136.16	71.07	40.58	133.33	12.72	11.93	1	0	0	0	1	**1**
46	1	0	97.75	3	122.9	71.07	72.04	114.97	23.1	18.69	2	0	0	0	1	**2**
47	1	0	159.3	2	108.03	56.19	38.86	105.97	9.62	9.07	1	0	0	0	1	**1**
48	0	0	0	0	0	0	0	0	0	0	0	0	1	1482.35	1	**1**
49	1	1	108.96	3	162.69	100.8	85.54	147.99	39.48	28.98	1	0	0	0	1	**1**
50	1	1	52.97	3	109.7	89.31	118.66	91.12	50.87	25.95	2	0	0	0	1	**2**
51	1	1	86.66	3	141.7	92.42	93.68	126.44	40.04	27.38	2	0	0	0	1	**2**
52	1	0	57.92	3	79.5	47.45	78.65	73.41	16.95	13.11	2	0	0	0	1	**2**
53	1	1	83.21	3	58.61	30.58	40.35	57.4	5.44	5.11	1	0	0	0	1	**2**
54	1	1	64.49	3	132.49	106.72	117.71	110.39	60.2	31.14	2	0	0	0	1	**2**
55	0	1	0	0	0	0	0	0	0	0	1	0	1	201.41	1	**1**
56	1	1	68.27	3	66.07	35.88	55.45	63.52	8.85	7.84	1	0	0	0	1	**2**
57	0	0	0	0	0	0	0	0	0	0	1	0	1	184.9	1	**1**
58	1	1	34.87	3	58.53	38.84	96.16	51.9	17.33	11.57	2	0	0	0	1	**2**
59	0	0	0	0	0	0	0	0	0	0	1	0	1	181.31	1	**1**
60	1	0	100	3	25.83	12.99	14.8	25.76	0.84	0.83	1	0	0	0	1	**2**
61	1	1	100	3	19.11	9.58	10.95	19.08	0.46	0.46	2	0	0	0	1	**2**
62	1	0	238.73	2	33.37	16.71	8.01	33.35	0.58	0.58	1	0	0	0	1	**1**
63	1	1	258.01	1	252.76	137.56	56.13	242.77	34.38	30.34	1	0	0	0	1	**2**
64	1	1	54.29	3	123.75	118.05	130.61	98.64	75.65	31.61	2	0	0	0	1	**2**
65	1	1	45.41	3	94.95	78.34	119.81	78.57	45.14	22.64	2	0	0	0	1	**2**
66	1	1	87.06	3	167.88	125.47	110.49	143.05	65.66	37.43	2	0	0	0	1	**2**
67	1	1	54.59	3	71.2	41.68	74.73	66.26	14.09	11.2	2	0	0	0	1	**2**
68	1	1	63.45	3	103.51	67.41	93.46	92.41	29.12	19.96	2	0	0	0	1	**2**
69	0	0	0	0	0	0	0	0	0	0	0	1	0	511.33	1	**1**
70	1	1	80.15	3	38.89	19.83	27.8	38.51	2.42	2.35	2	0	0	0	1	**2**
71	1	1	29.25	3	40.17	23.98	78.7	37.09	8.58	6.63	2	0	0	0	1	**2**
72	1	1	36.25	3	80.18	72.27	126.72	64.81	44.6	20	2	0	0	0	1	**2**
73	1	1	35.83	3	102.66	257.52	164.16	70.98	224.17	30.89	2	0	0	0	1	**2**
74	1	0	71.06	3	59.48	31.61	47.95	57.76	6.71	6.13	1	0	0	0	1	**2**
75	1	1	69.33	3	66.66	36.16	55.09	64.12	8.86	7.86	2	0	0	0	1	**2**
76	0	0	0	0	0	0	0	0	0	0	1	1	0	184.98	1	**1**
77	0	0	0	0	0	0	0	0	0	0	1	0	1	250.39	1	**1**
78	1	0	42.26	3	33.63	17.76	45.59	32.75	3.58	3.3	1	0	0	0	0	**2**
79	0	0	0	0	0	0	0	0	0	0	1	1	0	97	1	**1**
80	1	1	81.08	3	93.08	52.43	65.78	88.05	15.47	12.99	1	0	0	0	1	**2**

81	0	0	0	0	0	0	0	0	0	0	0	0	1	1905.82	1	**1**
82	0	0	0	0	0	0	0	0	0	0	0	0	1	547.19	1	**1**
83	1	0	196.76	2	202.65	111.35	59.01	193.81	29.32	25.52	1	0	0	0	1	**1**
84	0	0	0	0	0	0	0	0	0	0	0	0	1	1482.35	1	**1**
85	1	0	43.9	3	39.35	21.11	51.36	38.04	4.81	4.34	1	0	0	0	0	**2**
86	1	1	39.3	3	61.51	39.08	89.69	55.42	16.13	11.43	2	0	0	0	1	**2**
87	1	0	196.93	2	166.98	88.88	48.58	162.03	19.13	17.44	1	0	0	0	1	**1**
88	0	0	0	0	0	0	0	0	0	0	0	0	1	2749.44	1	**1**
89	0	0	0	0	0	0	0	0	0	0	1	0	1	228.49	0	**1**
90	1	1	91.36	3	169.65	122.12	106.4	146.3	61.15	36.63	1	0	0	0	1	**2**
91	1	0	62.99	3	36.88	18.99	33.55	36.36	2.8	2.68	2	0	0	0	1	**2**
92	1	0	59.42	3	85.88	52.4	82.81	78.6	19.8	14.85	2	0	0	0	1	**2**
93	1	0	192.44	2	192.25	105.01	57.24	184.35	26.78	23.51	1	0	0	0	1	**1**
94	0	0	0	0	0	0	0	0	0	0	1	0	1	420.37	0	**1**
95	1	1	34.87	3	58.53	38.84	96.16	51.9	17.33	11.57	2	0	0	0	1	**2**
96	1	0	258.01	1	252.76	137.56	56.13	242.77	34.38	30.34	1	0	0	0	1	**1**
97	1	0	258.01	1	252.76	137.56	56.13	242.77	34.38	30.34	1	0	0	0	1	**1**
98	0	0	0	0	0	0	0	0	0	0	1	0	1	412.55	1	**1**
99	1	1	65.42	3	119.89	85.27	105.01	103.8	42.05	25.6	2	0	0	0	1	**2**
100	0	0	0	0	0	0	0	0	0	0	0	1	0	5814.45	1	**1**
101	1	1	145.07	3	159.51	88.9	63	151.6	25.07	21.38	1	0	0	0	1	**1**
102	1	1	44.28	3	73.77	48.71	95.46	65.53	21.55	14.49	2	0	0	0	1	**2**
103	0	0	0	0	0	0	0	0	0	0	0	1	0	1238.11	1	**1**
104	0	0	0	0	0	0	0	0	0	0	0	0	1	2807.25	1	**1**
105	1	1	36.71	3	65.46	45.47	102.17	57.12	21.73	13.65	2	0	0	0	1	**2**
106	0	0	0	0	0	0	0	0	0	0	0	0	1	1482.35	1	**1**
107	1	1	46.92	3	109.34	109.27	133.52	86.22	72	28.41	2	0	0	0	0	**2**
108	1	1	39.3	3	61.51	39.08	89.69	55.42	16.13	11.43	2	0	0	0	1	**2**
109	1	0	283.52	1	165.08	84.95	33.36	162.76	12.45	11.93	1	0	0	0	1	**1**
110	1	1	58.07	3	74.33	43.24	73.34	69.36	14.33	11.49	1	0	0	0	1	**2**
111	1	0	144.43	3	164.33	92.35	65.19	155.61	27	22.75	1	0	0	0	1	**2**
112	1	1	21.65	3	40.37	29.18	106.84	34.77	14.68	8.75	2	0	0	0	1	**2**
113	1	1	58.34	3	145.73	175.01	143.13	110.69	126.14	39.89	1	0	0	0	1	**2**
114	1	1	173.02	2	72.43	36.75	23.98	71.9	3.86	3.78	1	0	0	0	1	**1**
115	0	1	29.13	3	54.14	39	106.49	46.68	19.55	11.7	2	0	0	0	1	**2**
116	0	0	0	0	0	0	0	0	0	0	1	1	0	210.5	1	**1**
117	1	1	29.13	3	54.14	39	106.49	46.68	19.55	11.7	2	0	0	0	1	**2**
118	0	0	0	0	0	0	0	0	0	0	1	1	0	318.21	1	**1**
119	0	0	0	0	0	0	0	0	0	0	0	0	1	2215.02	1	**1**
120	1	1	54.15	3	44.82	23.78	47.42	43.55	4.99	4.57	2	0	0	0	1	**2**
121	0	0	66.58	3	56.85	30.29	48.93	55.14	6.57	5.98	1	0	0	0	0	**2**
122	1	0	213.62	2	48.47	24.34	13	48.37	1.38	1.37	0	0	0	0	1	**1**
123	1	0	125.24	3	101.96	53.99	46.64	99.16	11.14	10.23	1	0	0	0	1	**2**
124	1	1	116.29	3	220.68	162.22	108.73	189.03	83.3	48.54	2	0	0	0	1	**2**
125	1	1	80.88	3	115.76	70.32	82.01	106.13	26.29	19.84	1	0	0	0	1	**2**
126	0	0	0	0	0	0	0	0	0	0	1	0	1	212.22	1	**1**
127	1	0	119.36	3	196.32	128.54	94.24	174.93	56.05	38.14	2	0	0	0	1	**2**
128	0	0	0	0	0	0	0	0	0	0	0	1	0	538.51	1	**1**
129	0	0	0	0	0	0	0	0	0	0	0	0	1	2807.25	1	**1**
130	0	1	104.99	3	71.95	37.45	39.27	70.55	6.48	6.1	1	0	0	0	1	**2**
131	0	0	0	0	0	0	0	0	0	0	1	0	1	1523.1	1	**1**
132	0	0	0	0	0	0	0	0	0	0	0	0	1	1905.82	1	**1**
133	0	0	0	0	0	0	0	0	0	0	0	1	0	667.29	1	**1**
134	0	0	1052.79	1	71.16	35.59	3.87	71.14	0.6	0.6	0	3	3	0	1	**1**
135	1	0	33.07	3	26.69	14.12	46.24	25.97	2.89	2.66	1	0	0	0	1	**2**
136	0	0	0	0	0	0	0	0	0	0	1	1	0	154.01	1	**1**
137	0	0	0	0	0	0	0	0	0	0	1	1	0	404.81	1	**1**
138	0	0	0	0	0	0	0	0	0	0	2	1	0	89.18	1	**1**
139	1	0	149.81	3	74.27	37.91	28.4	73.51	4.72	4.58	0	0	0	0	0	**1**
140	1	1	62.85	3	153.17	171	139.64	117.98	119.34	41.17	2	0	0	0	1	**2**
141	0	0	0	0	0	0	0	0	0	0	1	0	1	112.25	0	**1**
142	1	1	55.1	3	141.22	185.1	146.85	105.62	138.03	39.38	1	0	0	0	1	**1**
143	0	0	0	0	0	0	0	0	0	0	1	1	0	350.5	1	**1**
144	1	0	136.19	3	88.41	45.83	37.19	86.87	7.5	7.11	1	0	0	0	1	**1**
145	1	1	87.06	3	167.88	125.47	110.49	143.05	65.66	37.43	2	0	0	0	1	**2**
146	1	0	87.81	3	42.47	21.66	27.71	42.06	2.63	2.56	1	0	0	0	1	**2**
147	0	0	0	0	0	0	0	0	0	0	0	1	0	5814.45	1	**1**
148	0	0	0	0	0	0	0	0	0	0	1	0	1	412.55	1	**1**
149	0	0	0	0	0	0	0	0	0	0	0	1	0	5814.45	1	**1**
150	1	0	1654.81	1	193.23	96.72	6.69	193.12	2.82	2.82	0	0	0	0	1	**1**
151	0	0	0	0	0	0	0	0	0	0	0	1	0	975.89	1	**1**
152	1	1	101.94	3	90.33	48.37	50.77	87.4	10.89	9.84	2	0	1	0	1	**2**
153	0	0	0	0	0	0	0	0	0	0	1	1	0	145.78	1	**1**
154	0	0	0	0	0	0	0	0	0	0	1	1	0	121.83	1	**1**
155	0	0	0	0	0	0	0	0	0	0	1	0	1	378.16	1	**1**
156	0	0	0	0	0	0	0	0	0	0	0	0	1	2215.02	1	**1**
157	1	1	44.65	3	64.33	39.19	82.55	58.91	14.76	11.09	2	0	0	0	1	**2**
158	0	0	0	0	0	0	0	0	0	0	0	1	0	676.08	1	**1**
159	0	0	0	0	0	0	0	0	0	0	0	1	0	5814.45	1	**1**

151	0	0	0	0	0	0	0	0	0	0	0	1	0	975.89	1	**1**
152	1	1	101.94	3	90.33	48.37	50.77	87.4	10.89	9.84	2	0	1	0	1	**2**
153	0	0	0	0	0	0	0	0	0	0	1	1	0	145.78	1	**1**
154	0	0	0	0	0	0	0	0	0	0	1	1	0	121.83	1	**1**
155	0	0	0	0	0	0	0	0	0	0	1	0	1	378.16	1	**1**
156	0	0	0	0	0	0	0	0	0	0	0	0	1	2215.02	1	**1**
157	1	1	44.65	3	64.33	39.19	82.55	58.91	14.76	11.09	2	0	0	0	1	**2**
158	0	0	0	0	0	0	0	0	0	0	0	1	0	676.08	1	**1**
159	0	0	0	0	0	0	0	0	0	0	0	1	0	5814.45	1	**1**
160	1	1	45.41	3	94.95	78.34	119.81	78.57	45.14	22.64	2	0	0	0	1	**2**
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2938	1	1	39	3	61.94	39.68	90.99	55.63	16.64	11.66	2	0	0	0	1	**2**
2939	0	0	0	0	0	0	0	0	0	0	2	0	1	52.81	1	**2**
2940	1	1	46.02	3	31.49	16.39	39.21	30.88	2.83	2.67	2	0	0	0	1	**2**
2941	0	0	0	0	0	0	0	0	0	0	2	1	0	67.49	1	**2**
2942	1	1	115.49	3	52.06	26.48	25.83	51.62	3	2.92	2	0	0	0	1	**2**
2943	0	0	0	0	0	0	0	0	0	0	1	0	1	242.52	1	**1**
2944	1	1	38.05	3	64.59	43.2	97.25	57.11	19.52	12.9	2	0	0	0	1	**2**
2945	0	0	0	0	0	0	0	0	0	0	2	0	1	55.01	1	**2**
2946	1	1	61.51	3	24.55	12.44	22.86	24.38	1.24	1.22	2	0	0	0	1	**2**
2947	0	0	0	0	0	0	0	0	0	0	2	1	0	8.57	1	**2**
2948	1	1	165.38	2	75.21	38.27	26.06	74.56	4.37	4.26	1	0	0	0	1	**1**
2949	0	0	0	0	0	0	0	0	0	0	1	1	0	148.83	1	**1**
2950	1	1	56.79	3	53.02	28.62	53.49	51.11	6.8	6.08	2	0	0	0	1	**2**
2951	0	0	0	0	0	0	0	0	0	0	2	1	0	47.45	1	**2**
2952	1	1	29.23	3	46.44	29.76	91.02	41.71	12.48	8.75	2	0	0	0	1	**2**
2953	0	0	0	0	0	0	0	0	0	0	1	0	1	113.96	1	**1**
2954	1	1	28.06	3	49.02	33.5	100.1	43.02	15.64	10.04	2	0	0	0	1	**2**
2955	0	0	0	0	0	0	0	0	0	0	1	1	0	113.96	1	**1**
2956	1	1	77.69	3	36.5	18.59	26.92	36.16	2.19	2.13	2	0	0	0	1	**2**
2957	0	0	0	0	0	0	0	0	0	0	1	1	0	206.32	1	**1**
2958	1	1	124.24	3	28.96	14.55	13.36	28.9	0.85	0.84	2	0	0	0	1	**2**
2959	0	0	0	0	0	0	0	0	0	0	2	1	0	46.55	1	**2**
2960	1	1	58.93	3	59.3	32.44	57.66	56.83	8.34	7.3	2	0	0	0	1	**2**
2961	0	0	0	0	0	0	0	0	0	0	1	1	0	202.8	1	**1**
2962	1	1	173.02	2	72.43	36.75	23.98	71.9	3.86	3.78	1	0	0	0	1	**1**
2963	0	0	0	0	0	0	0	0	0	0	2	1	0	91.59	1	**2**
2964	1	1	89.9	3	55.83	28.85	35.58	54.93	4.52	4.3	2	0	0	0	1	**2**
2965	0	0	0	0	0	0	0	0	0	0	2	0	1	21.76	1	**2**
2966	1	1	58.19	3	55.27	29.92	54.42	53.22	7.24	6.44	2	0	0	0	1	**2**
2967	0	0	0	0	0	0	0	0	0	0	2	1	0	21.76	1	**2**
2968	1	1	35.83	3	102.66	257.52	164.16	70.98	224.17	30.89	2	0	0	0	1	**2**
2969	0	0	0	0	0	0	0	0	0	0	1	1	0	114.47	1	**1**
2970	1	1	19.27	3	49.01	62.43	145.69	36.83	46.06	13.59	2	0	0	0	1	**2**
2971	0	0	0	0	0	0	0	0	0	0	2	1	0	79.64	1	**2**
2972	1	1	73.55	3	49.26	25.59	38.37	48.34	4.33	4.09	2	0	0	0	1	**2**
2973	0	0	0	0	0	0	0	0	0	0	2	1	0	39.12	1	**2**
2974	1	1	23.02	3	38.21	25.17	95.11	33.97	11.09	7.48	2	0	0	0	1	**2**
2975	0	0	0	0	0	0	0	0	0	0	2	1	0	94.67	1	**2**
2976	1	1	42.19	3	80.56	59.59	109.41	68.87	30.83	17.81	2	0	0	0	1	**2**
2977	0	0	0	0	0	0	0	0	0	0	2	1	0	46.19	1	**2**
2978	1	1	47.01	3	46.03	25.05	56.1	44.21	6.26	5.52	2	0	0	0	1	**2**
2979	0	0	0	0	0	0	0	0	0	0	2	1	0	67.16	1	**2**
2980	1	1	30.81	3	70.18	66.87	130.53	55.96	42.82	17.92	2	0	0	0	1	**2**
2981	0	0	0	0	0	0	0	0	0	0	2	1	0	34.52	1	**2**
2982	1	1	63.66	3	85.77	50.81	77.19	79.43	17.79	13.91	2	0	0	0	1	**2**
2983	0	0	0	0	0	0	0	0	0	0	2	0	1	26	1	**2**
2984	1	1	147.98	3	27.91	14	10.81	27.87	0.66	0.66	2	0	0	0	1	**2**
2985	0	0	0	0	0	0	0	0	0	0	1	1	0	171.46	1	**1**
2986	1	1	474.41	1	132.06	66.46	15.95	131.64	4.63	4.59	1	0	0	0	1	**0**
2987	0	0	0	0	0	0	0	0	0	0	2	0	0	40.81	1	**2**
2988	0	0	0	0	0	0	0	0	0	0	2	0	0	68.71	1	**2**
2989	0	0	0	0	0	0	0	0	0	0	2	0	0	32.83	1	**2**
2990	0	0	0	0	0	0	0	0	0	0	2	0	0	35.15	1	**2**
2991	0	0	0	0	0	0	0	0	0	0	2	0	0	44.3	1	**2**
2992	0	0	0	0	0	0	0	0	0	0	2	0	0	78.94	1	**2**
2993	0	0	0	0	0	0	0	0	0	0	2	0	0	26.06	1	**2**
2994	0	0	0	0	0	0	0	0	0	0	2	0	0	43.53	1	**2**
2995	0	0	0	0	0	0	0	0	0	0	1	0	0	263.98	1	**1**
2996	0	0	0	0	0	0	0	0	0	0	2	0	0	66.87	1	**2**
2997	0	0	0	0	0	0	0	0	0	0	0	0	0	652.19	1	**0**

6.5 Desarrollo de código de RNA N°01.

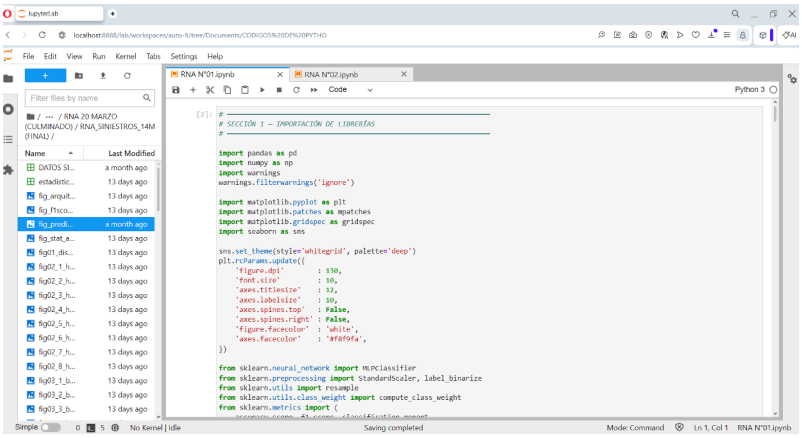


Figura 237 Importacion de librerias.
Fuente: Elaboración propia

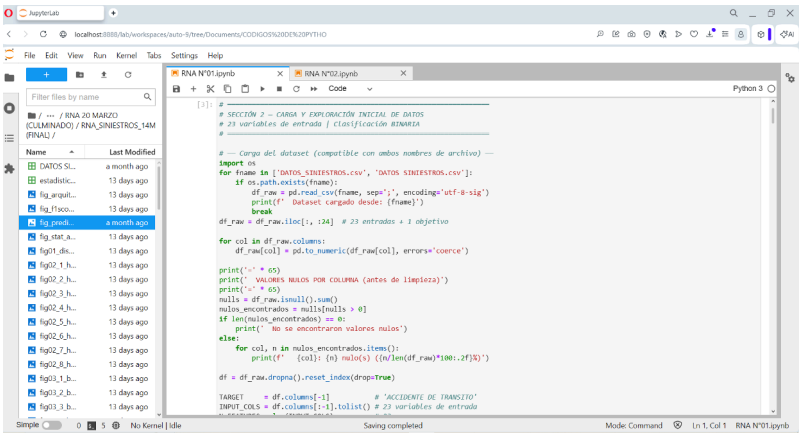


Figura 238 Importacion de base de datos.
Fuente: Elaboración propia

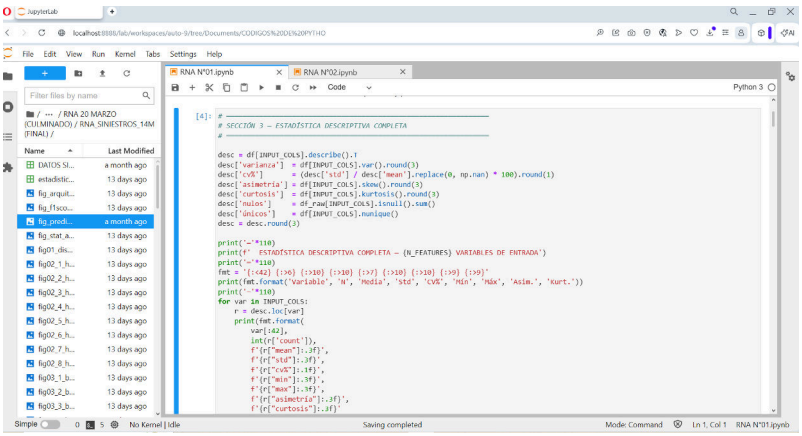
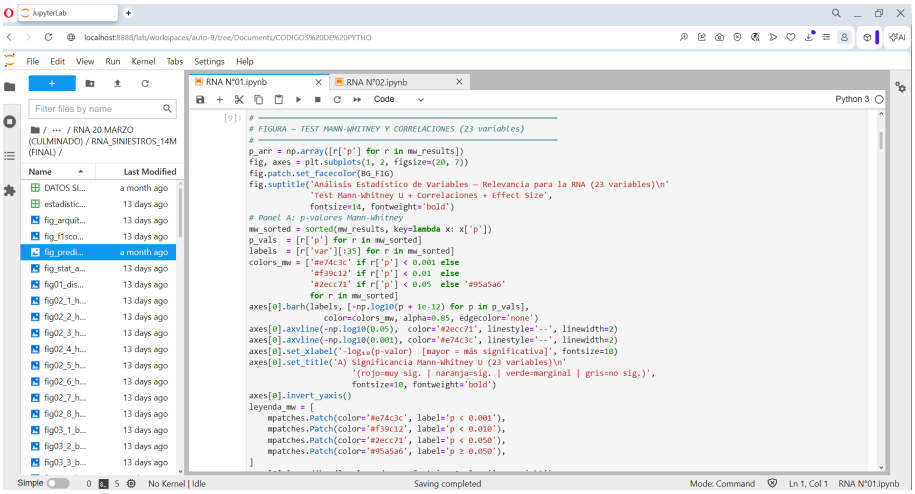
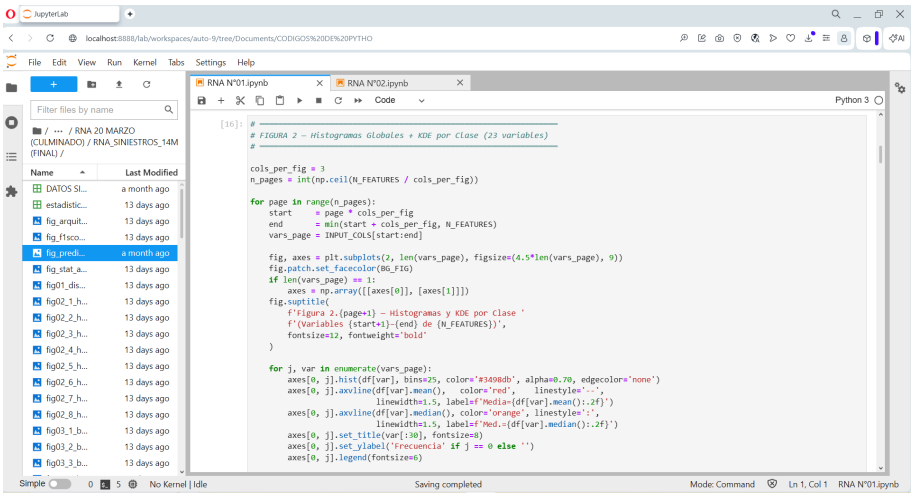


Figura 239 Análisis estadistico.
Fuente: Elaboración propia



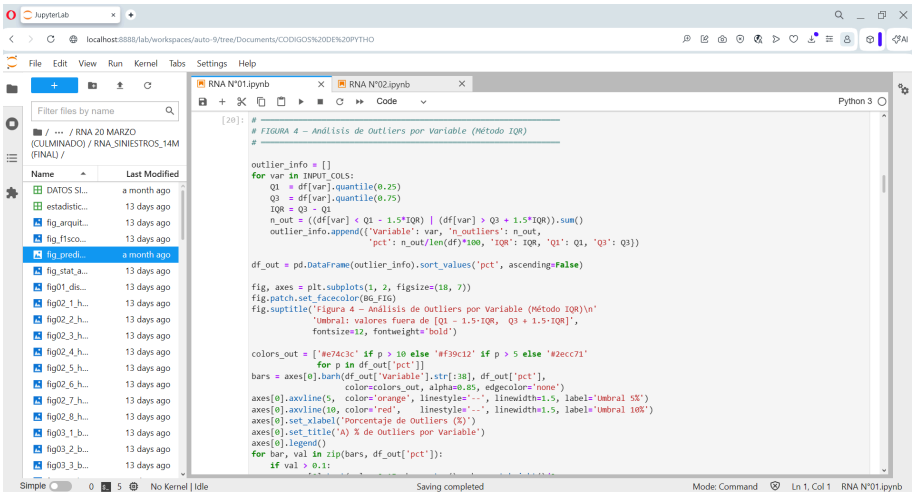
```
# ──────────────────────────────────────────────────────────────
# FIGURA — TEST MANN-WHITNEY Y CORRELACIONES (23 variables)
# ──────────────────────────────────────────────────────────────
p_arr = np.array([r['p'] for r in mw_results])
fig, axes = plt.subplots(1, 2, figsize=(20, 7))
fig.patch.set_facecolor(BG_FIG)
fig.suptitle('Análisis Estadístico de Variables — Relevancia para la RNA (23 variables)\n'
             'Test Mann-Whitney U + Correlaciones + Effect Size',
             fontsize=14, fontweight='bold')
# Panel A: p-valores Mann-Whitney
mw_sorted = sorted(mw_results, key=lambda x: x['p'])
p_vals  = [r['p'] for r in mw_sorted]
labels  = [r['var'][:35] for r in mw_sorted]
colors_mw = ['#e74c3c' if r['p'] < 0.001 else
             '#f39c12' if r['p'] < 0.01  else
             '#2ecc71' if r['p'] < 0.05  else '#95a5a6'
             for r in mw_sorted]
axes[0].barh(labels, [-np.log10(p + 1e-12) for p in p_vals],
             color=colors_mw, alpha=0.85, edgecolor='none')
axes[0].axvline(-np.log10(0.05),  color='#2ecc71', linestyle='--', linewidth=2)
axes[0].axvline(-np.log10(0.001), color='#e74c3c', linestyle='--', linewidth=2)
axes[0].set_xlabel('-log₁₀(p-valor)  [mayor = más significativa]', fontsize=10)
axes[0].set_title('A) Significancia Mann-Whitney U (23 variables)\n'
                  '(rojo=muy sig. | naranja=sig. | verde=marginal | gris=no sig.)',
                  fontsize=10, fontweight='bold')
axes[0].invert_yaxis()
leyenda_mw = [
    mpatches.Patch(color='#e74c3c', label='p < 0.001'),
    mpatches.Patch(color='#f39c12', label='p < 0.010'),
    mpatches.Patch(color='#2ecc71', label='p < 0.050'),
    mpatches.Patch(color='#95a5a6', label='p ≥ 0.050'),
]
```

Figura 240 Test de mann whittney y correlaciones.
Fuente: Elaboración propia



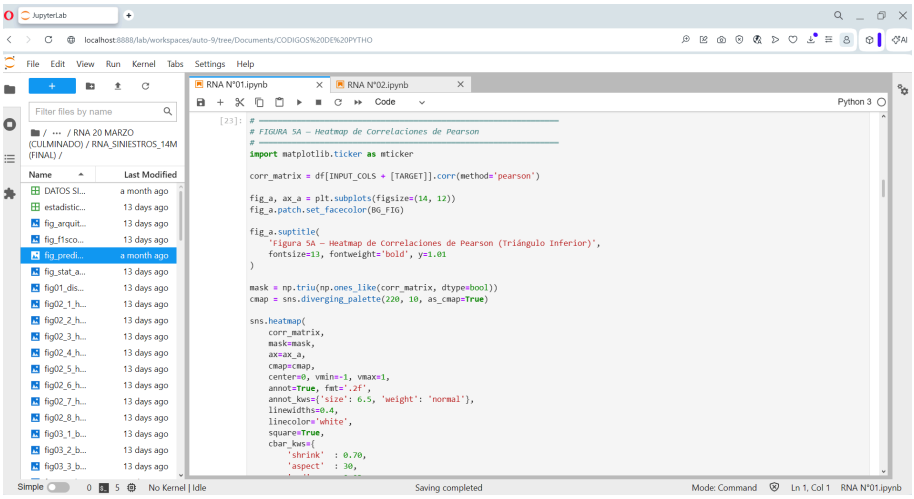
```
# ──────────────────────────────────────────────────────────────
# FIGURA 2 — Histogramas Globales + KDE por Clase (23 variables)
# ──────────────────────────────────────────────────────────────

cols_per_fig = 3
n_pages = int(np.ceil(N_FEATURES / cols_per_fig))

for page in range(n_pages):
    start     = page * cols_per_fig
    end       = min(start + cols_per_fig, N_FEATURES)
    vars_page = INPUT_COLS[start:end]

    fig, axes = plt.subplots(2, len(vars_page), figsize=(4.5*len(vars_page), 9))
    fig.patch.set_facecolor(BG_FIG)
    if len(vars_page) == 1:
        axes = np.array([[axes[0]], [axes[1]]])
    fig.suptitle(
        f'Figura 2.{page+1} — Histogramas y KDE por Clase '
        f'(Variables {start+1}–{end} de {N_FEATURES})',
        fontsize=12, fontweight='bold'
    )

    for j, var in enumerate(vars_page):
        axes[0, j].hist(df[var], bins=25, color='#3498db', alpha=0.70, edgecolor='none')
        axes[0, j].axvline(df[var].mean(),   color='red',    linestyle='--',
                           linewidth=1.5, label=f'Media={df[var].mean():.2f}')
        axes[0, j].axvline(df[var].median(), color='orange', linestyle=':',
                           linewidth=1.5, label=f'Med.={df[var].median():.2f}')
        axes[0, j].set_title(var[:30], fontsize=8)
        axes[0, j].set_ylabel('Frecuencia' if j == 0 else '')
        axes[0, j].legend(fontsize=6)
```

Figura 241 Histogramas
Fuente: Elaboración propia



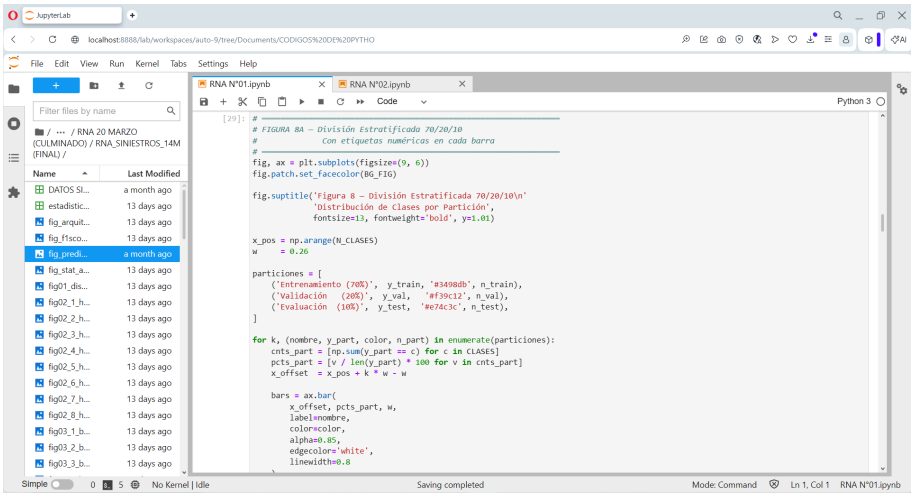
```
# ──────────────────────────────────────────────────────────────
# FIGURA 4 — Análisis de Outliers por Variable (Método IQR)
# ──────────────────────────────────────────────────────────────

outlier_info = []
for var in INPUT_COLS:
    Q1  = df[var].quantile(0.25)
    Q3  = df[var].quantile(0.75)
    IQR = Q3 - Q1
    n_out = ((df[var] < Q1 - 1.5*IQR) | (df[var] > Q3 + 1.5*IQR)).sum()
    outlier_info.append({'Variable': var, 'n_outliers': n_out,
                         'pct': n_out/len(df)*100, 'IQR': IQR, 'Q1': Q1, 'Q3': Q3})

df_out = pd.DataFrame(outlier_info).sort_values('pct', ascending=False)

fig, axes = plt.subplots(1, 2, figsize=(18, 7))
fig.patch.set_facecolor(BG_FIG)
fig.suptitle('Figura 4 — Análisis de Outliers por Variable (Método IQR)\n'
             'Umbral: valores fuera de [Q1 - 1.5·IQR,  Q3 + 1.5·IQR]',
             fontsize=12, fontweight='bold')

colors_out = ['#e74c3c' if p > 10 else '#f39c12' if p > 5 else '#2ecc71'
              for p in df_out['pct']]
bars = axes[0].barh(df_out['Variable'].str[:38], df_out['pct'],
                    color=colors_out, alpha=0.85, edgecolor='none')
axes[0].axvline(5,  color='orange', linestyle='--', linewidth=1.5, label='Umbral 5%')
axes[0].axvline(10, color='red',    linestyle='--', linewidth=1.5, label='Umbral 10%')
axes[0].set_xlabel('Porcentaje de Outliers (%)')
axes[0].set_title('A) % de Outliers por Variable')
axes[0].legend()
for bar, val in zip(bars, df_out['pct']):
    if val > 0.1:
```

Figura 242 Análsis de outilers por variable.
Fuente: Elaboración propia



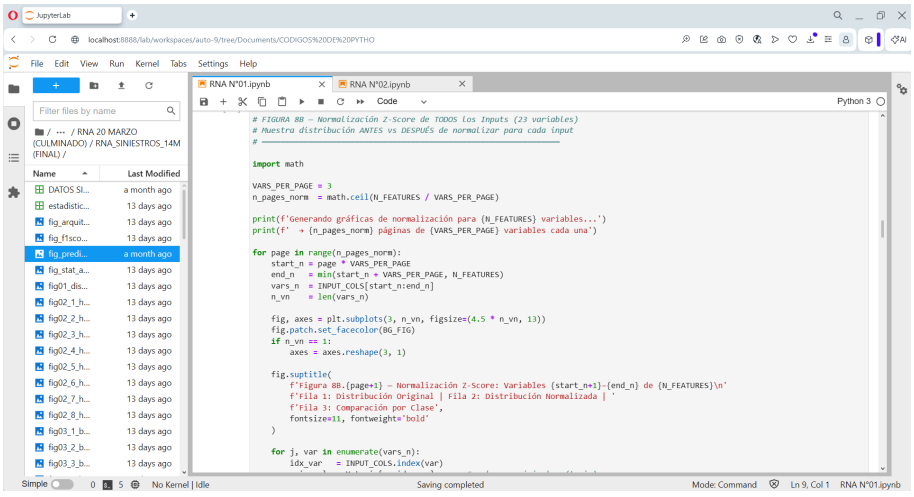
```
# ──────────────────────────────────────────────────────────────
# FIGURA 5A — Heatmap de Correlaciones de Pearson
# ──────────────────────────────────────────────────────────────
import matplotlib.ticker as mticker

corr_matrix = df[INPUT_COLS + [TARGET]].corr(method='pearson')

fig_a, ax_a = plt.subplots(figsize=(14, 12))
fig_a.patch.set_facecolor(BG_FIG)

fig_a.suptitle(
    'Figura 5A — Heatmap de Correlaciones de Pearson (Triángulo Inferior)',
    fontsize=13, fontweight='bold', y=1.01
)

mask = np.triu(np.ones_like(corr_matrix, dtype=bool))
cmap = sns.diverging_palette(220, 10, as_cmap=True)

sns.heatmap(
    corr_matrix,
    mask=mask,
    ax=ax_a,
    cmap=cmap,
    center=0, vmin=-1, vmax=1,
    annot=True, fmt='.2f',
    annot_kws={'size': 6.5, 'weight': 'normal'},
    linewidths=0.4,
    linecolor='white',
    square=True,
    cbar_kws={
        'shrink'  : 0.70,
        'aspect'  : 30,
```

Figura 243 Correlación de Pearson.
Fuente: Elaboración propia



```
# ──────────────────────────────────────────────────────────────
# FIGURA 8A — División Estratificada 70/20/10
#             Con etiquetas numéricas en cada barra
# ──────────────────────────────────────────────────────────────
fig, ax = plt.subplots(figsize=(9, 6))
fig.patch.set_facecolor(BG_FIG)

fig.suptitle('Figura 8 — División Estratificada 70/20/10\n'
             'Distribución de Clases por Partición',
             fontsize=13, fontweight='bold', y=1.01)

x_pos = np.arange(N_CLASES)
w     = 0.26

particiones = [
    ('Entrenamiento (70%)',  y_train, '#3498db', n_train),
    ('Validación   (20%)',   y_val,   '#f39c12', n_val),
    ('Evaluación  (10%)',    y_test,  '#e74c3c', n_test),
]

for k, (nombre, y_part, color, n_part) in enumerate(particiones):
    cnts_part = [np.sum(y_part == c) for c in CLASES]
    pcts_part = [v / len(y_part) * 100 for v in cnts_part]
    x_offset  = x_pos + k * w - w

    bars = ax.bar(
        x_offset, pcts_part, w,
        label=nombre,
        color=color,
        alpha=0.85,
        edgecolor='white',
        linewidth=0.8
```

Figura 244 Division de base de datos.
Fuente: Elaboración propia



```
# FIGURA 8B — Normalización Z-Score de TODOS los Inputs (23 variables)
# Muestra distribución ANTES vs DESPUÉS de normalizar para cada input
# ──────────────────────────────────────────────────────────────

import math

VARS_PER_PAGE = 3
n_pages_norm  = math.ceil(N_FEATURES / VARS_PER_PAGE)

print(f'Generando gráficas de normalización para {N_FEATURES} variables...')
print(f'  → {n_pages_norm} páginas de {VARS_PER_PAGE} variables cada una')

for page in range(n_pages_norm):
    start_n = page * VARS_PER_PAGE
    end_n   = min(start_n + VARS_PER_PAGE, N_FEATURES)
    vars_n  = INPUT_COLS[start_n:end_n]
    n_vn    = len(vars_n)

    fig, axes = plt.subplots(3, n_vn, figsize=(4.5 * n_vn, 13))
    fig.patch.set_facecolor(BG_FIG)
    if n_vn == 1:
        axes = axes.reshape(3, 1)

    fig.suptitle(
        f'Figura 8B.{page+1} — Normalización Z-Score: Variables {start_n+1}–{end_n} de {N_FEATURES}\n'
        f'Fila 1: Distribución Original | Fila 2: Distribución Normalizada | '
        f'Fila 3: Comparación por Clase',
        fontsize=11, fontweight='bold'
    )

    for j, var in enumerate(vars_n):
        idx_var   = INPUT_COLS.index(var)
```

Figura 245 Normalización de datos.
Fuente: Elaboración propia

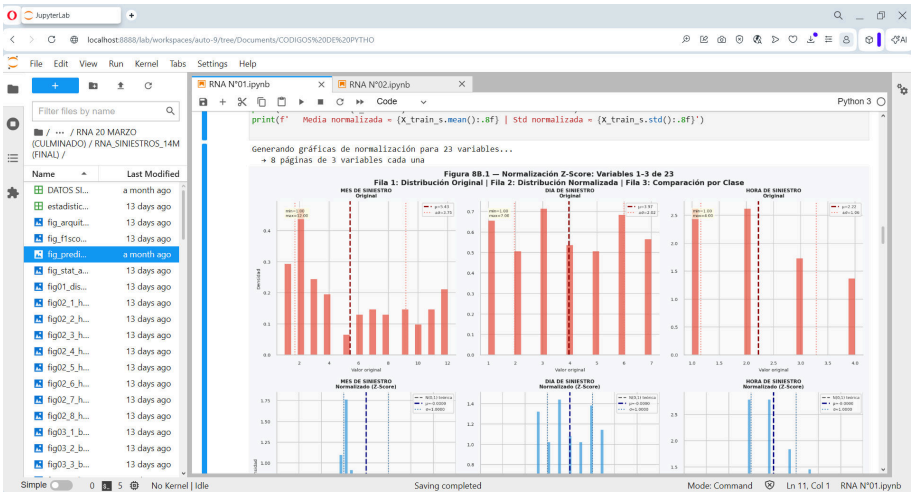


Figura 246 Graficas de normalización.
Fuente: Elaboración propia

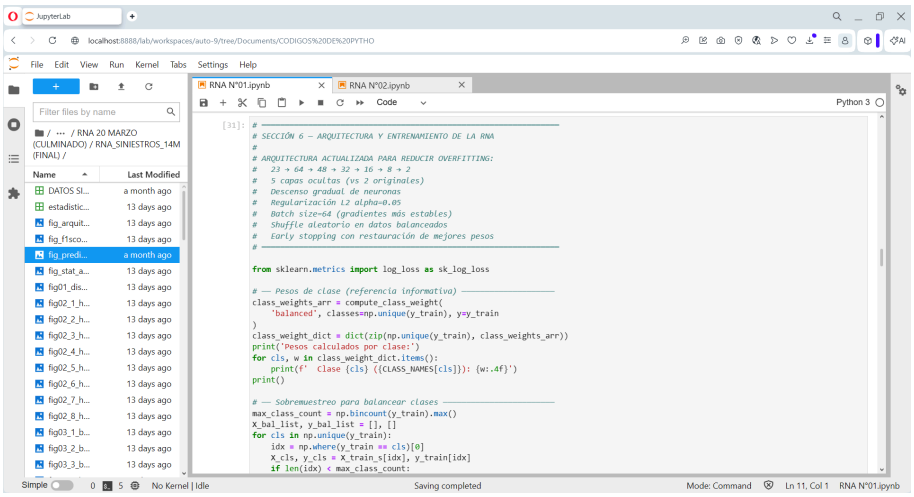


Figura 247 Arquitectura de la RNA N°01.
Fuente: Elaboración propia

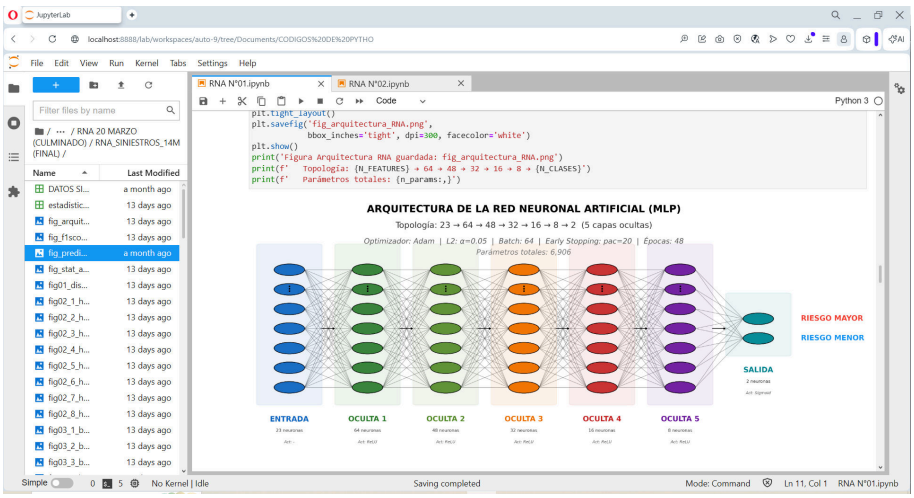


Figura 248 Grafica de arquitectura de la RNA.
Fuente: Elaboración propia

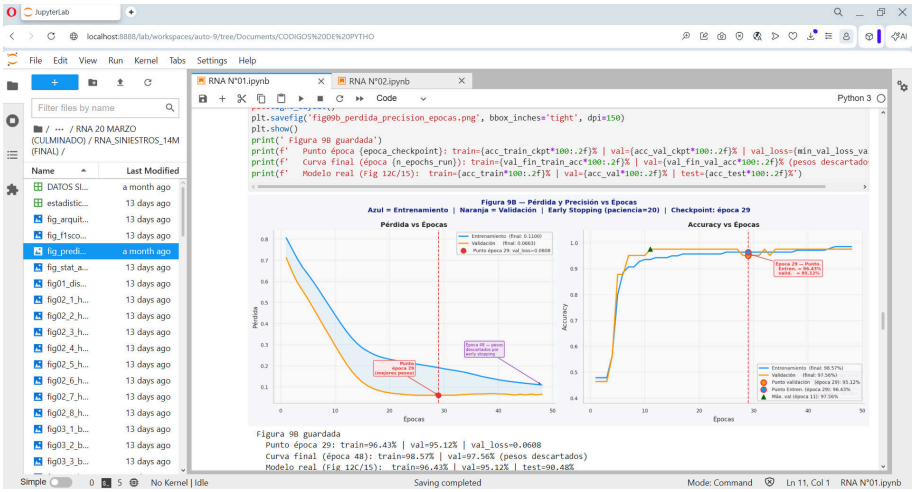


Figura 249 Graficas de perdida y Accuracy.
Fuente: Elaboración propia

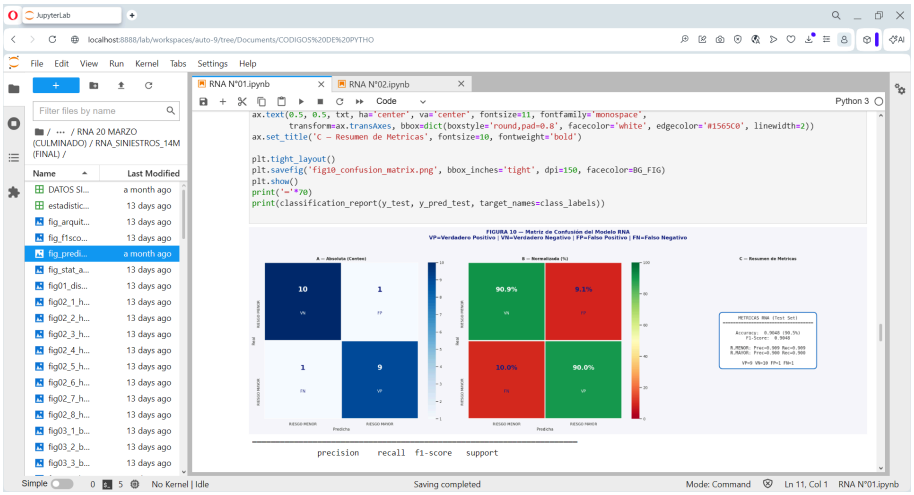


Figura 250 Matriz de confusión.
Fuente: Elaboración propia

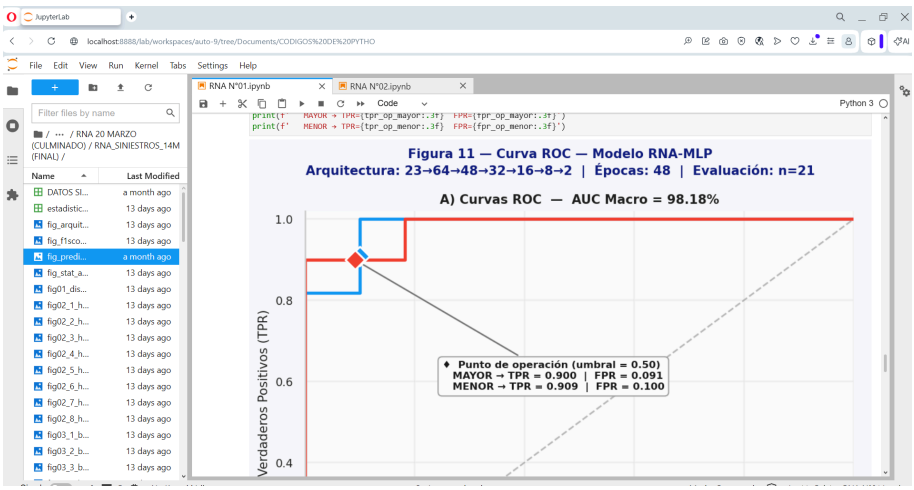


Figura 251 Curva Roc- AUC.
Fuente: Elaboración propia

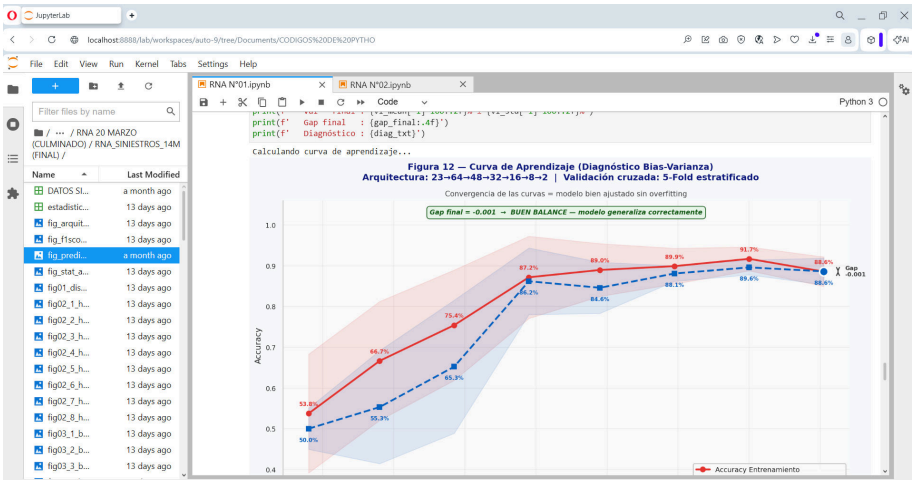


Figura 252 Curva de aprendizaje.
Fuente: Elaboración propia

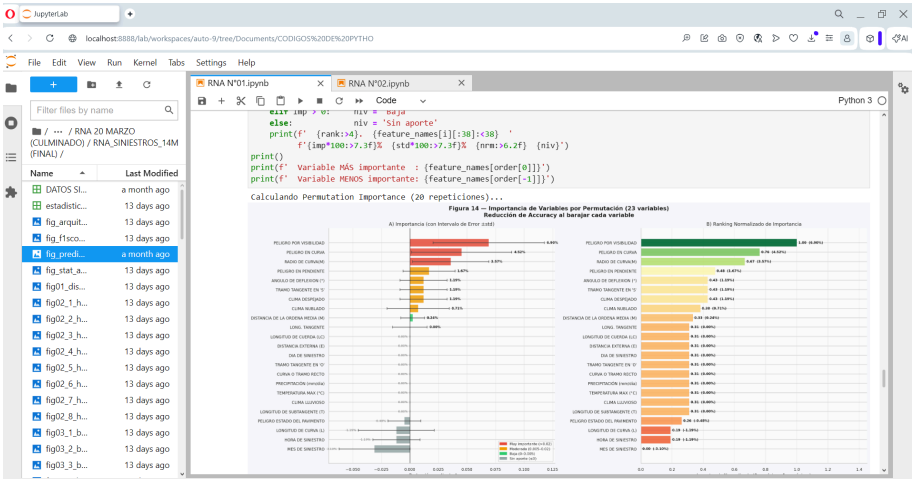


Figura 253 Importancia de variables.
Fuente: Elaboración propia

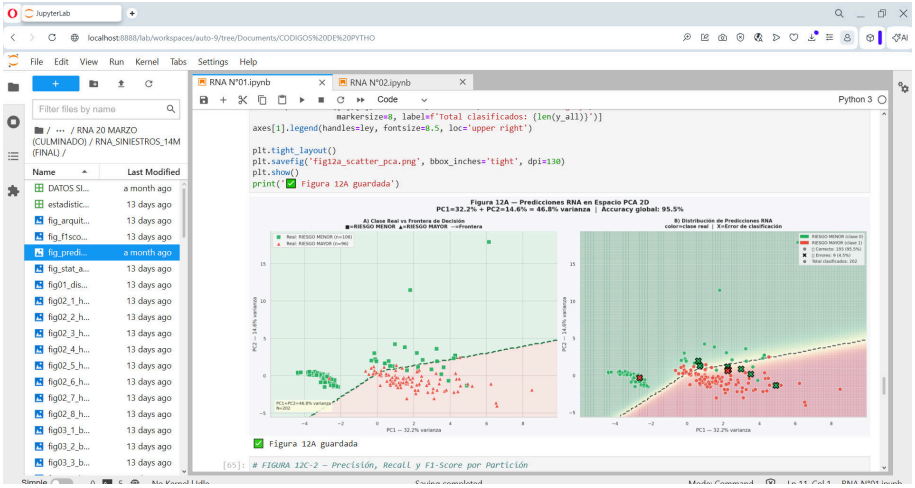


Figura 254 Frontera de decisión.
Fuente: Elaboración propia

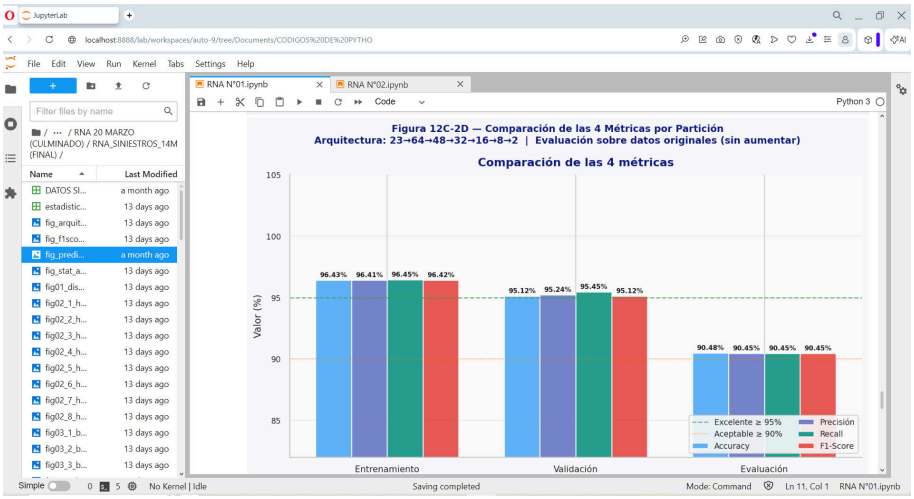


Figura 255 Resumen de metricas de evaluación de la RNA.
Fuente: Elaboración propia

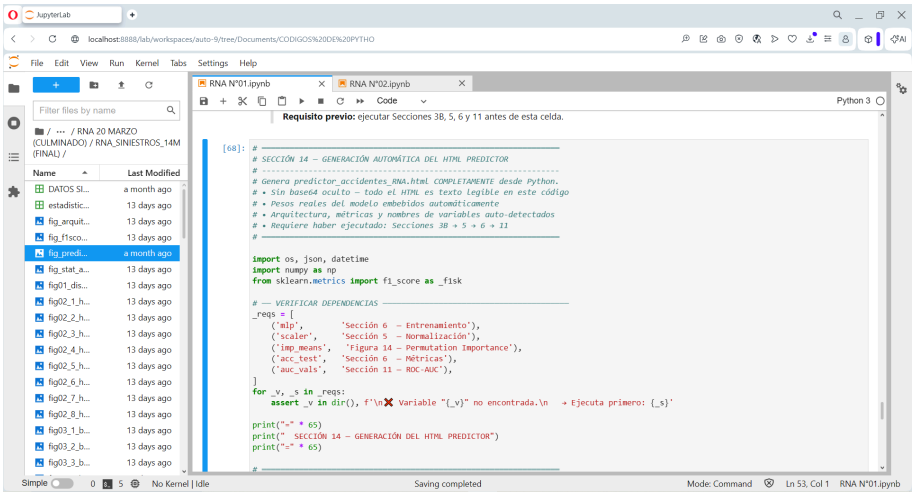


Figura 256 Generación de archivo html.
Fuente: Elaboración propia

6.6 Desarrollo de código de RNA N°02.

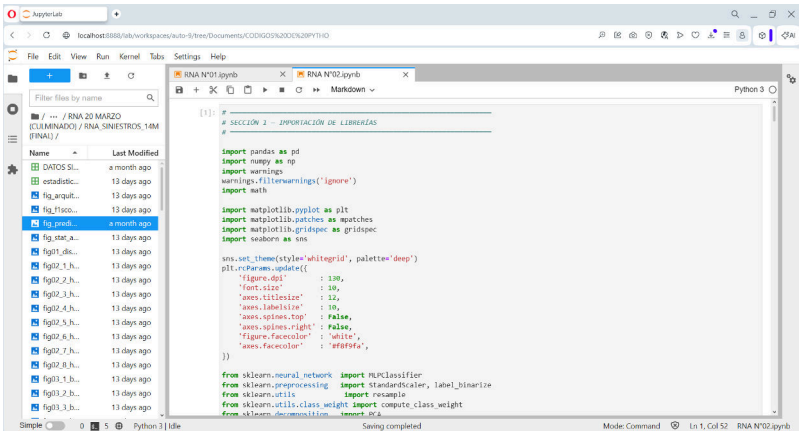


Figura 257 Importación de librerias.
Fuente: Elaboración propia

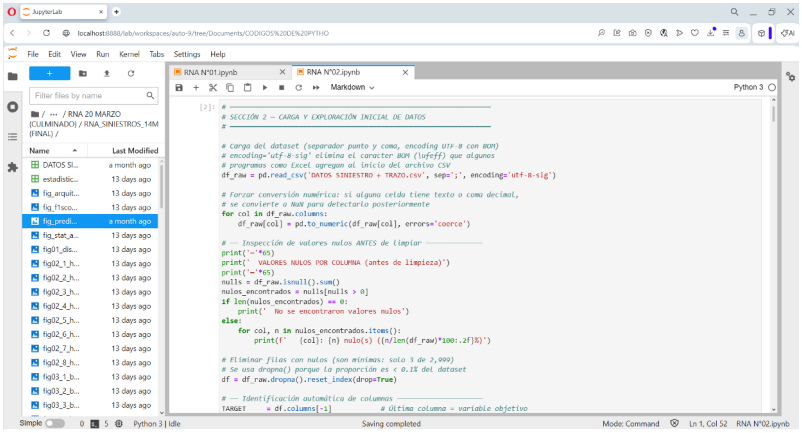


Figura 258 Importación de base de datos.
Fuente: Elaboración propia

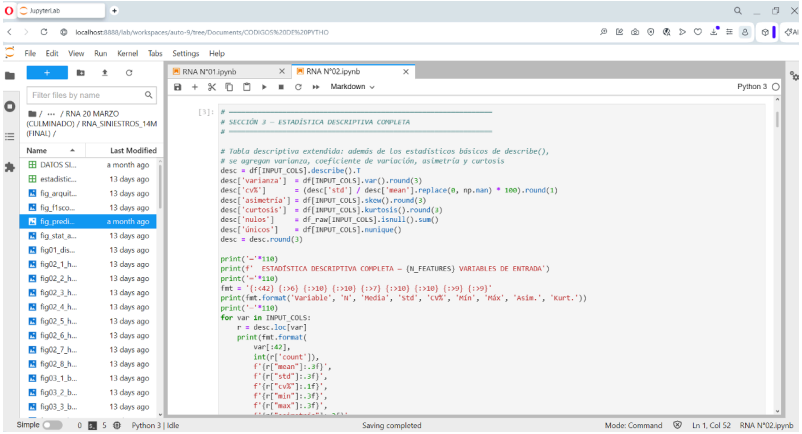


Figura 259 Estadistica de datos.
Fuente: Elaboración propia

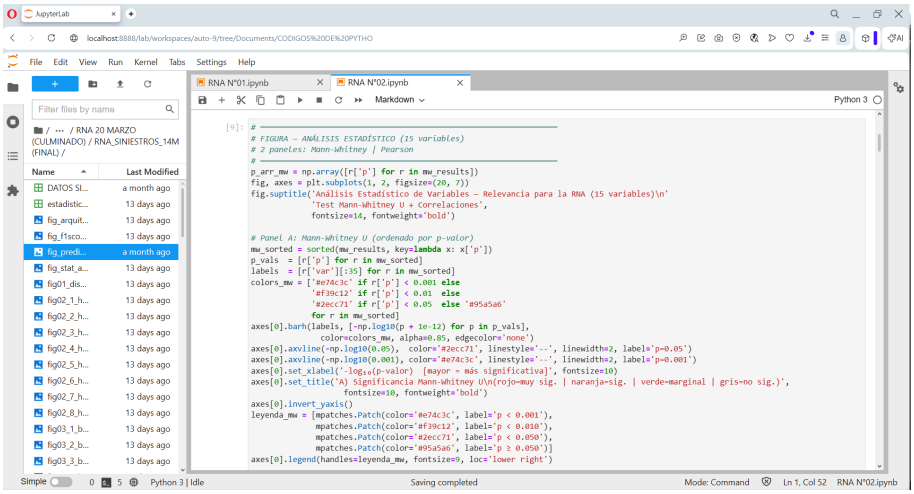


Figura 260 Estadistica de datos de Pearson.
Fuente: Elaboración propia

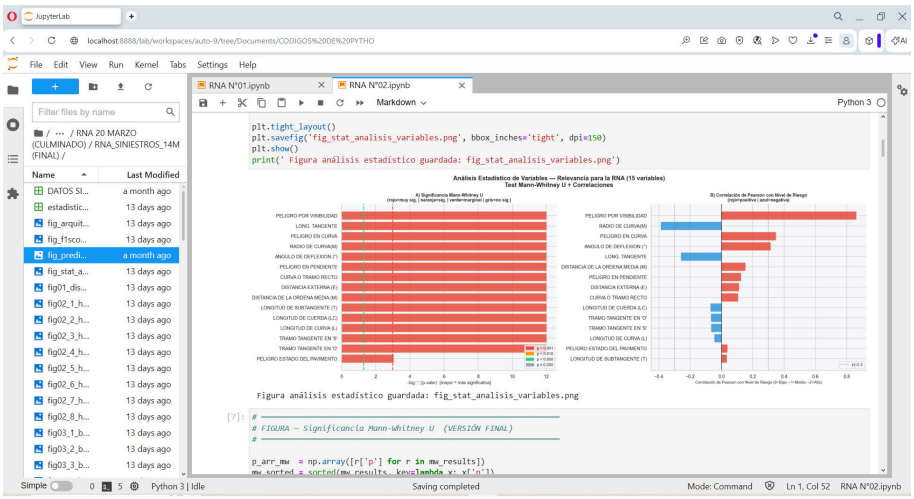


Figura 261 Test estadistico de variables de Mann Whitney.
Fuente: Elaboración propia

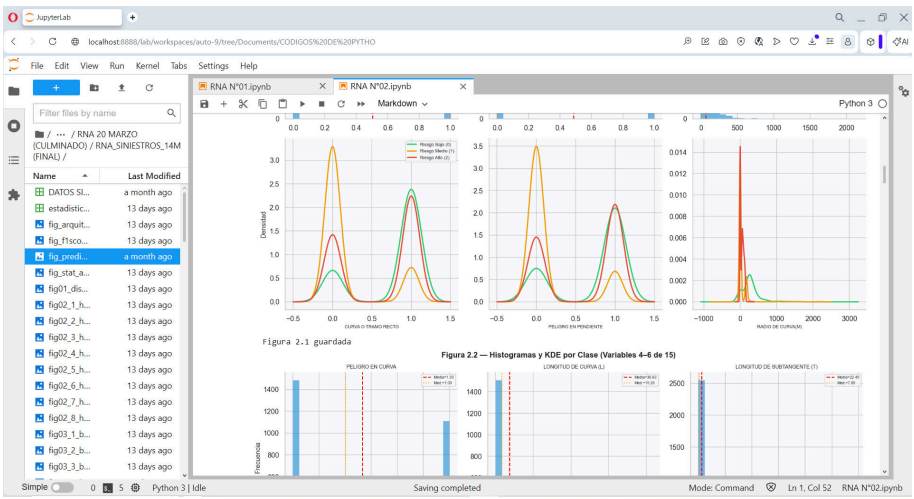


Figura 262 Grafica de histogramas.
Fuente: Elaboración propia

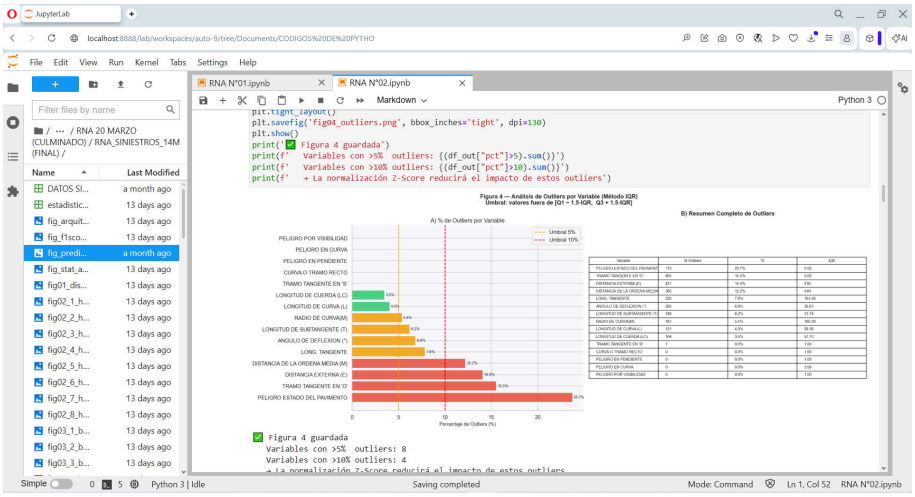


Figura 263 Análisis de Outilers.

Fuente: Elaboración propia

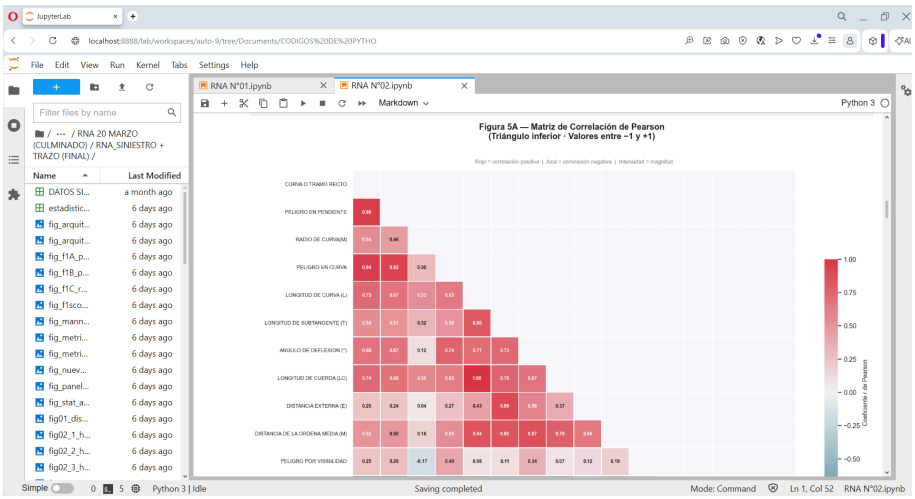


Figura 264 Matriz de correlación de pearson.

Fuente: Elaboración propia

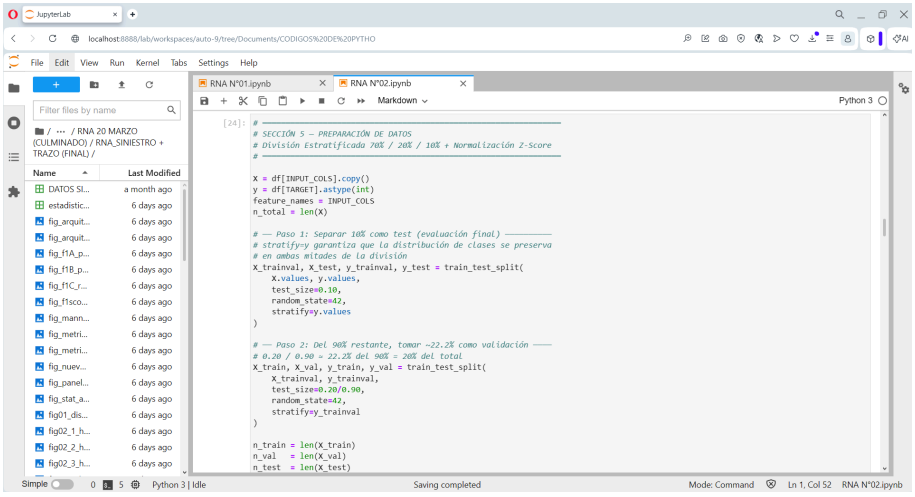


Figura 265 División de base de datos.

Fuente: Elaboración propia

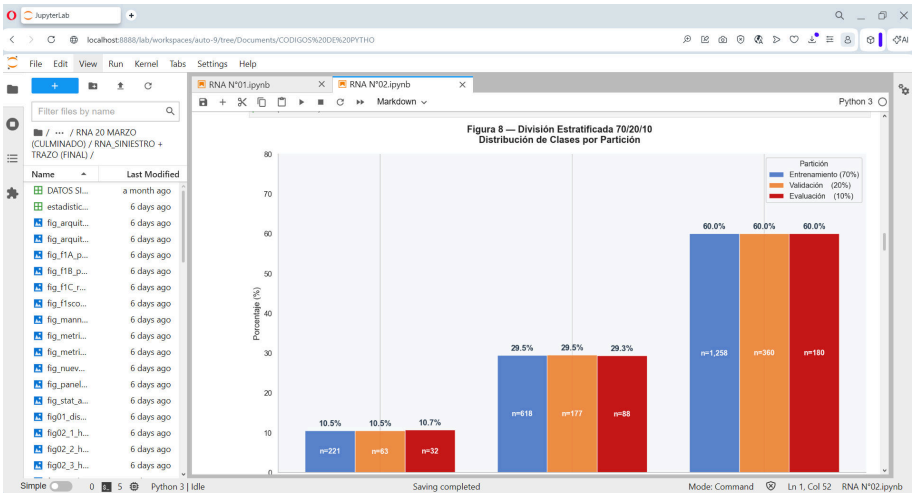


Figura 266 División de base de datos.
Fuente: Elaboración propia

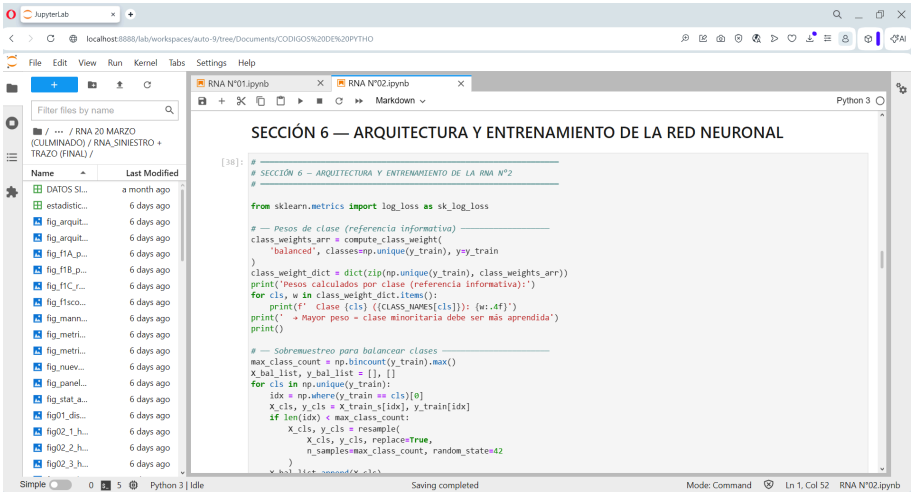


Figura 267 Arquitectura de la RNA N°02.
Fuente: Elaboración propia

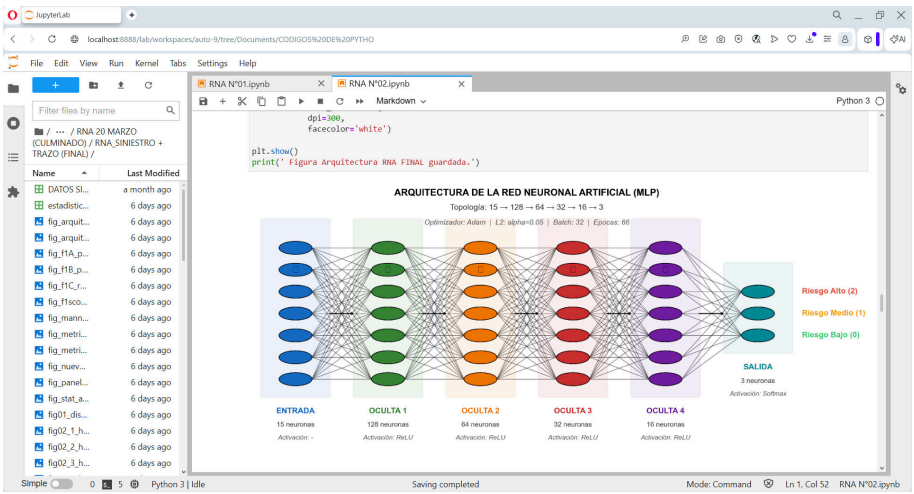


Figura 268 Grafica de arquitectura de la RNA.
Fuente: Elaboración propia

Figura 269 Grafica de perdida vs epocas.
Fuente: Elaboración propia

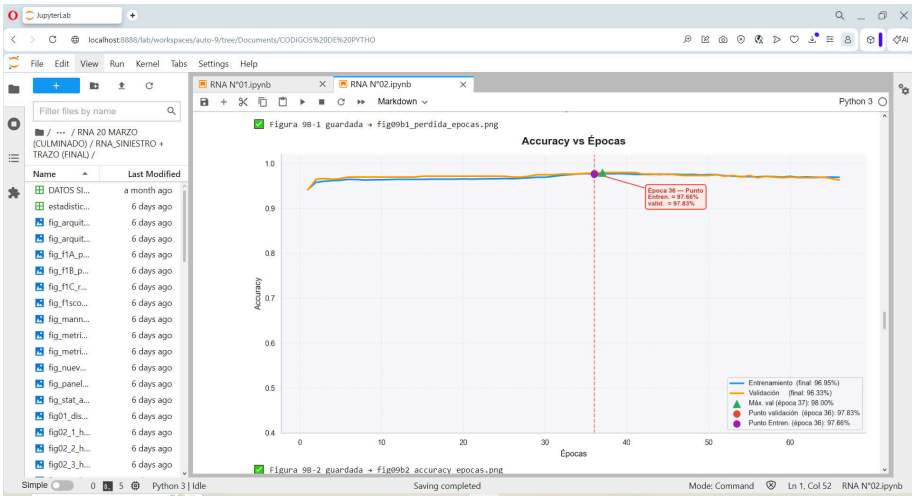


Figura 270 Gráfica de Accuracy vs epocas.
Fuente: Elaboración propia

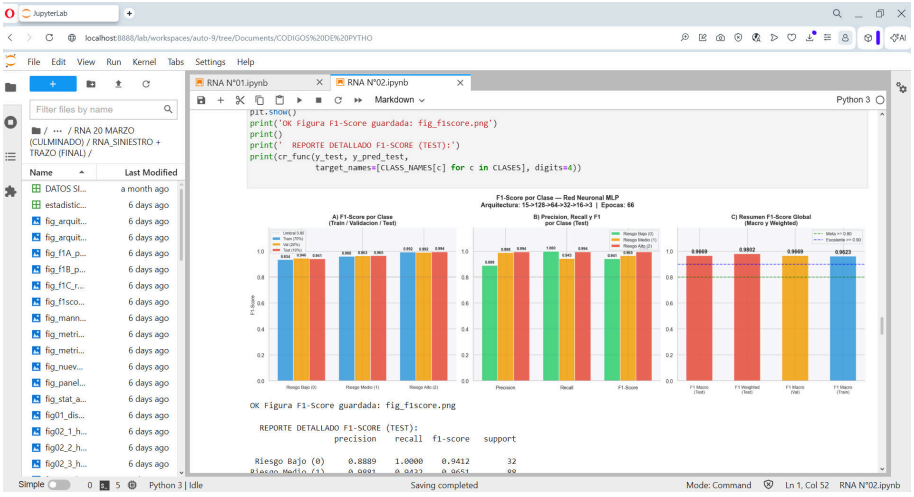


Figura 271 Gráfica de F1 - Score.
Fuente: Elaboración propia

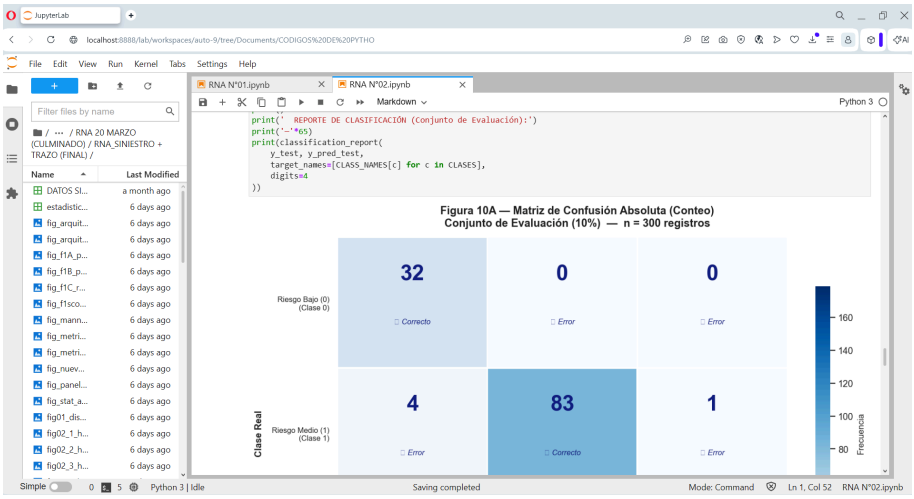


Figura 272 Matriz de confusión.
Fuente: Elaboración propia

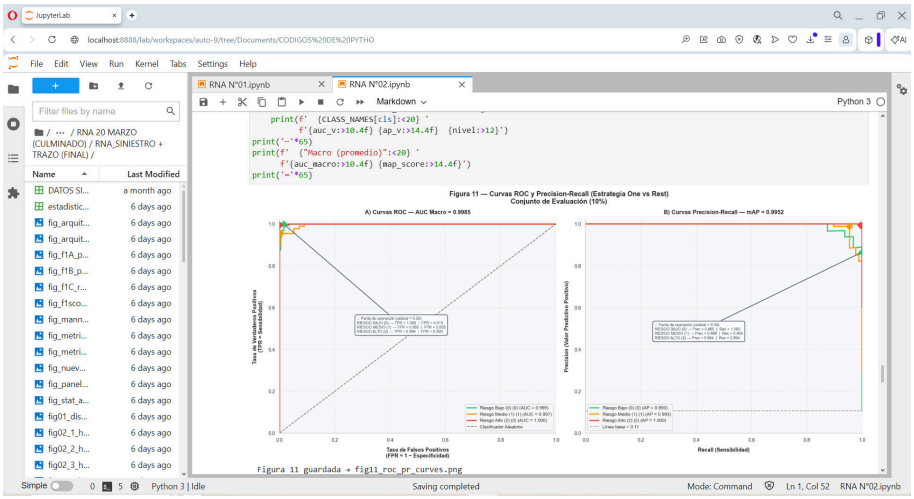


Figura 273 Curva ROC - AUC.
Fuente: Elaboración propia

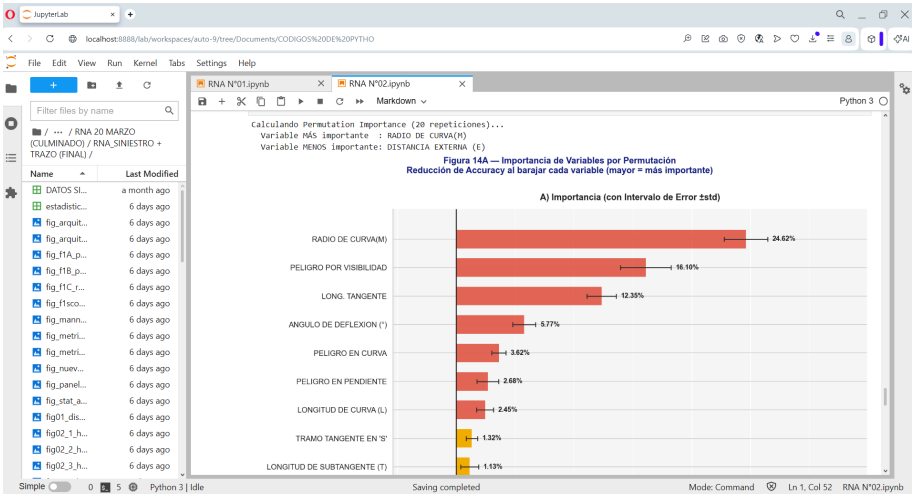


Figura 274 Importancia de variables.
Fuente: Elaboración propia

Figura 275 Gráfica de frontera de decisión.
Fuente: Elaboración propia

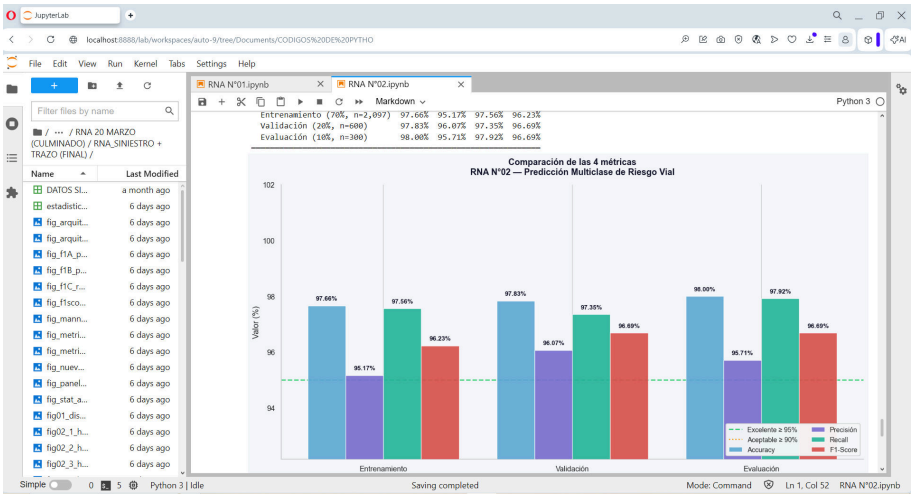


Figura 276 Resumen de metricas de la RNA.
Fuente: Elaboración propia

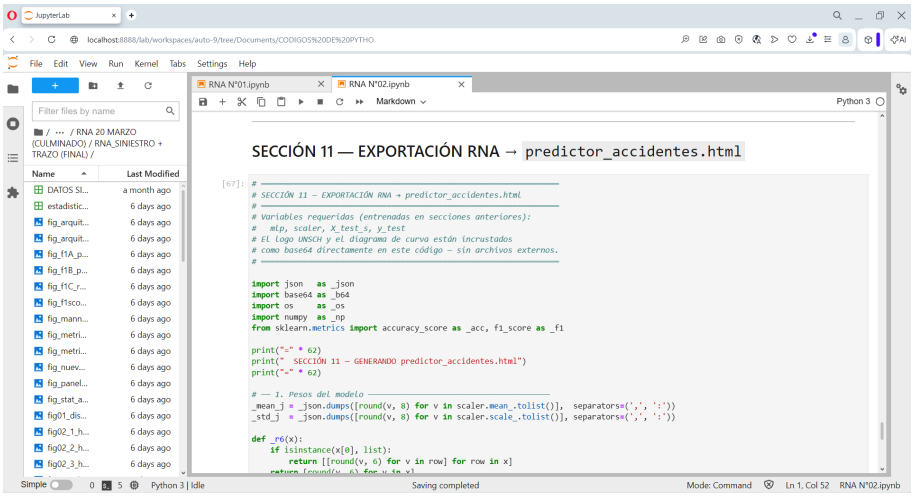


Figura 277 Generación del programa html.
Fuente: Elaboración propia

UNSCH FACULTAD DE INGENIERÍA DE MINAS, GEOLOGÍA Y CIVIL

ACTA DE SUSTENTACIÓN DE TESIS Nº 46-2026-FIMGC

PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO CIVIL

En la Universidad Nacional de San Cristóbal de Huamanga, en la ciudad de Ayacucho, en cumplimiento a la **Resolución Decanal No 260-2026-FIMGC-D**, a los **catorce días del mes de julio de 2026**, siendo las **10:26 a.m.**, reunidos en el **Auditorio de la Escuela Profesional de Ingeniería Civil**, bajo la presidencia del **MSc. José Ernesto ESTRADA CARDENAS**, y los miembros: **Ing. Hemerson LIZARBE ALARCON** y **Ing. Adolfo LINARES FLORES**, actuando como secretario docente el **Ing. Saul Walter RETAMOZO FERNANDEZ**, para proceder a la sustentación de tesis para optar el **Título Profesional de Ingeniero Civil**, del Bachiller en Ciencias de la Ingeniería Civil:

SOCRATES ALLCCAHUAMAN ANDIA

Quien presentó la tesis denominada:

Clasificación del peligro de accidentes de tránsito mediante redes neuronales artificiales y optimización del diseño geométrico según normas internacionales y DG-2018 en la vía Los Libertadores (PE-28A), Ayacucho

Los señores miembros del jurado luego de expuesta la tesis y absueltas las preguntas, deliberaron y declararon:

Aprobado con 16 (dieciséis)

Siendo las **11:52 a.m.** del día **11 de julio del 2026**, culmina el acto de sustentación de tesis, y en conformidad de lo actuado los miembros del jurado firmamos al pie del presente.

MSc. José Ernesto ESTRADA CARDENAS
Presidente

Ing. Adolfo LINARES FLORES
Miembro

Ing. Hemerson LIZARBE ALARCON
Miembro - Asesor

Ing. Saul Walter RETAMOZO FERNANDEZ
Secretario

FACULTAD DE INGENIERÍA
DE MINAS Y CIVIL
Av. Independencia S/N
Ciudad Universitaria
Central Tel. 066 312510
Anexo 151

CONSTANCIA DE ORIGINALIDAD DE TRABAJO DE INVESTIGACIÓN

CONSTANCIA N° 22-2026-FIMGC/ASIH

El que suscribe; responsable verificador de originalidad de trabajos de tesis de pregrado con el software Turnitin, de la Escuelas Profesional de **Ingeniería Civil** de la **Facultad de Ingeniería de Minas, Geología y Civil**; en cumplimiento a la **Resolución de Consejo Universitario N° 039-2021-UNSCH-CU**, Reglamento de Originalidad de Trabajos de Investigación de la Universidad Nacional San Cristóbal de Huamanga y **Resolución Decanal N° 697-2024-FIMGC-UNSCH-D**, dejo constancia de originalidad de trabajo de investigación, que el/la Sr./Srta.

Apellidos y Nombres	:	Socrates ALLCCAHUAMAN ANDIA
Escuela Profesional	:	INGENIERÍA CIVIL
Título de la Tesis	:	Clasificación del peligro de accidentes de tránsito mediante redes neuronales artificiales y optimización del diseño geométrico según normas internacionales y DG-2018 en la vía Los Libertadores (PE-28A), Ayacucho
Evaluación de la Originalidad	:	7 % Índice de Similitud
Identificador de la entrega	:	3005563353

Por tanto, según los Artículos 12, 13 y 17 del Reglamento de Originalidad de Trabajos de Investigación, es **PROCEDENTE** otorgar la **Constancia de Originalidad** para los fines que crea conveniente.

En señal de conformidad y verificación se firma la presente constancia

Ayacucho, 31 de julio del 2026

UNIVERSIDAD NACIONAL DE SAN CRISTÓBAL DE HUAMANGA
Facultad de Ingeniería de Minas, Geología y Civil

Mg. Ing. Alex Sander IRCAÑAUPA HUAMANI
Verificador de Originalidad de Trabajos de Tesis de Pregrado
Escuela de Formación Profesional de Ingeniería Civil

Con depósito para Tramites
Cc. Archivo.

Clasificación del peligro de accidentes de tránsito mediante redes neuronales artificiales y optimización del diseño geométrico según normas internacionales y DG-2018 en la vía Los Libertadores (PE-28A), Ayacucho

por Socrates ALLCCAHUAMAN ANDIA

Fecha de entrega: 30-jul-2026 07:46a. m. (UTC-0500)
Identificador de la entrega: 3005563353
Nombre del archivo: Te_Socrates_Allccahuaman_Andia.pdf (15.5M)
Total de palabras: 75595
Total de caracteres: 386602

Clasificación del peligro de accidentes de tránsito mediante redes neuronales artificiales y optimización del diseño geométrico según normas internacionales y DG-2018 en la vía Los Libertadores (PE-28A), Ayacucho

INFORME DE ORIGINALIDAD

7%	7%	2%	4%
INDICE DE SIMILITUD	FUENTES DE INTERNET	PUBLICACIONES	TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	Submitted to Universidad Nacional de San Cristóbal de Huamanga Trabajo del estudiante	3%
2	hdl.handle.net Fuente de Internet	1%
3	repositorio.unc.edu.pe Fuente de Internet	<1%
4	repositorio.urp.edu.pe Fuente de Internet	<1%
5	cdn.www.gob.pe Fuente de Internet	<1%
6	repositorio.ucv.edu.pe Fuente de Internet	<1%
7	Submitted to Universidad Americana Trabajo del estudiante	<1%
8	Submitted to Universidad Francisco de Vitoria Trabajo del estudiante	<1%
9	repositorio.uncp.edu.pe Fuente de Internet	<1%
10	vitela.javerianacali.edu.co Fuente de Internet	<1%
11	Submitted to Uniminuto Virtual Trabajo del estudiante	<1%
12	repositorio.uandina.edu.pe Fuente de Internet	<1%
13	repositorio.usm.cl Fuente de Internet	<1%
14	repository.udistrital.edu.co Fuente de Internet	<1%

15	Cristian Eduardo Cedeño Macías, Jessenia Aracely Reyes Chonillo. "Prediction of traffic accidents in the province of Manabí-Ecuador using machine learning algorithms", 2025 Fifth International Conference on Information Systems and Software Technologies (ICI2ST), 2025 Publicación	<1 %
16	oa.upm.es Fuente de Internet	<1 %
17	arxiv.org Fuente de Internet	<1 %
18	repositorio.continental.edu.pe Fuente de Internet	<1 %
19	repositorio.utn.edu.ec Fuente de Internet	<1 %
20	archive.org Fuente de Internet	<1 %
21	Solis, Claudia Stefanny Coyllo. "Modelo Prolab: Servicio Tecnológico de Monitoreo, supervisión y Asistencia de Condiciones de Manejo Para Usuarios de Vehículos Particulares – Cleverdriver", Pontificia Universidad Catolica del Peru (Peru), 2024 Publicación	<1 %
22	Submitted to Universidad Politécnica de Madrid Trabajo del estudiante	<1 %
23	César Corrales, Juan Carlos Rubio-Romero, María del Carmen Pardo-Ferreira. "Analysis of the Severity of Road Accidents Using Combined Data Mining Techniques", Sustainability, 2026 Publicación	<1 %
24	repositorio.undac.edu.pe Fuente de Internet	<1 %
25	www.theibfr.com Fuente de Internet	<1 %
26	repositorio.uti.edu.ec Fuente de Internet	<1 %
27	www.saludcapital.gov.co Fuente de Internet	<1 %
28	Submitted to Universidad Nacional Federico Villarreal Trabajo del estudiante	<1 %
29	Submitted to Universidad Torcuato di Tella - Tii Trabajo del estudiante	<1 %

30	Submitted to Universitat Politècnica de València Trabajo del estudiante	<1 %
31	repositorio.comillas.edu Fuente de Internet	<1 %
32	ri-ng.uaq.mx Fuente de Internet	<1 %
33	Submitted to Centro Universitario de la Guardia Civil (CUGC) Trabajo del estudiante	<1 %
34	repositorio.udh.edu.pe Fuente de Internet	<1 %
35	Submitted to Weatherford College Trabajo del estudiante	<1 %
36	repositorio.ucp.edu.pe Fuente de Internet	<1 %
37	tesisdigitales.umich.mx Fuente de Internet	<1 %
38	vsip.info Fuente de Internet	<1 %
39	Submitted to Universidad Tecnológica Centroamericana UNITEC Trabajo del estudiante	<1 %
40	laccei.org Fuente de Internet	<1 %
41	www.slideshare.net Fuente de Internet	<1 %
42	Javier Fernández-Río, Jose A. Cecchini, Antonio Méndez-Giménez. "Effects of cooperative learning on perceived competence, motivation, social goals, effort and boredom in prospective Primary Education teachers / Efectos del aprendizaje cooperativo sobre la competencia percibida, la motivación, las relaciones sociales, el esfuerzo y el aburrimiento de futuros docentes de Educación Primaria", Infancia y Aprendizaje, 2014 Publicación	<1 %
43	www.perulng.com Fuente de Internet	<1 %
44	Submitted to Universidad Andina Nestor Caceres Velasquez Trabajo del estudiante	<1 %

45	Quispe Lujano, Alfredo. "Formulación de un modelo de Red Neuronal Artificial para el pronóstico de concentración de oxígeno disuelto y clorofila en la bahía interior de Puno", Universidad Nacional del Altiplano de Puno (Peru) Publicación	<1%
46	Submitted to Unviersidad de Granada Trabajo del estudiante	<1%

Excluir citas	Activo	Excluir coincidencias	< 30 words
Excluir bibliografía	Activo		