



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ingeniería Electrónica y Eléctrica

Escuela Profesional de Ingeniería Biomédica

**Modelo computacional basado en datos para el
mantenimiento predictivo con soporte preventivo de
ambulancias médicas urbanas tipo II en Lima
Metropolitana**

TESIS

Para optar el Título Profesional de Ingeniero Biomédico

AUTOR

Diego Jesús OSORES AGUILAR

ASESOR

Dr. Franklin Alfredo CABEZAS HUERTA

Lima, Perú

2026



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Osores, D. (2026). *Modelo computacional basado en datos para el mantenimiento predictivo con soporte preventivo de ambulancias médicas urbanas tipo II en Lima Metropolitana*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería Electrónica y Eléctrica, Escuela Profesional de Ingeniería Biomédica]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Diego Jesús Osoreo Aguilar
Tipo de documento de identidad	DNI
Número de documento de identidad	71029001
URL de ORCID	https://orcid.org/0000-0001-8967-4156
Datos de asesor	
Nombres y apellidos	Franklin Alfredo Cabezas Huerta
Tipo de documento de identidad	DNI
Número de documento de identidad	43809849
URL de ORCID	https://orcid.org/0000-0001-7106-1087
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Jorge Elías Moscoso Sánchez
Tipo de documento	DNI
Número de documento de identidad	07206008
Miembro del jurado 1	
Nombres y apellidos	Arlich Joel Portillo Allende
Tipo de documento	DNI
Número de documento de identidad	10125495
Miembro del jurado 2	
Nombres y apellidos	Roberto Florentino Unsihuay Tovar
Tipo de documento	DNI
Número de documento de identidad	19944586
Datos de investigación	
Línea de investigación	C.19.0.26 Inteligencia artificial
Grupo de investigación	No Aplica
Agencia de financiamiento	No Aplica
Ubicación geográfica de la investigación	Ámbito: Lima Metropolitana y Callao País: Perú Departamento/Región: Lima y Provincia

	Constitucional del Callao Provincia: Lima y Callao Distrito: No aplica (ámbito metropolitano) Latitud: 11°45'–12°30' S Longitud: 76°45'–77°10' O
Año o rango de años en que se realizó la investigación	Enero 2024 – Diciembre 2025
URL de disciplinas OCDE	Ingeniería médica https://purl.org/pe-repo/ocde/ford#2.06.01 Ciencias de la computación e información https://purl.org/pe-repo/ocde/ford#1.02.00 Ingeniería eléctrica, Ingeniería electrónica https://purl.org/pe-repo/ocde/ford#2.02.01



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
(Universidad del Perú, DECANA DE AMÉRICA)
FACULTAD DE INGENIERÍA ELECTRÓNICA Y ELÉCTRICA
Teléfono 619-7000 Anexo 4226
Calle Germán Amezaga 375 – Lima 1 – Perú



ACTA DE SUSTENTACIÓN

TESIS

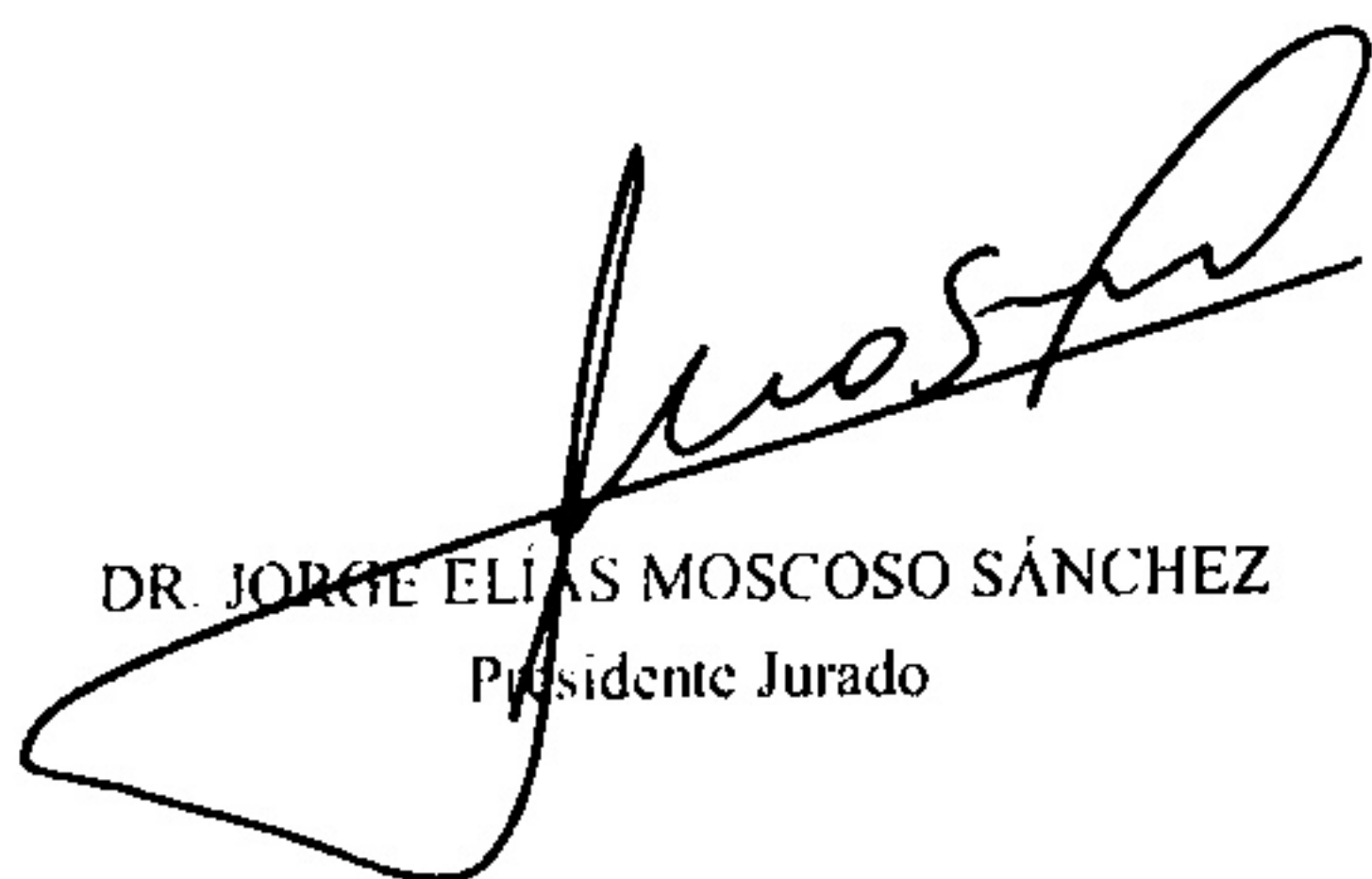
ACTA N°60/FIEE-PPP/2026

Los suscritos Miembros del Jurado, docentes permanentes de las Escuelas Profesionales de la Facultad de Ingeniería Electrónica y Eléctrica, reunidos en la fecha 28 de Agosto del 2026, como Presidente de Jurado el **DR. JORGE ELÍAS MOSCOSO SÁNCHEZ**, integrado por el Miembro de Jurado el **MG. ARLICH JOEL PORTILLO ALLENDE**, el Miembro de Jurado el **MG. ROBERTO FLORENTINO UNSIHUAY TOVAR** y Miembro Asesor el **DR. FRANKLIN ALFREDO CABEZAS HUERTA**.

Después de escuchar la Sustentación Presencial de la Tesis del Bachiller **DIEGO JESÚS OSORES AGUILAR** con código 19190119 que para optar el Título Profesional de Ingeniero Biomédico sustentó la Tesis titulada **MODELO COMPUTACIONAL BASADO EN DATOS PARA EL MANTENIMIENTO PREDICTIVO CON SOPORTE PREVENTIVO DE AMBULANCIAS MÉDICAS URBANAS TIPO II EN LIMA METROPOLITANA**.

El jurado examinador procedió a formular las preguntas reglamentarias y luego de una deliberación en privado, decidió otorgarle el calificativo de 14 (catorce).

Ciudad Universitaria, 28 de Agosto del 2026


DR. JORGE ELÍAS MOSCOSO SÁNCHEZ
Presidente Jurado


MG. ARLICH JOEL PORTILLO ALLENDE
Miembro Jurado


MG. ROBERTO FLORENTINO UNSIHUAY TOVAR
Miembro Jurado


DR. FRANKLIN ALFREDO CABEZAS HUERTA
Miembro Asesor

CERTIFICADO DE SIMILITUD

Yo, **FRANKLIN ALFREDO CABEZAS HUERTA** en mi condición de asesor acreditado con el Dictamen N°62/FIEE-PPP/2026 de la Tesis titulada **MODELO COMPUTACIONAL BASADO EN DATOS PARA EL MANTENIMIENTO PREDICTIVO CON SOPORTE PREVENTIVO DE AMBULANCIAS MÉDICAS URBANAS TIPO II EN LIMA METROPOLITANA**, presentado por el bachiller **DIEGO JESÚS OSORES AGUILAR**, para optar el título profesional de Ingeniero Biomédico. CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el 5% de similitud, nivel PERMITIDO para continuar con los trámites correspondientes y para su publicación en el repositorio institucional.

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del título profesional de Ingeniero Biomédico.

Lima, 28 de Agosto del 2026



Firma del Asesor

DNI: 43809849

Nombres y apellidos del asesor:

FRANKLIN ALFREDO CABEZAS HUERTA

DEDICATORIA

A Dios. A mis padres.

De mi madre aprendí que el conocimiento es la herramienta más poderosa que alguien puede entregar a otro. De mi padre aprendí que sin valores, ninguna herramienta sirve de mucho.

Al ingeniero que hoy culmina esta etapa, obra de ambos.

A mis abuelos, que partieron antes de ver este momento.

Sé que lo ven.

AGRADECIMIENTOS

A mis padres, por ser el origen de todo lo que soy.

A los docentes y mentores que encontré dentro y fuera de la universidad, quienes me enseñaron que la ingeniería biomédica no es solo una disciplina técnica sino una forma de entender el mundo y de servir a las personas.

A mis amigos y compañeros de promoción, con quienes compartí el reto de abrir camino como primera generación de ingenieros biomédicos de nuestra universidad. Gracias por el apoyo, los aprendizajes y por ser parte de esta historia.

A ustedes les dejo este mensaje: Siempre siéntanse orgullosos de ser ingenieros biomédicos, de ser sanmarquinos, de haber escrito esta historia, porque la escribimos juntos.

RESUMEN

La inoperatividad de ambulancias médicas urbanas Tipo II limita la capacidad de respuesta del sistema prehospitalario peruano en Lima Metropolitana. El mantenimiento preventivo establecido en la Norma Técnica de Salud N.º 051-MINSA opera principalmente sobre intervalos fijos, sin incorporar sistemáticamente información histórica sobre uso, mantenimiento e inoperatividad de las unidades.

Se desarrolló un modelo computacional de clasificación binaria supervisada para estimar la probabilidad de inoperatividad de ambulancias en un horizonte de 14 días. Se construyó un dataset sintético de 3102 observaciones mediante simulación paramétrica-estocástica que corresponde a una flota simulada de 33 ambulancias durante 2024-2025. La simulación fue calibrada con parámetros documentados sobre la flota SAMU de Lima Metropolitana, la ficha de homologación de ambulancia urbana Tipo II, la normativa nacional aplicable y literatura técnica de mantenimiento predictivo. Se evaluaron cuatro algoritmos mediante validación retrospectiva temporal: Regresión Logística, Árboles de Decisión, *Random Forest* y *Gradient Boosting*.

Al usar *Random Forest* se obtuvo el mejor desempeño relativo para la priorización preventiva con sensibilidad de 91,86% frente al 30,52% de sensibilidad del esquema preventivo tradicional. Los resultados permitieron estructurar lineamientos técnicos, criterios de alerta temprana y una plantilla semanal institucional para apoyar la toma de decisiones de mantenimiento.

Palabras clave: mantenimiento predictivo, ambulancias Tipo II, *Random Forest*, aprendizaje supervisado, disponibilidad operativa.

ABSTRACT

The inoperability of Type II urban medical ambulances limits the response capacity of the Peruvian prehospital system in Metropolitan Lima. The preventive maintenance approach established in Health Technical Standard No. 051-MINSA operates mainly on fixed intervals, without systematically incorporating historical information on unit usage, maintenance, and inoperability.

A supervised binary classification computational model was developed to estimate the probability of ambulance inoperability over a 14-day horizon. A synthetic dataset of 3,102 observations was constructed through parametric-stochastic simulation, corresponding to a simulated fleet of 33 ambulances during 2024-2025. The simulation was calibrated using documented parameters on the SAMU fleet in Metropolitan Lima, the Type II urban ambulance homologation sheet, applicable national regulations, and technical literature on predictive maintenance. Four algorithms were evaluated through temporal retrospective validation: Logistic Regression, Decision Trees, Random Forest, and Gradient Boosting.

Random Forest achieved the best relative performance for preventive prioritization, with a sensitivity of 91.86%, compared with 30.52% sensitivity for the traditional preventive scheme. The results enabled the structuring of technical guidelines, early warning criteria, and an institutional weekly template to support maintenance decision-making.

Keywords: predictive maintenance, Type II ambulances, Random Forest, supervised learning, operational availability.

ÍNDICE GENERAL

CARÁTULA	i
RESUMEN	iv
ABSTRACT	v
ÍNDICE GENERAL.....	vi
LISTA DE FIGURAS	xii
LISTA DE TABLAS	xvi
CAPÍTULO I : INTRODUCCIÓN	1
I.1. Introducción	1
I.2. Planteamiento del problema	2
I.2.1. Determinación del problema	2
I.2.2. Formulación del problema	5
I.3. Objetivos	6
I.3.1. Objetivo General	6
I.3.2. Objetivos Específicos	6
I.4. Importancia y alcance de la investigación	7
I.4.1. Justificación sanitaria y social	7
I.4.2. Justificación tecnológica y científica	8
I.4.3. Justificación normativa e institucional	9
I.5. Limitaciones de la Investigación	10
CAPÍTULO II : REVISIÓN DE LITERATURA.....	12
II.1. Marco Teórico	12
II.1.1. Atención prehospitalaria y sistema de ambulancias	12

	vii
II.1.2. Clasificación normativa de las ambulancias médicas	15
II.1.3. Gestión del mantenimiento en ambulancias médicas	23
II.1.4. Mantenimiento predictivo: evolución y fundamentos conceptuales	27
II.1.5. Mantenimiento predictivo basado en datos	32
II.1.6. Modelos computacionales para mantenimiento predictivo	39
II.1.7. Integración del modelo predictivo como soporte al mantenimiento preventivo	66
II.1.8. Validación de modelos predictivos en mantenimiento	71
II.1.9. Relación del marco teórico con la investigación	78
II.2. Antecedentes del estudio	80
II.2.1. Antecedente nacional 1	80
II.2.2. Antecedente nacional 2	81
II.2.3. Antecedente internacional 1	82
II.2.4. Antecedente internacional 2	83
II.3. Bases teóricas	84
II.3.1. Primer eje teórico: El mantenimiento predictivo data-driven como enfoque metodológico	85
II.3.2. Segundo eje teórico: El modelo computacional de clasificación binaria supervisada	86
II.3.3. Tercer eje teórico: La validación retrospectiva como estrategia metodológica	87
II.4. Definición de términos	87
CAPÍTULO III : HIPÓTESIS Y VARIABLES	92
III.1. Hipótesis	92
III.1.1. Hipótesis General	92
III.1.2. Hipótesis Específicas	92
III.2. Variables	93

	viii
III.2.1. Variable independiente	93
III.2.2. Variable dependiente	93
III.2.3. Variables de control	94
III.3. Operacionalización de las variables	94
CAPÍTULO IV : MATERIALES Y MÉTODOS.....	96
IV.1. Área de estudio	96
IV.2. Diseño de investigación	97
IV.3. Población y muestra	97
IV.3.1. Ámbito del estudio	97
IV.3.2. Población	98
IV.3.3. Unidad de análisis	98
IV.4. Procedimientos, técnicas e instrumentos de recolección de información	99
IV.4.1. Fuentes de datos	99
IV.4.2. Materiales y herramientas de software	112
IV.4.3. Parámetros temporales del modelo	113
IV.4.4. Variable objetivo (dependiente)	114
IV.4.5. Variables explicativas (independientes)	114
IV.4.6. Valoración técnica complementaria mediante juicio de especialistas	115
IV.5. Análisis estadístico	116
Fase 1: Generación y preparación de datos simulados	117
Fase 2: Construcción del dataset por ventanas temporales	119
Fase 3: Modelado computacional basado en datos	121
Fase 4: Validación retrospectiva (backtesting temporal)	123
Fase 5: Generación de salidas para soporte preventivo	125
CAPÍTULO V : RESULTADOS	127
V.1. Presentación y análisis de los resultados	127

V.1.1. Descripción del dataset de observaciones	127
V.1.2. Análisis exploratorio de variables explicativas	132
V.1.3. Desempeño del modelo computacional - Métricas de validación	136
V.1.4. Importancia de variables explicativas	149
V.1.5. Prototipo de soporte preventivo	152
V.1.6. Lineamientos técnicos derivados del modelo	156
V.1.7. Valoración técnica complementaria mediante juicio de especialistas	161
CAPÍTULO VI : DISCUSIÓN.....	163
VI.1. Sobre el enfoque metodológico y la viabilidad del modelo <i>data-driven</i>	163
VI.2. Sobre el desempeño del modelo y su comparación con la línea base	164
VI.3. Sobre la importancia de variables y sus implicancias para la gestión	169
VI.4. Sobre el soporte preventivo y su integración al sistema prehospitalario	170
VI.5. Sobre la contrastación de hipótesis	173
VI.6. Limitaciones y perspectivas de mejora	174
CAPÍTULO VII : CONCLUSIONES	176
VII.1. Conclusión 1 - Sobre el objetivo específico 1: Análisis de factores técnicos y operativos	176
VII.2. Conclusión 2 - Sobre el objetivo específico 2: Diseño de la base de datos	177
VII.3. Conclusión 3 - Sobre el objetivo específico 3: Implementación del modelo computacional	177
VII.4. Conclusión 4 - Sobre el objetivo específico 4: Validación del modelo	178

VII.5. Conclusión 5 - Sobre el objetivo específico 5: Lineamientos técnicos y operativos	178
VII.6. Conclusión general	179
CAPÍTULO VIII : RECOMENDACIONES.....	181
VIII.1. Recomendación 1 - Al SAMU: gestión y estandarización de registros	181
VIII.2. Recomendación 2 - Al SAMU: implementación piloto del modelo	181
VIII.3. Recomendación 3 - A la comunidad académica: replicación con datos reales	182
VIII.4. Recomendación 4 - A la comunidad académica: enriquecimiento del modelo	182
VIII.5. Recomendación 5 - A la Escuela Profesional de Ingeniería Biomédica	183
CAPÍTULO IX : BIBLIOGRAFÍA.....	184
CAPÍTULO X : ANEXOS.....	191
ANEXO A - Estructura del <i>dataset</i> simulado y diccionario de variables	191
A.1. Descripción general del dataset	191
A.2. Diccionario de variables del dataset de observaciones	192
ANEXO B - Código fuente del generador de <i>dataset</i> simulado	193
ANEXO C - Código fuente del modelo computacional predictivo	195
ANEXO D - Código fuente de lineamientos técnicos, soporte preventivo y plantilla semanal	197
D.1. Script de lineamientos técnicos y soporte preventivo	197
D.2. Script de generación de plantilla semanal	199
ANEXO E - Código fuente del análisis inferencial complementario de clasificadores	201
ANEXO F - Resultados completos de la validación retrospectiva	205
F.1. Tabla completa de métricas por modelo	205

F.2. Importancia de variables completa	206
ANEXO G - Protocolo operativo completo de uso del modelo	206
ANEXO H - Criterios de alerta temprana por variable crítica	207
ANEXO I - Plantilla semanal de programación de mantenimiento con soporte predictivo.	208
ANEXO J - Matriz de consistencia de la investigación	213
ANEXO K - Valoración técnica complementaria mediante juicio de especialistas	216
K.1. Cuestionario aplicado	216
K.2. Perfil de los profesionales participantes	231
K.3. Matriz consolidada de respuestas	231
K.4. Resultados consolidados por dimensión	233
K.5. Observaciones y recomendaciones profesionales	234

LISTA DE FIGURAS

<i>Figura 1.</i> Representación de los indicadores técnicos de mantenimiento sobre la línea de tiempo operativa de una ambulancia Tipo II.	26
<i>Figura 2.</i> Flujo conceptual del mantenimiento predictivo basado en datos históricos.	38
<i>Figura 3.</i> Función sigmoide empleada en la regresión logística para la transformación de la combinación lineal de variables en probabilidad.	44
<i>Figura 4.</i> Distribución de clases antes y después del balanceo mediante oversampling aleatorio con reemplazo.	47
<i>Figura 5.</i> Estructura ilustrativa de un árbol de decisión aplicado al problema de clasificación de inoperatividad en ambulancias Tipo II	49
<i>Figura 6.</i> Esquema de bagging en Random Forest: muestreo con reemplazo, entrenamiento de K árboles independientes y agregación por votación.	51
<i>Figura 7.</i> Proceso secuencial del Gradient Boosting: cada árbol débil corrige los errores del modelo anterior mediante el gradiente de la función de pérdida.	55
<i>Figura 8.</i> Estructura de la matriz de confusión y métricas derivadas para la evaluación de modelos de clasificación binaria.	57
<i>Figura 9.</i> Curvas ROC comparativas de los modelos evaluados en la validación retrospectiva.	61
<i>Figura 10.</i> Relación entre precisión y sensibilidad según el umbral de clasificación y comparativa de sensibilidad entre modelos.	63
<i>Figura 11.</i> Relación entre el sistema real, los datos históricos y el modelo computacional predictivo.	65

<i>Figura 12.</i> Integración del modelo computacional predictivo como soporte al mantenimiento preventivo de ambulancias médicas urbanas Tipo II.	71
<i>Figura 13.</i> Esquema conceptual del proceso de validación retrospectiva (backtesting temporal).	78
<i>Figura 14.</i> Decodificación del identificador de ambulancia empleado en el dataset sintético.	101
<i>Figura 15.</i> Muestra de la tabla de mantenimiento generada por el script generar_dataset_ambulancias.py.	103
<i>Figura 16.</i> Muestra de la tabla de inoperatividad (downtime) generada por el script generar_dataset_ambulancias.py.	104
<i>Figura 17.</i> Muestra de la tabla de uso operativo generada por el script generar_dataset_ambulancias.py.	105
<i>Figura 18.</i> Arquitectura del simulador paramétrico-estocástico para la generación de tablas fuente sintéticas.	118
<i>Figura 19.</i> Construcción del dataset de observaciones mediante ventanas temporales deslizantes.	120
<i>Figura 20.</i> Diagrama de flujo del pipeline de entrenamiento de los modelos de clasificación supervisada.	122
<i>Figura 21.</i> Lógica de construcción del dataset mediante ventanas temporales deslizantes.	128
<i>Figura 22.</i> Estructura del dataset de observaciones.	129
<i>Figura 23.</i> Distribución de clases de la variable objetivo.	131
<i>Figura 24.</i> Distribución de la variable objetivo por conjunto de datos. Prevalencia de inoperatividad ($Y = 1$) en cada período.	131
<i>Figura 25.</i> Mapa de calor de correlaciones entre las 14 variables explicativas. Dataset de observaciones. $n = 3\ 102$. Período 2024-2025.	133
<i>Figura 26.</i> Diagramas de dispersión entre variables explicativas relevantes y eventos de inoperatividad.	134

<i>Figura 27.</i> Distribución de variables explicativas principales según la clase objetivo.	135
<i>Figura 28.</i> Matrices de confusión comparativas de los cinco enfoques evaluados. Período de validación 2025. $n = 1\ 650$. Umbral $\tau = 0,30$.	137
<i>Figura 29.</i> Perfil de desempeño de los modelos evaluados. Período de validación 2025. Umbral $\tau = 0,30$.	139
<i>Figura 30.</i> Curvas ROC comparativas - Validación retrospectiva período 2025.	140
<i>Figura 31.</i> Distribución de probabilidades predichas por clase - Modelo Random Forest.	141
<i>Figura 32.</i> Curvas Precision-Recall comparativas - Validación retrospectiva período 2025.	148
<i>Figura 33.</i> Importancia de variables explicativas - Modelo Random Forest. Top 10 variables.	151
<i>Figura 34.</i> Estructura del archivo de salida del prototipo de soporte preventivo. Primeras filas con la probabilidad estimada y el nivel de riesgo por unidad.	153
<i>Figura 35.</i> Lineamientos técnicos y operativos para la priorización de mantenimiento con soporte predictivo en ambulancias Tipo II durante el período de validación 2025.	154
<i>Figura 36.</i> Protocolo semanal de uso del modelo computacional predictivo para soporte al mantenimiento preventivo de ambulancias médicas urbanas Tipo II.	157
<i>Figura 37.</i> Primera hoja de la plantilla semanal de programación de mantenimiento con soporte predictivo.	159
<i>Figura 38.</i> Segunda hoja de la plantilla semanal de programación de mantenimiento con soporte predictivo.	160
<i>Figura B.1.</i> Código QR de acceso al repositorio digital de la investigación: generador del dataset simulado.	195

<i>Figura C.1.</i> Código QR de acceso al repositorio digital de la investigación: modelo computacional predictivo y validación retrospectiva.	197
<i>Figura D.1.</i> Código QR de acceso al repositorio digital de la investigación: lineamientos técnicos y soporte preventivo.	199
<i>Figura D.2.</i> Código QR de acceso al repositorio digital de la investigación: generador de la plantilla semanal de programación de mantenimiento.	201
<i>Figura E.1.</i> Código QR de acceso al repositorio digital de la investigación: análisis inferencial complementario de clasificadores.	204
<i>Figura I.1.</i> Vista ampliada de la primera hoja de la plantilla semanal.	209
<i>Figura I.2.</i> Vista ampliada del detalle operativo de unidades priorizadas.	210
<i>Figura I.3.</i> Vista ampliada del control y seguimiento semanal	211
<i>Figura I.4.</i> Vista ampliada de la segunda hoja de datos completos	212

LISTA DE TABLAS

<i>Tabla 1</i> Clasificación de ambulancias según UNE-EN 1789:2020	15
<i>Tabla 2</i> Criterios funcionales de clasificación de ambulancias según normas técnicas internacionales y regionales	16
<i>Tabla 3</i> Clasificación de ambulancias terrestres según el lugar donde prestan el servicio	17
<i>Tabla 4</i> Clasificación de ambulancias según equipamiento y tipo de atención (NTS N.º 051-MINSA/OGDN-V.01)	18
<i>Tabla 5</i> Configuración técnico-operativa y alcance funcional de la ambulancia médica Tipo II	22
<i>Tabla 6</i> Comparación entre enfoques de mantenimiento preventivo y correctivo en ambulancias médicas	24
<i>Tabla 7</i> Indicadores técnicos utilizados en la gestión del mantenimiento de ambulancias médicas	25
<i>Tabla 8</i> Comparación conceptual entre mantenimiento preventivo y mantenimiento predictivo	30
<i>Tabla 9</i> Diferenciación práctica entre mantenimiento predictivo basado en datos históricos y basado en monitoreo continuo	36
<i>Tabla 10</i> Tipologías de modelos computacionales en mantenimiento predictivo	64
<i>Tabla 11</i> Relación entre mantenimiento preventivo tradicional y mantenimiento con soporte predictivo	70
<i>Tabla 12</i> Métricas de desempeño para la validación del modelo computacional predictivo	75
<i>Tabla 13</i> Enfoques de validación en modelos predictivos de mantenimiento	76
<i>Tabla 14</i> Operacionalización de variables de la investigación	95
<i>Tabla 15</i> Parámetros del mecanismo jerárquico de generación del riesgo de inoperatividad	110

<i>Tabla 16</i> Métricas de desempeño comparativas - Validación retrospectiva (período 2025)	136
<i>Tabla 17</i> Comparación estadística de clasificadores mediante la prueba de McNemar - Validación retrospectiva 2025.	144
<i>Tabla 18</i> Comparación del AUC-ROC mediante la prueba de DeLong para clasificadores correlacionados - Validación retrospectiva 2025.	146
<i>Tabla 19</i> Importancia de variables explicativas - Modelo Random Forest, Top 10	150
<i>Tabla 20.</i> Resultados de la valoración técnica complementaria mediante juicio de especialistas	162
<i>Tabla A.1</i> <i>Parámetros de simulación del dataset</i>	192
<i>Tabla A.2</i> <i>Diccionario de variables - Dataset de observaciones</i>	192
<i>Tabla F.1</i> <i>Resultados completos de la validación retrospectiva - Periodo 2025</i>	205
<i>Tabla F.2</i> <i>Importancia de variables completa - Modelo Random Forest</i>	206
<i>Tabla G.1</i> <i>Protocolo semanal de uso del modelo computacional predictivo</i>	207
<i>Tabla H.1</i> <i>Criterios de alerta temprana basados en importancia de variables del modelo</i>	208

CAPÍTULO I : INTRODUCCIÓN

I.1. Introducción

La gestión eficiente de los activos críticos en el sistema de salud constituye un desafío de creciente relevancia para la ingeniería biomédica, particularmente en contextos donde la disponibilidad operativa de los equipos y unidades médicas incide directamente sobre la capacidad de respuesta ante emergencias (WHO, 2011; *National Academies of Sciences, Engineering, and Medicine*, 2020). En este escenario, las ambulancias médicas urbanas ocupan un lugar central como plataformas de atención prehospitalaria móvil, cuya operatividad sostenida es condición necesaria para garantizar la continuidad asistencial del paciente crítico desde el lugar del evento hasta el establecimiento de salud (WHO, 2005).

Los enfoques tradicionales de mantenimiento, basados en intervenciones programadas por intervalos fijos de tiempo o uso, presentan limitaciones reconocidas en la literatura cuando se aplican a sistemas con patrones de operación heterogéneos y alta criticidad funcional (Moubray, 1997; Molęda et al., 2023). El desarrollo de modelos computacionales basados en datos históricos, conocidos como enfoques *data-driven*, ofrece una alternativa metodológicamente fundamentada para anticipar el comportamiento de los activos y fortalecer la planificación del mantenimiento sin requerir instrumentación adicional, lo que los hace especialmente pertinentes en entornos con restricciones presupuestales y tecnológicas (Taoufyq et al., 2025; Prytz, 2014). En el caso de las ambulancias médicas, esta aproximación resulta especialmente relevante debido a que la unidad constituye un sistema técnico integrado por componentes vehiculares,

eléctricos y biomédico-asistenciales. Por ello, la gestión de su mantenimiento requiere criterios que relacionen el historial de uso, las intervenciones técnicas y los eventos de inoperatividad con la disponibilidad operativa.

La presente investigación se inscribe en esta perspectiva y propone el desarrollo de un modelo computacional de mantenimiento predictivo con soporte preventivo para ambulancias médicas urbanas Tipo II en Lima Metropolitana, sustentado en el análisis de registros simulados de operación y mantenimiento, estructurados bajo una lógica histórica, temporal y reproducible. El documento está organizado en ocho capítulos de desarrollo: el Capítulo I establece el planteamiento del problema, los objetivos, la importancia y las limitaciones; el Capítulo II desarrolla el marco teórico, los antecedentes, las bases teóricas y la definición de términos; el Capítulo III presenta las hipótesis y variables; el Capítulo IV describe los materiales y métodos; y los Capítulos V, VI, VII y VIII presentan los resultados, la discusión, las conclusiones y las recomendaciones de la investigación.

I.2. Planteamiento del problema

I.2.1. Determinación del problema

En el sistema de salud, la atención prehospitalaria constituye el primer eslabón de la cadena de respuesta inmediata ante emergencias, cuya calidad determina directamente la supervivencia y recuperación de los pacientes, siendo esencial para la reducción de la morbilidad y mortalidad (Cimino & Braun, 2023). Bajo este contexto, las ambulancias son el recurso que articula la atención inicial con la atención hospitalaria, trasladando a los pacientes hacia los centros de salud y facilitando el contacto con personal especializado y equipamiento de soporte vital. Por esta razón, la disponibilidad y el correcto funcionamiento de estas unidades son determinantes para garantizar una respuesta médica rápida y eficiente.

En el Perú, la atención prehospitalaria está a cargo del Sistema de Atención

Móvil de Urgencias (SAMU), dependiente del Ministerio de Salud. En enero de 2025, el MINSA publicó que el SAMU registró 37 107 atenciones médicas prehospitalarias a nivel nacional durante el año 2024 (Ministerio de Salud del Perú [MINSA], 2025), cifra que evidencia la alta demanda y la importancia operativa de este servicio en el país. Este volumen de atenciones refleja la dependencia del sistema de salud respecto a la disponibilidad permanente de las ambulancias para garantizar la continuidad de la respuesta prehospitalaria. En este contexto, cada unidad fuera de servicio representa una reducción directa de la capacidad operativa del sistema, incrementando la presión sobre la flota disponible y comprometiendo la oportunidad de atención durante situaciones de emergencia.

Conforme a la Norma Técnica de Salud N.º 051-MINSA/DGSP-V.01 (Ministerio de Salud del Perú, 2006), las ambulancias se clasifican en función de su complejidad de atención médica y el equipamiento que incorporan, distinguiéndose tres categorías: Tipo I, Tipo II y Tipo III. La ambulancia Tipo II está destinada a la atención de pacientes que requieren monitorización y soporte vital básico, por lo que debe contar con sistemas eléctricos, de oxigenoterapia, de aspiración y de monitorización de signos vitales. En consecuencia, su mantenimiento no se limita a los componentes vehiculares, sino que abarca también los equipos y dispositivos médicos que contienen, lo que confiere a su gestión técnica una complejidad mayor respecto a otros tipos de unidades. Esta integración de sistemas vehiculares y biomédicos implica que las decisiones de mantenimiento deban considerar simultáneamente aspectos mecánicos, eléctricos y funcionales de los dispositivos médicos incorporados. Como consecuencia, la indisponibilidad de cualquiera de estos subsistemas puede afectar la operatividad integral de la ambulancia y limitar su capacidad para brindar atención segura durante la respuesta prehospitalaria.

Sin embargo, el servicio presenta diversas limitaciones operativas, especialmente en Lima Metropolitana, donde la demanda supera la capacidad de respuesta disponible. La Defensoría del Pueblo (2023a)

verificó que parte de las ambulancias asignadas al SAMU se encuentran en estado inoperativo o con fallas técnicas activas, con impacto directo en la disponibilidad y los tiempos de atención. Por ejemplo, en diversas regiones del país se han documentado demoras en la gestión del mantenimiento y reparación de unidades móviles (Defensoría del Pueblo, 2023b). Estas condiciones evidencian la necesidad de fortalecer la gestión técnica y la disponibilidad operativa de las ambulancias médicas urbanas para garantizar una atención prehospitalaria adecuada y segura.

Actualmente, el mantenimiento de las ambulancias se lleva a cabo bajo un enfoque preventivo programado, establecido en función del kilometraje recorrido o el tiempo de uso, conforme a la Norma Técnica de Salud (NTS) N.º 051 (MINSA, 2006). Sin embargo, este esquema no considera el estado real de los componentes ni permite anticipar fallas antes de que deriven en inoperatividad, lo que evidencia la necesidad de incorporar un enfoque predictivo basado en análisis de datos. Boppana (2023) demostró que el uso de analítica de datos e inteligencia artificial para el mantenimiento predictivo en el sector salud permite identificar fallas con anticipación, favoreciendo la reducción de costos de mantenimiento, la mejora de la disponibilidad operativa y el incremento de la seguridad del paciente. La aplicación de estos principios al contexto vehicular y biomédico propio de las ambulancias médicas urbanas representa una vía técnicamente viable para superar las limitaciones del modelo preventivo vigente.

En el contexto de países de ingresos bajos y medios, los estándares internacionales recomiendan contar con al menos una ambulancia por cada 50 000 habitantes para garantizar una atención prehospitalaria efectiva (Mishra et al., 2020). Esta referencia adquiere relevancia particular en Lima Metropolitana, donde la brecha entre la dotación requerida y la disponibilidad operativa real de las unidades del SAMU se ve agravada por las fallas técnicas y los periodos de inoperatividad documentados (Defensoría del Pueblo, 2023a). Esta situación se explica, en parte, por deficiencias en la gestión técnica del mantenimiento y por la ausencia de herramientas de seguimiento que permitan anticipar fallas antes de que deriven en

inoperatividad. Frente a este escenario, la presente investigación propone el desarrollo de un modelo computacional de mantenimiento predictivo con soporte preventivo que fortalezca la gestión técnica y apoye la toma de decisiones orientada a mejorar la disponibilidad operativa de las ambulancias médicas urbanas Tipo II en Lima Metropolitana.

1.2.2. Formulación del problema

Problema General

¿Cómo desarrollar un modelo computacional basado en datos para el mantenimiento predictivo con soporte preventivo que estime la probabilidad de inoperatividad y apoye la gestión técnica de las ambulancias médicas urbanas Tipo II en Lima Metropolitana durante 2024-2025?

Problemas Específicos

- ¿Cuáles son los principales factores técnicos y operativos que influyen en la ocurrencia de fallas, eventos de inoperatividad y requerimientos de mantenimiento de las ambulancias médicas urbanas Tipo II en Lima Metropolitana?
- ¿Cómo se puede estructurar una base de datos que integre variables operativas, técnicas, de mantenimiento e inoperatividad, de modo que permita el análisis predictivo del desempeño de las ambulancias médicas urbanas Tipo II?
- ¿Qué tipo de modelo computacional basado en datos permite estimar la probabilidad de inoperatividad de las ambulancias médicas urbanas Tipo II, integrando variables de uso operativo, mantenimiento, *downtime* y equipamiento funcional?
- ¿Qué indicadores de desempeño: exactitud, especificidad, precisión, sensibilidad, F1-Score, error medio absoluto (MAE), AUC-ROC y disponibilidad operativa proyectada permiten validar el desempeño del modelo predictivo propuesto frente al esquema de mantenimiento

preventivo programado como línea base de comparación?

- ¿Qué lineamientos técnicos y operativos pueden derivarse del modelo computacional propuesto para apoyar la priorización preventiva, fortalecer la gestión del mantenimiento y orientar la mejora de la disponibilidad operativa de las ambulancias médicas urbanas en Lima Metropolitana?

I.3. Objetivos

I.3.1. Objetivo General

Desarrollar un modelo computacional basado en datos para el mantenimiento predictivo con soporte preventivo de ambulancias médicas urbanas Tipo II en Lima Metropolitana durante el período 2024-2025, orientado a estimar la probabilidad de inoperatividad y apoyar la gestión técnica del sistema prehospitalario.

I.3.2. Objetivos Específicos

- Analizar los principales factores técnicos y operativos que influyen en la ocurrencia de fallas, eventos de inoperatividad y requerimientos de mantenimiento de las ambulancias médicas urbanas Tipo II, cuya configuración técnica está definida en la NTS N.º 051 (MINSA, 2006).
- Diseñar una base de datos estructurada que integre variables operativas, técnicas, de mantenimiento e inoperatividad representativas, permitiendo el almacenamiento y análisis de información para el entrenamiento del modelo computacional.
- Implementar un modelo computacional basado en datos que combine componentes de mantenimiento predictivo y preventivo, mediante

algoritmos de clasificación supervisada aplicados a registros históricos simulados de operación, mantenimiento e inoperatividad, para estimar la probabilidad de inoperatividad de las ambulancias médicas urbanas Tipo II.

- Validar el desempeño del modelo computacional mediante métricas cuantitativas de clasificación: exactitud, especificidad, precisión, sensibilidad, F1-Score, error medio absoluto (MAE), AUC-ROC y disponibilidad operativa proyectada, comparando sus resultados frente al esquema de mantenimiento preventivo programado como línea base.
- Proponer lineamientos técnicos y operativos para la adopción del modelo computacional en la gestión del mantenimiento de ambulancias médicas urbanas Tipo II, orientados a apoyar la priorización preventiva, reducir eventos de inoperatividad y fortalecer la toma de decisiones técnicas del sistema prehospitalario en Lima Metropolitana.

I.4. Importancia y alcance de la investigación

I.4.1. Justificación sanitaria y social

El fortalecimiento de la atención prehospitalaria se ha consolidado como una prioridad en los sistemas de salud a nivel global, dado que representa la primera línea de respuesta ante emergencias médicas. La atención prehospitalaria efectiva se traduce en reducción de la mortalidad y mejora de los resultados clínicos de los pacientes (Cimino & Braun, 2023). Sin embargo, cuando los recursos disponibles son insuficientes para gestionar eficientemente las ambulancias y ejecutar el mantenimiento técnico en el momento oportuno, ello incide negativamente en la capacidad de respuesta ante emergencias (Bhattarai et al., 2023). El Perú enfrenta precisamente esta situación: una alta demanda de atenciones

prehospitalarias, documentada en los reportes del SAMU, coexiste con una disponibilidad operativa limitada por fallas mecánicas y tiempos prolongados de inoperatividad (Defensoría del Pueblo, 2023a).

Desde una perspectiva social, cada ambulancia inoperativa dificulta o imposibilita que la atención médica llegue a los pacientes en el momento oportuno, lo que puede derivar en complicaciones médicas graves o incluso en la muerte. Esto evidencia que la operatividad de las ambulancias y la gestión de su mantenimiento son factores clave para el acceso equitativo a los servicios de emergencia (Adem et al., 2024). Por tanto, garantizar la disponibilidad operativa de estas unidades no constituye únicamente un desafío técnico, sino también una intervención estratégica de alto impacto para la equidad en salud y el acceso universal a la atención de emergencias.

1.4.2. Justificación tecnológica y científica

Actualmente, de acuerdo con la Norma Técnica de Salud N.º 051-MINSA/DGSP-V.01 (Ministerio de Salud del Perú, 2006), el mantenimiento de las ambulancias se lleva a cabo bajo un enfoque preventivo y programado. Sin embargo, este tipo de mantenimiento consiste en actividades periódicas que no consideran necesariamente las condiciones reales de operación de los recursos vehiculares y biomédicos, lo que limita la capacidad para detectar fallas o degradaciones funcionales de forma anticipada.

Monks et al. (2023) demostraron que los modelos predictivos basados en datos históricos pueden mejorar la capacidad de planificación operativa y reducir los errores de estimación de demanda en servicios de ambulancias. Por otro lado, Ersöz et al. (2022) recopilaron evidencia de que las estrategias de mantenimiento predictivo en el sector vehicular, apoyadas en aprendizaje automático, pueden aumentar la confiabilidad y reducir las interrupciones no planificadas, lo que respalda que el enfoque predictivo basado en datos puede contribuir a la detección temprana de fallas y a una mejor planificación

de recursos y tiempos de intervención.

Desde la perspectiva de la ingeniería, la presente investigación representa un aporte científico en el campo de la Ingeniería Biomédica al integrar el mantenimiento predictivo basado en datos en el contexto específico de las ambulancias médicas urbanas, en el que convergen componentes vehiculares y equipamiento biomédico. El desarrollo de un modelo computacional orientado a este tipo de unidades constituye una contribución metodológica concreta para apoyar la gestión de la disponibilidad y la confiabilidad operativa del sistema prehospitario de Lima Metropolitana.

1.4.3. Justificación normativa e institucional

El presente estudio se enmarca en la Norma Técnica de Salud N.º 051-MINSA/DGSP-V.01 (Ministerio de Salud del Perú, 2006), que establece los lineamientos para el transporte asistido de pacientes por vía terrestre, clasifica las ambulancias y define los requisitos técnicos y de equipamiento que deben cumplir, incluyendo la realización de mantenimientos preventivos de los sistemas vehiculares y biomédicos. Sin embargo, esta normativa no contempla criterios de mantenimiento predictivo ni mecanismos de seguimiento digital del estado operativo de las unidades, lo que abre un espacio concreto para la innovación tecnológica en la gestión del mantenimiento asistencial. Esta brecha normativa permite justificar la incorporación de herramientas computacionales que complementen el mantenimiento preventivo establecido, sin contradecir sus lineamientos operativos.

En este contexto, la Defensoría del Pueblo (2023a) ha señalado que un porcentaje significativo de las ambulancias del sistema se encuentra inoperativo o con fallas técnicas, afectando directamente la capacidad de respuesta ante emergencias. La posible implementación de modelos predictivos basados en datos representaría una vía técnicamente viable para fortalecer la gestión de estas unidades y orientar la mejora de su

disponibilidad y continuidad operativa.

A nivel internacional, esta investigación se alinea con los Objetivos de Desarrollo Sostenible de las Naciones Unidas (2015), específicamente con el ODS 3 (salud y bienestar) y el ODS 9 (innovación e infraestructura), en tanto propone una solución tecnológica orientada a mejorar la capacidad de respuesta del sistema prehospitalario mediante herramientas computacionales de bajo costo y alta reproducibilidad.

En el ámbito institucional, el desarrollo de este modelo computacional contribuye a fortalecer el perfil del ingeniero biomédico egresado de la UNMSM, al integrar la gestión técnica sanitaria con el análisis de datos y el modelado predictivo aplicado a sistemas críticos de salud.

I.5. Limitaciones de la Investigación

La presente investigación trabaja exclusivamente con datos operativos y técnicos de mantenimiento, sin incluir información clínica, datos personales de pacientes ni información sensible de carácter sanitario. Esta delimitación permite orientar el análisis hacia la disponibilidad operativa de las ambulancias y no hacia variables asistenciales individuales, preservando el enfoque técnico de la investigación.

La obtención de registros institucionales se gestiona en el marco de la Ley N.º 27806 de Transparencia y Acceso a la Información Pública (Congreso de la República del Perú, 2002), garantizando el uso exclusivamente académico de la información y su manejo con criterios de confidencialidad institucional.

Dado que los registros históricos institucionales del SAMU no estuvieron disponibles durante el período de ejecución de la investigación, se empleó una base de datos simulada construida a partir de los parámetros operativos documentados en la NTS N.º 051-MINSA/OGDN-V.01 y en la literatura técnica de referencia, condición declarada explícitamente en la metodología y en los resultados del presente documento.

La limitación más relevante del estudio deriva de la indisponibilidad de registros históricos institucionales del SAMU durante el período de ejecución de la investigación, condición que determinó el uso de un *dataset* simulado como estrategia metodológica de sustitución. Esta estrategia, documentada en la literatura de mantenimiento predictivo para contextos con restricciones de acceso a datos reales (Taoufyq et al., 2025; Prytz, 2014), permite demostrar la viabilidad del marco computacional propuesto, aunque el *dataset* resultante no captura la totalidad de la complejidad operativa auténtica del SAMU en Lima Metropolitana, incluyendo la variabilidad por zona geográfica, las condiciones de tráfico urbano, la antigüedad heterogénea de la flota y los patrones de demanda estacional.

El alcance demostrado por la investigación corresponde, por tanto, a la validez interna del enfoque metodológico, siendo la replicación con datos reales de la institución el paso siguiente para establecer el desempeño del modelo en condiciones operativas auténticas.

Una segunda limitación es la posible heterogeneidad en la forma de registro de eventos de mantenimiento entre distintas unidades o períodos, que podría afectar la calidad y completitud de los datos en una implementación real. Esta limitación se aborda en los lineamientos técnicos mediante recomendaciones específicas de estandarización de registros.

Los resultados del modelo son específicos para el contexto de Lima Metropolitana durante el período 2024-2025 y no son directamente generalizables a otras regiones o períodos sin un proceso de reentrenamiento y validación equivalente.

Finalmente, la validación retrospectiva empleada, metodológicamente documentada y respaldada por la ISO 17359:2018 y la literatura especializada (Si et al., 2011; Taoufyq et al., 2025), corresponde a una evaluación de la validez interna del modelo sobre datos históricos; su extensión a condiciones operativas reales, mediante una implementación piloto con datos del SAMU, constituye el paso metodológico siguiente del programa de investigación.

CAPÍTULO II : REVISIÓN DE LITERATURA

II.1. Marco Teórico

II.1.1. Atención prehospitalaria y sistema de ambulancias

Desde un enfoque de salud pública, la Organización Mundial de la Salud (OMS) establece que la atención prehospitalaria constituye un componente central dentro de los sistemas de salud, dado que representa el primer contacto asistencial organizado frente a situaciones de emergencia médica (*World Health Organization* [WHO], 2005). Su objetivo primordial es ofrecer atención oportuna, segura y continua al paciente desde el lugar del evento hasta su ingreso a un establecimiento de salud, disminuyendo la morbilidad relacionada con los eventos críticos y mejorando los resultados clínicos posteriores. Este primer contacto asistencial exige recursos móviles disponibles, funcionales y técnicamente confiables.

Los sistemas de atención prehospitalaria no solo dependen de la formación del personal, sino también de la disponibilidad y la calidad de los recursos logísticos que los componen, entre los cuales las ambulancias médicas desempeñan un papel fundamental (WHO, 2005). En este sentido, la atención prehospitalaria no puede entenderse como una actividad aislada, sino como un eslabón dentro de una cadena de atención en la que cada componente resulta determinante para garantizar la continuidad asistencial. Por ello, la disponibilidad de las ambulancias constituye una condición operativa esencial del sistema.

Las ambulancias médicas constituyen un elemento crítico de esta cadena, al

permitir la estabilización inicial del paciente, la monitorización y el soporte durante el traslado y la realización de intervenciones básicas que pueden ser decisivas para su supervivencia. Su función trasciende el simple transporte, dado que estas unidades integran sistemas vehiculares, eléctricos y biomédicos destinados a mantener condiciones adecuadas de soporte vital durante la atención prehospitalaria (Ministerio de Salud del Perú [MINSA], 2006). Esta integración técnica incrementa la relevancia del mantenimiento como soporte de la continuidad asistencial.

Diversos estudios han evidenciado que la disponibilidad oportuna de ambulancias y la rapidez en la respuesta ante una emergencia se asocian directamente con mejores resultados clínicos (Blackwell & Kaufman, 2002; Pons & Markovchick, 2002). Blackwell y Kaufman (2002) demostraron en un sistema urbano de servicios médicos de emergencia que los tiempos de respuesta influyen de manera significativa en la supervivencia de los pacientes, especialmente en situaciones críticas. Por su parte, Pons y Markovchick (2002) cuestionaron la generalización del criterio de tiempo de respuesta como indicador universal de calidad asistencial, señalando que su efectividad depende del tipo de evento y las condiciones clínicas del paciente.

En el estudio de Blackwell y Kaufman (2002) se registró que el riesgo de mortalidad fue de 1,58% para los pacientes con tiempos de respuesta superiores a 5 minutos, frente a 0,51% para aquellos atendidos en menos de ese umbral, evidenciando una asociación estadísticamente significativa entre el tiempo de respuesta y la supervivencia. A partir de estos hallazgos, se infiere que cualquier condición que prolongue el tiempo de respuesta, incluyendo la inoperatividad de las unidades, puede incidir negativamente en los resultados clínicos del paciente. En este sentido, la inoperatividad de las ambulancias médicas representa un factor que reduce directamente la capacidad de respuesta ante una emergencia.

Desde el marco normativo nacional, la Norma Técnica de Salud N.º 051-MINSA/OGDN-V.01 (Ministerio de Salud del Perú, 2006) establece los lineamientos para la operación y el mantenimiento de las ambulancias

médicas en el Perú. La norma define los requisitos técnicos mínimos que deben cumplir las unidades destinadas a la atención prehospitalaria, incluyendo la realización de mantenimientos preventivos periódicos de los sistemas vehiculares y del equipamiento biomédico.

Sin embargo, estos lineamientos se basan en criterios de intervalos fijos de tiempo o kilometraje, sin contemplar mecanismos de seguimiento del estado real de los componentes ni herramientas de análisis predictivo que permitan anticipar fallas, lo que representa una limitación técnica relevante para la gestión operativa de las unidades. Esta limitación abre la posibilidad de complementar el esquema preventivo con herramientas predictivas basadas en datos.

Desde una perspectiva sistémica, los informes sobre gestión de crisis en salud señalan que la capacidad de respuesta de los sistemas sanitarios depende críticamente de la disponibilidad operativa de sus recursos, incluyendo las unidades de atención prehospitalaria (*National Academies of Sciences, Engineering, and Medicine*, 2020). Este enfoque subraya que la inoperatividad de ambulancias no constituye únicamente un problema técnico aislado, sino una variable con impacto directo sobre la capacidad del sistema para cumplir sus estándares de atención, especialmente en contextos de alta demanda o recursos limitados.

En conjunto, el sistema de atención prehospitalaria y sus unidades móviles constituyen componentes críticos de la respuesta sanitaria, cuya efectividad depende directamente de la disponibilidad operativa de las ambulancias. La problemática derivada de la inoperatividad de estas unidades tiene implicancias tanto técnicas como sanitarias, y justifica la necesidad de investigaciones orientadas a fortalecer la gestión del mantenimiento mediante enfoques basados en datos (Defensoría del Pueblo, 2023a).

Este es precisamente el marco en el que se inscribe la presente investigación, que propone un modelo computacional de mantenimiento predictivo con soporte preventivo para las ambulancias médicas urbanas

Tipo II en Lima Metropolitana.

II.1.2. Clasificación normativa de las ambulancias médicas

II.1.2.1. Clasificación general de las ambulancias.

La clasificación de las ambulancias constituye un elemento fundamental para la organización eficiente del sistema de atención de emergencias, dado que permite distribuir los recursos disponibles en función de la complejidad del evento atendido y de las necesidades clínicas del paciente durante el transporte asistido. Desde una perspectiva técnica, las definiciones y guías internacionales coinciden en que dicha clasificación debe sustentarse en criterios funcionales, asistenciales y operativos, y no en características exclusivamente vehiculares (*World Health Organization*, 2005).

En este sentido, la norma europea UNE-EN 1789:2020 establece una clasificación funcional de referencia internacional, en la cual las ambulancias se agrupan según el nivel de atención que pueden proporcionar al paciente (Comité Europeo de Normalización, 2020). Esta norma distingue cuatro clases funcionales: A1, A2, B y C, que se agrupan según el nivel de asistencia que pueden proporcionar al paciente durante el traslado, tal como se detalla en la Tabla 1.

Tabla 1 Clasificación de ambulancias según UNE-EN 1789:2020

Clase	Definición	Nivel de asistencia
A1	Transporte sanitario de un paciente en camilla, no asistencial	No urgente
A2	Transporte sanitario de varios pacientes, con camilla o silla, no asistencial	No urgente
B	Ambulancia asistencial para Soporte Vital Básico (SVB)	Emergencias básicas
C	Ambulancia asistencial para Soporte Vital Avanzado (SVA) o Unidad de Vigilancia Intensiva (UVI) móvil	Emergencias críticas

Fuente. Elaboración propia a partir del Comité Europeo de Normalización (2020).

En la misma línea, la Norma Oficial Mexicana NOM-034-SSA3-2013 adopta un enfoque funcional similar, clasificando las ambulancias según el nivel de atención prehospitalaria que ofrecen: ambulancias de traslado, de urgencias básicas, de urgencias avanzadas y de cuidados intensivos (Secretaría de Salud de México, 2013). La convergencia entre estas clasificaciones internacionales y la normativa peruana refleja un consenso técnico en torno al principio de que la categoría de una ambulancia debe determinarse por su capacidad asistencial y no por sus características vehiculares, criterio que fundamenta la delimitación de la ambulancia Tipo II como unidad de análisis de la presente investigación. Esta coincidencia normativa fortalece la consistencia técnica del objeto de estudio adoptado en la investigación.

Normas técnicas de otros países de la región, como la Norma Técnica Colombiana NTC 3729 (Instituto Colombiano de Normas Técnicas y Certificación [ICONTEC], 2007) y la Norma Técnica Ecuatoriana INEN 2512 (Instituto Ecuatoriano de Normalización [INEN], 2009), refuerzan este enfoque, evidenciando un consenso técnico regional en torno a la clasificación de ambulancias según el nivel de soporte vital y la capacidad asistencial durante el transporte. La Tabla 2 sintetiza los criterios funcionales compartidos por estas normativas internacionales y regionales.

Tabla 2 Criterios funcionales de clasificación de ambulancias según normas técnicas internacionales y regionales

Criterio técnico	Transporte simple	Soporte vital básico	Soporte vital avanzado
Finalidad principal	Traslado sin atención médica	Atención prehospitalaria básica	Atención prehospitalaria avanzada
Nivel de soporte vital	No aplica	SVB	SVA
Capacidad de intervención	Nula	Intervenciones básicas	Intervenciones avanzadas/invasivas
Complejidad del equipamiento	Solo transporte	Equipamiento biomédico básico	Equipamiento biomédico avanzado
Normas de referencia	UNE-EN 1789: Tipo A; NOM-034: traslado	UNE-EN 1789: Tipo B; NOM-034: urgencias básicas	UNE-EN 1789: Tipo C; NOM-034: urgencias avanzadas e intensivas

Fuente. Elaboración propia a partir de Comité Europeo de Normalización (2020), Secretaría de Salud de México (2013), ICONTEC (2007) e INEN (2009).

II.1.2.2. Clasificación de las ambulancias en el marco normativo nacional

II.1.2.2.1. Clasificación según el ámbito de operación.

En el marco normativo peruano, la NTS N.º 051-MINSA/OGDN-V.01 establece que las ambulancias destinadas al transporte asistido de pacientes por vía terrestre se clasifican en dos tipos según el lugar donde prestan el servicio: ambulancias urbanas y ambulancias rurales (MINSA, 2006).

Esta clasificación no responde únicamente a una denominación geográfica, sino a la necesidad de adecuar las condiciones técnicas mínimas del vehículo al entorno operativo. La norma establece requisitos de configuración mecánica diferenciados: la ambulancia urbana requiere tracción 4×2, mientras que la ambulancia rural exige tracción 4×4 de doble tracción para operar en terrenos de accesibilidad limitada (MINSA, 2006). Esta diferenciación permite que la unidad sea seleccionada según las condiciones reales de accesibilidad y desplazamiento del entorno de operación.

En consecuencia, la distinción entre ambulancia urbana y rural constituye un criterio normativo que determina directamente los requerimientos técnicos del componente vehicular y la capacidad operativa esperada para cada entorno de servicio, tal como se resume en la Tabla 3.

Tabla 3 Clasificación de ambulancias terrestres según el lugar donde prestan el servicio

Tipo de ambulancia	Ámbito de operación	Tracción	Enfoque operativo
Ambulancia urbana	Zonas urbanas y periurbanas	4×2	Atención prehospitalaria en entornos con infraestructura vial
Ambulancia rural	Zonas rurales o de difícil acceso	4×4 (todo terreno)	Atención prehospitalaria en entornos con accesibilidad limitada

Fuente. Elaboración propia a partir de la NTS N.º 051-MINSA/OGDN-V.01 (MINSA, 2006).

II.1.2.2.2. Clasificación según equipamiento y nivel de atención.

De manera complementaria a la clasificación por ámbito de operación, la NTS N.º 051-MINSA/OGDN-V.01 establece una clasificación de las ambulancias terrestres según su equipamiento y el tipo de atención que prestan durante el traslado asistido, definiendo tres niveles de complejidad creciente: Tipo I, Tipo II y Tipo III (MINSA, 2006). La Ambulancia Tipo I está destinada al transporte de pacientes estables con equipamiento básico; la Tipo II incorpora equipamiento biomédico adicional para la monitorización y atención de pacientes críticos durante el traslado; y la Tipo III, de mayor complejidad, está orientada a pacientes críticos inestables que requieren soporte avanzado y equipamiento especializado. El detalle de esta clasificación se presenta en la Tabla 4.

Tabla 4 Clasificación de ambulancias según equipamiento y tipo de atención (NTS N.º 051-MINSA/OGDN-V.01)

Tipo de ambulancia	Condición del paciente	Nivel de atención durante el traslado	Equipamiento característico
Tipo I	Paciente estable	Atención básica de traslado	Equipamiento básico
Tipo II	Paciente crítico	Atención médica durante el traslado	- Equipamiento básico - Monitor desfibrilador portátil - Oxímetro de pulso
Tipo III	Paciente crítico inestable	Atención médica especializada	- Equipamiento Tipo II - Ventilador mecánico portátil - Bomba de infusión - Equipamiento especializado

Fuente. Elaboración propia a partir de la NTS N.º 051-MINSA/OGDN-V.01 (MINSA, 2006).

Esta estructura escalonada permite asignar la unidad más adecuada según la condición clínica del paciente y las necesidades asistenciales durante el transporte. En el marco de la presente investigación, la ambulancia Tipo II constituye la unidad de análisis, dado que representa un nivel de complejidad relevante dentro del sistema de atención prehospitalaria urbana del SAMU en Lima Metropolitana (Ministerio de Salud del Perú, 2006; Defensoría del Pueblo, 2023a), y concentra tanto componentes vehiculares

como equipamiento biomédico, lo que confiere a su gestión de mantenimiento una complejidad técnica que justifica el desarrollo de un modelo predictivo basado en datos.

II.1.2.2.3. Ambulancia médica Tipo II: definición, alcance y configuración técnico-operativa.

En el marco de la Norma Técnica de Salud N.º 051-MINSA/OGDN-V.01, la ambulancia médica Tipo II se define como una unidad destinada al transporte asistido de pacientes en estado crítico con capacidad de asistencia médica durante el traslado, constituyendo una unidad de complejidad intermedia dentro del sistema de atención prehospitalaria nacional (Ministerio de Salud del Perú, 2006).

Desde el punto de vista normativo y funcional, la ambulancia Tipo II tiene como propósito principal garantizar la atención médica continua del paciente crítico durante el transporte asistido, permitiendo su vigilancia clínica, estabilización inicial y traslado seguro hacia un establecimiento de salud (MINSA, 2006). A diferencia de las unidades de transporte simple, esta ambulancia no se limita al desplazamiento del paciente, sino que constituye un entorno asistencial móvil en el que se desarrollan intervenciones médicas acordes a su nivel de complejidad durante el trayecto. Esta capacidad asistencial determina el nivel de complejidad funcional que caracteriza a este tipo de unidad.

El alcance asistencial de la ambulancia médica Tipo II está directamente condicionado por su configuración técnico-operativa, la cual integra de manera coordinada componentes vehiculares, eléctricos y biomédicos. Desde la perspectiva vehicular, se trata de una ambulancia terrestre diseñada para operar predominantemente en entornos urbanos o periurbanos con una estructura que permite el acondicionamiento interno necesario para la atención clínica durante el traslado. El diseño interior debe facilitar la ubicación segura de la camilla, el acceso del personal de salud al paciente y la instalación del equipamiento médico, manteniendo condiciones mínimas de ergonomía y seguridad durante la operación (MINSA, 2006).

El sistema eléctrico constituye un subsistema crítico dentro de la ambulancia Tipo II, ya que de él depende el funcionamiento continuo del equipamiento biomédico durante el transporte asistido. La NTS N.º 051-MINSA/OGDN-V.01 establece que la unidad debe contar con un suministro energético confiable que permita la operación de los dispositivos médicos, independientemente del estado del motor del vehículo (MINSA, 2006). Esta dependencia energética convierte al sistema eléctrico en un componente de alta criticidad operativa, dado que una falla en este subsistema puede comprometer directamente la capacidad asistencial de la unidad durante una emergencia.

En cuanto al equipamiento biomédico, la NTS 051 dispone que la ambulancia Tipo II incorpore, además del equipamiento básico común a todas las ambulancias, dispositivos que habiliten la atención médica activa durante el traslado. Entre estos equipos se incluyen, de manera obligatoria, el monitor desfibrilador portátil y el oxímetro de pulso portátil, los cuales permiten la monitorización continua de parámetros fisiológicos esenciales y la respuesta inmediata ante eventos cardiovasculares durante el transporte asistido (MINSA, 2006). La presencia de estos dispositivos define el carácter asistencial de la ambulancia Tipo II y delimita su capacidad de intervención clínica.

Desde una perspectiva operativa, la ambulancia médica Tipo II constituye un recurso asistencial relevante en los sistemas urbanos de atención prehospitalaria, dado que permite brindar atención médica durante el traslado sin corresponder necesariamente al nivel de cuidados intensivos especializados (MINSA, 2006; Defensoría del Pueblo, 2023a). Su adecuada disponibilidad y funcionamiento resultan determinantes para sostener la capacidad de respuesta del sistema, particularmente en contextos de alta demanda y presión asistencial (WHO, 2011; *National Academies of Sciences, Engineering, and Medicine*, 2020).

La configuración técnico-operativa de la ambulancia Tipo II permite comprenderla como un sistema técnico integrado, cuya capacidad asistencial depende del funcionamiento coordinado de componentes

vehiculares, eléctricos, electrónicos, comunicacionales y biomédico-asistenciales. La ficha de homologación prepublicada por PERÚ COMPRAS de la ambulancia urbana Tipo II establece que la unidad incorpora, además del sistema automotriz y la cabina de atención médica, un sistema eléctrico independiente, equipos electrónicos y de comunicación, sistema de oxigenoterapia, aspirador de secreciones, camilla, monitor desfibrilador, oxímetro de pulso y otros dispositivos destinados a la atención del paciente durante el traslado (Central de Compras Públicas, 2024).

De forma concordante, el estándar CAAS GVS para ambulancias terrestres organiza los requisitos técnicos de estas unidades en bloques funcionales vinculados con el diseño vehicular, chasis, componentes del vehículo, sistema eléctrico, iluminación, cabina, compartimiento del paciente, almacenamiento, sistema de oxígeno, succión, parámetros ambientales y comunicaciones. Esta estructura respalda el análisis de la ambulancia como un activo técnico compuesto por subsistemas interdependientes, cuya confiabilidad funcional depende de la operación conjunta de sus componentes críticos (*Commission on Accreditation of Ambulance Services* [CAAS], 2016).

Esta lectura por bloques funcionales resulta especialmente importante para el mantenimiento, debido a que la falla de un subsistema puede afectar la disponibilidad completa de la unidad, aun cuando los demás componentes se encuentren operativos. Por ejemplo, una deficiencia en el sistema eléctrico puede impedir el uso seguro del equipamiento biomédico, mientras que una falla vehicular puede imposibilitar el traslado aunque los dispositivos médicos se encuentren funcionales. Por ello, el análisis técnico de la ambulancia Tipo II requiere una visión integrada de sus componentes y no una evaluación aislada de cada elemento.

A partir de esta visión integrada, se organiza los componentes técnicos de la ambulancia Tipo II en tres macro-subsistemas analíticos: vehicular, eléctrico/electrónico y biomédico-asistencial. El macro-subsistema vehicular agrupa los elementos asociados al chasis, motor, transmisión, frenos, suspensión, neumáticos, carrocería y condiciones mecánicas de operación.

El macro-subsistema eléctrico/electrónico comprende el sistema eléctrico automotriz y asistencial, alimentación energética, iluminación, señalización, balizas, sirena, sistema de carga, GPS, intercomunicador y radiocomunicación. El macro-subsistema biomédico-asistencial integra el equipamiento médico y los elementos directamente vinculados con la atención del paciente, incluyendo oxigenoterapia, aspiración, mecánica de gases, monitor desfibrilador, oxímetro, camilla y accesorios clínicos. Esta agrupación permite representar funcionalmente los componentes relevantes para el análisis de mantenimiento, inoperatividad y disponibilidad operativa, manteniendo coherencia con la ficha de homologación nacional prepublicada por PERÚ COMPRAS y con el enfoque sistémico del estándar CAAS GVS (Central de Compras Públicas, 2024; CAAS, 2016).

Esta condición convierte a la ambulancia médica Tipo II en un elemento crítico dentro del sistema de atención prehospitalaria urbana. Desde la perspectiva de la ingeniería biomédica, su análisis como sistema técnico complejo, susceptible de gestión de mantenimiento y evaluación de disponibilidad operativa, constituye el fundamento técnico sobre el que se construye la presente investigación, la cual se detalla en la Tabla 5.

Tabla 5 Configuración técnico-operativa y alcance funcional de la ambulancia médica Tipo II

Dimensión	Descripción
Tipo de unidad	Ambulancia terrestre asistencial
Tipo de paciente	Paciente crítico
Nivel de atención	Atención médica durante el traslado
Entorno de operación	Urbano y periurbano
Macro-subsistema vehicular	Chasis, motor, transmisión, frenos, suspensión, neumáticos, carrocería, cabina de conducción y acondicionamiento físico de la unidad
Macro-subsistema eléctrico/electrónico	Sistema eléctrico automotriz y asistencial, batería, alternador, inversor, tomacorrientes, iluminación, señalización, balizas, sirena, sistema de carga, GPS, intercomunicador y radiocomunicación

Macro-subsistema biomédico-asistencial	Equipamiento médico y elementos de soporte clínico: oxigenoterapia, aspiración, mecánica de gases, monitor desfibrilador, oxímetro, camilla y accesorios asistenciales
Capacidad asistencial	Intervenciones médicas básicas e intermedias durante el traslado
Limitaciones	No orientada a cuidados intensivos especializados
Dependencia operativa	Alta dependencia del estado funcional integrado de los macro-subsistemas técnicos

Fuente. Elaboración propia a partir de la NTS N.º 051-MINSA/OGDN-V.01 (Ministerio de Salud del Perú, 2006), Central de Compras Públicas (2024) y Commission on Accreditation of Ambulance Services (2016).

II.1.3. Gestión del mantenimiento en ambulancias médicas

La gestión del mantenimiento constituye un componente esencial en la administración de activos críticos dentro de los sistemas de salud. En el contexto de la atención prehospitalaria, su relevancia es especialmente significativa, dado que la disponibilidad operativa de las ambulancias médicas condiciona directamente la capacidad de respuesta ante emergencias. De acuerdo con la OMS (*World Health Organization* [WHO], 2011), la gestión del mantenimiento en el ámbito sanitario comprende el conjunto de actividades técnicas, administrativas y operativas orientadas a asegurar que los equipos médicos se mantengan en condiciones funcionales adecuadas durante todo su ciclo de vida, garantizando seguridad, confiabilidad y continuidad del servicio.

En el caso específico de las ambulancias médicas, el mantenimiento no se limita a la preservación de la integridad mecánica del vehículo, sino que comprende de forma integrada los sistemas vehiculares, eléctricos, biomédicos y de soporte, todos los cuales operan simultáneamente durante la atención y el traslado del paciente. La NTS N.º 051-MINSA/OGDN-V.01 establece que las ambulancias deben mantenerse en condiciones operativas adecuadas, considerando tanto el estado del vehículo como la funcionalidad del equipamiento médico incorporado, a fin de garantizar un traslado seguro y oportuno (Ministerio de Salud del Perú, 2006). Esta interdependencia entre subsistemas técnicos confiere a la gestión del mantenimiento de las

ambulancias una complejidad mayor que la de los vehículos convencionales, y fundamenta la necesidad de enfoques técnicos que aborden de manera integral su disponibilidad operativa.

Tradicionalmente, la gestión del mantenimiento en las ambulancias médicas se ha sustentado en dos enfoques principales: el preventivo y el correctivo. El mantenimiento preventivo comprende el conjunto de intervenciones planificadas que se ejecutan a intervalos predeterminados de tiempo o uso, con el objetivo de reducir la probabilidad de falla y prolongar la vida útil del activo (*World Health Organization*, 2011). Si bien este enfoque permite reducir la ocurrencia de fallas inesperadas, su efectividad está condicionada a la correcta definición de los intervalos de intervención y no considera el estado real de degradación de los componentes, lo que puede derivar en intervenciones innecesarias o en fallas no anticipadas (Moubray, 1997). En la Tabla 6 se resumen las características principales de ambos enfoques.

Tabla 6 Comparación entre enfoques de mantenimiento preventivo y correctivo en ambulancias médicas

Enfoque	Criterio de activación	Característica principal	Aplicación en ambulancias	Limitación principal
Preventivo	Intervalos de tiempo o uso	Intervenciones planificadas	Revisiones periódicas por tiempo o kilometraje	No considera el estado real de los componentes
Correctivo	Ocurrencia de falla	Intervención post-falla	Reparaciones no planificadas ante averías	Genera inoperatividad no programada

Fuente. Elaboración propia a partir de World Health Organization (2011) y Moubray (1997).

El mantenimiento correctivo, por su parte, consiste en las intervenciones ejecutadas tras la ocurrencia de una falla, con el fin de restablecer la funcionalidad del sistema. Si bien en ciertos contextos su ocurrencia es inherente a la operación, este enfoque se asocia a mayores tiempos de indisponibilidad, incremento de costos y riesgos operativos (Moubray, 1997). En el caso de las ambulancias médicas, donde una falla no planificada

puede comprometer directamente la atención de un paciente en situación crítica, la dependencia predominante del mantenimiento correctivo representa un riesgo asistencial significativo que refuerza la necesidad de enfoques más anticipativos.

La planificación y el control del mantenimiento se sustentan en indicadores cuantitativos que permiten evaluar el desempeño operativo de los activos y la eficacia de las estrategias implementadas. Estos indicadores incluyen el tiempo promedio entre fallas (MTBF, por sus siglas en inglés), el tiempo promedio de reparación (MTTR, por sus siglas en inglés) y la disponibilidad operativa, los cuales proporcionan información objetiva sobre el comportamiento histórico de los activos y constituyen herramientas fundamentales para la toma de decisiones técnicas en la gestión centrada en la confiabilidad (Moubray, 1997). En el contexto de las ambulancias médicas, el seguimiento sistemático de estos indicadores a partir de registros históricos representa la base para construir modelos predictivos orientados a apoyar la gestión del mantenimiento, tal como lo establece el marco de diagnóstico y pronóstico de la ISO 17359:2018. En la Tabla 7 se presentan los principales indicadores técnicos utilizados en la gestión del mantenimiento de ambulancias médicas junto con su definición e interpretación en el contexto operativo de las unidades.

Tabla 7 Indicadores técnicos utilizados en la gestión del mantenimiento de ambulancias médicas

Indicador	Definición técnica	Interpretación operativa
MTBF	Tiempo promedio de operación entre fallas sucesivas	Valores altos indican mayor confiabilidad operativa
MTTR	Tiempo promedio requerido para reparar y devolver el activo a operación	Valores bajos indican mayor eficiencia en mantenimiento correctivo
Disponibilidad operativa (A)	Proporción del tiempo en que el activo está disponible para su uso	Representa la capacidad real de respuesta del sistema
Downtime	Tiempo durante el cual la ambulancia no se encuentra operativa	Impacta directamente en la cobertura del servicio prehospitalario

Fuente. Elaboración propia a partir de Moubray (1997), World Health Organization (2011) e ISO 17359:2018.

Las expresiones matemáticas correspondientes a cada indicador se definen a continuación:

$$MTBF = \frac{\text{Tiempo operativo acumulado}}{\text{Número de fallas}} \quad (1)$$

$$MTTR = \frac{\text{Tiempo total de reparación}}{\text{Número de reparaciones}} \quad (2)$$

$$A = \frac{MTBF}{MTBF + MTTR} \quad (3)$$

$$\text{Downtime} = \text{Tiempo total} - \text{Tiempo operativo} \quad (4)$$

La Figura 1 ilustra la relación entre estos indicadores sobre la línea de tiempo operativa de una ambulancia Tipo II, mostrando cómo el MTBF captura el tiempo de operación efectiva entre fallas sucesivas, el MTTR refleja la duración de cada intervención correctiva, el *downtime* acumula los períodos de inoperatividad y la disponibilidad operativa sintetiza la proporción del tiempo en que la unidad está disponible para el servicio.

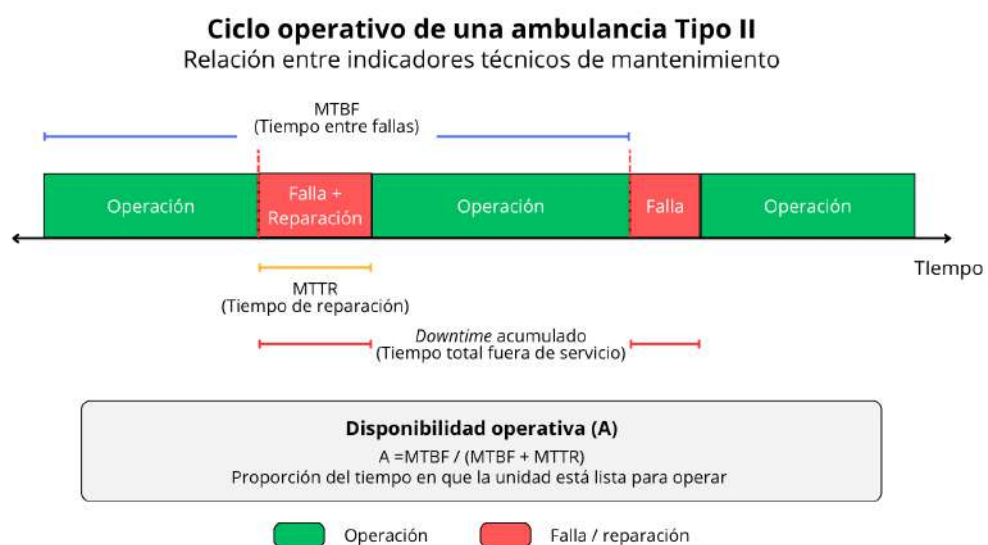


Figura 1. Representación de los indicadores técnicos de mantenimiento sobre la línea de tiempo operativa de una ambulancia Tipo II.

Fuente. Elaboración propia a partir de Moubray (1997) e ISO 17359:2018.

Si bien estos indicadores permiten describir y evaluar el desempeño histórico del sistema, su uso dentro de esquemas tradicionales de mantenimiento presenta limitaciones técnicas. Moubray (1997) señaló que los enfoques basados exclusivamente en intervalos fijos de tiempo o uso no consideran la variabilidad real de las condiciones operativas ni el comportamiento dinámico de las fallas, lo que puede conducir tanto a intervenciones innecesarias como a fallas no previstas. Esta limitación fundamenta la necesidad de transitar hacia enfoques que incorporen información histórica y operativa del activo como criterio de decisión, lo que constituye la base conceptual del mantenimiento predictivo desarrollado en los apartados siguientes. Por ello, la decisión de mantenimiento requiere complementar el criterio temporal con información operativa del activo.

En el contexto de las ambulancias médicas urbanas, estas limitaciones se acentúan por la heterogeneidad de los patrones de uso, las condiciones de operación y la criticidad del servicio prestado. La dependencia exclusiva de esquemas preventivos basados en tiempo o kilometraje resulta insuficiente para garantizar niveles adecuados de disponibilidad operativa, generando brechas entre la planificación del mantenimiento y el comportamiento real del sistema (Moubray, 1997; MINSA, 2006). Esta situación pone de manifiesto la necesidad de enfoques complementarios que permitan anticipar el comportamiento del activo a partir del análisis sistemático de información histórica, lo que establece las bases conceptuales para la incorporación del mantenimiento predictivo como eje central de la presente investigación.

II.1.4. Mantenimiento predictivo: evolución y fundamentos conceptuales

II.1.4.1. Evolución de los enfoques de mantenimiento

La evolución de los enfoques de mantenimiento ha estado estrechamente vinculada al incremento de la complejidad técnica de los activos y a la necesidad de mejorar su confiabilidad y disponibilidad operativa

(Molęda et al., 2023). En sus etapas iniciales, el mantenimiento se orientó principalmente hacia un enfoque correctivo, en el cual las intervenciones técnicas se realizaban únicamente tras la ocurrencia de una falla (Moubray, 1997). Si bien este enfoque resultaba aceptable en sistemas simples, evidenció limitaciones significativas en contextos donde la indisponibilidad de los activos generaba impactos operativos, económicos o de seguridad.

Como respuesta a estas limitaciones, surgió el mantenimiento preventivo, basado en la ejecución programada de intervenciones en función del tiempo o del uso acumulado del activo. Su adopción representó un cambio importante en la gestión del mantenimiento al incorporar una planificación sistemática de las actividades de inspección y sustitución de componentes, disminuyendo la dependencia de intervenciones exclusivamente reactivas. Este enfoque permitió reducir la frecuencia de fallas inesperadas, aunque introdujo nuevos desafíos asociados a la rigidez de los intervalos de mantenimiento y a la realización de intervenciones que no necesariamente respondían al estado real de degradación del activo (Moubray, 1997).

La literatura especializada identifica al mantenimiento predictivo como una etapa evolutiva posterior, orientada a superar las restricciones del mantenimiento preventivo tradicional. Molęda et al. (2023), en su revisión sistemática sobre la evolución del mantenimiento, describen este tránsito como un cambio desde enfoques reactivos y basados en intervalos de tiempo hacia estrategias que buscan anticipar el comportamiento futuro de los activos a partir del análisis de información histórica y operativa. En el contexto de las ambulancias médicas urbanas Tipo II, este tránsito resulta especialmente pertinente dado que las condiciones heterogéneas de operación y la criticidad del servicio hacen insuficientes los esquemas de mantenimiento programados por tiempo o kilometraje.

II.1.4.2. Conceptualización del mantenimiento predictivo

Desde un punto de vista conceptual, el mantenimiento predictivo constituye un enfoque de gestión orientado a anticipar la ocurrencia de fallas o

degradaciones funcionales, con el propósito de ejecutar intervenciones técnicas antes de que estas deriven en inoperatividad (Molęda et al., 2023; Taoufyq et al., 2025). A diferencia del mantenimiento preventivo, que se apoya en intervalos predefinidos de tiempo o uso, el mantenimiento predictivo se fundamenta en la interpretación de información relevante sobre el desempeño y el historial del activo.

Molęda et al. (2023) señalaron que el mantenimiento predictivo se caracteriza por el uso sistemático de datos para estimar el estado futuro del activo y apoyar la toma de decisiones de mantenimiento. Este enfoque no se limita necesariamente al uso de sensores o monitoreo en tiempo real, sino que puede sustentarse en datos históricos de fallas, registros de mantenimiento, condiciones de operación y patrones de uso, lo cual resulta especialmente pertinente en contextos donde la instrumentación avanzada no es viable o disponible. En estos escenarios, el valor del mantenimiento predictivo reside en la capacidad de transformar la información generada durante la operación cotidiana en evidencia útil para anticipar comportamientos de falla y orientar la planificación de las intervenciones técnicas.

Desde una perspectiva normativa, la ISO 17359:2018 establece un marco general para el monitoreo de condición y el diagnóstico de máquinas, introduciendo el concepto de pronóstico como el proceso mediante el cual se estima la evolución futura del estado de un activo a partir de información disponible. La norma no prescribe métodos específicos, pero establece que el análisis de tendencias y datos históricos constituye una base válida para la toma de decisiones de mantenimiento (*International Organization for Standardization* [ISO], 2018).

II.1.4.3. *Distinción conceptual entre mantenimiento preventivo y predictivo*

La distinción entre mantenimiento preventivo y mantenimiento predictivo es fundamental para comprender el aporte conceptual de este último dentro de la gestión técnica de activos críticos. Aunque ambos enfoques buscan

reducir la probabilidad de fallas y mejorar la disponibilidad operativa, difieren en la lógica que sustenta la decisión de intervención. Mientras el mantenimiento preventivo se apoya principalmente en criterios programados de tiempo o uso, el mantenimiento predictivo incorpora información histórica, operativa o de condición para estimar el comportamiento futuro del activo. Esta diferencia resulta relevante en sistemas con variabilidad operativa, como las ambulancias médicas urbanas, donde dos unidades con el mismo intervalo de mantenimiento pueden presentar niveles distintos de desgaste o riesgo funcional. En la Tabla 8 se sintetizan las principales diferencias conceptuales entre ambos enfoques.

Tabla 8 Comparación conceptual entre mantenimiento preventivo y mantenimiento predictivo

Criterio	Mantenimiento preventivo	Mantenimiento predictivo
Base de decisión	Intervalos de tiempo o uso	Análisis de información y datos
Enfoque	Programado y periódico	Anticipativo y orientado al estado
Consideración de variabilidad operativa	Limitada	Explícita
Riesgo de intervenciones innecesarias	Moderado - alto	Reducido
Dependencia de datos históricos	Baja	Alta
Objetivo principal	Prevenir fallas	Anticipar fallas

Fuente. Elaboración propia a partir de Molęda et al. (2023) e ISO 17359:2018.

Esta comparación pone de manifiesto que el mantenimiento predictivo no se define por una tecnología específica, sino por un cambio en el criterio de decisión, en el que la información disponible sobre el activo constituye el fundamento principal de la intervención técnica. En el caso de las ambulancias médicas, este cambio de criterio resulta especialmente viable dado que los registros históricos de operación y mantenimiento representan una fuente de información sistemática sobre el comportamiento del activo, sin requerir instrumentación adicional (Prytz, 2014; Taoufyq et al., 2025).

II.1.4.4. Rol del análisis de información en el mantenimiento predictivo

Un elemento central del mantenimiento predictivo es el análisis de información como base para la anticipación de fallas. La ISO 17359:2018 establece que los procesos de diagnóstico y pronóstico pueden sustentarse en el análisis de datos provenientes del historial de operación y mantenimiento, lo que permite identificar tendencias de degradación o patrones recurrentes que preceden a la falla funcional (ISO, 2018). Desde esta perspectiva, el mantenimiento predictivo puede implementarse mediante el uso de registros históricos, incluyendo el historial de fallas, la frecuencia y tipo de intervenciones realizadas, los tiempos de inoperatividad registrados y las condiciones de uso del activo (Taoufyq et al., 2025; Si et al., 2011).

Molęda et al. (2023) enfatizaron que el valor del mantenimiento predictivo reside en la capacidad de transformar datos históricos en información útil para la toma de decisiones, sin que ello implique necesariamente la incorporación de sistemas de monitoreo en tiempo real. Este enfoque resulta particularmente pertinente para el sistema de ambulancias médicas urbanas Tipo II en Lima Metropolitana, donde los registros de operación y mantenimiento pueden constituir una fuente relevante para anticipar eventos de inoperatividad y apoyar la gestión de la disponibilidad operativa de las unidades. De este modo, el análisis de información opera como vínculo entre el historial técnico del activo y la decisión de mantenimiento.

II.1.4.5. El mantenimiento predictivo en sistemas técnicos críticos

El mantenimiento predictivo resulta especialmente pertinente en sistemas técnicos considerados críticos, donde la indisponibilidad del activo puede generar consecuencias operativas y sanitarias significativas (Molęda et al., 2023). En este tipo de sistemas, la capacidad de anticipar fallas puede contribuir a mejorar la disponibilidad operativa, apoyar la asignación de recursos y reducir el impacto de intervenciones no planificadas (Taoufyq et al., 2025).

Molęda et al. (2023) destacaron que los enfoques predictivos resultan particularmente adecuados para activos con alta variabilidad operativa y elevada criticidad, características presentes en múltiples sectores, incluido el transporte y los servicios de emergencia. En estos contextos, el mantenimiento predictivo constituye un enfoque técnicamente fundamentado para apoyar decisiones de mantenimiento más informadas y alineadas con la realidad operativa del sistema. Esta adecuación se refuerza cuando la falla del activo afecta directamente la continuidad del servicio prestado.

En síntesis, el mantenimiento predictivo representa una evolución conceptual de los enfoques tradicionales, orientada a anticipar fallas mediante el análisis sistemático de datos históricos sobre el comportamiento de los activos, sin requerir necesariamente instrumentación adicional ni monitoreo en tiempo real. En el contexto de las ambulancias médicas urbanas Tipo II en Lima Metropolitana, esta evolución adquiere relevancia directa: la NTS N.º 051-MINSA/OGDN-V.01 establece un esquema de mantenimiento preventivo programado por intervalos fijos que no considera el estado real de los componentes (MINSA, 2006), lo que constituye la brecha técnica que la presente investigación busca abordar.

II.1.5. Mantenimiento predictivo basado en datos

El mantenimiento predictivo basado en datos (enfoque *data-driven*) se sustenta en la premisa de que los activos generan trazas observables durante su operación y mantenimiento: fallas registradas, intervenciones realizadas, tiempos fuera de servicio, condiciones de uso y variables de contexto. A diferencia de enfoques estrictamente basados en modelos físicos, los enfoques *data-driven* se orientan a inferir patrones y regularidades a partir de datos disponibles, sin requerir una representación completa del mecanismo físico de degradación (Taoufyq et al., 2025). Esta propiedad es particularmente útil en entornos con restricciones técnicas y presupuestales, donde la instrumentación adicional o el monitoreo en tiempo real no son viables.

Desde la perspectiva de la ingeniería de confiabilidad y pronósticos, la ISO 17359:2018 enmarca los procesos de monitoreo, diagnóstico y pronóstico como actividades sustentadas en información del activo y su desempeño. Aunque frecuentemente esta información proviene de mediciones en tiempo real, el principio central de la norma es que la toma de decisiones puede basarse en datos disponibles y evidencia técnica, incluyendo información histórica y tendencias observables (ISO, 2018). En consecuencia, para una flota de ambulancias médicas urbanas Tipo II, el enfoque *data-driven* puede plantearse con base en registros existentes de operación y mantenimiento, sin requerir el despliegue de sensores adicionales, lo que lo convierte en una alternativa técnicamente viable y adaptada a las restricciones del sistema prehospitario público.

II.1.5.1. Uso de datos históricos en mantenimiento predictivo

La literatura revisada describe que las estrategias de mantenimiento predictivo pueden construirse a partir de distintas fuentes de datos: datos de mantenimiento, tales como: órdenes de trabajo, historial de intervención y registros de repuestos; datos de fallas, tales como: modo de falla, criticidad y recurrencia; y datos de operación, tales como: tiempos de servicio, kilometraje y condiciones de uso. Estas fuentes pueden integrarse para modelar el comportamiento del activo y anticipar fallas funcionales (Taoufyq et al., 2025).

En el ámbito vehicular, Prytz (2014) documentó el uso de métodos de aprendizaje automático para mantenimiento predictivo empleando tanto datos embarcados (on-board) como datos históricos agregados provenientes de bases de datos externas (off-board). Este planteamiento es especialmente pertinente porque demuestra que el análisis predictivo puede sustentarse sobre información histórica ya recopilada por los sistemas de gestión y mantenimiento sin requerir instrumentación adicional en campo (Prytz, 2014).

De manera complementaria, Arena et al. (2022) señalaron que el crecimiento de datos y registros asociados a vehículos habilita la detección de patrones y

la estimación del estado del activo para apoyar decisiones de mantenimiento. Si bien parte de esta literatura enfatiza datos de monitoreo en tiempo real, el principio subyacente es que el valor predictivo se obtiene cuando la información disponible se estructura, depura e interpreta para anticipar fallas o degradación. En este sentido, el reto técnico central no es cómo instrumentar el activo, sino cómo transformar los registros existentes en información útil para la toma de decisiones, perspectiva que orienta directamente el enfoque metodológico de la presente investigación.

II.1.5.2. Ventajas operativas del enfoque data-driven

A nivel de gestión técnica, el enfoque *data-driven* ofrece ventajas concretas cuando el objetivo es mejorar la disponibilidad operativa sin inversiones adicionales en instrumentación. En primer lugar, permite aprovechar registros que ya se generan durante la operación y el mantenimiento del activo, reduciendo las barreras de implementación (Prytz, 2014; Taoufyq et al., 2025).

En segundo lugar, facilita la captura de diferencias entre unidades, rutas, turnos y condiciones de operación, lo que resulta especialmente relevante en flotas urbanas con alta variabilidad operativa (Arena et al., 2022). En tercer lugar, habilita la identificación de unidades o subsistemas con mayor probabilidad de falla, orientando los recursos de mantenimiento hacia donde generan mayor impacto (Taoufyq et al., 2025). Finalmente, permite introducir indicadores como la vida útil remanente para apoyar la planificación anticipada de intervenciones (Si et al., 2011). En conjunto, estas ventajas fortalecen la capacidad de planificación y priorización de los mantenimientos.

Estas ventajas no dependen de sensores en sí mismos, sino de tres condiciones fundamentales: la calidad mínima de los registros disponibles, la trazabilidad temporal de los datos y la consistencia en la codificación de fallas e intervenciones (Taoufyq et al., 2025). Estas condiciones pueden alcanzarse mediante una adecuada estructuración de la base de datos y criterios claros de gestión de la información, sin requerir infraestructura tecnológica adicional.

Aplicado al sistema de ambulancias médicas urbanas Tipo II, este enfoque resulta especialmente pertinente dado que la operación y el mantenimiento de estas unidades requieren registros técnicos y administrativos que, si se estructuran adecuadamente, pueden servir como insumo para anticipar eventos de inoperatividad y apoyar la toma de decisiones de mantenimiento (MINSA, 2006; Defensoría del Pueblo, 2023a). La presente investigación propone precisamente transformar esos registros existentes en información útil para la toma de decisiones técnicas sin requerir inversión adicional en infraestructura de monitoreo.

II.1.5.3. Estimación de fallas y vida útil remanente mediante enfoques data-driven

En mantenimiento predictivo, uno de los objetivos centrales es la estimación de la vida útil remanente (RUL, por sus siglas en inglés), entendida como el tiempo de operación que le resta a un activo antes de que ocurra una falla funcional (Si et al., 2011). La relevancia del RUL radica en que transforma el mantenimiento en un problema de decisión: intervenir demasiado pronto puede generar sobrecostos, mientras que intervenir tarde incrementa el riesgo de falla funcional.

Si et al. (2011) señalaron que la estimación del RUL es central en el mantenimiento basado en condición y en la disciplina de pronósticos, y revisaron enfoques estadísticos *data-driven* para su estimación, concluyendo que no existe un método universalmente óptimo y que los modelos dependen del tipo de información disponible y del contexto de aplicación. En una flota de ambulancias con datos históricos disponibles, este planteamiento implica que la estimación predictiva no busca replicar condiciones de laboratorio con mediciones de alta frecuencia, sino construir predicciones plausibles y útiles para la gestión técnica a partir de información real existente. En este sentido, el objetivo no es reemplazar el análisis técnico especializado, sino proporcionar una estimación que sirva como apoyo para priorizar unidades, optimizar la programación de intervenciones y fortalecer la toma de decisiones de mantenimiento basada en evidencia.

Asimismo, Liao y Köttig (2014) describieron la prognóstica como la disciplina orientada a predecir el desempeño futuro o el tiempo hasta falla de sistemas de ingeniería, e identificaron los modelos *data-driven* como una categoría formalmente reconocida para la predicción del RUL, junto con enfoques basados en experiencia y enfoques basados en física. Esta clasificación respalda el encuadre académico del enfoque *data-driven* como una línea consolidada de investigación y aplicación en ingeniería de mantenimiento (Liao & Köttig, 2014). La Tabla 9 sintetiza las diferencias prácticas entre el enfoque basado en datos históricos, adoptado en la presente investigación, y el enfoque basado en monitoreo continuo.

Tabla 9 Diferenciación práctica entre mantenimiento predictivo basado en datos históricos y basado en monitoreo continuo

Criterio	Mantenimiento Predictivo	
	Basado en datos históricos (enfoque de esta tesis)	Basado en monitoreo continuo
Fuente principal de información	Registros de fallas, órdenes de trabajo, tiempos de inoperatividad, kilometraje, uso.	Series de condición (vibración, temperatura, etc.) y telemetría.
Necesidad de despliegue en campo	No	Sí (típicamente)
Disponibilidad y madurez de datos	Depende de la calidad del registro histórico	Depende de instrumentación y continuidad de medición
Producto típico	Predicción de riesgo de falla, priorización y estimación temporal basada en históricos.	Detección temprana de anomalías y pronóstico con datos de condición.
Riesgo principal	Sesgos por subregistro, inconsistencias, cambios operativos.	Fallas del sensor, costos de infraestructura, ruido y mantenimiento del sistema de monitoreo.

Fuente. Elaboración propia a partir de Taoufyq et al. (2025), ISO 17359:2018, Si et al. (2011) y Liao y Köttig (2014).

II.1.5.4. Viabilidad del enfoque predictivo sin despliegue en campo

Para precisar el alcance metodológico del enfoque, resulta pertinente distinguir entre prescindir del despliegue en campo y prescindir de datos.

En flotas vehiculares, parte de la evidencia disponible muestra que la predicción puede sustentarse en información histórica agregada y en bases de datos externas, sin requerir la instalación de nuevos dispositivos. Prytz (2014) documentó el uso de datos históricos agregados (*historic and aggregated data from off-board databases*) como fuente para métodos predictivos, planteamiento metodológicamente aplicable a una flota de ambulancias. En este contexto, la disponibilidad de ese registro habilita el desarrollo de mantenimiento predictivo orientado a la priorización, el riesgo y la planificación, aun sin medición en tiempo real.

De manera complementaria, Arena et al. (2022) mostraron que la disponibilidad de datos en entornos reales habilita estrategias predictivas y de estimación de degradación, reconociendo la explotación de información disponible para el modelado y la toma de decisiones, y destacando la preparación y estructuración de datos como un paso crítico. En el contexto de la presente investigación, esto refuerza que el desafío técnico central no es instrumentar el activo, sino construir una base de datos estructurada y coherente de variables y eventos que permita aplicar análisis y validación retrospectiva.

Esta distinción es relevante porque permite sostener la viabilidad del enfoque predictivo en contextos donde no existe telemetría continua, sensores adicionales o infraestructura digital avanzada. En tales casos, el valor del modelo no depende de medir el activo en tiempo real, sino de organizar adecuadamente la información histórica disponible para identificar patrones asociados a fallas, mantenimiento e inoperatividad. Por ello, la ausencia de despliegue en campo no elimina la posibilidad de desarrollar un modelo predictivo, siempre que exista una base de datos estructurada, trazable y metodológicamente consistente.

La Figura 2 presenta el flujo conceptual del mantenimiento predictivo basado en datos históricos, estructurado en cinco etapas secuenciales que van desde la recolección e integración de registros de operación, mantenimiento e inoperatividad hasta la retroalimentación y mejora continua del sistema.

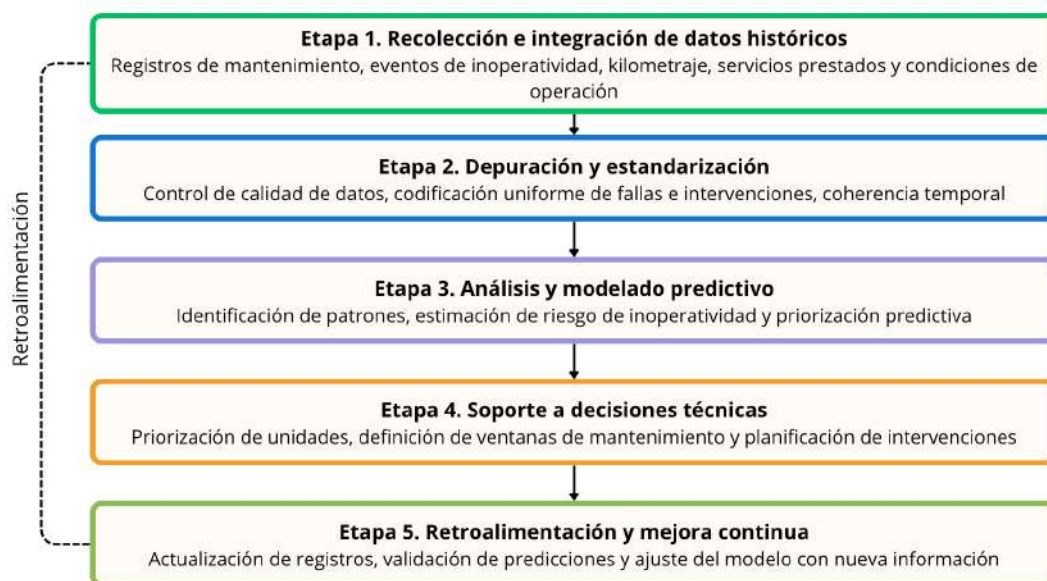


Figura 2. Flujo conceptual del mantenimiento predictivo basado en datos históricos.

Fuente. Elaboración propia a partir de International Organization for Standardization (2018), Taoufyq et al. (2025), Prytz (2014), Si et al. (2011) y Liao y Köttig (2014).

En síntesis, el mantenimiento predictivo basado en datos constituye un enfoque técnicamente coherente cuando se dispone de registros históricos y operativos suficientes, aun sin instrumentación adicional. La literatura revisada respalda que los modelos *data-driven* son una categoría formal de mantenimiento predictivo (Taoufyq et al., 2025) y que conceptos como la vida útil remanente pueden estimarse mediante enfoques estadísticos basados en datos, condicionados por la disponibilidad y calidad de la información (Si et al., 2011). Esta condición refuerza la pertinencia del enfoque para escenarios institucionales con limitaciones de información en tiempo real.

Asimismo, el marco de diagnóstico y pronóstico de la ISO 17359:2018 permite encuadrar el rol del análisis de información en la toma de decisiones de mantenimiento (ISO, 2018). Por tanto, para una flota urbana de ambulancias Tipo II, el aprovechamiento de datos históricos se plantea como una vía metodológicamente fundamentada para estimar riesgo de inoperatividad, fortalecer la gestión técnica y apoyar decisiones orientadas a la disponibilidad operativa, estableciendo las bases para la discusión de modelos computacionales en el apartado siguiente.

II.1.6. Modelos computacionales para mantenimiento predictivo

II.1.6.1. Modelo computacional: definición y rol en ingeniería

En el ámbito de la ingeniería, un modelo computacional se define como una representación abstracta y formal de un sistema real, construida con el propósito de analizar, comprender y anticipar su comportamiento bajo determinadas condiciones (Montáns et al., 2019). A diferencia del sistema físico, el modelo no busca reproducir la totalidad de sus características, sino capturar los aspectos esenciales que permiten explicar o predecir su desempeño. En este sentido, los modelos computacionales constituyen herramientas fundamentales para el estudio de sistemas complejos, donde la experimentación directa puede resultar costosa, riesgosa o técnicamente inviable.

Desde una perspectiva contemporánea, la literatura reconoce que los modelos computacionales cumplen un rol central en la toma de decisiones técnicas, al permitir evaluar escenarios, estimar resultados futuros y apoyar procesos de planificación. Montáns et al. (2019) señalaron que el modelado computacional en ingeniería ha evolucionado desde enfoques puramente deterministas hacia esquemas más flexibles que incorporan datos observacionales y evidencia empírica para describir sistemas reales. Bajo este marco, el modelo computacional actúa como intermediario entre el sistema físico y el proceso de decisión, facilitando el análisis predictivo sin intervenir directamente sobre el activo. En este contexto, este rol se concreta en el desarrollo de un modelo orientado a estimar la probabilidad de inoperatividad de las ambulancias Tipo II a partir de sus registros históricos de operación y mantenimiento.

II.1.6.2. Modelos computacionales predictivos en sistemas de ingeniería

Un modelo computacional es predictivo cuando su finalidad principal es estimar el comportamiento futuro de un sistema a partir de información

disponible en el presente o en el pasado. En ingeniería, estos modelos se emplean para anticipar eventos, evaluar riesgos y apoyar la planificación de acciones preventivas o correctivas. Su finalidad consiste en proporcionar información prospectiva que contribuya a una toma de decisiones más informada. La predicción, como tal, no implica certeza absoluta, sino una estimación fundamentada que permite reducir la incertidumbre asociada a la operación de sistemas complejos, orientando la toma de decisiones técnicas hacia intervenciones más oportunas y eficientes (Montáns et al., 2019).

La literatura distingue entre modelos descriptivos, explicativos y predictivos. Mientras los primeros se orientan a caracterizar el estado actual de un sistema y los segundos a comprender sus mecanismos internos, los modelos predictivos se centran en estimar la evolución futura de variables de interés. Montáns et al. (2019) destacaron que la capacidad predictiva puede lograrse incluso sin un conocimiento completo de las leyes físicas que gobiernan el sistema, siempre que exista información suficiente para inferir patrones relevantes a partir de los datos disponibles. Este atributo cobra especial importancia en sistemas reales de ingeniería con alta variabilidad operativa, como las flotas de ambulancias médicas urbanas, donde la complejidad del comportamiento dificulta la formulación de modelos físicos exhaustivos, mientras que los registros históricos de operación y mantenimiento, cuando se encuentran estructurados y disponibles, pueden ofrecer una base suficiente para el modelado predictivo.

II.1.6.3. Modelos computacionales aplicados al mantenimiento predictivo

En el campo del mantenimiento, los modelos computacionales predictivos se emplean como herramientas de apoyo para anticipar fallas, estimar la degradación de los activos y fortalecer la planificación de las intervenciones técnicas (Taoufyq et al., 2025). Su objetivo principal no es ejecutar acciones de mantenimiento de manera automática, sino proporcionar información analítica que permita a los gestores técnicos tomar decisiones más informadas respecto a cuándo y cómo intervenir sobre un activo.

La literatura especializada reconoce que estos modelos pueden orientarse a distintos objetivos, entre ellos la predicción de la ocurrencia de fallas, la estimación del riesgo operativo o el cálculo de la vida útil remanente de componentes y sistemas. Taoufyq et al. (2025) señalaron que los modelos predictivos en mantenimiento constituyen un pilar de las estrategias modernas de gestión de activos, al permitir transitar desde enfoques reactivos hacia esquemas anticipativos basados en evidencia. En el plano normativo, la ISO 17359:2018 enmarca el pronóstico como una actividad clave dentro del monitoreo de condición y el diagnóstico, legitimando el uso de modelos para apoyar la toma de decisiones de mantenimiento a partir de información disponible (ISO, 2018). En el contexto del presente estudio, este marco conceptual y normativo sustenta el desarrollo de un modelo computacional orientado a estimar la probabilidad de inoperatividad de las ambulancias Tipo II a partir de sus registros históricos, integrando los principios del mantenimiento predictivo basado en datos con los lineamientos de la ingeniería de pronósticos.

II.1.6.4. Modelos data-driven como tipología del modelado computacional

Dentro de las tipologías de modelado computacional, los modelos *data-driven* se caracterizan por extraer relaciones directamente de los datos observados, sin requerir una formulación explícita de las leyes físicas que gobiernan el sistema (Montáns et al., 2019; Taoufyq et al., 2025). Este enfoque se fundamenta en el análisis de información empírica para identificar patrones, tendencias y asociaciones relevantes que permitan anticipar fallas y apoyar la gestión técnica del activo.

Montáns et al. (2019) describieron los modelos *data-driven* como una respuesta al incremento de la disponibilidad de datos en sistemas de ingeniería y destacaron su capacidad para abordar problemas complejos donde el conocimiento físico es incompleto o difícil de formalizar. Taoufyq et al. (2025) clasificaron los modelos *data-driven* como una de las principales categorías del mantenimiento predictivo, distinguiéndolos de los enfoques basados en modelos físicos y de los enfoques híbridos, y

resaltaron su aplicabilidad en entornos reales con restricciones técnicas y presupuestales. En este contexto, el modelo computacional propuesto se inscribe en esta categoría, dado que su construcción se sustenta en registros históricos de operación y mantenimiento de las ambulancias Tipo II, sin requerir instrumentación adicional ni modelado físico de los mecanismos de degradación.

II.1.6.5. Rol del aprendizaje supervisado en el modelado predictivo

El aprendizaje supervisado constituye un marco metodológico consolidado dentro de los modelos computacionales predictivos basados en datos, caracterizado por el uso de conjuntos de datos que contienen variables de entrada y una variable objetivo conocida, a partir de los cuales se identifican relaciones estadísticas que pueden emplearse para generar predicciones sobre nuevos datos (Taoufyq et al., 2025).

En mantenimiento predictivo, el aprendizaje supervisado se utiliza para identificar patrones que preceden a eventos de falla, estimar la probabilidad de ocurrencia de determinados eventos o anticipar la degradación de los activos. Taoufyq et al. (2025) precisaron que este enfoque no debe entenderse como una técnica específica, sino como una categoría metodológica que engloba diversas formas de modelado predictivo. Si et al. (2011) señalaron que los enfoques supervisados son especialmente apropiados para la estimación de variables pronósticas, como la vida útil remanente, siempre que se disponga de datos históricos suficientemente representativos. En la presente investigación, el aprendizaje supervisado se delimita a un problema de clasificación binaria, orientado a estimar la probabilidad de inoperatividad de las ambulancias Tipo II en un horizonte de 14 días a partir de registros históricos de operación y mantenimiento.

II.1.6.5.1. Regresión Logística

La regresión logística es un modelo de clasificación supervisada que estima la probabilidad de pertenencia a una clase binaria a partir de un conjunto de variables explicativas. A diferencia de la regresión lineal, cuya salida es un

valor continuo no acotado, la regresión logística transforma la combinación lineal de las variables de entrada mediante una función que restringe la salida al intervalo (0, 1), interpretable directamente como probabilidad (Bishop, 2006).

En el contexto de esta investigación, la regresión logística se emplea como modelo base interpretable para representar la relación entre las variables explicativas y la probabilidad de inoperatividad de una ambulancia Tipo II. Su utilidad radica en que permite expresar la salida del modelo como una probabilidad, lo que facilita la comparación con un umbral de decisión y la generación de una clasificación final. Además, al estar basada en coeficientes asociados a cada variable, ofrece una primera aproximación a la dirección e intensidad de la relación entre los factores operativos y el evento de inoperatividad. Esta característica la convierte en un punto de referencia adecuado frente a modelos no lineales más complejos.

Función logística o sigmoide

El núcleo matemático del modelo es la función sigmoide. Bishop (2006) la define como:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (5)$$

donde z representa la combinación lineal de las variables explicativas:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p = \beta_0 + \sum_{j=1}^p \beta_j x_j \quad (6)$$

En esta expresión, $x_1, x_2, \dots, x_p$ son las variables explicativas del modelo, $\beta_1, \beta_2, \dots, \beta_p$ son los coeficientes asociados a cada variable y β_0 es el término independiente o sesgo. La función sigmoide produce valores cercanos a 1 cuando z es positivo y grande, y valores cercanos a 0 cuando z es negativo y grande, con una transición suave centrada en z igual a cero como se ilustra en la Figura 3.

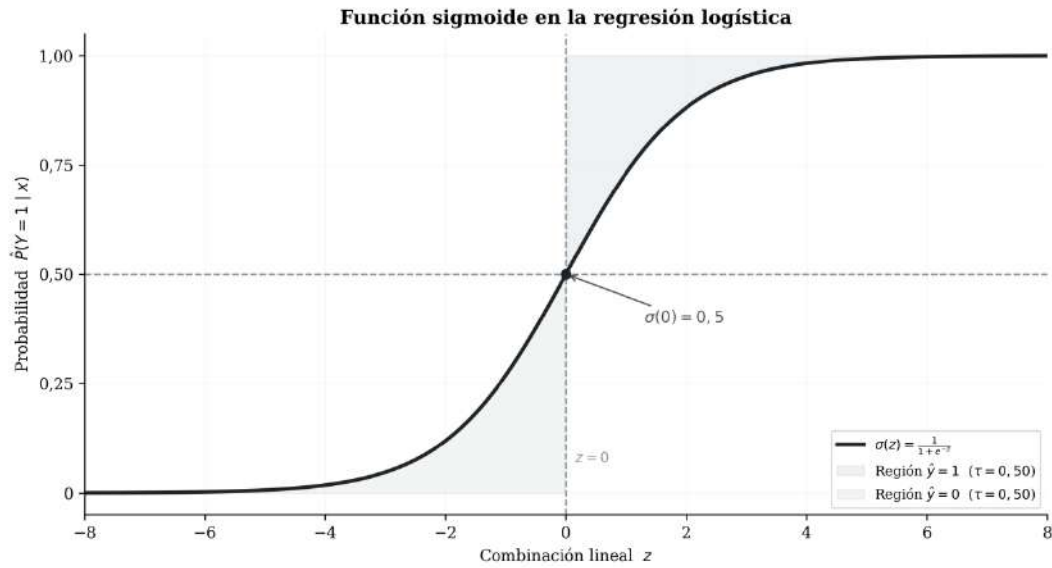


Figura 3. Función sigmoide empleada en la regresión logística para la transformación de la combinación lineal de variables en probabilidad.

Fuente. Elaboración propia.

La probabilidad estimada de que la variable objetivo tome el valor 1, dado el vector de variables explicativas $\mathbf{x}$, se expresa formalmente como (Bishop, 2006):

$$\hat{P}(y = 1 | \mathbf{x}) = \sigma(\mathbf{x}^T \boldsymbol{\beta}) = \frac{1}{1 + e^{-\mathbf{x}^T \boldsymbol{\beta}}} \quad (7)$$

La clasificación final se obtiene comparando esta probabilidad con un umbral τ , generalmente igual a 0,50 en condiciones de clases balanceadas (Bishop, 2006):

$$\hat{y} = \begin{cases} 1 & \text{si } \hat{P}(y = 1 | \mathbf{x}) \geq \tau \\ 0 & \text{si } \hat{P}(y = 1 | \mathbf{x}) < \tau \end{cases} \quad (8)$$

En problemas con desbalance de clases, el umbral τ se ajusta a valores inferiores a 0,50 para favorecer la detección de la clase minoritaria. Para esta investigación se empleó un umbral $\tau = 0,30$, seleccionado para priorizar la sensibilidad sobre la precisión, dado que el costo operativo de una inoperatividad no anticipada supera al costo de una intervención preventiva innecesaria.

Función de pérdida y estimación de parámetros

Los coeficientes β se estiman maximizando la función de verosimilitud del modelo, lo que equivale a minimizar la función de pérdida de entropía cruzada binaria (*binary cross-entropy*). Bishop (2006) la define como:

$$\mathcal{L}(\beta) = -\frac{1}{n} \sum_{i=1}^n [y_i \log \hat{P}_i + (1 - y_i) \log(1 - \hat{P}_i)] \quad (9)$$

donde n es el número de observaciones de entrenamiento, y_i es el valor real de la variable objetivo para la observación i y $\hat{P}_i$ es la probabilidad estimada por el modelo para esa misma observación. Esta función penaliza con mayor intensidad las predicciones erróneas cuando el modelo asigna alta confianza a la clase incorrecta. En consecuencia, el proceso de entrenamiento busca encontrar el conjunto de coeficientes que minimiza esta penalización global, favoreciendo predicciones cuya probabilidad estimada se aproxime al comportamiento observado en los datos de entrenamiento.

La minimización se realiza mediante el algoritmo de descenso por gradiente, cuya regla de actualización de parámetros en cada iteración es (Bishop, 2006):

$$\beta^{(t+1)} = \beta^{(t)} - \alpha \nabla_{\beta} \mathcal{L}(\beta^{(t)}) \quad (10)$$

donde α es la tasa de aprendizaje que controla el tamaño del paso en cada actualización y $\nabla_{\beta} \mathcal{L}$ es el gradiente de la función de pérdida respecto a los coeficientes. Cada iteración del algoritmo ajusta los coeficientes en la dirección que reduce la función de pérdida, permitiendo que el modelo converja progresivamente hacia una solución que represente de mejor manera la relación entre las variables explicativas y la variable objetivo.

No obstante, la optimización exclusiva de la función de pérdida puede conducir a modelos excesivamente ajustados a los datos de entrenamiento cuando existe un número elevado de variables explicativas o relaciones de

dependencia entre ellas. Por esta razón, resulta necesario incorporar mecanismos que favorezcan la capacidad de generalización del modelo.

Regularización

Para evitar el sobreajuste en presencia de muchas variables o multicolinealidad, se incorpora un término de regularización a la función de pérdida. La regularización L2, también denominada Ridge, añade una penalización proporcional al cuadrado de los coeficientes (Bishop, 2006):

$$\mathcal{L}_{L2}(\beta) = -\frac{1}{n} \sum_{i=1}^n [y_i \log \hat{P}_i + (1 - y_i) \log(1 - \hat{P}_i)] + \frac{1}{2C} \sum_{j=1}^p \beta_j^2 \quad (11)$$

donde C es el parámetro de regularización inversa: valores pequeños de C implican mayor penalización y coeficientes más pequeños. En la presente investigación se utilizó C igual a 0,1, lo que corresponde a una regularización moderada que previene el sobreajuste sin eliminar el aporte de las variables explicativas. En términos prácticos, la regularización favorece la capacidad de generalización del modelo al reducir su sensibilidad frente a patrones particulares del conjunto de entrenamiento. Así, se busca mantener un desempeño más estable cuando el modelo se aplica a nuevos registros.

Manejo del desbalance de clases

Cuando las clases presentan distribución desigual, como ocurre en el dataset de esta investigación, el conjunto de entrenamiento presenta un desbalance aproximado de 5:1 entre observaciones negativas y positivas. Por ello, la función de pérdida se ajustó asignando mayor peso a los errores de clasificación de la clase minoritaria. Esto evita que el modelo favorezca de manera predominante la clase mayoritaria durante el entrenamiento.

La Figura 4 ilustra la distribución de clases en el *dataset* de entrenamiento antes y después de aplicar el oversampling aleatorio con reemplazo, mostrando el efecto del balanceo sobre la proporción relativa de ambas clases.

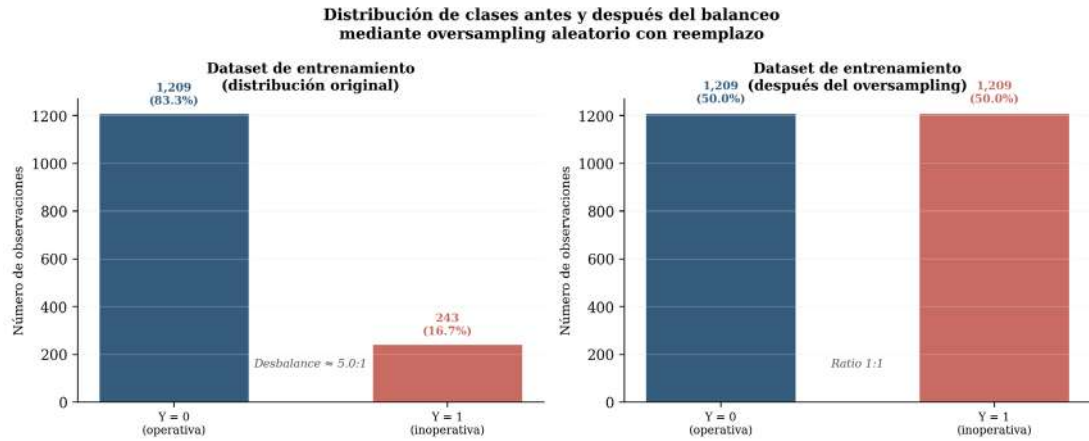


Figura 4. Distribución de clases antes y después del balanceo mediante oversampling aleatorio con reemplazo.

Fuente. Elaboración propia.

Breiman et al. (1984) fundamentan este principio en el marco de los árboles de clasificación y su extensión al caso logístico se implementa mediante el parámetro *class_weight* igual a *'balanced'*, que calcula automáticamente los pesos de cada clase como:

$$w_k = \frac{n}{K \cdot n_k} \quad (12)$$

donde n es el total de observaciones, K es el número de clases y n_k es el número de observaciones de la clase k . Este mecanismo hace que los errores sobre la clase minoritaria (inoperatividad, $Y=1$) penalicen proporcionalmente más en la función de pérdida durante el entrenamiento.

Aplicación en la presente investigación

En el contexto de la tesis, la regresión logística estima la probabilidad de que una ambulancia Tipo II presente inoperatividad en los 14 días siguientes al corte temporal t_0 , a partir de las 14 variables explicativas calculadas en la ventana histórica de 60 días. El modelo opera como referencia interpretable dentro del experimento comparativo, dado que sus coeficientes permiten cuantificar la dirección e intensidad de la relación entre cada variable explicativa y la probabilidad de inoperatividad, a diferencia de los modelos de

ensamble que no ofrecen esa interpretabilidad directa (Prytz, 2014). Sin embargo, ante la aplicación de un umbral conservador de $\tau = 0,30$ sobre datos con desbalance de clases, la regresión logística tendió a clasificar la totalidad de las observaciones como positivas, lo que limitó su utilidad interpretativa en la práctica, conforme se analiza en la Sección VI.2.

II.1.6.5.2. Árboles de Decisión

Un árbol de decisión es un modelo de aprendizaje supervisado que particiona el espacio de variables explicativas mediante una secuencia de reglas de división binaria, construyendo una estructura jerárquica de nodos de decisión y nodos hoja (Breiman et al., 1984). Cada nodo de decisión evalúa una condición sobre una variable explicativa y dirige las observaciones hacia una de dos ramas según el resultado. Los nodos hoja contienen la predicción final del modelo, que en problemas de clasificación binaria corresponde a la clase mayoritaria de las observaciones que alcanzaron ese nodo durante el entrenamiento. La secuencia de particiones continúa hasta alcanzar un criterio de detención previamente establecido o hasta que los nodos presenten un nivel suficiente de homogeneidad.

Criterio de división: impureza de Gini

El proceso de construcción del árbol selecciona en cada nodo la variable y el umbral de división que minimizan la impureza del nodo resultante. El criterio de impureza de Gini, utilizado por el algoritmo CART (*Classification and Regression Trees*), fue introducido por Breiman et al. (1984) y se define como:

$$G(t) = 1 - \sum_{k=0}^{K-1} p_k^2 \quad (13)$$

donde p_k es la proporción de observaciones de la clase k en el nodo t y K es el número de clases. Un nodo con impureza de Gini igual a 0 es un nodo puro, es decir, todas sus observaciones son de la misma clase. Un nodo con impureza máxima de 0,5 contiene igual proporción de ambas clases.

La ganancia de información obtenida al dividir el nodo padre t en dos nodos hijos t_L y t_R se calcula como (Breiman et al., 1984):

$$\Delta G = G(t) - \frac{n_L}{n} G(t_L) - \frac{n_R}{n} G(t_R) \quad (14)$$

donde n es el número total de observaciones en el nodo padre, y n_L y n_R son los tamaños de los nodos hijo izquierdo y derecho respectivamente. El algoritmo selecciona en cada nodo la división que maximiza ΔG entre todas las posibles combinaciones de variable y umbral disponibles. Esto permite que cada división contribuya progresivamente a separar las observaciones según su comportamiento respecto a la variable objetivo.

La Figura 5 presenta un esquema ilustrativo de la lógica de particionamiento de un árbol de decisión aplicado al problema de clasificación de inoperatividad en ambulancias Tipo II. La figura emplea variables representativas del problema, pero no corresponde a un árbol individual extraído del modelo Random Forest final.

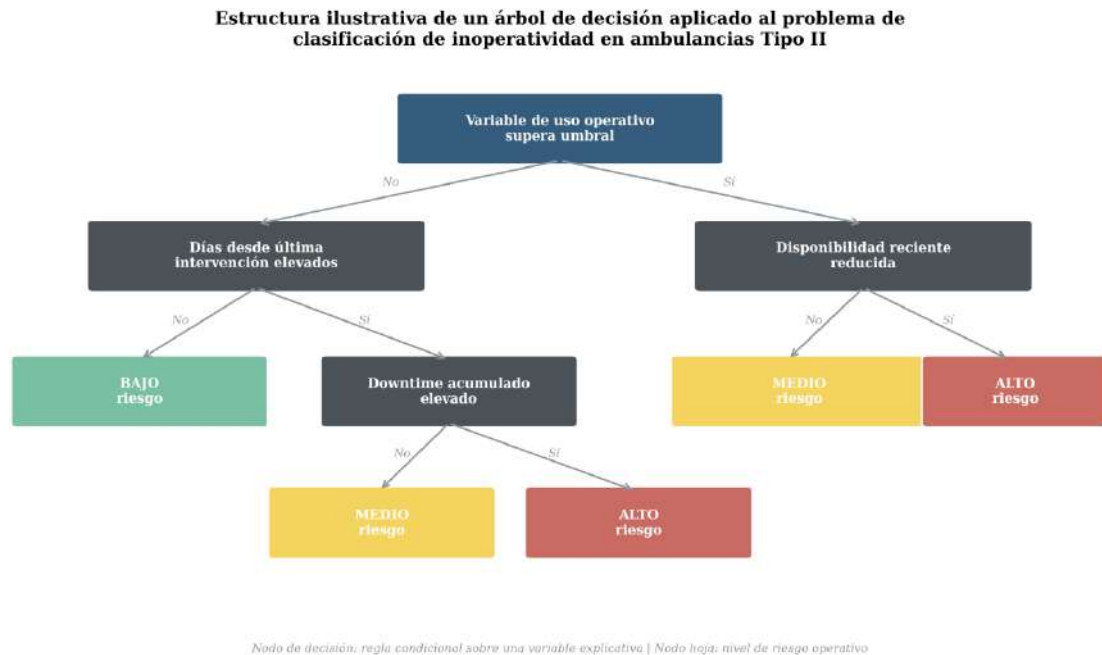


Figura 5. Estructura ilustrativa de un árbol de decisión aplicado al problema de clasificación de inoperatividad en ambulancias Tipo II
Fuente. Elaboración propia.

En la presente investigación, el Árbol de Decisión se incorporó como modelo comparativo individual, con el propósito de evaluar el desempeño de una estructura de clasificación basada en reglas jerárquicas antes de analizar modelos de ensamble derivados de árboles. Este modelo permite establecer una referencia intermedia entre la Regresión Logística y los métodos de ensamble, como Random Forest y Gradient Boosting.

II.1.6.5.3. *Random Forest*

Random Forest: *bagging* y agregación de árboles

Un árbol de decisión individual es susceptible al sobreajuste, ya que puede memorizar los patrones del conjunto de entrenamiento sin generalizar adecuadamente a nuevas observaciones. El algoritmo *Random Forest* aborda esta limitación mediante el método de *bagging* (*bootstrap aggregating*), que entrena un conjunto de K árboles de decisión independientes sobre submuestras distintas del conjunto de entrenamiento y agrega sus predicciones mediante votación mayoritaria (Breiman, 2001).

El proceso de construcción de un *Random Forest* sigue los pasos formales siguientes. Para cada árbol k con $k = 1, 2, \dots, K$:

- Primero, se genera una submuestra D_k de tamaño n mediante muestreo con reemplazo (*bootstrapping*) del conjunto de entrenamiento original de n observaciones, de modo que cada submuestra incluye aproximadamente el 63,2% de las observaciones originales únicas.
- Segundo, se entrena un árbol de decisión $h_k(\mathbf{x})$ sobre D_k , seleccionando en cada nodo el mejor umbral de división entre un subconjunto aleatorio de $\sqrt{p}$ variables, donde p es el número total de variables explicativas. Esta restricción introduce diversidad entre los árboles y reduce la correlación entre sus predicciones.
- Tercero, el árbol crece hasta alcanzar la profundidad máxima definida o hasta que los nodos hoja contienen menos del número mínimo de observaciones requerido.

La predicción final del modelo para una nueva observación $\mathbf{x}$ se obtiene mediante votación mayoritaria sobre las predicciones de los K árboles (Breiman, 2001):

$$\hat{y} = \text{mode}\{h_1(\mathbf{x}), h_2(\mathbf{x}), \dots, h_K(\mathbf{x})\} \quad (15)$$

Para estimaciones de probabilidad, necesarias para la evaluación mediante AUC-ROC, la probabilidad de la clase positiva se estima como la proporción de árboles que predicen $y = 1$ (Breiman, 2001). Esta agregación es una de los principales mecanismos que mejora la estabilidad de las predicciones:

$$\hat{P}(y = 1 | \mathbf{x}) = \frac{1}{K} \sum_{k=1}^K 1[h_k(\mathbf{x}) = 1] \quad (16)$$

La Figura 6 ilustra el esquema de *bagging* en *Random Forest* con la generación de submuestras *bootstrap*, el entrenamiento de árboles independientes y la agregación por votación mayoritaria.

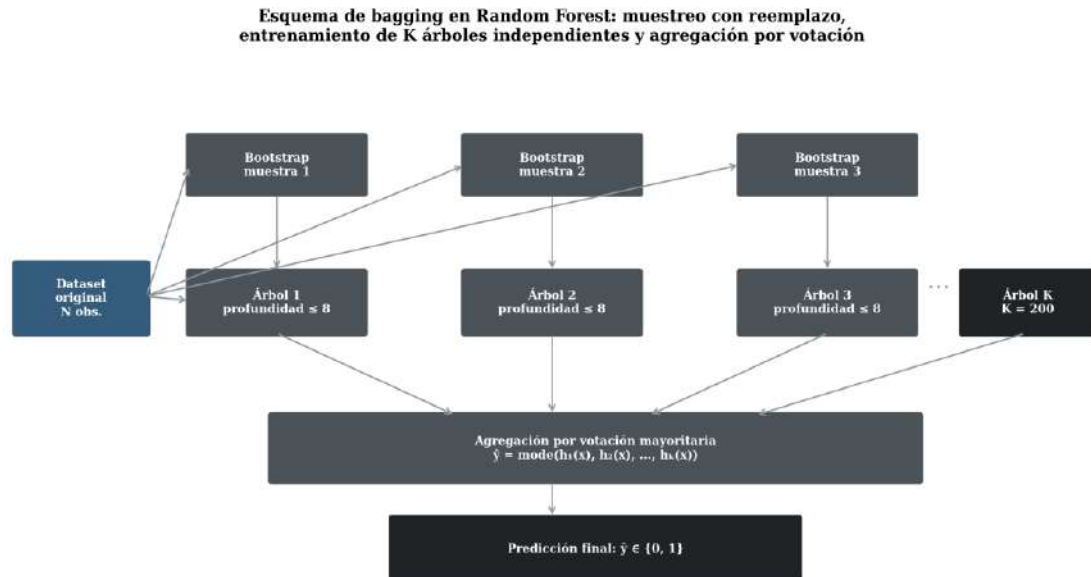


Figura 6. Esquema de bagging en Random Forest: muestreo con reemplazo, entrenamiento de K árboles independientes y agregación por votación.

Fuente. Elaboración propia.

Importancia de variables

Una de las ventajas principales del *Random Forest* para la presente investigación es su capacidad para estimar la importancia de cada variable explicativa en la predicción. La importancia de una variable x_j se calcula como la reducción total ponderada de impureza de Gini acumulada a través de todos los nodos donde esa variable fue seleccionada como criterio de división, promediada sobre los K árboles del bosque (Breiman, 2001):

$$\text{Imp}(x_j) = \frac{1}{K} \sum_{k=1}^K \sum_{\substack{t \in T_k \\ v(t)=j}} \frac{n_t}{n} \Delta G(t) \quad (17)$$

donde T_k es el conjunto de nodos del árbol k , $v(t)$ es la variable de división en el nodo t y n_t es el número de observaciones que alcanzan ese nodo. Las importancias normalizadas suman 1 y representan la contribución relativa de cada variable al poder predictivo total del modelo. Este criterio permite interpretar el modelo no solo por su desempeño predictivo, sino también por la relevancia relativa de los factores técnicos y operativos considerados.

En la presente investigación, este análisis reveló que las variables de uso operativo concentran aproximadamente el 35,4% de la importancia relativa total del modelo, siendo el kilometraje acumulado en la ventana W 18,52% y los servicios prestados en W 16,91% los predictores de mayor peso entre las 14 variables explicativas consideradas.

Hiperparámetros empleados en la presente investigación

El modelo *Random Forest* fue configurado con los parámetros siguientes: número de árboles K igual a 200, profundidad máxima de cada árbol igual a 8 niveles, número mínimo de observaciones por hoja igual a 10, ponderación de clases igual a '*balanced*' como medida adicional de robustez frente al desbalance original de la clase minoritaria, y semilla aleatoria fija igual a 42 para garantizar la reproducibilidad de los resultados.

II.1.6.5.4.

Gradient Boosting

El *Gradient Boosting* es un algoritmo de aprendizaje supervisado basado en el principio del *boosting*, que construye un modelo predictivo fuerte mediante la combinación secuencial de múltiples modelos débiles, generalmente árboles de decisión de profundidad reducida (Friedman, 2001). A diferencia del *Random Forest*, donde los árboles se entrenan de forma independiente y en paralelo, el *Gradient Boosting* entrena cada árbol para corregir los errores residuales del modelo acumulado en la iteración anterior, orientando el aprendizaje en la dirección del gradiente negativo de la función de pérdida. Esta lógica secuencial permite concentrar el aprendizaje en los errores residuales del modelo previo, reforzando gradualmente la capacidad predictiva del conjunto.

Formulación del modelo aditivo

El modelo de *Gradient Boosting* construye una aproximación $F_M(\mathbf{x})$ mediante la suma de M modelos débiles $h_m(\mathbf{x})$, partiendo de una estimación inicial $F_0(\mathbf{x})$ (Friedman, 2001):

$$F_M(\mathbf{x}) = F_0(\mathbf{x}) + \sum_{m=1}^M \eta \cdot h_m(\mathbf{x}) \quad (18)$$

donde η es la tasa de aprendizaje (*learning rate*) que controla la contribución de cada árbol al modelo acumulado y $h_m(\mathbf{x})$ es el árbol débil entrenado en la iteración m . La estimación inicial $F_0(\mathbf{x})$ corresponde a la predicción constante que minimiza la función de pérdida sobre el conjunto de entrenamiento, que para problemas de clasificación binaria con entropía cruzada es el logaritmo de la razón de probabilidades de las clases:

$$F_0(\mathbf{x}) = \log \frac{\bar{y}}{1 - \bar{y}} \quad (19)$$

donde $\bar{y}$ es la prevalencia de la clase positiva en el conjunto de entrenamiento.

Algoritmo de entrenamiento: descenso por gradiente funcional

La contribución central de Friedman (2001) es la formulación del *boosting* como un proceso de descenso por gradiente en el espacio funcional. En cada iteración m , el algoritmo calcula los residuos negativos o pseudo-residuales r_{im} , que representan el gradiente negativo de la función de pérdida evaluado en las predicciones del modelo actual:

$$r_{im} = - \left[\frac{\partial \mathcal{L}(y_i, F(\mathbf{x}_i))}{\partial F(\mathbf{x}_i)} \right]_{F=F_{m-1}} \quad (20)$$

Donde $\mathcal{L}(y_i, F(\mathbf{x}_i))$ es la función de pérdida evaluada para la observación i . Para el caso de clasificación binaria con función de pérdida de entropía cruzada, el pseudo-residual de cada observación se simplifica como (Friedman, 2001):

$$r_{im} = y_i - \hat{P}_{m-1}(y_i = 1 | \mathbf{x}_i) \quad (21)$$

Es decir, la diferencia entre el valor real de la variable objetivo y la probabilidad estimada por el modelo acumulado hasta la iteración anterior. Este pseudo-residual indica en qué dirección y con qué magnitud debe ajustarse la predicción del modelo.

El árbol débil $h_m(\mathbf{x})$ se entrena sobre los pseudo-residuales $\{r_{im}\}_{i=1}^n$ como variable objetivo, aproximando localmente el gradiente de la función de pérdida en cada región del espacio de variables. El modelo se actualiza entonces como (Friedman, 2001):

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \eta \cdot h_m(\mathbf{x}) \quad (22)$$

Este proceso se repite durante M iteraciones, construyendo de forma aditiva un modelo que converge progresivamente hacia la función que minimiza la función de pérdida sobre el conjunto de entrenamiento. La Figura 7 ilustra el proceso del modelo con la corrección iterativa de errores residuales.

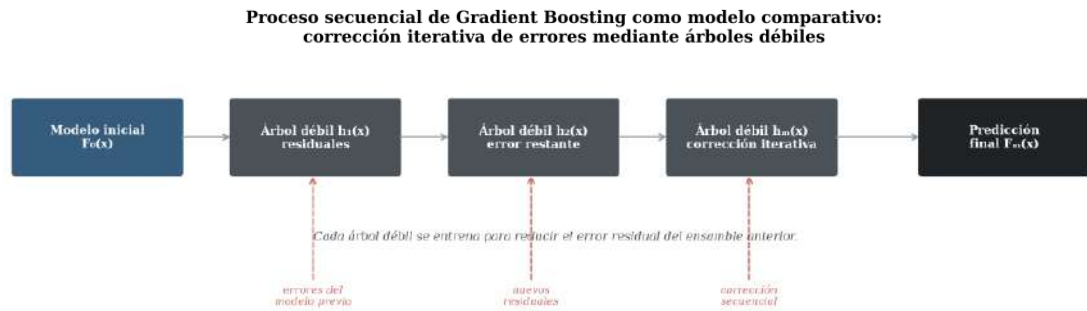


Figura 7. Proceso secuencial del Gradient Boosting: cada árbol débil corrige los errores del modelo anterior mediante el gradiente de la función de pérdida.

Fuente. Elaboración propia.

Regularización y control del sobreajuste

El *Gradient Boosting* incorpora tres mecanismos de regularización que controlan la complejidad del modelo y reducen el riesgo de sobreajuste (Friedman, 2001).

El primero es la tasa de aprendizaje η , que escala la contribución de cada árbol al modelo acumulado. Valores pequeños de η requieren un mayor número de iteraciones M para converger, pero generan modelos con mejor capacidad de generalización. Existe una relación de compromiso entre η y M : reducir η a la mitad requiere aproximadamente duplicar M para obtener el mismo error de entrenamiento.

El segundo es la profundidad máxima de cada árbol débil, que limita la complejidad de las interacciones entre variables que el modelo puede capturar. Árboles de profundidad 1, denominados *stumps*, capturan únicamente efectos principales de una sola variable. La presente investigación utilizó una profundidad máxima de 4, lo que permite capturar interacciones de hasta cuatro variables sin generar estructuras excesivamente complejas.

El tercero es el *subsampling*, que entrena cada árbol sobre una fracción aleatoria del conjunto de entrenamiento sin reemplazo. Friedman (2001) demostró que valores de *subsampling* entre 0,5 y 0,8 mejoran la precisión

del modelo al introducir aleatoriedad que reduce la correlación entre los árboles sucesivos. En la presente investigación se empleó un valor de *subsampling* de 0,8.

Predicción final para clasificación binaria

Una vez completadas las M iteraciones, el modelo $F_M(\mathbf{x})$ produce un valor real no acotado. Para obtener la probabilidad estimada de inoperatividad se aplica la función sigmoide definida en la Ecuación (5):

$$\hat{P}(y = 1 | \mathbf{x}) = \sigma(F_M(\mathbf{x})) = \frac{1}{1 + e^{-F_M(\mathbf{x})}} \quad (23)$$

La clasificación final se obtiene comparando esta probabilidad con el umbral τ definido en la Ecuación (8). Esta formulación establece una conexión directa entre el *Gradient Boosting* y la regresión logística: ambos modelos estiman la probabilidad de la clase positiva mediante la función sigmoide, pero mientras la regresión logística ajusta una función lineal en el espacio de variables, el *Gradient Boosting* ajusta una función no lineal mediante la combinación de múltiples árboles. Esta diferencia permite representar relaciones más complejas entre las variables explicativas y la probabilidad de inoperatividad, sin asumir una dependencia estrictamente lineal entre ellas.

Hiperparámetros empleados en la presente investigación

El modelo *Gradient Boosting* fue configurado con los parámetros siguientes: número de estimadores M igual a 150, profundidad máxima de cada árbol igual a 4 niveles, tasa de aprendizaje η igual a 0,05, fracción de *subsampling* igual a 0,8 y semilla aleatoria fija igual a 42 para garantizar la reproducibilidad de los resultados. Esta configuración implementa las recomendaciones de regularización de Friedman (2001) combinando una tasa de aprendizaje baja con *subsampling* para reducir el sobreajuste sobre el conjunto de entrenamiento simulado. Con ello se prioriza un equilibrio entre capacidad de ajuste, estabilidad del modelo y control de complejidad.

II.1.6.5.5. Métricas de evaluación para clasificación binaria desbalanceada

La evaluación del desempeño de un modelo de clasificación binaria requiere métricas que capturen su capacidad discriminativa sobre datos no vistos durante el entrenamiento. En problemas con clases desbalanceadas, como este estudio donde la clase positiva es 20,8% de las observaciones en el conjunto de validación 2025, las métricas tradicionales como la exactitud global resultan insuficientes porque un clasificador trivial que prediga siempre la clase negativa alcanzaría una exactitud aproximada de 79,2% sin capacidad real para detectar eventos de inoperatividad (Bishop, 2006). Por esta razón, la evaluación del modelo se fundamenta en métricas derivadas directamente de la matriz de confusión y en el área bajo la curva ROC.

Matriz de confusión

La matriz de confusión es la estructura fundamental a partir de la cual se derivan todas las métricas de evaluación en clasificación binaria (Breiman et al., 1984). Para un clasificador binario con clases positiva, inoperatividad, y negativa, operatividad, la matriz de confusión organiza las predicciones del modelo en cuatro categorías. La Figura 8 presenta esta estructura y resume la interpretación operativa de cada celda en el contexto de la predicción de inoperatividad en ambulancias Tipo II.



Figura 8. Estructura de la matriz de confusión y métricas derivadas para la evaluación de modelos de clasificación binaria.

Fuente. Elaboración propia.

Donde VP son las observaciones correctamente identificadas como próximas a la inoperatividad, FN son las observaciones correspondientes a ambulancias que presentarán inoperatividad pero que el modelo no anticipó, FP son las observaciones clasificadas como de riesgo que finalmente no presentarán inoperatividad y VN son las observaciones correctamente identificadas como operativas. En el contexto operativo de la presente investigación, los FN representan el error de mayor consecuencia porque corresponden a inoperatividades no anticipadas que comprometen la capacidad de respuesta del sistema prehospitalario.

Por esta razón, la evaluación del modelo no puede limitarse a una única métrica de desempeño, sino que requiere analizar de manera conjunta indicadores que reflejen distintos tipos de aciertos y errores de clasificación, de acuerdo con las implicancias operativas de cada uno.

Precisión

La precisión mide la proporción de predicciones positivas que son correctas, es decir, de las observaciones que el modelo señala como próximas a la inoperatividad, cuántas efectivamente corresponden a eventos positivos (Bishop, 2006):

$$\text{Precisión} = \frac{VP}{VP + FP} \quad (24)$$

Un valor bajo de precisión indica un alto número de falsas alarmas. En problemas de clases desbalanceadas con umbral de clasificación conservador, la precisión tiende a ser baja porque el modelo clasifica muchas observaciones como positivas para maximizar la detección de la clase minoritaria.

Sensibilidad o Recall

La sensibilidad, también denominada *recall* o tasa de verdaderos positivos, mide la proporción de eventos de inoperatividad reales que el modelo identifica correctamente (Bishop, 2006):

$$\text{Sensibilidad} = \frac{VP}{VP + FN} \quad (25)$$

La sensibilidad es la métrica de mayor relevancia operativa en la presente investigación porque cuantifica directamente la capacidad del modelo para anticipar inoperatividades antes de que ocurran. En la presente investigación, un valor de sensibilidad de 91,86% en el modelo *Random Forest* significa que, de cada 100 eventos de inoperatividad reales del conjunto de validación, el modelo anticipa correctamente aproximadamente 92. En un sistema crítico como el prehospitalario, esta métrica adquiere prioridad operativa debido a que el costo de una inoperatividad no anticipada puede ser mayor que el costo de una inspección preventiva adicional.

Especificidad

La especificidad mide la proporción de ambulancias operativas que el modelo clasifica correctamente como tales (Bishop, 2006):

$$\text{Especificidad} = \frac{VN}{VN + FP} \quad (26)$$

Un valor bajo de especificidad implica un número elevado de intervenciones preventivas innecesarias, lo que tiene implicancias sobre la eficiencia en la asignación de recursos de mantenimiento. Por ello, esta métrica resulta especialmente útil para evaluar el impacto operativo de las falsas alarmas sobre la planificación de intervenciones, la carga del equipo técnico y la utilización eficiente de los recursos disponibles. Su interpretación complementa a la sensibilidad, ya que permite valorar el equilibrio entre anticipar fallas y evitar intervenciones innecesarias.

Puntuación F1

La puntuación F1 es la media armónica entre la precisión y la sensibilidad, y provee una métrica única que equilibra ambas dimensiones del desempeño (Bishop, 2006):

$$F_1 = 2 \cdot \frac{\text{Precisión} \times \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}} = \frac{2 \cdot VP}{2 \cdot VP + FP + FN} \quad (27)$$

La media armónica penaliza con mayor intensidad los valores bajos: un modelo con precisión alta pero sensibilidad muy baja, o viceversa, obtendrá un F1 bajo. Esta propiedad la hace especialmente adecuada para evaluar modelos sobre conjuntos desbalanceados donde uno de los dos indicadores tiende a ser bajo.

Error medio absoluto

El error medio absoluto (MAE) sobre las probabilidades estimadas mide la distancia promedio entre las probabilidades predichas por el modelo y los valores reales de la variable objetivo (Bishop, 2006):

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{p}_i| \quad (28)$$

Donde $\hat{p}_i$ es la probabilidad estimada de inoperatividad para la observación i y y_i es el valor real binario. A diferencia de las métricas basadas en clasificación con umbral fijo, el MAE evalúa la calidad de las probabilidades estimadas independientemente del umbral seleccionado.

Área bajo la curva ROC

La curva ROC (*Receiver Operating Characteristic*) representa la relación entre la sensibilidad y la tasa de falsos positivos del clasificador para todos los posibles valores del umbral de clasificación τ (Bishop, 2006). La tasa de falsos positivos se define como:

$$\text{TFP} = \frac{FP}{FP + VN} = 1 - \text{Especificidad} \quad (29)$$

El área bajo la curva ROC (AUC-ROC) resume en un único valor escalar el desempeño discriminativo del clasificador independientemente del umbral.

Se define como (Bishop, 2006):

$$\text{AUC-ROC} = \int_0^1 \text{Sensibilidad (TFP)} d(\text{TFP}) \quad (30)$$

La Figura 9 ilustra conceptualmente la estructura de las curvas ROC y la interpretación del AUC-ROC como métrica de evaluación.

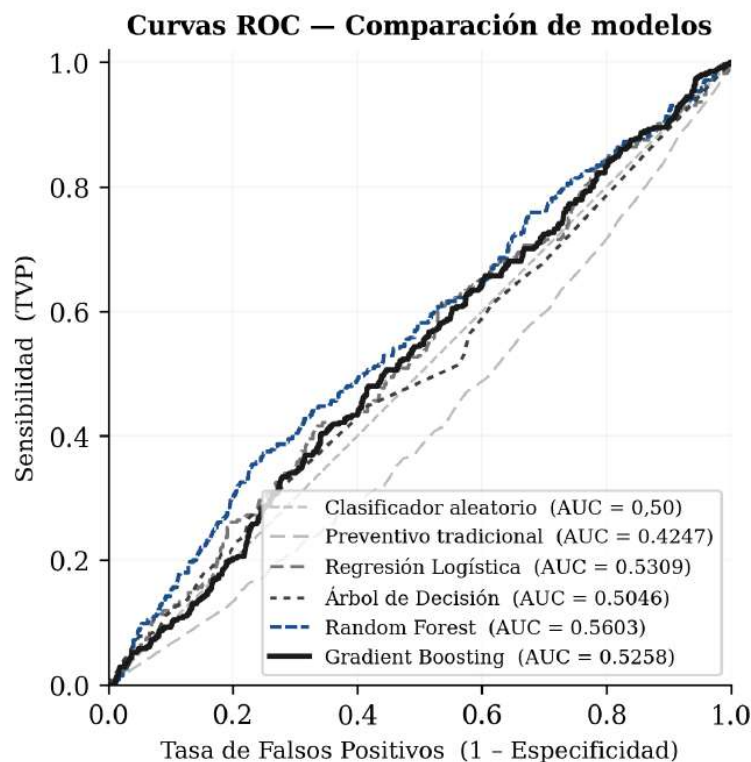


Figura 9. Curvas ROC comparativas de los modelos evaluados en la validación retrospectiva.

Nota. La figura corresponde a los resultados del conjunto de validación 2025, calculados con las probabilidades estimadas por cada modelo.

Fuente. Elaboración propia.

Un AUC-ROC de 1,0 indica discriminación perfecta, mientras que un valor de 0,5 indica que el modelo no tiene capacidad discriminativa superior al azar. El AUC-ROC es una métrica útil para problemas desbalanceados porque evalúa el desempeño discriminativo del modelo en todo el espectro de umbrales posibles, sin depender de la elección de un umbral específico.

En la presente investigación el AUC-ROC constituye el indicador principal de

comparación entre modelos, complementado por la sensibilidad como indicador de relevancia operativa. Esta elección es coherente con la recomendación de Prytz (2014) para la evaluación de modelos predictivos aplicados a flotas vehiculares con datos históricos desbalanceados.

Disponibilidad operativa proyectada

Adicionalmente a las métricas estándar de clasificación, la presente investigación incorpora la disponibilidad operativa proyectada como métrica de relevancia directa para la gestión del mantenimiento. Esta métrica cuantifica la proporción de observaciones que el modelo clasifica como operativas en el horizonte de predicción T:

$$\text{Disp. proyectada} = \frac{VN + FN}{n} = \frac{\sum_{i=1}^n 1[\hat{y}_i = 0]}{n} \quad (31)$$

La diferencia entre la disponibilidad proyectada por el modelo y la disponibilidad real observada refleja el conservadurismo del umbral de clasificación: un umbral bajo genera mayor número de intervenciones preventivas y menor disponibilidad proyectada, mientras que un umbral alto genera menor número de intervenciones y mayor disponibilidad proyectada pero con mayor riesgo de inoperatividades no anticipadas.

La Figura 10 ilustra la relación inversa entre precisión y sensibilidad al variar el umbral τ , evidenciando el compromiso que debe gestionar todo clasificador binario en problemas con clases desbalanceadas.

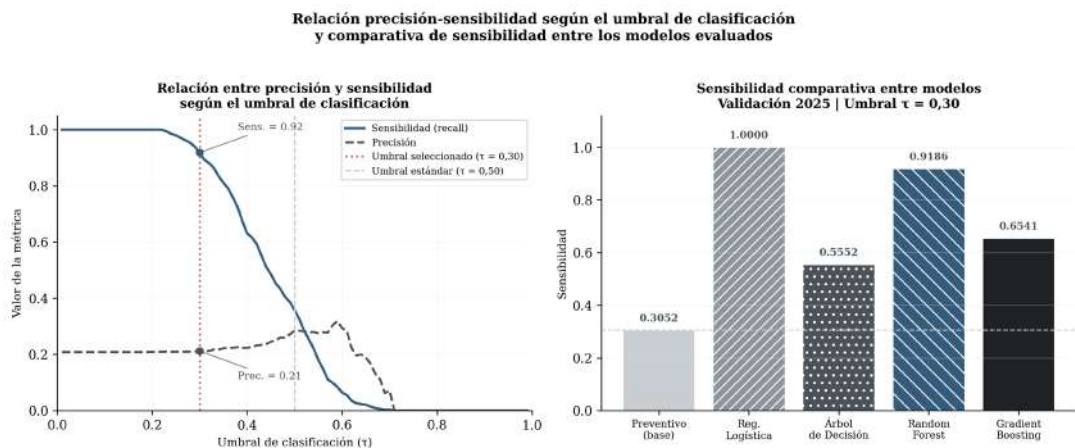


Figura 10. Relación entre precisión y sensibilidad según el umbral de clasificación y comparativa de sensibilidad entre modelos.

Fuente. Elaboración propia.

En problemas de clasificación binaria desbalanceada, la sensibilidad adquiere especial relevancia cuando la clase positiva representa eventos críticos cuya omisión puede generar consecuencias operativas significativas. Branco et al. (2016) señalan que, en este tipo de dominios, los eventos menos frecuentes suelen corresponder precisamente a los casos de mayor interés para el usuario final, por lo que métricas globales pueden no reflejar adecuadamente la capacidad del modelo para detectarlos. En el ámbito del mantenimiento predictivo, esta consideración resulta particularmente importante, dado que la anomalía o evento positivo suele representar una falla potencial que se busca identificar antes de que afecte la disponibilidad del sistema (Carrasco et al., 2021). En consecuencia, la sensibilidad constituye una métrica especialmente útil para evaluar la capacidad de detección temprana de eventos críticos, complementándose con indicadores como precisión, especificidad y exactitud para obtener una valoración integral del desempeño del modelo.

II.1.6.6. El modelo computacional data-driven como herramienta para el mantenimiento predictivo

En el marco de la presente investigación, la variable independiente se define como un modelo computacional basado en datos orientado al mantenimiento predictivo de ambulancias médicas urbanas Tipo II. La revisión teórica presentada sustenta que los modelos computacionales *data-driven* constituyen una herramienta metodológicamente válida en ingeniería para anticipar el comportamiento futuro de sistemas reales, especialmente en contextos donde la disponibilidad de datos históricos supera la viabilidad de implementar sistemas de monitoreo en tiempo real.

Este enfoque es consistente con las características del sistema de estudio, al permitir el aprovechamiento de registros existentes de operación y mantenimiento para apoyar decisiones técnicas relacionadas con la

disponibilidad y confiabilidad de las unidades. En consecuencia, el modelo computacional propuesto no busca reemplazar los procesos tradicionales de mantenimiento, sino complementarlos mediante un soporte analítico que apoye la planificación y priorización de intervenciones. En el caso de las ambulancias Tipo II, esta elección metodológica resulta pertinente porque el comportamiento operativo de la unidad depende de múltiples factores que no siempre pueden ser representados mediante ecuaciones físicas explícitas. Por ello, el enfoque *data-driven* permite modelar relaciones empíricas entre uso, mantenimiento e inoperatividad, manteniendo una lógica aplicable a contextos con información histórica estructurada. Esto refuerza su utilidad como herramienta de apoyo para la gestión del mantenimiento y la toma de decisiones técnicas. La Tabla 10 sintetiza las diferencias entre los modelos basados en física y los modelos *data-driven*, destacando las características que justifican la selección de este último enfoque para esta investigación.

Tabla 10 Tipologías de modelos computacionales en mantenimiento predictivo

Criterio	Modelos basados en física	Modelos <i>data-driven</i> (enfoque de la tesis)
Base del modelo	Ecuaciones matemáticas y leyes físicas	Datos históricos y observacionales
Conocimiento del sistema	Alto y explícito	Parcial o implícito
Requerimientos de información	Moderados	Elevados
Adaptabilidad a sistemas reales	Limitada	Alta
Uso en mantenimiento predictivo	Condicionado	Ampliamente documentado
Dependencia de sensores	Frecuente	No necesaria

Fuente. Elaboración propia a partir de Montáns et al. (2019), Taoufyq et al. (2025) e ISO 17359:2018.

La Figura 11 ilustra la relación estructural entre el sistema real, los datos históricos disponibles y el modelo computacional predictivo, mostrando cómo

los registros generados por la operación y el mantenimiento de las ambulancias Tipo II alimentan el modelo sin requerir instrumentación adicional, y cómo las salidas del modelo se traducen en soporte para la toma de decisiones técnicas.



Figura 11. Relación entre el sistema real, los datos históricos y el modelo computacional predictivo.

Fuente. Elaboración propia a partir de Montáns et al. (2019), Prytz (2014), International Organization for Standardization (2018) y Taoufyq et al. (2025).

En síntesis, los modelos computacionales *data-driven* constituyen un pilar conceptual del mantenimiento predictivo moderno, al permitir anticipar el comportamiento de los activos mediante el análisis de datos históricos y operativos. Su fundamentación en la literatura de ingeniería y su posible aplicación sin sensorización en tiempo real los posiciona como una alternativa metodológicamente viable para sistemas reales, como las ambulancias médicas urbanas (Taoufyq et al., 2025; Prytz, 2014). Además, su flexibilidad permite adaptar el análisis a distintos niveles de disponibilidad y calidad de datos.

Esta base teórica habilita el desarrollo del modelo computacional propuesto en la presente investigación, orientado a estimar la probabilidad de inoperatividad, fortalecer la gestión técnica y apoyar decisiones orientadas a

la disponibilidad operativa de las ambulancias. En ese sentido, el modelo se concibe como una herramienta de soporte analítico que transforma registros históricos en información útil para la priorización preventiva. Asimismo, da un criterio objetivo para complementar la planificación del mantenimiento. Esta formulación establece las bases para el análisis del soporte preventivo y la validación del modelo en los apartados siguientes.

II.1.7. Integración del modelo predictivo como soporte al mantenimiento preventivo

II.1.7.1. El mantenimiento preventivo como base operativa del sistema

El mantenimiento preventivo constituye el eje tradicional sobre el cual se estructura la gestión técnica de los activos en los sistemas de salud. De acuerdo con la Organización Mundial de la Salud, este enfoque se basa en la ejecución planificada de actividades de inspección, ajuste y sustitución de componentes con el objetivo de reducir la probabilidad de fallas y asegurar la continuidad operativa de los equipos médicos y las unidades de transporte sanitario (WHO, 2011). En el contexto de la atención prehospitalaria, este enfoque cobra especial importancia dado que la indisponibilidad de una ambulancia puede comprometer directamente la capacidad de respuesta del sistema ante situaciones de emergencia.

Desde una perspectiva operativa, el mantenimiento preventivo presenta ventajas operativas documentadas: permite programar intervenciones con antelación, facilita la asignación de recursos y contribuye al cumplimiento de requisitos normativos y de seguridad (WHO, 2011; Moubray, 1997). Asimismo, constituye el enfoque predominante en instituciones públicas, donde la estandarización de procedimientos y la previsibilidad de costos son elementos determinantes para la gestión. El mantenimiento preventivo constituye, por tanto, una base operativa necesaria sobre la cual pueden articularse estrategias más avanzadas de gestión técnica, como el enfoque

predictivo propuesto en la presente investigación para las ambulancias médicas urbanas Tipo II.

II.1.7.2. Limitaciones operativas del mantenimiento preventivo tradicional

A pesar de sus ventajas, la literatura especializada reconoce que el mantenimiento preventivo tradicional presenta limitaciones cuando se aplica de manera rígida en sistemas caracterizados por alta variabilidad operativa. Moubray (1997) señaló que los programas basados exclusivamente en intervalos de tiempo o kilometraje asumen implícitamente que el deterioro de los activos sigue patrones uniformes y predecibles, supuesto que no siempre se verifica en la práctica. En el caso de las ambulancias médicas, factores como la intensidad de uso, las condiciones del entorno urbano, el tipo de emergencias atendidas y la carga operativa generan patrones de desgaste heterogéneos entre unidades.

Estas limitaciones pueden traducirse en intervenciones innecesarias sobre componentes que aún conservan condiciones operativas adecuadas, así como en la omisión de fallas incipientes que no se manifiestan dentro de los intervalos establecidos (Molęda et al., 2023). Desde el punto de vista de la gestión, ello implica un uso subóptimo de recursos técnicos y económicos, y una capacidad limitada para priorizar las intervenciones en función del riesgo real asociado a cada activo. En el caso de las ambulancias médicas urbanas del SAMU en Lima Metropolitana, donde los recursos técnicos y presupuestales son limitados (Defensoría del Pueblo, 2023a), estas restricciones evidencian la necesidad de complementar el mantenimiento preventivo con enfoques que incorporen información sobre el estado real del activo, función que el modelo computacional propuesto en la presente investigación busca cumplir mediante el análisis de datos históricos de las ambulancias Tipo II.

II.1.7.3. Rol del modelo predictivo como soporte al mantenimiento preventivo

En el presente estudio, los modelos predictivos basados en datos se configuran como una herramienta de soporte al mantenimiento preventivo, sin pretender sustituirlo. Su función principal consiste en aportar información analítica que permita anticipar comportamientos futuros del activo a partir del análisis de datos históricos de operación y mantenimiento. Tal como señalaron Taoufyq et al. (2025), el mantenimiento predictivo puede integrarse a los esquemas existentes como apoyo a la planificación, orientado a reducir la incertidumbre asociada a la ocurrencia de fallas.

El concepto de soporte preventivo implica que el modelo predictivo no define de manera autónoma las intervenciones de mantenimiento, sino que complementa los programas preventivos al proporcionar estimaciones sobre el estado probable de los activos (Taoufyq et al., 2025; ISO, 2018). En la práctica, esto permite identificar unidades con mayor probabilidad de presentar fallas, ajustar prioridades de atención y reforzar la vigilancia técnica sobre aquellos elementos que representan un mayor riesgo para la continuidad operativa. En el contexto del sistema prehospitalario de Lima Metropolitana, el modelo computacional propuesto constituye una herramienta de apoyo a la decisión alineada con los principios de gestión del mantenimiento y con las restricciones técnicas y presupuestales propias del SAMU, donde la disponibilidad operativa de las ambulancias Tipo II es determinante para la capacidad de respuesta ante emergencias.

II.1.7.4. Priorización preventiva y planificación de intervenciones

Uno de los principales aportes del soporte predictivo al mantenimiento preventivo radica en la posibilidad de ajustar las prioridades de intervención y orientar la planificación de acciones técnicas. Moubray (1997) destacó que la efectividad del mantenimiento no depende únicamente de la frecuencia de las intervenciones, sino de su adecuación al riesgo real de falla. En este sentido, la información proporcionada por los modelos predictivos permite introducir flexibilidad en los programas preventivos, ajustando las intervenciones en función del comportamiento observado de los activos (Taoufyq et al., 2025).

Desde una perspectiva operativa, esta optimización se traduce en una mejor asignación de recursos, al concentrar los esfuerzos de mantenimiento en aquellas unidades que presentan mayores probabilidades de falla o mayor impacto potencial sobre la disponibilidad del sistema (Mołęda et al., 2023). En el caso de las ambulancias médicas urbanas del SAMU en Lima Metropolitana, donde los recursos técnicos y presupuestales son limitados (Defensoría del Pueblo, 2023a), la priorización basada en información predictiva puede contribuir a orientar la mejora de la disponibilidad operativa, aunque su implementación requiere evaluar la carga de mantenimiento generada por las alertas y los falsos positivos. De este modo, el mantenimiento preventivo puede evolucionar desde un proceso estrictamente calendarizado hacia un esquema más adaptativo, informado por el análisis sistemático de datos históricos.

II.1.7.5. Enfoque complementario: integración predictivo - preventivo

La integración del mantenimiento predictivo como soporte al mantenimiento preventivo constituye un enfoque complementario consistente con la práctica real de la gestión de activos en el sector salud. La ISO 17359:2018 reconoce que las actividades de monitoreo, diagnóstico y pronóstico deben orientarse a respaldar la toma de decisiones de mantenimiento, en lugar de sustituir los procesos establecidos (ISO, 2018). Desde esta perspectiva, el enfoque predictivo se incorpora como una capa adicional de análisis que fortalece los programas preventivos existentes.

Este enfoque integrado permite mantener la estructura organizativa y normativa del mantenimiento preventivo, al tiempo que introduce herramientas analíticas que mejoran la eficiencia y la capacidad de anticipación del sistema (Mołęda et al., 2023). En consecuencia, el mantenimiento predictivo con soporte preventivo representa una evolución gradual y metodológicamente fundamentada de la gestión técnica, especialmente adecuada para sistemas críticos como el de ambulancias médicas urbanas, donde la continuidad operativa y la seguridad del paciente son determinantes para la calidad del servicio (WHO, 2011). Esta

complementariedad permite conservar los criterios programados del mantenimiento preventivo, incorporando al mismo tiempo información predictiva para ajustar prioridades según el riesgo estimado de inoperatividad.

La Tabla 11 resume las diferencias entre el mantenimiento preventivo tradicional y el mantenimiento con soporte predictivo, y la Figura 12 ilustra conceptualmente su integración, estableciendo las bases para el análisis de los criterios de validación del modelo en los apartados siguientes.

Tabla 11 Relación entre mantenimiento preventivo tradicional y mantenimiento con soporte predictivo

Aspecto	Mantenimiento preventivo tradicional	Mantenimiento con soporte predictivo
Criterio de intervención	Período de tiempo o kilometraje acumulado	Período de tiempo complementado con información predictiva
Flexibilidad del programa	Baja	Alta
Priorización de activos	Limitada	Basada en riesgo estimado de inoperatividad
Uso de datos históricos	Mínimo	Intensivo
Objetivo principal	Cumplimiento del plan de mantenimiento	Priorización preventiva orientada a la disponibilidad operativa
Rol del modelo computacional	No aplica	Soporte a la toma de decisiones

Fuente. Elaboración propia a partir de WHO (2011), Moubray (1997), ISO 17359:2018 y Taoufyq et al. (2025).

La Figura 12 ilustra conceptualmente la integración del modelo predictivo como soporte al mantenimiento preventivo, mostrando que ambos enfoques coexisten y se complementan sin que el segundo sustituya al primero, orientándose hacia una gestión preventiva más informada y alineada con la disponibilidad operativa de las ambulancias Tipo II.

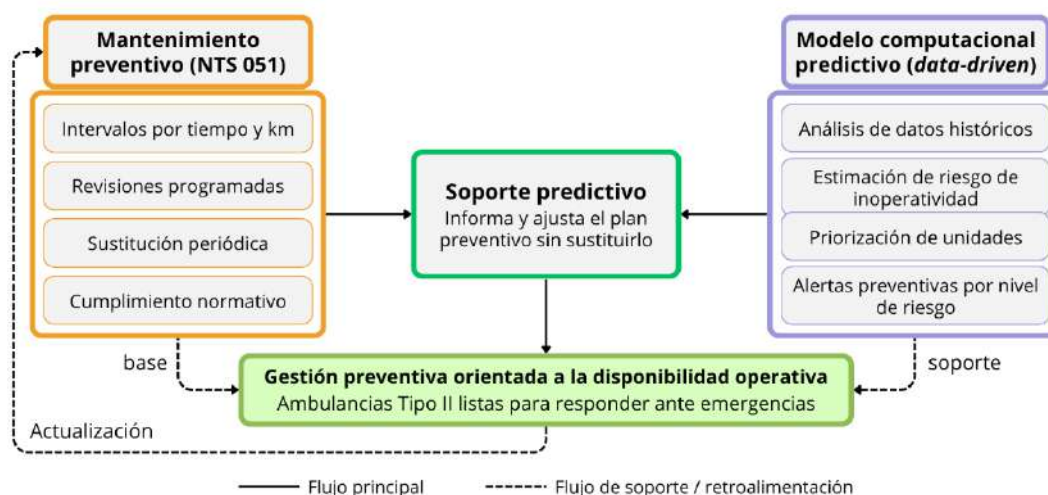


Figura 12. Integración del modelo computacional predictivo como soporte al mantenimiento preventivo de ambulancias médicas urbanas Tipo II.

Fuente. Elaboración propia a partir de Moubray (1997), ISO 17359:2018 y Taoufyq et al. (2025).

En síntesis, el soporte preventivo a partir de modelos predictivos permite fortalecer los programas tradicionales de mantenimiento sin alterar su estructura fundamental. Al incorporar información predictiva basada en datos históricos, se fortalece la capacidad de priorización preventiva, se orienta la planificación de intervenciones y se reduce la incertidumbre asociada a la gestión de activos críticos (Taoufyq et al., 2025). Este enfoque complementario es especialmente valioso para el sistema de ambulancias médicas urbanas Tipo II, donde la disponibilidad operativa constituye un factor determinante para la atención oportuna ante emergencias (WHO, 2011).

II.1.8. Validación de modelos predictivos en mantenimiento

II.1.8.1. Concepto de validación en modelos predictivos de mantenimiento

La validación constituye una etapa esencial en el desarrollo de modelos predictivos aplicados al mantenimiento, ya que permite evaluar el grado en

que las estimaciones generadas por el modelo representan adecuadamente el comportamiento real del sistema bajo análisis. En el ámbito del mantenimiento predictivo, la validación no persigue únicamente comprobar la exactitud numérica de una predicción. Su propósito central es determinar la utilidad operativa del modelo para apoyar la toma de decisiones técnicas y de gestión (ISO, 2018).

A diferencia de los enfoques deterministas tradicionales, los modelos predictivos basados en datos trabajan con información histórica y con fenómenos inherentemente variables, por lo que su evaluación requiere metodologías que consideren explícitamente la dimensión temporal y la incertidumbre asociada a los procesos de degradación. En este sentido, Si et al. (2011) y Taoufyq et al. (2025) coinciden en que la validación debe realizarse mediante procedimientos que preserven la coherencia cronológica de los datos y permitan contrastar las predicciones con eventos reales ocurridos en un período posterior al de entrenamiento. En la presente investigación, este principio se operacionaliza mediante una estrategia de backtesting temporal, en la que los registros del período 2024 constituyen el conjunto de entrenamiento y los datos de 2025 se utilizan como conjunto de validación.

II.1.8.2. Validación retrospectiva y enfoque de backtesting

La validación retrospectiva, técnicamente denominada backtesting, constituye un enfoque metodológicamente establecido para evaluar modelos predictivos cuando no es viable realizar pruebas en tiempo real o despliegues experimentales en campo. Este enfoque se basa en entrenar el modelo utilizando un subconjunto de datos históricos y evaluar su desempeño sobre un período posterior no utilizado durante el proceso de ajuste, respetando estrictamente el orden temporal de los registros (Si et al., 2011).

Taoufyq et al. (2025) y Si et al. (2011) señalaron que el backtesting permite reproducir de forma controlada las condiciones bajo las cuales el modelo habría operado en un escenario real, posibilitando la comparación entre las

predicciones generadas y los eventos de falla efectivamente observados. Este tipo de validación es metodológicamente apropiado en investigaciones aplicadas, especialmente en sistemas críticos donde la experimentación directa puede implicar riesgos operativos o restricciones presupuestales.

En el contexto de la gestión del mantenimiento, el enfoque retrospectivo es especialmente adecuado cuando los registros históricos de fallas, intervenciones y condiciones de operación constituyen una fuente de información sistemática y verificable para evaluar la capacidad del modelo de anticipar comportamientos futuros (Si et al., 2011). El backtesting no debe interpretarse como una validación incompleta, sino como una estrategia metodológica coherente con las condiciones reales de operación (Taoufyq et al., 2025; Si et al., 2011). En el caso de las ambulancias médicas urbanas Tipo II del SAMU en Lima Metropolitana, esta estrategia permite evaluar si el modelo computacional propuesto habría anticipado correctamente los eventos de inoperatividad observados en el conjunto de validación temporal, sin requerir intervención sobre el sistema en operación.

II.1.8.3. Validación temporal y coherencia cronológica de los datos

Un aspecto crítico en la validación de modelos predictivos es la preservación de la coherencia temporal de los datos utilizados. A diferencia de enfoques donde se aplican particiones aleatorias del conjunto de datos, en el mantenimiento predictivo la información presenta una dependencia temporal inherente, asociada a la evolución progresiva del estado de los activos (ISO, 2018).

Si et al. (2011) enfatizaron que la validación debe respetar la secuencia cronológica de los registros, utilizando datos históricos anteriores para el entrenamiento y datos posteriores para la evaluación del desempeño. Este principio evita el uso de información futura durante el proceso de validación y garantiza que los resultados reflejen un escenario realista de aplicación del modelo. En consecuencia, la coherencia temporal constituye un requisito metodológico indispensable para asegurar la credibilidad de los modelos predictivos en mantenimiento, principio que orienta el diseño de

validación adoptado en la presente investigación mediante la estrategia de backtesting temporal descrita en la sección de metodología.

II.1.8.4. Métricas de desempeño en la validación predictiva

La evaluación del desempeño de un modelo predictivo requiere el uso de métricas que permitan cuantificar su capacidad para anticipar eventos relevantes de mantenimiento. En el campo del mantenimiento predictivo, las métricas no se limitan a indicadores estadísticos abstractos, sino que deben vincularse con la utilidad práctica del modelo para la gestión de activos (Taoufyq et al., 2025).

Taoufyq et al. (2025) identificaron entre los criterios de evaluación más relevantes la capacidad del modelo para identificar correctamente las unidades con alta probabilidad de falla, la anticipación temporal con la que se generan las alertas y la reducción de intervenciones innecesarias respecto a enfoques tradicionales. En términos técnicos, estas capacidades se operacionalizan mediante métricas cuantitativas como la precisión, la sensibilidad, el error medio y la disponibilidad operativa proyectada, las cuales permiten evaluar tanto la exactitud del modelo como su contribución a la optimización de la gestión de activos. La Tabla 12 presenta la definición técnica e interpretación operativa de cada una de estas métricas, en el contexto de su aplicación al modelo computacional propuesto en la presente investigación.

A las métricas de clasificación ya descritas, se incorporan dos indicadores adicionales de especial relevancia para la evaluación completa del desempeño del modelo: la exactitud y la especificidad. La exactitud (*accuracy*) mide la proporción global de predicciones correctas sobre el total de observaciones, considerando tanto la clase positiva como la negativa. Aunque es una métrica intuitiva, su interpretación en *datasets* desbalanceados debe hacerse con cautela: podría alcanzar una exactitud aproximada de 79,2% en el conjunto de validación 2025 sin detectar ningún evento real de inoperatividad. Por este motivo, la exactitud se interpreta de forma complementaria con la sensibilidad y la especificidad, no como métrica

principal de desempeño. La especificidad (*specificity*) mide la capacidad del modelo para identificar correctamente las observaciones de la clase negativa, es decir, la proporción de períodos operativos que el modelo clasifica correctamente como tales. Esta métrica es relevante en el contexto del mantenimiento de ambulancias porque una baja especificidad implica un alto número de falsos positivos, lo que se traduce en intervenciones preventivas innecesarias.

El balance entre sensibilidad y especificidad es, por tanto, una variable de diseño operativo que debe ajustarse en función del costo relativo de cada tipo de error en el sistema prehospitalario. Se señala que las cuatro métricas principales: exactitud, especificidad, precisión y sensibilidad, son conceptualmente distintas entre sí y no deben confundirse. La Tabla 12 resume las diferencias clave:

Tabla 12 Métricas de desempeño para la validación del modelo computacional predictivo

Métrica	Definición técnica	Expresión matemática	Interpretación operativa
Precisión	Proporción de predicciones positivas que son correctas	$\text{Precisión} = \frac{VP}{(VP + FP)}$	Indica qué proporción de las alertas generadas corresponde a eventos reales de inoperatividad
Sensibilidad (Recall)	Proporción de fallas reales correctamente identificadas	$\text{Sensibilidad} = \frac{VP}{(VP + FN)}$	Indica qué proporción de los eventos reales de inoperatividad fue anticipada por el modelo
Exactitud	Proporción global de predicciones correctas sobre el total de observaciones	$\text{Exactitud} = \frac{(VP + VN)}{n}$	Mide predicciones correctas totales
Especificidad	Proporción de períodos operativos que el modelo clasifica correctamente	$\text{Especificidad} = \frac{VN}{(VN + FP)}$	Mide períodos operativos detectados
Error medio (MAE)	Promedio de las diferencias absolutas entre valores predichos y reales	$\text{MAE} = \frac{\sum y_i - \hat{y}_i }{n}$	Valores bajos indican menor diferencia promedio entre la probabilidad estimada y el valor real observado

Disponibilidad proyectada	Proporción de unidades clasificadas como operativas en el período de validación	$A_{\text{proj}} = \frac{\sum_{i=1}^n 1[\hat{y}_i = 0]}{n}$	Proporción de observaciones clasificadas como operativas en el horizonte de predicción
----------------------------------	---	---	--

Nota. VP = verdaderos positivos; FP = falsos positivos; FN = falsos negativos; y_i = valor real observado; $\hat{y}_i$ = clase predicha; $\hat{P}_i$ = probabilidad estimada de inoperatividad; n = número de observaciones.

Fuente. Elaboración propia a partir de Bishop (2006), International Organization for Standardization (2018) y Taoufyq et al. (2025).

II.1.8.5. Sustento metodológico de la validación retrospectiva

La ausencia de pruebas en tiempo real o de despliegues experimentales no invalida, por sí misma, un modelo predictivo desarrollado en el ámbito del mantenimiento. En estos casos, la evidencia obtenida a partir de registros históricos permite evaluar el comportamiento del modelo bajo condiciones operativas. Si et al. (2011) y Taoufyq et al. (2025) reconocieron que la validación retrospectiva constituye una práctica metodológicamente aceptada cuando existen restricciones técnicas, económicas o de seguridad que impiden la experimentación directa.

La ISO 17359:2018 establece que la evaluación de métodos de diagnóstico y pronóstico puede realizarse mediante el análisis de datos históricos, siempre que se garantice la trazabilidad de los resultados y la coherencia del proceso de validación (ISO, 2018). En coherencia con este marco normativo y científico, la validación retrospectiva se presenta como una alternativa metodológica sólida, especialmente adecuada para estudios aplicados orientados a la mejora de la gestión del mantenimiento en sistemas críticos como las ambulancias médicas urbanas Tipo II. La Tabla 13 sintetiza los principales enfoques de validación en modelos predictivos de mantenimiento y su aplicabilidad en distintos contextos.

Tabla 13 Enfoques de validación en modelos predictivos de mantenimiento

Enfoque de validación	Descripción	Aplicabilidad en contextos con restricciones
-----------------------	-------------	--

Validación retrospectiva (backtesting)	Evaluación del modelo sobre datos históricos posteriores al período de entrenamiento	Viable sin restricciones adicionales
Validación temporal	Separación de datos respetando estrictamente la secuencia cronológica	Viable sin restricciones adicionales
Validación en tiempo real	Evaluación del modelo durante la operación activa del sistema	Viable con restricciones operativas
Validación experimental en campo	Pruebas controladas en entornos reales con intervención directa	Requiere condiciones controladas y recursos adicionales

Fuente. Elaboración propia a partir de ISO 17359:2018, Si et al. (2011) y Taoufyq et al. (2025).

Entre estos enfoques, la validación retrospectiva representa la alternativa más adecuada cuando se dispone de registros históricos estructurados y no es posible implementar evaluaciones prospectivas durante la operación del sistema. Esta condición coincide con el contexto de la investigación, en el que el desempeño del modelo se evalúa con información registrada y cronológicamente diferenciada. Además, permite reproducir de manera controlada el escenario en el que el modelo habría generado predicciones antes de la ocurrencia de los eventos evaluados.

La Figura 13 ilustra esquemáticamente el proceso de validación retrospectiva adoptado en la presente investigación, mostrando la separación temporal entre el conjunto de entrenamiento correspondiente al período 2024 y el conjunto de validación del período 2025, así como el proceso de comparación entre las predicciones del modelo y los eventos de inoperatividad efectivamente registrados.

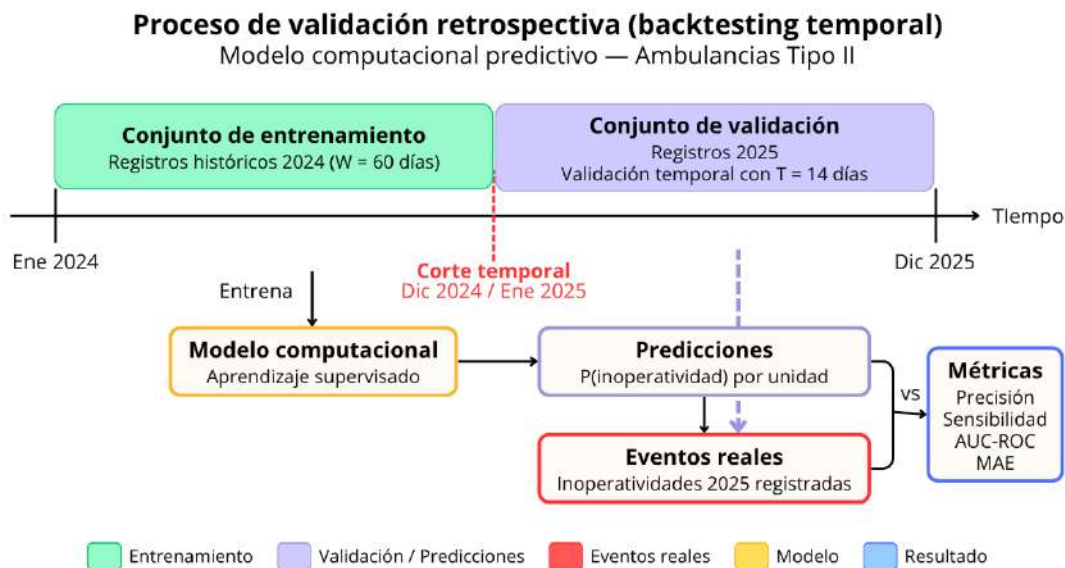


Figura 13. Esquema conceptual del proceso de validación retrospectiva (backtesting temporal).

Fuente. Elaboración propia a partir de Si et al. (2011) y Taoufyq et al. (2025).

En conjunto, la validación de modelos predictivos en mantenimiento debe entenderse como un proceso metodológico orientado a demostrar la consistencia, utilidad y coherencia temporal de las estimaciones generadas. El enfoque de validación retrospectiva, sustentado en el análisis de datos históricos y en la preservación de la secuencia cronológica, está documentado en la literatura especializada y respaldado por marcos normativos internacionales (Si et al., 2011; Taoufyq et al., 2025; ISO, 2018). En este sentido, la validación no solo mide el desempeño estadístico del modelo, sino también su pertinencia como herramienta de apoyo para decisiones técnicas en condiciones operativas. Este sustento teórico permite respaldar la validez del enfoque metodológico adoptado en la investigación, estableciendo bases para evaluar el modelo computacional propuesto.

II.1.9. Relación del marco teórico con la investigación

El marco teórico desarrollado establece una base conceptual coherente que permite comprender la problemática abordada desde una perspectiva sistémica y técnica. A partir del análisis de la atención prehospitalaria y del

rol crítico de las ambulancias médicas urbanas Tipo II, se establece un contexto en el que la disponibilidad operativa de estas unidades se reconoce como un elemento determinante para la capacidad de respuesta del sistema de salud. Esta comprensión inicial delimita el alcance de la investigación y orienta el análisis hacia la gestión técnica del mantenimiento como variable central para la continuidad operativa del servicio.

Asimismo, la revisión de los enfoques tradicionales de mantenimiento y de sus limitaciones operativas permite identificar al mantenimiento predictivo como una evolución metodológicamente fundamentada dentro de la gestión de activos, particularmente cuando se integra como soporte al mantenimiento preventivo. Desde esta perspectiva, el uso de modelos predictivos basados en datos históricos se presenta como un enfoque conceptualmente consistente con las restricciones operativas y presupuestales propias de los sistemas públicos de salud, evitando la dependencia de tecnologías de sensorización o monitoreo en tiempo real (Taoufyq et al., 2025; Prytz, 2014). El marco teórico establece así los límites conceptuales del enfoque de la investigación, sustentando la relación entre el modelo computacional propuesto y el fortalecimiento de la gestión del mantenimiento.

Finalmente, el análisis de los criterios de validación de modelos predictivos en mantenimiento proporciona el sustento teórico necesario para respaldar el enfoque metodológico adoptado en la investigación. El respaldo de la validación retrospectiva como práctica metodológicamente documentada, basada en la coherencia temporal de los datos y en la evaluación objetiva del desempeño predictivo, permite articular de manera consistente el marco teórico con la estrategia de análisis planteada. En conjunto, el marco teórico no solo fundamenta conceptualmente las variables de estudio, sino que orienta y respalda las decisiones metodológicas de la investigación, asegurando su coherencia interna y su alineación con los objetivos propuestos.

II.2. Antecedentes del estudio

II.2.1. Antecedente nacional 1

González Otárola (2020) desarrolló una investigación orientada a la implementación de un plan de mantenimiento preventivo con el propósito de mejorar la confiabilidad y disponibilidad de las ambulancias del Hospital Víctor Ramos Guardia de Huaraz. El estudio se enmarca en el contexto de la atención prehospitalaria, reconociendo que la inoperatividad de las ambulancias constituye un factor crítico que afecta la capacidad de respuesta del sistema de salud ante emergencias médicas.

La investigación adoptó un enfoque aplicado, sustentado en el análisis de indicadores clásicos de mantenimiento, tales como el tiempo promedio entre fallas (MTBF), el tiempo promedio de reparación (MTTR) y la disponibilidad operativa. Estos indicadores permitieron caracterizar el desempeño histórico de la flota y establecer una línea base para evaluar el efecto de las intervenciones de mantenimiento implementadas. A partir de un diagnóstico inicial del estado de la flota, se diseñó e implementó un programa de mantenimiento preventivo estructurado, orientado principalmente a reducir la ocurrencia de fallas no planificadas y a mejorar la continuidad operativa de las unidades.

Los resultados evidenciaron un incremento en los niveles de confiabilidad y disponibilidad de las ambulancias, así como una reducción en los tiempos de inoperatividad asociados a fallas técnicas. La autora concluye que la aplicación sistemática de actividades de mantenimiento preventivo permite fortalecer la gestión técnica de las ambulancias y contribuir a la sostenibilidad del servicio de atención prehospitalaria.

Este antecedente guarda estrecha relación con la presente investigación, dado que aborda directamente la problemática de la operatividad de ambulancias en un contexto hospitalario peruano y establece una relación

explícita entre la gestión del mantenimiento y la continuidad del servicio asistencial. No obstante, el estudio se limita a un enfoque preventivo tradicional, basado en intervenciones programadas, sin incorporar modelos predictivos ni análisis sistemático de datos históricos orientados a estimar el riesgo de inoperatividad futura, lo que justifica la exploración de enfoques complementarios como el propuesto en la presente tesis.

II.2.2. Antecedente nacional 2

Aguilar Sota (2024) realizó una investigación centrada en la aplicación de la gestión de mantenimiento basado en la confiabilidad (*Reliability-Centered Maintenance*, RCM), con el objetivo de incrementar la disponibilidad de la flota vehicular de una red de servicios de salud. El estudio se desarrolló considerando que la operatividad de los vehículos constituye un componente esencial para el apoyo a las actividades asistenciales y logísticas del sistema de salud.

La investigación empleó un enfoque aplicado, fundamentado en los principios del RCM, mediante la identificación de funciones, modos de falla y consecuencias asociadas a la indisponibilidad de las unidades vehiculares. A partir de este análisis, se establecieron estrategias de mantenimiento orientadas a priorizar los componentes críticos y fortalecer la disponibilidad de la flota, considerando el impacto operativo de cada falla sobre el servicio.

Los resultados del estudio evidenciaron que la aplicación del mantenimiento basado en la confiabilidad contribuye a mejorar la disponibilidad vehicular y a fortalecer la toma de decisiones en la gestión del mantenimiento, particularmente en contextos de recursos limitados. Se destaca que el enfoque RCM permite una asignación más eficiente de los recursos de mantenimiento, al focalizar las intervenciones según criticidad.

Este antecedente es pertinente para la presente investigación en tanto aborda la gestión del mantenimiento de flotas vehiculares vinculadas a servicios de salud y utiliza la confiabilidad como criterio central de análisis.

Sin embargo, el enfoque propuesto se basa principalmente en análisis funcionales y cualitativos, sin integrar modelos computacionales predictivos ni técnicas de análisis de datos históricos que permitan estimar el riesgo de inoperatividad futura de las unidades, aspecto que constituye el eje central de la presente investigación.

II.2.3. Antecedente internacional 1

Prytz (2014) desarrolló una investigación académica orientada a la aplicación de métodos de aprendizaje automático para el mantenimiento predictivo de vehículos, utilizando tanto datos generados a bordo del vehículo como registros históricos externos de mantenimiento. El estudio se desarrolló en el contexto de la gestión de flotas vehiculares, donde la anticipación de fallas técnicas constituye un elemento determinante para mejorar la confiabilidad operativa y reducir los tiempos de inactividad no planificados. Esta combinación de fuentes permitió analizar el comportamiento del activo desde una perspectiva operacional e histórica, coherente con los enfoques data-driven aplicados al mantenimiento predictivo.

El objetivo principal de la investigación fue evaluar la viabilidad de modelos computacionales basados en datos para la detección temprana y predicción de fallas en sistemas vehiculares complejos, superando las limitaciones de los enfoques tradicionales de mantenimiento preventivo basados exclusivamente en intervalos fijos. Para ello, se empleó una metodología cuantitativa sustentada en el análisis de datos históricos provenientes de flotas reales de vehículos, integrando información operativa, registros de mantenimiento y variables técnicas asociadas al desempeño de los sistemas. Este enfoque resulta pertinente para sistemas donde la información histórica se usa como insumo principal del modelado predictivo.

Los resultados del estudio evidenciaron que los modelos predictivos basados en datos históricos pueden contribuir a la detección temprana de patrones

asociados a fallas vehiculares, apoyando una planificación más informada de las intervenciones de mantenimiento y la gestión de la disponibilidad operativa de la flota (Prytz, 2014).

Este antecedente es pertinente para la presente investigación, ya que valida el uso de modelos computacionales predictivos aplicados a sistemas vehiculares reales, sustentados en el análisis de datos históricos y sin requerir necesariamente la incorporación de nueva instrumentación. No obstante, el estudio se enfoca en flotas vehiculares de uso general y no aborda las particularidades operativas, normativas y asistenciales propias de las ambulancias médicas, lo que justifica la necesidad de investigaciones específicas orientadas a este tipo de unidades críticas, como la propuesta en la presente investigación.

II.2.4. Antecedente internacional 2

Ersöz et al. (2022) desarrollaron una revisión sistemática de literatura orientada a mapear el estado del arte de las técnicas de mantenimiento predictivo aplicadas a sistemas de transporte, con énfasis en la industria automotriz y vehicular durante el período 2017-2022. El estudio fue publicado en la revista Sustainability del editor MDPI, indexada en Scopus y Web of Science, y se enmarca en la convergencia entre el avance de las tecnologías de red, el análisis de datos y la inteligencia artificial como habilitadores del mantenimiento predictivo en sistemas de transporte.

La investigación adoptó la metodología de revisión sistemática de literatura (Systematic Literature Review, SLR), aplicando criterios de inclusión y exclusión explícitos sobre cuatro bases de datos científicas: IEEE Xplore, ResearchGate, ScienceDirect y YokTez. A partir de cadenas de búsqueda específicas para los dominios automotriz, ferroviario, aéreo, marítimo y vehicular, se seleccionaron y analizaron en profundidad 52 estudios. La revisión caracterizó las tendencias de publicación, los enfoques algorítmicos predominantes, los tipos de datos utilizados como entrada y los

resultados esperados de las aplicaciones de mantenimiento predictivo en el sector.

Los resultados evidenciaron que los enfoques basados en aprendizaje automático, en particular los algoritmos supervisados como *Random Forest*, redes neuronales y máquinas de soporte vectorial, son los más frecuentemente empleados en aplicaciones de mantenimiento predictivo vehicular, dada su capacidad para identificar patrones de degradación a partir de datos históricos de operación y condición. Los autores documentaron que la disponibilidad de datos estructurados y la calidad de los registros históricos son los factores determinantes para el desempeño de estos modelos, e identificaron como área prioritaria de investigación futura la aplicación de mantenimiento predictivo a flotas vehiculares en contextos con restricciones tecnológicas y de instrumentación. Asimismo, resaltan la importancia de adaptar estos modelos a las características operativas propias de cada sistema de transporte.

Este antecedente es pertinente para la presente investigación en tanto sistematiza la evidencia disponible sobre el uso de modelos de aprendizaje automático para el mantenimiento predictivo en sistemas vehiculares, validando la viabilidad del enfoque *data-driven* adoptado en la presente tesis. Sin embargo, el estudio de Ersöz et al. (2022) tiene un alcance bibliométrico y de revisión, sin implementar ni validar un modelo computacional específico, y no aborda las particularidades operativas, normativas ni asistenciales propias de las ambulancias médicas urbanas.

La presente investigación extiende este enfoque hacia un contexto de alta criticidad sanitaria, proponiendo un modelo computacional concreto, evaluado cuantitativamente mediante validación retrospectiva, y orientado a un sistema de salud prehospitalaria donde la disponibilidad, calidad y estructuración de datos constituyen condiciones críticas para una futura implementación operativa.

II.3. Bases teóricas

Las bases teóricas de la presente investigación se articulan en torno a tres ejes teóricos que sustentan directamente el diseño y validación del modelo computacional propuesto: el mantenimiento predictivo basado en datos históricos, el modelado computacional mediante aprendizaje supervisado y la validación retrospectiva por *backtesting* temporal.

II.3.1. *Primer eje teórico: El mantenimiento predictivo data-driven como enfoque metodológico*

El mantenimiento predictivo constituye una evolución metodológicamente fundamentada de los enfoques tradicionales de gestión de activos, orientada a anticipar fallas mediante el análisis sistemático de información sobre el comportamiento del sistema (Molęda et al., 2023). En su modalidad *data-driven*, este enfoque extrae relaciones predictivas directamente de los datos históricos disponibles, sin requerir la formulación explícita de las leyes físicas que gobiernan el deterioro del activo (Montáns et al., 2019). Esta característica lo hace especialmente adecuado para sistemas técnicos complejos como las ambulancias médicas urbanas Tipo II, donde la interacción entre múltiples macro-subsistemas: vehicular, eléctrico/electrónico y biomédico-asistencial, dificulta la construcción de modelos físicos deterministas. En estos contextos, los modelos basados en datos permiten representar el comportamiento del sistema a partir de patrones observados durante su operación, sin depender de una caracterización físico-matemática completa de cada uno de sus componentes.

El uso de datos históricos de operación y mantenimiento como fuente principal del modelo puede reducir la necesidad de instrumentación adicional y permite aprovechar información que las instituciones podrían generar durante el curso normal de sus operaciones, siempre que esta sea registrada, estructurada y trazable (Prytz, 2014; Taoufyq et al., 2025). La calidad predictiva del modelo depende de tres condiciones fundamentales: la calidad mínima de los registros disponibles, la trazabilidad

temporal de los datos y la consistencia en la codificación de fallas e intervenciones (Taoufyq et al., 2025), condiciones que orientan los lineamientos técnicos propuestos en la presente investigación.

II.3.2. Segundo eje teórico: El modelo computacional de clasificación binaria supervisada

Un modelo computacional se define como una representación abstracta y formal de un sistema real, construida con el propósito de analizar, comprender y anticipar su comportamiento bajo determinadas condiciones (Montáns et al., 2019). En el contexto del mantenimiento predictivo, los modelos computacionales *data-driven* se fundamentan en el aprendizaje supervisado, enfoque metodológico que utiliza conjuntos de datos con variables de entrada y una variable objetivo conocida para identificar relaciones estadísticas aplicables a nuevas observaciones (Taoufyq et al., 2025). Este enfoque resulta adecuado cuando se dispone de registros estructurados que permiten vincular condiciones previas del activo con eventos posteriores de falla o inoperatividad.

En la presente investigación, el problema se formula como una clasificación binaria supervisada: dado el estado de una ambulancia Tipo II en un corte temporal t_0 , el modelo estima la probabilidad de que la unidad presente inoperatividad en el horizonte de 14 días siguientes. Esta formulación permite transformar el problema de mantenimiento en una tarea de clasificación supervisada, donde cada observación se asocia a una decisión predictiva basada en el comportamiento histórico del activo. Esta formulación es coherente con el enfoque adoptado por Prytz (2014) para la predicción de fallas en flotas vehiculares mediante datos históricos, y con los fundamentos del aprendizaje supervisado aplicado al mantenimiento predictivo descritos por Taoufyq et al. (2025). El algoritmo *Random Forest* fue considerado como modelo principal por su capacidad para capturar relaciones no lineales entre variables, estimar la importancia relativa de los predictores y ofrecer el mejor desempeño relativo en la validación retrospectiva realizada.

II.3.3. Tercer eje teórico: La validación retrospectiva como estrategia metodológica

La validación de un modelo predictivo constituye la etapa que determina su utilidad operativa para la toma de decisiones técnicas y de gestión, más allá de su exactitud numérica (ISO, 2018). Para modelos basados en datos históricos, la estrategia de validación más adecuada es la validación retrospectiva o backtesting temporal, que consiste en entrenar el modelo con un subconjunto de datos históricos y evaluar su desempeño sobre un período posterior no utilizado durante el entrenamiento, respetando estrictamente el orden cronológico de los registros (Si et al., 2011). Esta metodología permite aproximar las condiciones bajo las cuales el modelo habría sido utilizado en un escenario operativo real, utilizando únicamente la información que habría estado disponible en el momento de generar la predicción. De este modo, se evita la incorporación de información futura durante el entrenamiento y se obtiene una estimación más representativa de la capacidad de generalización del modelo.

Este enfoque reproduce de forma controlada las condiciones bajo las cuales el modelo operaría en un escenario real, permitiendo contrastar las predicciones generadas con los eventos de inoperatividad observados en el conjunto de validación temporal. La coherencia temporal de los datos, es decir, la preservación del orden cronológico en la partición entrenamiento-validación, constituye un requisito metodológico indispensable para garantizar la credibilidad de los resultados (Si et al., 2011; ISO, 2018). En la presente investigación, este principio se operacionaliza mediante la separación entre datos del período 2024 (entrenamiento) y datos del período 2025 (validación), conforme a los fundamentos del backtesting documentados por Taoufyq et al. (2025).

II.4. Definición de términos

Los términos técnicos empleados con mayor frecuencia en la presente investigación se definen a continuación con el propósito de garantizar la precisión conceptual del documento y facilitar su comprensión por parte de lectores de distintas disciplinas.

Ambulancia médica urbana Tipo II. Unidad de transporte asistido de pacientes que, además del equipamiento básico, incorpora monitor desfibrilador portátil y oxímetro de pulso portátil, y está destinada a la atención médica continua del paciente crítico durante el traslado en el ámbito urbano (MINSA, 2006).

Aprendizaje supervisado. Marco metodológico de la inteligencia artificial que utiliza conjuntos de datos con variables de entrada y una variable objetivo conocida para identificar relaciones estadísticas aplicables a la predicción sobre nuevos datos (Taoufyq et al., 2025).

AUC-ROC (*Area Under the Receiver Operating Characteristic Curve*). Métrica de evaluación de modelos de clasificación binaria que mide la capacidad discriminativa del modelo independientemente del umbral de clasificación. Valores cercanos a 1,0 indican alta capacidad discriminativa, mientras que valores cercanos a 0,5 indican un comportamiento equivalente al azar (Bishop, 2006).

Backtesting temporal. Estrategia de validación retrospectiva que consiste en entrenar un modelo con datos históricos de un período anterior y evaluar su desempeño sobre datos de un período posterior, respetando estrictamente la coherencia cronológica. Es el enfoque de validación adoptado en la presente investigación (Si et al., 2011).

Clasificación binaria. Tipo de problema de aprendizaje supervisado en el que la variable objetivo toma únicamente dos valores: positivo (1) o negativo (0). En esta investigación, la variable objetivo corresponde a la ocurrencia o no de inoperatividad en el horizonte de predicción $T = 14$ días (Bishop, 2006).

Disponibilidad operativa. Proporción del tiempo durante el cual una ambulancia se encuentra en condiciones funcionales de prestar el servicio de atención prehospitalaria para el que fue diseñada. Se expresa como la razón entre el tiempo operativo y el tiempo total del período analizado (WHO, 2011).

Downtime. Período durante el cual una ambulancia se encuentra fuera de servicio por causas técnicas, administrativas o de mantenimiento, sin capacidad de responder ante emergencias. En el contexto de mantenimiento, corresponde al tiempo de indisponibilidad del activo dentro del período analizado (Moubray, 1997; World Health Organization, 2011).

Gradient Boosting. Algoritmo de aprendizaje automático basado en el ensamble secuencial de árboles de decisión, donde cada árbol corrige progresivamente los errores residuales del modelo anterior mediante el gradiente de la función de pérdida (Friedman, 2001). En la presente investigación, se empleó como modelo comparativo de ensamble por su capacidad para capturar relaciones no lineales entre variables explicativas y eventos de inoperatividad.

Horizonte de predicción (T). Período futuro respecto al corte temporal de evaluación t_0 durante el cual se evalúa si ocurre un evento de inoperatividad. En la presente investigación $T = 14$ días (Si et al., 2011).

Inoperatividad. Condición de una ambulancia médica en la que la unidad no se encuentra disponible para prestar el servicio de atención prehospitalaria, independientemente de la causa que la origina. En esta investigación, se operacionaliza como la ocurrencia de al menos un evento de inoperatividad dentro del horizonte de predicción de 14 días, en coherencia con la necesidad normativa y operativa de mantener las ambulancias en condiciones funcionales para el transporte asistido de pacientes (Ministerio de Salud del Perú, 2006; World Health Organization, 2011).

Mantenimiento correctivo. Intervención técnica ejecutada tras la ocurrencia de una falla, orientada a restablecer la operatividad de la unidad (Moubray,

1997).

Mantenimiento predictivo. Enfoque de gestión del mantenimiento orientado a anticipar la ocurrencia de fallas o degradaciones funcionales mediante el análisis de datos históricos o de condición del activo, con el propósito de ejecutar intervenciones técnicas antes de que se produzca la inoperatividad (Molęda et al., 2023; Taoufyq et al., 2025).

Mantenimiento preventivo. Intervención técnica planificada y ejecutada a intervalos predeterminados de tiempo o uso, con el objetivo de reducir la probabilidad de falla y prolongar la vida útil del activo, independientemente de su estado real de degradación (Moubray, 1997).

Modelo computacional *data-driven*. Modelo que extrae relaciones predictivas directamente de los datos observados, sin requerir la formulación explícita de las leyes físicas que gobiernan el sistema. Su entrenamiento se basa en el análisis estadístico de información histórica u observacional estructurada (Montáns et al., 2019; Taoufyq et al., 2025).

MTBF (*Mean Time Between Failures*). Tiempo promedio entre fallas consecutivas de una unidad. Es un indicador de confiabilidad que refleja la frecuencia con que un activo presenta eventos de falla o inoperatividad durante su operación (Moubray, 1997).

MTTR (*Mean Time To Repair*). Tiempo promedio requerido para restablecer la operatividad de una unidad tras la ocurrencia de una falla. Es un indicador de mantenibilidad del sistema, dado que expresa el tiempo medio necesario para reparar el activo y devolverlo a condición operativa (Moubray, 1997).

Random Forest. Algoritmo de aprendizaje automático basado en el ensamble de múltiples árboles de decisión entrenados de forma independiente sobre submuestras del conjunto de datos mediante *bagging* o agregación *bootstrap*. La predicción final se obtiene por votación mayoritaria o por promedio de probabilidades, lo que reduce la varianza del modelo y mejora su capacidad de generalización frente a un árbol individual (Breiman,

2001). En la presente investigación, *Random Forest* fue seleccionado como modelo predictivo principal por su mejor desempeño relativo en la validación retrospectiva, su alta sensibilidad para anticipar eventos de inoperatividad y su utilidad para estimar la importancia relativa de las variables explicativas.

Sensibilidad (*Recall*). Proporción de observaciones positivas, es decir, con evento de inoperatividad, que el modelo identifica correctamente. Se calcula como $VP / (VP + FN)$, donde VP son los verdaderos positivos y FN los falsos negativos. En sistemas críticos como el prehospitalario, la sensibilidad es una métrica de alta relevancia operativa porque cuantifica la capacidad de detección de eventos críticos (Bishop, 2006).

Soporte preventivo. Uso del modelo computacional predictivo como herramienta complementaria al plan de mantenimiento preventivo establecido, proporcionando probabilidades estimadas de inoperatividad que permiten priorizar intervenciones sin sustituir el esquema preventivo normativo (Taoufyq et al., 2025; ISO, 2018).

Ventana histórica (W). Período de datos históricos previos al corte temporal de evaluación t_0 utilizado para calcular las variables explicativas del modelo. En la presente investigación, se definió $W = 60$ días como parámetro metodológico para resumir el comportamiento reciente de operación, mantenimiento e inoperatividad de cada ambulancia, en coherencia con el uso de registros históricos en modelos predictivos de mantenimiento (Prytz, 2014; Taoufyq et al., 2025).

CAPÍTULO III : HIPÓTESIS Y VARIABLES

III.1. Hipótesis

III.1.1. Hipótesis General

El desarrollo de un modelo computacional basado en datos para el mantenimiento predictivo con soporte preventivo permite estimar la probabilidad de inoperatividad y apoyar la gestión técnica de las ambulancias médicas urbanas Tipo II en Lima Metropolitana durante el período 2024-2025.

III.1.2. Hipótesis Específicas

- **H1.** La identificación de los factores técnicos y operativos que influyen en las fallas, eventos de inoperatividad y requerimientos de mantenimiento de las ambulancias médicas urbanas Tipo II permitirá determinar las variables críticas que deben incorporarse en el modelo predictivo.
- **H2.** La estructuración de una base de datos que integre variables operativas, técnicas, de mantenimiento e inoperatividad posibilitará el almacenamiento y análisis de la información necesaria para el entrenamiento del modelo computacional y la estimación predictiva del riesgo de inoperatividad de las ambulancias médicas urbanas Tipo II.

- **H3.** La implementación de un modelo computacional basado en datos, mediante algoritmos de clasificación supervisada, permitirá estimar la probabilidad de inoperatividad de las ambulancias médicas urbanas Tipo II en un horizonte de 14 días, integrando variables de uso operativo, mantenimiento, *downtime* y equipamiento funcional.
- **H4.** El modelo computacional propuesto presentará un desempeño evaluable mediante indicadores cuantitativos de exactitud, especificidad, precisión, sensibilidad, F1-Score, error medio absoluto (MAE), AUC-ROC y disponibilidad operativa proyectada, permitiendo comparar su capacidad de detección de eventos de inoperatividad frente al esquema de mantenimiento preventivo programado como línea base.
- **H5.** Los lineamientos técnicos y operativos derivados del modelo computacional permitirán apoyar la priorización preventiva, orientar los procedimientos de mantenimiento y fortalecer la toma de decisiones en la gestión técnica de las ambulancias médicas urbanas Tipo II.

III.2. Variables

III.2.1. Variable independiente

Modelo computacional basado en datos para el mantenimiento predictivo con soporte preventivo.

III.2.2. Variable dependiente

Inoperatividad de las ambulancias médicas urbanas Tipo II, estimada mediante la probabilidad de ocurrencia de un evento de inoperatividad en un horizonte temporal de 14 días. Como indicador complementario, se

considera la disponibilidad operativa proyectada, entendida como la proporción de observaciones clasificadas como operativas por el modelo.

III.2.3. Variables de control

- Período de análisis: 2024-2025.
- Registros históricos simulados sobre operación, mantenimiento, *downtime*, inoperatividad y funcionalidad del equipamiento de las unidades.
- Tipo de ambulancia: unidad médica urbana Tipo II, conforme a la NTS N.º 051 (MINSA, 2006).
- Ventana histórica de análisis: 60 días previos al punto de evaluación ($W = 60$ días).
- Horizonte de predicción: 14 días posteriores al punto de evaluación ($T = 14$ días).

III.3. Operacionalización de las variables

La operacionalización de las variables establece el puente entre las definiciones conceptuales declaradas en la Sección III.2 y los indicadores cuantitativos empleados en el modelo computacional para su medición y evaluación. De este modo, se asegura la correspondencia entre el planteamiento teórico de la investigación y su tratamiento metodológico. Además, facilita la trazabilidad entre las variables definidas y los resultados obtenidos.

La Tabla 14 presenta la operacionalización de las variables de la investigación. La articulación integral entre el problema, los objetivos, las hipótesis, las variables y la metodología se presenta en el **Anexo J**.

Tabla 14 Operacionalización de variables de la investigación

Variable	Definición conceptual	Definición operacional	Indicadores	Escala
Modelo computacional data-driven (Variable independiente)	Representación formal que estima el comportamiento futuro de un sistema a partir de datos históricos sin requerir formulación física explícita (Montáns et al., 2019)	Modelo computacional de clasificación binaria supervisada entrenado con 14 variables explicativas calculadas en $W = 60$ días, orientado a estimar la probabilidad de inoperatividad en $T = 14$ días. Se evaluaron Regresión Logística, Árbol de Decisión, Random Forest y Gradient Boosting, seleccionándose Random Forest por su mejor desempeño relativo.	Precisión, sensibilidad, especificidad, exactitud, F1-Score, AUC-ROC, MAE y disponibilidad operativa proyectada	Razón
Inoperatividad de ambulancias Tipo II (Variable dependiente)	Condición en la que una ambulancia no se encuentra disponible para prestar el servicio de atención prehospitalaria para el que fue diseñada, afectando la continuidad operativa del sistema (Ministerio de Salud del Perú, 2006; World Health Organization, 2011).	Variable binaria $y_{inoperativa_14}$, donde $Y = 1$ indica la ocurrencia de al menos un evento de inoperatividad en los 14 días posteriores al corte temporal t_0 , y $Y = 0$ indica ausencia de inoperatividad en dicho horizonte.	Probabilidad estimada de inoperatividad, clase predicha, eventos positivos $Y = 1$, disponibilidad operativa proyectada, <i>downtime</i>	Razón
Período de análisis (Variable de control)	Marco temporal que delimita el alcance de los registros históricos del estudio	Registros 2024-01-01 a 2025-12-31 con partición entrenamiento 2024 y validación 2025	Fecha de inicio y fin	Intervalo
Tipo de ambulancia (Variable de control)	Clasificación normativa que determina la configuración técnica y asistencial de la unidad (MINSA, 2006)	Unidades Tipo II con equipamiento completo según NTS 051: equipo básico, monitor desfibrilador y oxímetro de pulso	Cumplimiento NTS 051	Nominal
Ventana histórica W (Variable de control)	Período de datos históricos previos al corte t_0 utilizado para calcular las variables explicativas	$W = 60$ días previos al corte t_0	Días de historia por observación	Razón
Horizonte de predicción T (Variable de control)	Período futuro respecto al corte t_0 durante el cual se evalúa la ocurrencia de inoperatividad	$T = 14$ días posteriores al corte t_0	Días de anticipación	Razón

Fuente. Elaboración propia a partir de Montáns et al. (2019), World Health Organization (2011), Ministerio de Salud del Perú (2006), Prytz (2014), International Organization for Standardization (2018) y Taoufyq et al. (2025).

CAPÍTULO IV : MATERIALES Y MÉTODOS

IV.1. Área de estudio

El área de estudio de la presente investigación comprende el ámbito de cobertura operativa del Sistema de Atención Móvil de Urgencias (SAMU) en Lima Metropolitana, Perú. Geográficamente, esta área se circunscribe a la provincia de Lima y la Provincia Constitucional del Callao, delimitada entre las coordenadas 11°45' y 12°30' de latitud sur y 76°45' y 77°10' de longitud oeste, abarcando el espacio urbano y periurbano de Lima Metropolitana y el Callao.

El SAMU opera en este ámbito como el sistema de respuesta prehospitalaria de referencia del Ministerio de Salud del Perú, con ambulancias médicas destinadas a la atención prehospitalaria distribuidas estratégicamente en bases operativas a lo largo del área metropolitana (MINSA, 2025). Esta distribución territorial condiciona los patrones de uso, desplazamiento y disponibilidad operativa de las unidades. La delimitación geográfica responde a la relevancia sanitaria y operativa de Lima Metropolitana dentro del sistema de atención prehospitalaria, así como a la pertinencia de estudiar ambulancias médicas urbanas Tipo II en un entorno de alta demanda asistencial. No obstante, debido a que los registros históricos institucionales del SAMU no estuvieron disponibles durante el período de ejecución de la investigación, el análisis computacional se desarrolló sobre una flota simulada de 33 unidades, construida mediante parámetros de referencia a partir de la NTS N.º 051-MINSA y la literatura técnica.

El análisis se circunscribe al período comprendido entre el 1 de enero de

2024 y el 31 de diciembre de 2025, que corresponde al horizonte temporal de los registros históricos empleados para el entrenamiento y validación del modelo computacional propuesto.

IV.2. Diseño de investigación

La investigación se desarrolla bajo un enfoque **cuantitativo**, ya que se basa en el procesamiento y análisis de registros históricos simulados de operación y mantenimiento de ambulancias médicas urbanas Tipo II en Lima Metropolitana, correspondientes al período 2024-2025. Los datos se analizan mediante métricas objetivas de mantenimiento y disponibilidad, con el propósito de estimar el riesgo de inoperatividad futura.

Por su finalidad, el estudio es **aplicado**, debido a que propone un modelo computacional basado en datos que genere un resultado utilizable para la gestión del mantenimiento (priorización, alerta temprana y soporte preventivo). El alcance es **descriptivo-predictivo**, dado que primero caracteriza el comportamiento histórico de la operatividad y el mantenimiento y, posteriormente, estima la probabilidad de inoperatividad en un horizonte definido.

El diseño es **no experimental y longitudinal**, ya que no se manipulan variables de manera deliberada y se analizan series de eventos en el tiempo, estructuradas en ventanas históricas y horizontes de predicción.

IV.3. Población y muestra

IV.3.1. Ámbito del estudio

El estudio se circunscribe a ambulancias médicas urbanas Tipo II que operan en Lima Metropolitana, durante el período 2024-2025, considerando registros históricos de mantenimiento, eventos de indisponibilidad

(*downtime*) y uso operativo (por ejemplo, kilometraje o número de servicios prestados).

IV.3.2. Población

La población está conformada por las ambulancias urbanas Tipo II bajo análisis en el período de estudio. Conforme a la NTS N.º 051-MINSA/OGDN-V.01, una ambulancia Tipo II se caracteriza por contar con el equipamiento básico y adicionalmente con monitor desfibrilador portátil y oxímetro de pulso portátil (MINSA, 2006).

Para efectos del desarrollo computacional, la muestra de trabajo estuvo constituida por una flota simulada de 33 ambulancias médicas urbanas Tipo II, evaluadas semanalmente durante el período 2024-2025. Esta simulación generó 3 102 observaciones temporales mediante ventanas deslizantes, cada una correspondiente al estado de una ambulancia en un corte temporal específico.

Por tanto, la muestra no corresponde a registros institucionales reales del SAMU, sino a un escenario computacional reproducible diseñado para demostrar la viabilidad metodológica del modelo propuesto. Cada observación representa una instancia independiente para el entrenamiento, validación y evaluación del modelo computacional.

IV.3.3. Unidad de análisis

La unidad de análisis es la ambulancia Tipo II como sistema, entendida como una unidad técnica y asistencial integrada por los siguientes componentes:

- **Macro-subsistema vehicular:** está conformado por los componentes asociados al chasis, motor, transmisión, frenos, suspensión, neumáticos, carrocería, cabina de conducción y condiciones

mecánicas de operación de la unidad.

- **Macro-subsistema eléctrico/electrónico:** comprende el sistema eléctrico automotriz y asistencial, alimentación energética, iluminación, señalización, balizas, sirena, sistema de carga, GPS, intercomunicador, radiocomunicación y dispositivos electrónicos auxiliares.
- **Macro-subsistema biomédico-asistencial:** integra el equipamiento médico y los elementos directamente vinculados con la atención del paciente, incluyendo oxigenoterapia, aspiración, mecánica de gases, monitor desfibrilador, oxímetro de pulso, camilla y accesorios clínicos, cuya disponibilidad funcional forma parte de la condición operativa de la unidad (Ministerio de Salud del Perú, 2006; Central de Compras Públicas, 2024).
- **Gestión de mantenimiento:** Se refiere a la planificación y ejecución del mantenimiento, tanto de intervenciones programadas como de eventos correctivos no planificados, gestionando los tiempos de intervención y retorno a servicio.

Esta definición integral es relevante porque la inoperatividad de la unidad puede tener origen en cualquiera de estos subsistemas, y no exclusivamente en el equipamiento médico.

IV.4. Procedimientos, técnicas e instrumentos de recolección de información

IV.4.1. Fuentes de datos

La investigación emplea datos simulados estructurados de forma histórica con registros operativos y técnicos, definidos conforme a los criterios establecidos en la NTS N.º 051-MINSA/OGDN-V.01 (MINSA, 2006),

incluyendo como mínimo:

- **Eventos de mantenimiento:** preventivo o correctivo, planificado o no planificado, subsistema afectado, fecha de inicio y fin, observaciones, y registro de si el evento derivó en condición de inoperatividad.
- **Registro de inoperatividad (*downtime*):** períodos en los que la ambulancia estuvo fuera de servicio, con fecha de inicio, fin y duración.
- **Uso operativo:** indicadores como kilometraje recorrido, número de servicios o cualquier indicador disponible de intensidad de uso.
- **Verificación funcional del equipamiento obligatorio:** registro de disponibilidad funcional del equipamiento básico, monitor desfibrilador y oxímetro de pulso, entendido como control de condición operativa y no como variabilidad de inventario.

Dado que los registros históricos institucionales del SAMU no estuvieron disponibles durante el período de ejecución de la investigación, se construyó un dataset sintético mediante simulación paramétrica implementada en código Python propio. Esta estrategia, documentada en la literatura de mantenimiento predictivo para contextos con restricciones de acceso a datos reales (Taoufyq et al., 2025; Prytz, 2014), permite demostrar la viabilidad del marco computacional propuesto bajo condiciones controladas y reproducibles. El dataset no proviene de ningún repositorio público ni de registros institucionales externos; fue generado íntegramente mediante el script “generar_dataset_ambulancias.py”, disponible en el repositorio GitHub de la investigación (<https://github.com/DJOA5799/tesis-mantenimiento-predictivo-ambulancias>). Se emplearon principalmente las bibliotecas **numpy** y **pandas** para la generación y estructuración de datos, y **scikit-learn** para el entrenamiento y evaluación de los modelos, manteniendo una semilla de reproducibilidad fijada mediante **np.random.seed(42)**, lo que garantiza que cualquier investigador pueda reproducir el dataset exacto ejecutando el mismo código bajo las mismas condiciones computacionales.

La simulación utiliza un mecanismo paramétrico-estocástico reproducible, implementado en Python, con semilla fija `np.random.seed(42)`. La flota simulada estuvo conformada por 33 ambulancias Tipo II, en coherencia con la flota SAMU Lima Metropolitana reportada por la Defensoría del Pueblo. El mantenimiento preventivo fue modelado con una periodicidad semestral de 180 ± 15 días, conforme a la ficha de homologación de ambulancia urbana Tipo II prepublicada por PERÚ COMPRAS, que establece mantenimiento preventivo cada seis meses o según manual del fabricante.

La referencia de inoperatividad base fue de 21,2%, correspondiente a 7 unidades inoperativas de un total de 33 ambulancias reportadas. Los eventos técnicos fueron agrupados en tres macro-subsistemas analíticos: vehicular, eléctrico/electrónico y biomédico-asistencial. Esta agrupación no pretende sustituir la desagregación técnica completa de la ambulancia, sino organizar funcionalmente los componentes relevantes para el análisis de mantenimiento y disponibilidad operativa. La estructura completa del dataset, incluyendo el diccionario de las 14 variables explicativas y la variable objetivo, se presenta en el **Anexo A**. El código fuente completo del script de generación se documenta en el **Anexo B**.

La Figura 14 presenta la decodificación del identificador asignado a cada unidad simulada en el dataset.



Figura 14. Decodificación del identificador de ambulancia empleado en el dataset sintético.

Nota: El segmento T2 corresponde a la clasificación establecida en la NTS N.º 051-MINSA/OGDN-V.01 (MINSA, 2006).

Fuente. Elaboración propia.

IV.4.1.1. *Arquitectura de generación del dataset sintético*

La primera capa corresponde a la tabla de mantenimiento, que registra eventos preventivos y correctivos para cada ambulancia durante el período 2024-2025. Los intervalos entre mantenimientos preventivos se generan mediante una distribución Normal con media de 180 días y desviación estándar de 15 días, en coherencia con la periodicidad semestral de mantenimiento preventivo establecida en la ficha de homologación, prepublicada por PERÚ COMPRAS, de ambulancia urbana Tipo II que señala mantenimiento preventivo cada seis meses o según el manual del fabricante.

Esta capa permite representar la trazabilidad básica del mantenimiento de cada unidad, diferenciando intervenciones planificadas de eventos asociados a fallas o requerimientos no programados. En el contexto del modelo, dicha información resulta necesaria porque el historial de mantenimiento constituye una señal relevante del estado técnico acumulado de la ambulancia. Asimismo, la periodicidad semestral introduce una estructura temporal coherente con el esquema preventivo vigente, mientras que la variabilidad incorporada mediante la desviación estándar evita asumir un comportamiento estrictamente uniforme entre unidades. Esto permite calcular variables derivadas consistentes para el modelo predictivo.

Los eventos correctivos se generan mediante una distribución de Poisson cuya tasa depende de un factor de riesgo individual asignado a cada ambulancia mediante una distribución Uniforme $U(0,70; 1,50)$, lo que introduce heterogeneidad operativa entre unidades. El macro-subsistema afectado por cada evento se asigna mediante selección probabilística ponderada entre las categorías vehicular, eléctrico/electrónico y biomédico-asistencial. Esta ponderación fue definida como un supuesto paramétrico de simulación para representar la mayor frecuencia relativa esperada de eventos vehiculares frente a eventos eléctricos/electrónicos y biomédico-asistenciales, manteniendo coherencia funcional con la configuración técnica de la ambulancia Tipo II. La Figura 15 ilustra una muestra de las primeras filas generadas por este proceso.

1	id_ambulancia	fecha_inicio	fecha_fin	tipo_mantenimiento	subsistema	deriva_inoperatividad	duracion_dias	observaciones
2	AMB-T2-001	2024-01-07	2024-01-08	preventivo	vehicular	False	1.7	Inspección del sistema de refrigeración del motor
3	AMB-T2-001	2024-02-03	2024-02-07	correctivo	electrico_electronico	True	4.1	Reemplazo de fusibles por cortocircuito
4	AMB-T2-001	2024-02-12	2024-02-16	correctivo	biomedico_asistencial	False	4.7	Reparación de oxímetro de pulso portátil
5	AMB-T2-001	2024-05-10	2024-05-12	correctivo	vehicular	True	2.9	Cambio de correa de distribución por desgaste
6	AMB-T2-001	2024-06-06	2024-06-08	correctivo	electrico_electronico	True	2.4	Reparación del sistema de balizas y sirena
7	AMB-T2-001	2024-07-02	2024-07-03	preventivo	vehicular	False	1.5	Revisión general del motor y transmisión
8	AMB-T2-001	2024-10-23	2024-10-24	correctivo	vehicular	False	1.8	Corrección de falla en sistema de dirección
9	AMB-T2-001	2024-10-23	2024-10-24	correctivo	vehicular	False	1.8	Corrección de falla en sistema de dirección
10	AMB-T2-001	2024-10-23	2024-10-25	correctivo	biomedico_asistencial	True	2.3	Reemplazo de monitor desfibrilador por falla técnica
11	AMB-T2-001	2024-10-31	2024-10-31	correctivo	biomedico_asistencial	True	0.5	Reemplazo de componente de equipamiento básico
12	AMB-T2-001	2024-11-28	2024-11-28	correctivo	biomedico_asistencial	True	2.7	Reparación de oxímetro de pulso portátil
13	AMB-T2-001	2024-12-01	2024-12-04	correctivo	biomedico_asistencial	True	3.2	Corrección de falla en equipo de monitorización
14	AMB-T2-001	2024-12-09	2024-12-10	correctivo	electrico_electronico	True	1.9	Reparación del sistema de balizas y sirena
15	AMB-T2-001	2025-01-13	2025-01-14	preventivo	biomedico_asistencial	False	1.3	Verificación funcional del monitor desfibrilador
16	AMB-T2-001	2025-03-14	2025-03-18	correctivo	vehicular	True	4.0	Reparación de fuga en sistema de frenos
17	AMB-T2-001	2025-04-19	2025-04-20	correctivo	vehicular	False	1.7	Reemplazo de batería vehicular por falla
18	AMB-T2-001	2025-04-22	2025-04-24	correctivo	vehicular	True	2.4	Reparación de fuga en sistema de frenos

Figura 15. Muestra de la tabla de mantenimiento generada por el script `generar_dataset_ambulancias.py`.

Nota. Se presentan las primeras filas de cada tabla fuente como ilustración de la estructura de datos generada. Los valores corresponden a la ejecución del script con semilla `np.random.seed(42)`.

Fuente. Elaboración propia.

La segunda capa corresponde a la tabla de inoperatividad (*downtime*). Esta capa permite representar los períodos en los que la ambulancia no se encuentra disponible para la atención prehospitalaria. Su incorporación es fundamental porque la variable de inoperatividad constituye el evento objetivo que el modelo busca anticipar. Además, diferencia los episodios asociados a fallas correctivas de aquellos que pueden surgir por condiciones no programadas.

Los episodios de inoperatividad se generan a partir de dos fuentes: eventos derivados de los registros correctivos (aquellos marcados como generadores de *downtime* mediante un proceso de Bernoulli con probabilidad $0,65 \times \text{factor_riesgo}$) y episodios espontáneos adicionales por ambulancia, generados mediante una distribución de Poisson. La duración de los episodios derivados de correctivos sigue una distribución Normal $N(3,5; 2,0)$ días, mientras que los episodios espontáneos tienen duraciones distribuidas según una función Exponencial con media 4,0 días. La Figura 16 ilustra una muestra de las primeras filas de esta tabla.

1	id_ambulancia	fecha_inicio_downtime	fecha_fin_downtime	duracion_dias	causa
2	AMB-T2-001	2024-02-03	2024-02-04	1.4	falla_electrica
3	AMB-T2-001	2024-05-10	2024-05-14	4.1	falla_mecanica
4	AMB-T2-001	2024-06-06	2024-06-06	0.5	falla_electrica
5	AMB-T2-001	2024-10-23	2024-10-25	2.3	falla_equipamiento
6	AMB-T2-001	2024-10-31	2024-11-01	1.5	falla_equipamiento
7	AMB-T2-001	2024-11-26	2024-11-26	0.7	falla_equipamiento
8	AMB-T2-001	2024-12-01	2024-12-02	1.5	falla_equipamiento
9	AMB-T2-001	2024-12-09	2024-12-13	4.7	falla_electrica
10	AMB-T2-001	2025-03-14	2025-03-15	1.5	falla_mecanica
11	AMB-T2-001	2025-04-22	2025-04-23	1.5	falla_mecanica
12	AMB-T2-001	2025-10-06	2025-10-06	0.5	falla_equipamiento
13	AMB-T2-001	2025-10-22	2025-10-28	6.0	falla_electrica
14	AMB-T2-001	2025-11-04	2025-11-08	4.4	falla_equipamiento
15	AMB-T2-001	2025-12-07	2025-12-07	0.5	falla_mecanica
16	AMB-T2-002	2024-09-10	2024-09-11	1.5	falla_equipamiento
17	AMB-T2-002	2024-12-02	2024-12-06	4.6	falla_equipamiento
18	AMB-T2-002	2025-01-24	2025-01-30	6.1	falla_electrica

Figura 16. Muestra de la tabla de inoperatividad (*downtime*) generada por el script `generar_dataset_ambulancias.py`.

Nota. Se presentan las primeras filas de cada tabla fuente como ilustración de la estructura de datos generada. Los valores corresponden a la ejecución del script con semilla `np.random.seed(42)`.

Fuente. Elaboración propia.

La tercera capa corresponde a la tabla de uso operativo, que contiene registros mensuales de kilometraje y número de servicios prestados para cada ambulancia. Esta capa permite incorporar la intensidad de uso como dimensión operativa del modelo, dado que una mayor carga de trabajo puede asociarse con mayor desgaste acumulado y mayor probabilidad de requerimientos de mantenimiento. Además, el factor de uso individual permite representar diferencias entre ambulancias sometidas a distintos patrones de operación dentro de la flota.

El kilometraje mensual se genera mediante una distribución Normal con media 2 800 km y desviación estándar 600 km, y los servicios prestados mediante una distribución Normal con media 45 servicios y desviación estándar 12. Ambas variables se escalan por un factor de uso individual por

ambulancia ($U(0,70; 1,40)$), manteniendo la heterogeneidad de carga operativa entre unidades. La Figura 17 ilustra una muestra de las primeras filas correspondientes.

1	id_ambulancia	fecha	kilometraje_periodo	numero_servicios	km_acumulado
2	AMB-T2-001	2024-01-01	3157.0	75	80727.0
3	AMB-T2-001	2024-02-01	3909.0	54	84637.0
4	AMB-T2-001	2024-03-01	3461.0	44	88097.0
5	AMB-T2-001	2024-04-01	3502.0	58	91599.0
6	AMB-T2-001	2024-05-01	3187.0	60	94786.0
7	AMB-T2-001	2024-06-01	4107.0	58	98893.0
8	AMB-T2-001	2024-07-01	4221.0	51	103114.0
9	AMB-T2-001	2024-08-01	4425.0	62	107539.0
10	AMB-T2-001	2024-09-01	4751.0	90	112290.0
11	AMB-T2-001	2024-10-01	3541.0	56	115831.0
12	AMB-T2-001	2024-11-01	3676.0	68	119507.0
13	AMB-T2-001	2024-12-01	4619.0	53	124126.0
14	AMB-T2-001	2025-01-01	4052.0	59	128178.0
15	AMB-T2-001	2025-02-01	4194.0	79	132372.0
16	AMB-T2-001	2025-03-01	3679.0	41	136051.0
17	AMB-T2-001	2025-04-01	4270.0	51	140321.0
18	AMB-T2-001	2025-05-01	3738.0	70	144059.0

Figura 17. Muestra de la tabla de uso operativo generada por el script `generar_dataset_ambulancias.py`.

Nota. Se presentan las primeras filas de cada tabla fuente como ilustración de la estructura de datos generada. Los valores corresponden a la ejecución del script con semilla `np.random.seed(42)`.

Fuente. Elaboración propia.

A partir de estas tres tablas fuente, el proceso de ventanas temporales deslizantes genera para cada par (ambulancia $\times$ corte temporal t_0) una observación independiente: las 14 variables explicativas se calculan por agregación sobre el intervalo histórico $[t_0 - 60, t_0]$, y la variable objetivo se etiqueta según la ocurrencia de inoperatividad en el intervalo futuro $(t_0, t_0 + 14]$. Este diseño garantiza la separación temporal entre información histórica y etiqueta de predicción, eliminando el riesgo de incorporación de información futura (*data leakage*) en las variables explicativas.

IV.4.1.2. Mecanismo matemático de generación del riesgo de inoperatividad

Con el propósito de evitar que la variable objetivo fuera generada de manera independiente de las variables explicativas, el simulador incorporó un mecanismo generador jerárquico orientado a representar una relación técnicamente plausible entre uso operativo, mantenimiento, eventos correctivos e inoperatividad futura. Esta decisión metodológica responde al principio central del mantenimiento predictivo basado en datos: las variables históricas de uso, mantenimiento e inoperatividad deben contener información predictiva sobre el riesgo futuro del activo (Prytz, 2014; Taoufyq et al., 2025).

En primer lugar, el simulador asigna a cada ambulancia dos factores latentes: un factor de riesgo técnico y un factor de uso operativo. El factor de riesgo técnico representa la heterogeneidad no observada de la condición técnica de cada unidad. El factor de uso operativo representa diferencias en la intensidad de operación entre ambulancias. Este último no ingresa directamente en la función de riesgo de eventos correctivos, sino que se utiliza para generar las variables observables de carga operativa: kilometraje y servicios prestados. Posteriormente, estas variables observables sí ingresan en la función de riesgo de correctivos.

La generación del kilometraje mensual del período se expresa como:

$$km_t \sim \text{máx}[500, N(km_{\text{ref}} \times f_{\text{uso}}, \sigma_{km}^2)] \quad (32)$$

Donde:

- km_t : kilometraje generado para la ambulancia en el período.
- km_{ref} : kilometraje mensual de referencia.
- f_{uso} : factor latente de uso operativo de la ambulancia.
- σ_{km}^2 : varianza asociada al kilometraje mensual.
- $N()$: distribución normal.
- $\text{máx}[500, \dots]$: restricción para evitar valores operativamente irreales.

De manera equivalente, la generación del número mensual de servicios prestados se expresa como:

$$svc_t \sim \text{máx}[10, N(svc_{ref} \times f_{uso}, \sigma_{svc}^2)] \quad (33)$$

Donde:

- svc_t : número de servicios prestados en el período.
- svc_{ref} : número mensual de servicios de referencia.
- f_{uso} : factor latente de uso operativo.
- σ_{svc}^2 : varianza asociada al número mensual de servicios.
- $\text{máx}[10, \dots]$: restricción para evitar valores operativamente irreales.

Una vez generadas las variables de uso operativo, la tasa de eventos correctivos para cada ambulancia y período de exposición se define mediante una función de riesgo compuesta:

$$\lambda_{CM}(t) = \lambda_{base} \times f_{técnico} \times \left(\frac{km_t}{km_{ref}} \right)^\alpha \times \left(\frac{svc_t}{svc_{ref}} \right)^\beta \times e^{\left(\gamma \times \frac{d_{PM}}{T_{PM}} \right)} \times (1 + \delta \times n_{CM}) \quad (34)$$

Donde:

- $\lambda_{CM}(t)$: tasa esperada de eventos correctivos para la ambulancia en el período.
- λ_{base} : tasa base de eventos correctivos.
- $f_{técnico}$: factor latente de riesgo técnico de la ambulancia.
- km_t : kilometraje del período.
- km_{ref} : kilometraje de referencia.
- svc_t : servicios prestados en el período.
- svc_{ref} : servicios de referencia.
- α : elasticidad del riesgo respecto al kilometraje relativo.
- β : elasticidad del riesgo respecto a los servicios relativos.
- d_{PM} : días transcurridos desde el último mantenimiento preventivo.
- T_{PM} : intervalo preventivo de referencia.

- γ : coeficiente de incremento del riesgo por tiempo desde el último mantenimiento preventivo.
- n_{CM} : número de correctivos previos.
- δ : coeficiente de incremento del riesgo asociado al historial de correctivos.

A partir de esta tasa, el número de eventos correctivos se genera mediante una distribución de Poisson:

$$N_{CM} \sim \text{Poisson}[\lambda_{CM}(t)] \quad (35)$$

Donde:

- N_{CM} : número de eventos correctivos generados para la ambulancia en el período.
- $\lambda_{CM}(t)$: tasa esperada de eventos correctivos.

Esta formulación permite que las ambulancias con mayor carga operativa, mayor tiempo desde el último mantenimiento preventivo y mayor historial de correctivos presenten una mayor tasa esperada de eventos correctivos. La estructura multiplicativa de $\lambda_{CM}(t)$ se inspira en la lógica de los modelos de riesgos proporcionales, donde una tasa base es modulada por covariables asociadas al riesgo del sistema (Cox, 1972). En esta investigación, dicha formulación no se emplea como modelo inferencial de supervivencia, sino como mecanismo paramétrico de simulación para introducir una relación técnicamente plausible entre uso operativo, mantenimiento previo y ocurrencia futura de eventos técnicos. El incremento del riesgo con la carga de uso y el tiempo transcurrido desde la última intervención es coherente con los fundamentos de confiabilidad de activos físicos y con modelos de tasa de falla creciente asociados al desgaste operativo (Barlow & Proschan, 1965; Rausand & Høyland, 2004).

Los episodios de *downtime* se generan mediante dos mecanismos complementarios. El primero corresponde a *downtimes* derivados de eventos correctivos que afectan la operatividad de la unidad. En estos casos,

un evento correctivo puede generar inoperatividad de acuerdo con una probabilidad condicionada por el perfil de riesgo técnico de la ambulancia.

El segundo mecanismo corresponde a *downtimes* espontáneos, entendidos como episodios de inoperatividad no derivados directamente de una orden correctiva registrada. Para estos eventos se define una tasa específica:

$$\lambda_{ESP} = \text{PESO_ESPONTÁNEOS} \times 6 \times f_{\text{técnico}} \times f_{\text{uso}} \quad (36)$$

Donde:

- λ_{ESP} : tasa esperada de episodios espontáneos de *downtime*.
- PESO_ESPONTÁNEOS: peso relativo asignado a *downtimes* espontáneos.
- 6: factor de escala correspondiente al período bianual simulado.
- $f_{\text{técnico}}$: factor latente de riesgo técnico.
- f_{uso} : factor latente de uso operativo.

A partir de esta tasa, el número de episodios espontáneos de *downtime* se genera mediante:

$$N_{ESP} \sim \text{Poisson}(\lambda_{ESP}) \quad (37)$$

Donde:

- N_{ESP} : número de episodios espontáneos de *downtime*.
- λ_{ESP} : tasa esperada de episodios espontáneos.

Por tanto, PESO_ESPONTÁNEOS no forma parte de la función $\lambda_{CM}(t)$, sino del mecanismo complementario de generación de *downtimes* espontáneos. Esta diferenciación es relevante porque los eventos correctivos se generan mediante una tasa dependiente de variables operativas y de mantenimiento, mientras que los *downtimes* espontáneos representan inoperatividades adicionales condicionadas por el perfil técnico y operativo de cada ambulancia. Esta separación evita atribuir a una única función generadora todos los mecanismos de inoperatividad considerados en la simulación.

La Tabla 15 presenta los parámetros principales empleados en el mecanismo jerárquico de generación del riesgo de inoperatividad. La tabla diferencia los parámetros asociados a la generación de perfiles de ambulancia, carga operativa, eventos correctivos, *downtimes* espontáneos y construcción temporal del dataset, con el fin de evitar confundir los parámetros que forman parte de $\lambda_{CM}(t)$ con aquellos que pertenecen a otros componentes del simulador.

Tabla 15 Parámetros del mecanismo jerárquico de generación del riesgo de inoperatividad

Bloque del simulador	Parámetro	Valor	Interpretación	Sustento
Perfil técnico de la ambulancia	$f_{\text{técnico}}$	U(0,70; 1,50)	Factor latente de riesgo técnico asignado a cada ambulancia. Modula directamente la tasa de eventos correctivos $\lambda_{CM}(t)$ y la tasa de <i>downtimes</i> espontáneos λ_{ESP} .	Supuesto de simulación para representar heterogeneidad técnica entre unidades.
Perfil de uso operativo	f_{uso}	U(0,70; 1,40)	Factor latente de intensidad operativa. No ingresa directamente en $\lambda_{CM}(t)$, sino que se utiliza para generar km_t y svc_t .	Supuesto de simulación para representar diferencias de carga operativa entre unidades.
Generación de uso operativo	km_{ref}	2 800 km/mes	Kilometraje mensual de referencia empleado para generar km_t .	Supuesto paramétrico de simulación.
Generación de uso operativo	σ_{km}	600 km	Desviación estándar empleada para introducir variabilidad en el kilometraje mensual.	Supuesto paramétrico de simulación.
Generación de uso operativo	svc_{ref}	45 servicios/mes	Número mensual de servicios de referencia empleado para generar svc_t .	Supuesto paramétrico de simulación.
Generación de uso operativo	σ_{svc}	12 servicios	Desviación estándar empleada para introducir variabilidad en el número mensual de servicios.	Supuesto paramétrico de simulación.
Riesgo de correctivos	λ_{base}	1,2	Tasa base de eventos correctivos por ambulancia durante un período preventivo de referencia.	Supuesto de calibración para obtener una prevalencia positiva compatible con el análisis predictivo.
Riesgo de correctivos	α	0,6	Elasticidad de $\lambda_{CM}(t)$ respecto al kilometraje relativo acumulado.	Supuesto inspirado en el incremento del riesgo por carga de uso.
Riesgo de correctivos	β	0,5	Elasticidad de $\lambda_{CM}(t)$ respecto al número relativo de servicios prestados.	Supuesto análogo al kilometraje como indicador de intensidad operativa.

Riesgo de correctivos	d_{PM}	Variable	Días transcurridos desde el último mantenimiento preventivo.	Variable dinámica del simulador.
Riesgo de correctivos	T_{PM}	180 días	Intervalo preventivo de referencia utilizado para normalizar d_{PM} .	Ficha de homologación de ambulancia urbana Tipo II: mantenimiento preventivo semestral o según manual del fabricante.
Riesgo de correctivos	γ	0,8	Coefficiente de incremento del riesgo asociado al tiempo transcurrido desde el último mantenimiento preventivo.	Supuesto inspirado en modelos de tasa de falla creciente por degradación temporal.
Riesgo de correctivos	n_{CM}	Variable	Número de correctivos previos considerados como indicador de fragilidad operativa acumulada.	Variable dinámica del simulador.
Riesgo de correctivos	δ	0,25	Incremento relativo de la tasa de riesgo por correctivos previos acumulados.	Supuesto de fragilidad operativa acumulada.
Downtime espontáneo	PESO_ ESPONTÁNEOS	0,15	Peso relativo asignado a eventos de <i>downtime</i> no derivados directamente de correctivos registrados. Participa en λ_{ESP} , no en $\lambda_{CM}(t)$.	Supuesto operativo para representar inoperatividades no capturadas por una orden correctiva específica.
Construcción temporal del dataset	W	60 días	Ventana histórica utilizada para calcular las variables explicativas del modelo.	Decisión metodológica de la investigación.
Construcción temporal del dataset	T	14 días	Horizonte futuro utilizado para definir la variable objetivo de inoperatividad.	Decisión metodológica orientada a priorización preventiva de corto plazo.
Construcción temporal del dataset	Paso temporal	7 días	Frecuencia de generación de cortes temporales t_0 para construir observaciones.	Decisión metodológica para evaluación semanal del riesgo.

Nota. $U(a; b)$ = distribución uniforme entre a y b ; W = ventana histórica; T = horizonte de predicción; T_{PM} = intervalo preventivo de referencia; $\lambda_{CM}(t)$ = tasa de eventos correctivos; λ_{ESP} = tasa de *downtimes* espontáneos. Los parámetros α , β , γ , δ , λ_{base} y $PESO_ESPONTÁNEOS$ corresponden a supuestos paramétricos de simulación definidos para representar una relación técnicamente plausible entre carga operativa, degradación temporal, historial de correctivos e inoperatividad futura.

Fuente. Elaboración propia a partir de Cox (1972), Barlow y Proschan (1965), Rausand y Høyland (2004), Prytz (2014), Taoufyq et al. (2025), la ficha de homologación de ambulancia urbana Tipo II y los supuestos paramétricos definidos para la simulación.

Finalmente, esta formulación permitió que las variables explicativas calculadas en la ventana histórica $W = 60$ días estuvieran asociadas de manera plausible con la variable objetivo definida en el horizonte $T = 14$

días, sin introducir fuga de información temporal. Las variables del modelo se calcularon exclusivamente en el intervalo $(t_0 - 60, t_0]$, mientras que la etiqueta `y_inoperativa_14` se determinó únicamente a partir de eventos ocurridos en el intervalo $(t_0, t_0 + 14]$. Por tanto, la simulación conserva la separación cronológica entre información disponible al momento de predicción y eventos futuros evaluados durante la validación retrospectiva.

IV.4.2. *Materiales y herramientas de software*

Para la implementación del análisis y modelado se emplearon herramientas estándar de ciencia de datos, priorizando entornos reproducibles:

- **Lenguaje y entorno de análisis:** Python 3.12 (entorno *Jupyter* o *VS Code*).
- **Gestión y preparación de datos:** *pandas*, *numpy*.
- **Modelado y validación:** *scikit-learn*.
- **Visualización y análisis exploratorio:** *matplotlib* y/o herramientas equivalentes.
- **Gestión de versiones y reproducibilidad:** Git (repositorio local).
- **Almacenamiento y consolidación preliminar de datos:** hojas de cálculo (*Excel* o *Google Sheets*) únicamente como fuente de consolidación inicial.

La reproducibilidad integral del proceso de análisis está garantizada a través del repositorio digital de la investigación, alojado en GitHub (<https://github.com/DJOA5799/tesis-mantenimiento-predictivo-ambulancias>). El repositorio contiene cuatro scripts Python que implementan el flujo completo de la investigación:

- (1) **generar_dataset_ambulancias.py**, que ejecuta las Fases 1 y 2 y

- genera las tres tablas fuente y el dataset de observaciones;
- (2) **modelo_predictivo_ambulancias.py**, que implementa las Fases 3 y 4, incluyendo el entrenamiento, la validación retrospectiva y la generación de métricas y figuras;
 - (3) **lineamientos_tecnicos.py**, que implementa la Fase 5 y genera los lineamientos de priorización, el protocolo operativo y los criterios de alerta temprana; y
 - (4) **generar_plantilla_semanal.py**, que transforma las salidas del modelo predictivo y los lineamientos técnicos en una plantilla semanal editable de programación de mantenimiento.

La semilla de reproducibilidad `np.random.seed(42)` está fijada en los scripts de generación y modelamiento, lo que garantiza que la ejecución secuencial del flujo computacional reproduzca el dataset, los modelos y los resultados reportados en este documento.

IV.4.3. Parámetros temporales del modelo

Se establecieron los siguientes parámetros temporales para el estudio, seleccionados en función de los ciclos operativos típicos de flotas vehiculares de emergencia y de la necesidad de generar alertas con suficiente anticipación para programar intervenciones:

- **Ventana histórica ($W = 60$ días):** Intervalo retrospectivo utilizado para calcular las variables explicativas del sistema, que captura el historial reciente de operación y mantenimiento de cada unidad.
- **Horizonte de predicción ($T = 14$ días):** Intervalo prospectivo en el cual se evalúa si ocurre inoperatividad, seleccionado para permitir la programación anticipada de intervenciones preventivas.
- **Paso temporal semanal:** frecuencia de generación de cortes temporales t_0 , utilizada para construir observaciones sucesivas por ambulancia y simular un esquema operativo de revisión periódica del riesgo.

IV.4.4. **Variable objetivo (dependiente)**

La variable objetivo se define formalmente en la Ecuación 38, donde t_0 representa el corte temporal de evaluación y el horizonte de predicción es de 14 días.

$$y_{\text{inoperativa},14}(t_0) = \begin{cases} 1, & \text{si la ambulancia presenta inoperatividad en } (t_0, t_0 + 14) \\ 0, & \text{si no presenta inoperatividad en } (t_0, t_0 + 14) \end{cases} \quad (38)$$

La condición de inoperatividad se define de forma operativa y verificable mediante los registros de *downtime* o eventos técnicos que sacan de servicio a la unidad.

La variable objetivo toma valor 1 cuando, para una ambulancia identificada por '*id_ambulancia*' y un corte temporal determinado t_0 , se registra al menos un episodio de *downtime* en el intervalo $(t_0, t_0 + 14]$. Si en ese intervalo no se registra ningún episodio de inoperatividad, la variable toma valor 0. Esta definición implica que la etiqueta es dicotómica y está determinada únicamente por eventos futuros respecto al corte de evaluación. Es importante subrayar que existe una separación temporal estricta entre los datos utilizados para calcular las variables explicativas y los datos utilizados para construir la etiqueta. Las 14 variables explicativas se calculan exclusivamente sobre la ventana histórica $[t_0 - 60, t_0]$, utilizando únicamente información disponible en el momento del corte. La variable objetivo, en cambio, se determina a partir de eventos en el horizonte futuro $(t_0, t_0 + 14]$. Esta separación temporal garantiza la ausencia de fuga de información (*data leakage*): en ningún caso se incorpora información futura dentro del vector de características que el modelo utiliza para generar su predicción.

IV.4.5. **Variables explicativas (independientes)**

Para cada ambulancia y cada corte temporal t_0 , se calculan variables explicativas a partir de datos observados en el intervalo $(t_0 - 60, t_0)$

agrupadas en las siguientes categorías (Taoufyq et al., 2025; Prytz, 2014):

- **Historial de mantenimiento:** conteo de eventos preventivos o correctivos, recurrencias y tiempos desde la última intervención.
- **Indicadores de inoperatividad histórica:** *downtime* total en W, número de episodios de indisponibilidad y duración promedio.
- **Uso operativo:** kilometraje o número de servicios en W, como indicador de la carga de trabajo de la unidad.
- **Eventos por macro-subsistema:** distribución de eventos en los macro-subsistemas vehicular, eléctrico/electrónico y biomédico-asistencial.
- **Condición funcional del equipamiento obligatorio:** registros de disponibilidad funcional del equipamiento básico requerido por la NTS N.º 051, entendido como indicador de condición operativa y no como variabilidad de inventario.

IV.4.6. Valoración técnica complementaria mediante juicio de especialistas

De manera complementaria a la evaluación cuantitativa del modelo computacional, se realizó una valoración técnica mediante encuesta a especialistas con experiencia en el área orientada a examinar la pertinencia, coherencia y potencial aplicabilidad de los principales componentes operativos derivados de la propuesta de mantenimiento predictivo con soporte preventivo.

Se empleó una encuesta estructurada dirigida a 7 profesionales, seleccionados mediante criterio intencional en función de su formación y experiencia en áreas relacionadas con ingeniería biomédica, equipamiento médico, mantenimiento, gestión de activos, ambulancias médicas y transporte sanitario, ingeniería electromecánica, gestión de proyectos de

salud, adquisición de tecnología sanitaria o analítica de datos aplicada.

- El instrumento se organizó en cinco dimensiones de valoración:
- Pertinencia de las variables e información utilizada
- Coherencia del enfoque predictivo-preventivo
- Pertinencia de los parámetros temporales y criterios operativos
- Priorización y toma de decisiones
- Aplicabilidad del protocolo operativo

Adicionalmente, se incluyó una valoración global y preguntas abiertas destinadas a recoger observaciones y recomendaciones profesionales.

Los ítems cerrados fueron evaluados mediante una escala ordinal de cuatro categorías:

- 1 = no pertinente/no adecuado;
- 2 = poco pertinente/requiere cambios importantes;
- 3 = pertinente/adecuado con ajustes menores; y
- 4 = muy pertinente/adecuado.

Cuando correspondió, se permitió indicar que no se contaba con elementos suficientes para emitir una valoración.

Esta valoración tuvo carácter complementario y no se utilizó para establecer el desempeño predictivo del algoritmo, el cual fue evaluado mediante la validación retrospectiva y las métricas cuantitativas descritas previamente. Su propósito fue incorporar una apreciación profesional sobre la pertinencia técnica y aplicabilidad de los componentes propuestos para apoyar la gestión del mantenimiento de ambulancias médicas Tipo II.

IV.5. Análisis estadístico

El análisis estadístico de la investigación se desarrolló en cinco fases secuenciales. Cada fase corresponde a un componente específico del

proceso de modelado *data-driven* descrito en las bases teóricas, y su ejecución se realizó en entorno Python 3.12.10 con las bibliotecas declaradas en la Sección IV.4.2. La semilla de reproducibilidad (`np.random.seed(42)`) permite que los resultados sean replicables bajo las mismas condiciones computacionales. El código fuente completo que implementa las Fases 3 y 4 de la metodología se encuentra en el **Anexo C**. Esta organización por fases permitió una ejecución ordenada y reproducible de cada etapa metodológica.

Las fases son las siguientes:

Fase 1: Generación y preparación de datos simulados

La primera fase comprendió la generación, consolidación y preparación de los registros históricos simulados correspondientes a una flota de ambulancias Tipo II para el período 2024-2025. El objetivo fue construir un conjunto de datos sintético, reproducible y consistente que representara eventos de mantenimiento, inoperatividad y uso operativo.

La generación de los datos se realizó mediante un simulador paramétrico-estocástico compuesto por módulos independientes, cuyos parámetros fueron definidos a partir de referencias normativas e información documentada sobre la flota SAMU Lima Metropolitana. La Figura 18 presenta la arquitectura general del simulador implementado.

Las actividades desarrolladas en esta fase fueron las siguientes:

- Definición de los parámetros operativos y distribuciones probabilísticas utilizadas en la simulación.
- Generación de registros sintéticos de mantenimiento, inoperatividad y uso operativo para el período 2024 - 2025.
- Consolidación de las tablas fuente generadas por el simulador.
- Estandarización de identificadores de ambulancia.
- Normalización de fechas y horas.
- Verificación de reglas de consistencia temporal.

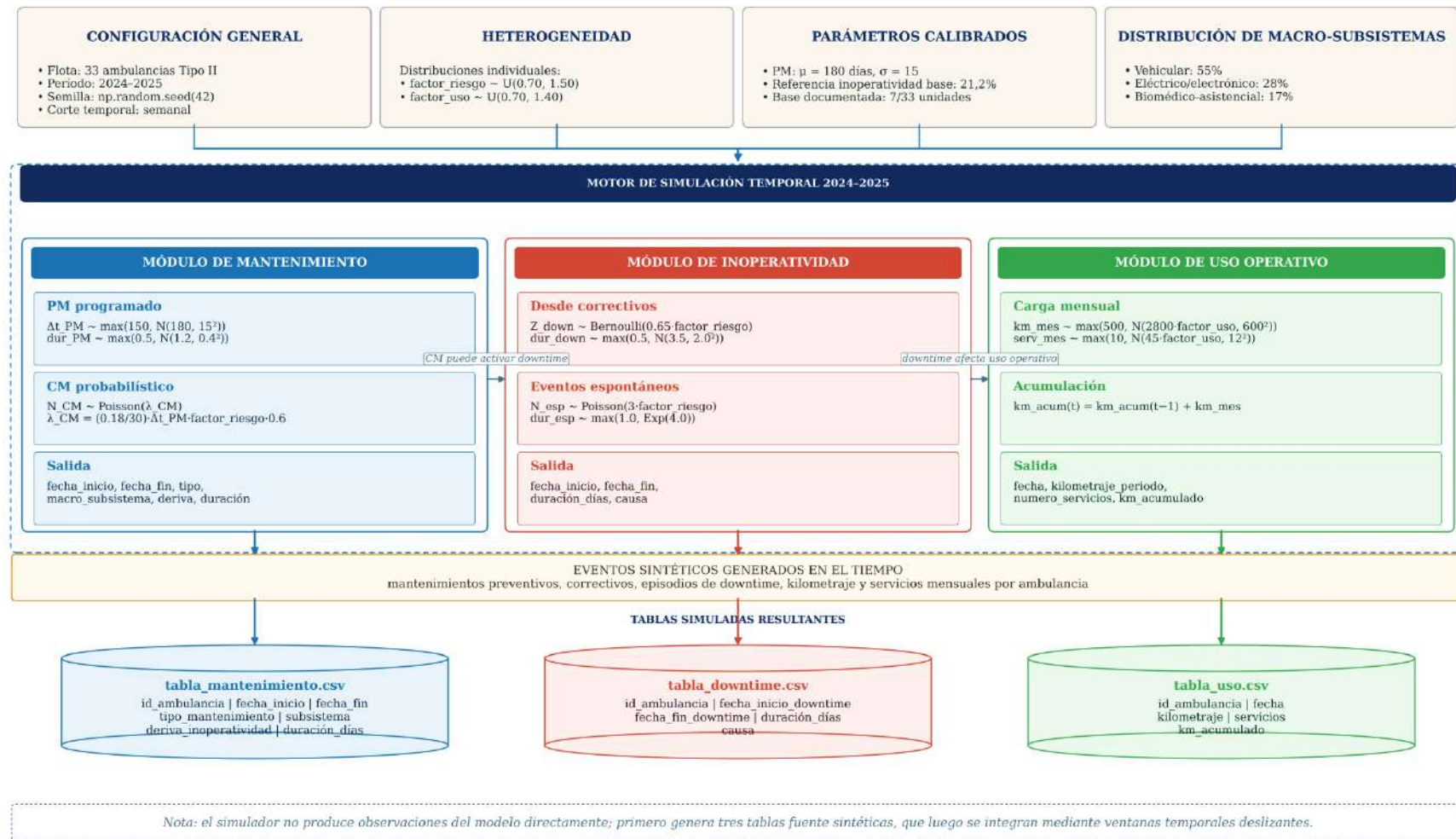


Figura 18. Arquitectura del simulador paramétrico-estocástico para la generación de tablas fuente sintéticas.

Fuente. Elaboración propia.

Posteriormente, las tablas fuente generadas fueron utilizadas para construir el dataset de observaciones empleado por los modelos supervisados, proceso descrito en la Fase 2.

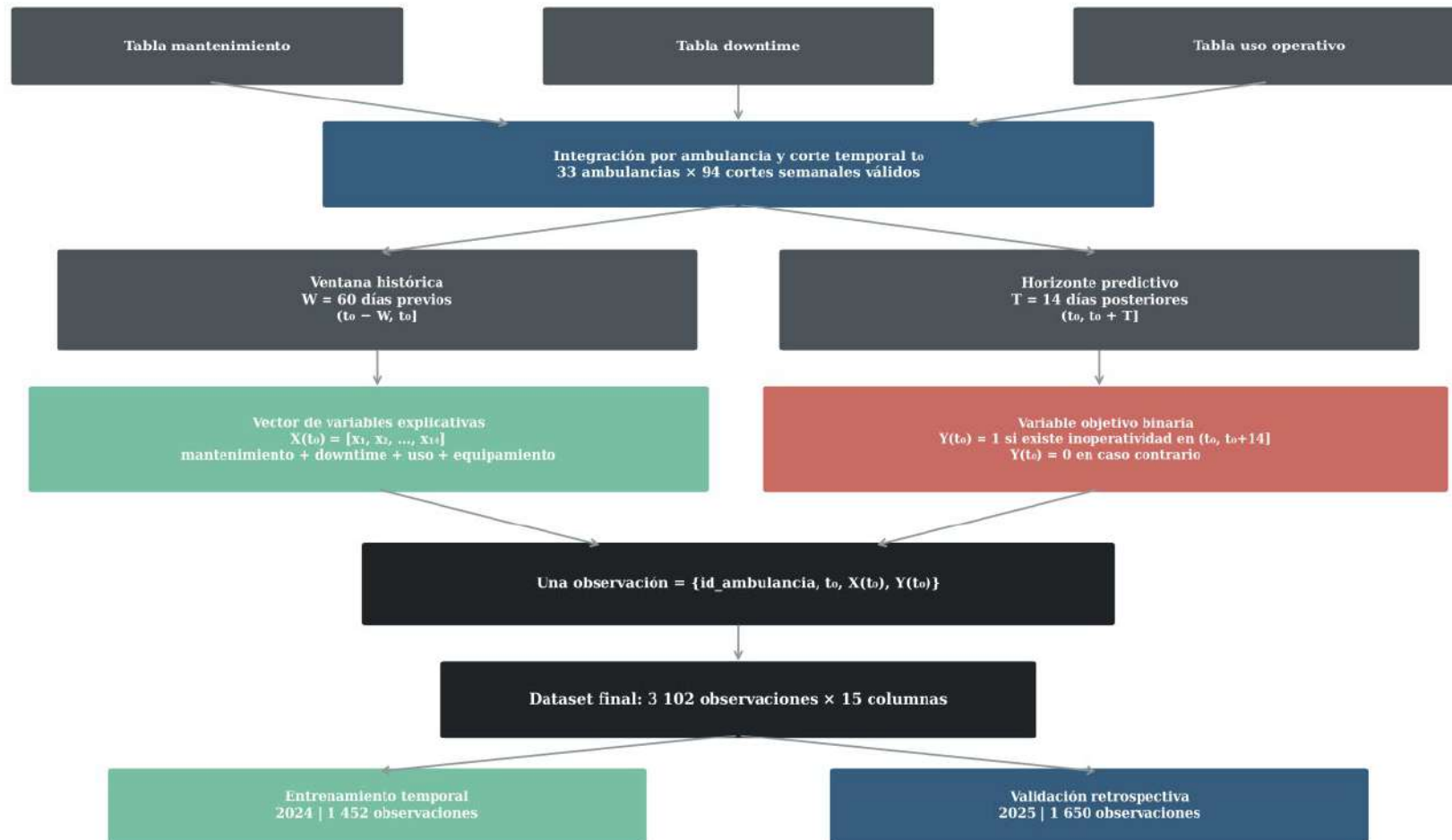
Fase 2: Construcción del dataset por ventanas temporales

La segunda fase transformó los registros históricos depurados en un *dataset* estructurado de observaciones temporales, aplicando el diseño de ventanas temporales deslizantes descrito en la Sección IV.4.3. Se reorganizó la información en una estructura compatible con algoritmos de aprendizaje supervisado. Así se trabajó con registros de mantenimiento, inoperatividad o uso operativo integrados en observaciones temporales comparables, asociadas a una ambulancia específica y a un punto de evaluación determinado. De este modo, cada observación representa el estado histórico de una unidad y su desenlace posterior dentro del horizonte de predicción.

Cada observación integra las variables explicativas calculadas en la ventana histórica W y la etiqueta de la variable objetivo en el horizonte T , constituyendo la unidad básica de aprendizaje del modelo. Los pasos ejecutados en esta fase fueron los siguientes:

- Definición de cortes temporales t_0 con periodicidad semanal para generar observaciones temporales por ambulancia.
- Cálculo de variables explicativas en la ventana histórica $W = 60$ días.
- Etiquetado de la variable objetivo en el horizonte de predicción $T = 14$ días.

La Figura 19 ilustra el diagrama de flujo del proceso de construcción del *dataset* mediante ventanas temporales deslizantes, desde las tres tablas fuente hasta la conformación del *dataset* final de 3 102 observaciones, conformado por 33 ambulancias evaluadas en 94 cortes temporales semanales válidos. El dataset final estuvo compuesto por 3 102 observaciones, de las cuales 1 452 correspondieron al entrenamiento temporal 2024 y 1 650 a la validación retrospectiva 2025.



Nota: cada fila del dataset representa una observación ambulancia-corte temporal; no representa una ambulancia distinta.

Figura 19. Construcción del dataset de observaciones mediante ventanas temporales deslizantes.
 Fuente. Elaboración propia.

Fase 3: Modelado computacional basado en datos

La tercera fase implementó el pipeline de entrenamiento de los cuatro modelos de clasificación supervisada descritos en el Capítulo II: Regresión Logística, Árbol de Decisión, *Random Forest* y *Gradient Boosting*. El proceso partió del *dataset* de entrenamiento del período 2024, aplicó la estrategia de balanceo de clases por *oversampling* aleatorio y ajustó los hiperparámetros de cada algoritmo según los criterios definidos en la Sección IV.4.2.

Esta fase permitió convertir el dataset estructurado en un conjunto de modelos predictivos comparables bajo una misma lógica de entrenamiento. La inclusión de algoritmos lineales, árboles individuales y métodos de ensamble permitió evaluar distintos niveles de complejidad e interpretabilidad. De este modo, la selección del modelo principal no dependió únicamente de una preferencia metodológica previa, sino del desempeño observado durante la validación temporal. Esto fortaleció la trazabilidad del proceso de selección del modelo final.

Las actividades desarrolladas fueron las siguientes:

- Formulación como problema de clasificación binaria supervisada.
- Entrenamiento del modelo en el subconjunto temporal de entrenamiento.
- Comparación de los modelos entrenados y selección del modelo principal en función de su desempeño relativo en la validación retrospectiva.

La Figura 20 presenta el diagrama de flujo del *pipeline* de entrenamiento de los modelos de clasificación supervisada, desde el dataset de entrenamiento del período 2024 hasta la selección del *Random Forest* como modelo predictivo principal. Si bien el *oversampling* igualó la distribución de clases en el dataset de entrenamiento, el parámetro *class_weight='balanced'* fue retenido en los modelos de Regresión Logística y *Random Forest* como medida adicional de robustez ante la posible reaparición del desbalance en subconjuntos internos durante el entrenamiento.

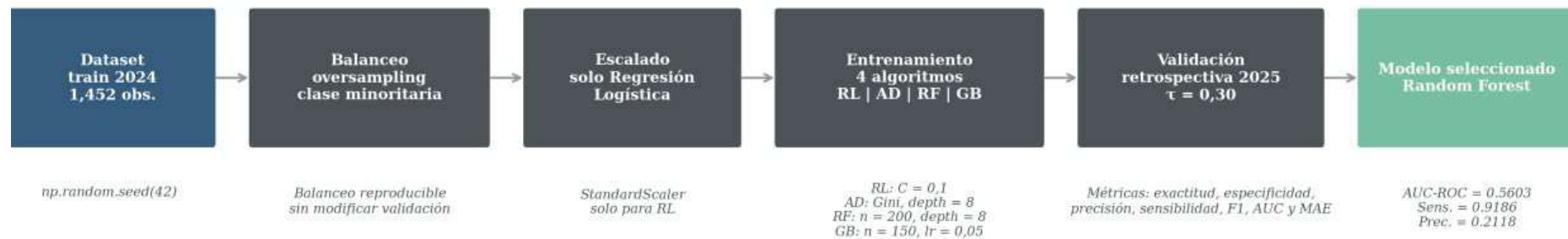


Figura 20. Diagrama de flujo del pipeline de entrenamiento de los modelos de clasificación supervisada.
 Fuente. Elaboración propia.

Los hiperparámetros finales se determinaron mediante exploración manual orientada por los rangos documentados en Breiman (2001) y Friedman (2001), priorizando profundidades moderadas, control del número mínimo de observaciones por hoja y ponderación de clases en los modelos de ensamble. Esta configuración buscó reducir el sobreajuste sobre el dataset simulado y favorecer la generalización durante la validación retrospectiva sobre el período 2025.

Fase 4: Validación retrospectiva (backtesting temporal)

La cuarta fase evaluó el desempeño de los modelos entrenados sobre el conjunto de validación del período 2025, estrictamente separado del conjunto de entrenamiento para garantizar la coherencia cronológica del *backtesting*. La evaluación se realizó sobre las métricas declaradas en el objetivo específico 4 de la investigación, con énfasis en la sensibilidad como indicador de alta relevancia operativa para el contexto prehospitalario, sin dejar de considerar la precisión, especificidad y disponibilidad proyectada para interpretar la carga potencial de falsos positivos. Las actividades comprendidas en esta fase fueron las siguientes:

- Separación temporal de los datos: entrenamiento en el período 2024 y evaluación en el período 2025.
- Evaluación mediante métricas de desempeño: exactitud, especificidad, precisión, sensibilidad, F1-Score, error medio absoluto (MAE), AUC-ROC y disponibilidad operativa proyectada.
- Análisis de errores y casos críticos, con énfasis en falsos negativos de alta relevancia operativa.

La validación se realizó mediante una partición temporal, utilizando el período 2024 para entrenamiento y el período 2025 para evaluación retrospectiva. Esta decisión buscó reproducir un escenario operativo real, en el cual el modelo se ajusta con observaciones de un período anterior y se evalúa sobre observaciones posteriores. En problemas con dependencia

temporal, la validación cruzada aleatoria puede resultar metodológicamente inadecuada, debido a que puede mezclar observaciones futuras con observaciones pasadas y generar estimaciones optimistas del desempeño. En este sentido, Cerqueira et al. (2020) señalan que, en escenarios reales de series temporales, los métodos fuera de muestra que preservan el orden temporal pueden producir estimaciones más adecuadas del desempeño predictivo. De manera similar, Bürkner et al. (2020) advierten que, cuando la tarea consiste en predecir el futuro a partir del pasado, ciertos esquemas de validación cruzada pueden permitir que información futura influya sobre predicciones pasadas.

Con el propósito de fortalecer el sustento matemático de la comparación entre los modelos evaluados, se incorporaron dos pruebas estadísticas inferenciales y una medida complementaria de desempeño bajo desbalance de clases, aplicadas sobre las predicciones generadas en la Fase 4 sobre el conjunto de validación 2025 ($n = 1\ 650$). Estos análisis no alteran el entrenamiento, los hiperparámetros, la semilla de reproducibilidad (`np.random.seed(42)`) ni el umbral de clasificación $\tau = 0,30$ empleados en las Fases 3 y 4; constituyen un procesamiento adicional de las predicciones ya generadas.

En primer lugar, se aplicó la prueba de McNemar (McNemar, 1947; Dietterich, 1998) para comparar, por pares de modelos, la proporción de observaciones del conjunto de validación clasificadas de forma discordante por *Random Forest* y cada uno de los demás enfoques evaluados. La selección de esta prueba responde a una condición estructural del diseño: *Random Forest* y los modelos de comparación se evalúan sobre las mismas 1 650 observaciones, lo que introduce una correlación entre sus resultados que invalida el uso de pruebas t convencionales para muestras independientes (Dietterich, 1998). Asimismo, dado que el diseño de validación retrospectiva adoptado corresponde a una partición temporal única (entrenamiento 2024 / validación 2025), sin repeticiones ni validación cruzada, se descartaron pruebas que requieren múltiples conjuntos de datos o particiones independientes, como el test de Friedman con post-hoc de

Nemenyi (Demšar, 2006), por no ser compatibles con el diseño adoptado.

En segundo lugar, se aplicó la prueba de DeLong (DeLong, DeLong & Clarke-Pearson, 1988), mediante el algoritmo de cálculo eficiente basado en midranks propuesto por Sun y Xu (2014), para comparar estadísticamente las áreas bajo la curva ROC (AUC-ROC) de pares de modelos correlacionados. Esta prueba permite establecer si las diferencias de AUC-ROC reportadas en la Tabla 16 son distinguibles de la variabilidad esperada bajo la hipótesis de igualdad de capacidad discriminativa entre dos clasificadores evaluados sobre el mismo conjunto de datos.

Finalmente, considerando que el AUC-ROC puede ofrecer una estimación optimista de la capacidad discriminativa de un clasificador cuando la clase positiva es minoritaria, debido a que la tasa de falsos positivos se calcula sobre una base muy amplia de verdaderos negativos (Saito & Rehmsmeier, 2015), se calculó la precisión media (*average precision*, AP) de cada modelo a partir de la curva *Precision-Recall*. Esta medida resume en un único valor el área bajo la curva *Precision-Recall* y resulta particularmente informativa en el contexto de la presente investigación, donde la prevalencia de la clase positiva ($y = 1$) es de 20,85% (344 de 1 650 observaciones), y donde la relación entre precisión y sensibilidad tiene una interpretación operativa directa: cada alerta generada por el modelo corresponde a una intervención preventiva, y cada evento detectado corresponde a una inoperatividad anticipada.

Estos tres análisis se implementaron como una extensión del script “modelo_predictivo_ambulancias.py”, documentada en el **Anexo E**, que reutiliza sin modificaciones las funciones de carga de datos, balanceo de clases, entrenamiento y evaluación descritas en esta sección IV.5 en las Fases 3 y 4.

Fase 5: Generación de salidas para soporte preventivo

La quinta fase tradujo las probabilidades de inoperatividad estimadas por el modelo *Random Forest* en salidas operativas accionables para el equipo de

mantenimiento. Esta fase corresponde al Entregable 2 de la investigación y su implementación se encuentra en el archivo “lineamientos_tecnicos.py”, complementado por el script “generar_plantilla_semanal.py” para la elaboración automática de la plantilla semanal de programación de mantenimiento. Las actividades realizadas fueron las siguientes:

- Transformación de probabilidades estimadas en niveles de riesgo operativo (bajo, medio y alto).
- Priorización de unidades para intervención preventiva basada en nivel de riesgo y criticidad del subsistema afectado.

Para la clasificación operativa de las unidades, las probabilidades estimadas se agruparon en tres niveles de riesgo:

- Alto cuando $P(\text{inoperatividad}) \geq 0,50$,
- Medio cuando $0,25 \leq P(\text{inoperatividad}) < 0,50$, y
- Bajo cuando $P(\text{inoperatividad}) < 0,25$.

Esta categorización permitió traducir la salida probabilística del modelo en criterios de priorización preventiva.

Como instrumento de operacionalización de dicho protocolo, se desarrolló adicionalmente una plantilla de programación semanal en formato Excel, generada automáticamente a partir de las salidas del modelo, que permite al responsable de mantenimiento tramitar las intervenciones con el taller autorizado de forma estructurada y trazable.

Complementariamente, los resultados de la valoración técnica por encuestas fueron analizadas mediante estadística descriptiva, considerando la distribución de frecuencias de las categorías de respuesta, la mediana de valoración por dimensión y proporción de valoraciones favorables, definida en conjunto por las categorías 3 (pertinente/adecuado con ajustes menores) y 4 (muy pertinente/adecuado). Las respuestas abiertas se revisaron de forma cualitativa para identificar recomendaciones técnicas recurrentes o consideraciones relevantes para una futura implementación institucional.

CAPÍTULO V : RESULTADOS

V.1. Presentación y análisis de los resultados

V.1.1. Descripción del dataset de observaciones

Antes de presentar las métricas del dataset, es necesario precisar la unidad de análisis. Cada una de las 3 102 observaciones no representa una ambulancia distinta, sino el estado observado de una ambulancia determinada en un corte temporal específico. Dado que el dataset se construye mediante ventanas temporales deslizantes con periodicidad semanal sobre una flota de 33 unidades durante dos años, cada ambulancia contribuye con múltiples observaciones a lo largo del período de análisis. En consecuencia, las clases ($Y = 0$ y $Y = 1$) no representan ambulancias únicas, sino observaciones clasificadas en función de si el horizonte de predicción de 14 días asociado a ese corte temporal incluye o no al menos un episodio de inoperatividad. La clase 0 indica que la unidad permaneció operativa durante los 14 días siguientes al corte; la clase 1 indica que presentó al menos un episodio de *downtime* en ese mismo horizonte. Esta precisión es importante porque evita interpretar la distribución de clases como una proporción directa de ambulancias operativas o inoperativas. En realidad, la clasificación refleja eventos temporales asociados a cada ventana de observación, por lo que una misma unidad puede aparecer como operativa en algunos cortes y como inoperativa en otros. Así, el dataset representa la evolución temporal del estado operativo de la flota. La Figura 21 ilustra la lógica de construcción del dataset mediante ventanas temporales deslizantes.

¿Por qué una ambulancia genera múltiples observaciones?

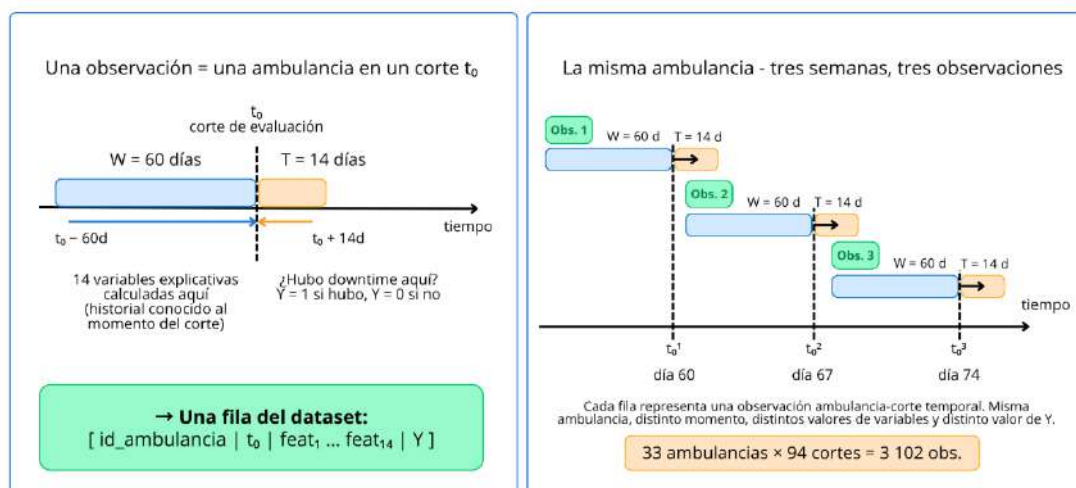


Figura 21. Lógica de construcción del dataset mediante ventanas temporales deslizantes.

Nota. Panel izquierdo: estructura de una observación individual a partir de la ventana histórica $W = 60$ días y el horizonte de predicción $T = 14$ días. Panel derecho: generación de observaciones temporales por deslizamiento semanal del corte t_0 sobre la misma ambulancia.

Fuente. Elaboración propia.

La investigación se sustentó en un dataset construido mediante ventanas temporales deslizantes aplicadas a tres tablas fuente: mantenimiento, inoperatividad y uso operativo, correspondientes a una flota simulada de 33 ambulancias médicas urbanas Tipo II durante 2024-2025. Sus parámetros fueron calibrados según la NTS N.º 051-MINSA/OGDN-V.01 (MINSA, 2006) y literatura técnica de referencia.

El *dataset* final estuvo conformado por 3 102 observaciones temporales generadas con cortes semanales, una ventana histórica de $W = 60$ días y un horizonte de predicción de $T = 14$ días. Cada observación integra 14 variables explicativas agrupadas en mantenimiento, inoperatividad, uso operativo y equipamiento obligatorio, además de la variable objetivo binaria. La Figura 22 presenta la estructura del dataset generado. Esta permite vincular el estado histórico de cada unidad con su desenlace operativo.

	id_ambulancia	t0	n_pm_w	n_cm_w	n_total_mant_w	dias_desde_ultima_interv	n_eventos_vehicular_w	n_eventos_electrico_w	n_eventos_equipamiento_w	n_episodios_downtime_w	downtime_total_dias_w	downtime_promedio_dias_w	disponibilidad_w	km_en_w	servicios_en_w	equipamiento_funcional	y_inoperativa_14
1	AMB-T2-001	2024-03-01	1	2	3	18	1	1	1	1	1.4	1.4	0.9767	7370.0	98	0	0
2	AMB-T2-001	2024-03-08	0	2	2	25	0	1	1	1	1.4	1.4	0.9767	7370.0	98	0	0
3	AMB-T2-001	2024-03-15	0	2	2	32	0	1	1	1	1.4	1.4	0.9767	7370.0	98	0	0
4	AMB-T2-001	2024-03-22	0	2	2	39	0	1	1	1	1.4	1.4	0.9767	7370.0	98	0	0
5	AMB-T2-001	2024-03-29	0	2	2	46	0	1	1	1	1.4	1.4	0.9767	7370.0	98	0	0
6	AMB-T2-001	2024-04-05	0	1	1	53	0	0	1	0	0.0	0.0	1.0	6963.0	102	0	0
7	AMB-T2-001	2024-04-12	0	0	0	60	0	0	0	0	0.0	0.0	1.0	6963.0	102	1	0
8	AMB-T2-001	2024-04-19	0	0	0	60	0	0	0	0	0.0	0.0	1.0	6963.0	102	1	0
9	AMB-T2-001	2024-04-26	0	0	0	60	0	0	0	0	0.0	0.0	1.0	6963.0	102	1	1
10	AMB-T2-001	2024-05-03	0	0	0	60	0	0	0	0	0.0	0.0	1.0	6689.0	118	1	1
11	AMB-T2-001	2024-05-10	0	1	1	0	1	0	0	1	4.1	4.1	0.9317	6689.0	118	1	0
12	AMB-T2-001	2024-05-17	0	1	1	7	1	0	0	1	4.1	4.1	0.9317	6689.0	118	1	0
13	AMB-T2-001	2024-05-24	0	1	1	14	1	0	0	1	4.1	4.1	0.9317	6689.0	118	1	1
14	AMB-T2-001	2024-05-31	0	1	1	21	1	0	0	1	4.1	4.1	0.9317	3187.0	60	1	1
15	AMB-T2-001	2024-06-07	0	2	2	1	1	1	0	2	4.6	2.3	0.9233	7294.0	118	1	0
16	AMB-T2-001	2024-06-14	0	2	2	8	1	1	0	2	4.6	2.3	0.9233	7294.0	118	1	0
17	AMB-T2-001	2024-06-21	0	2	2	15	1	1	0	2	4.6	2.3	0.9233	7294.0	118	1	0
18	AMB-T2-001	2024-06-28	0	2	2	22	1	1	0	2	4.6	2.3	0.9233	7294.0	118	1	0
19	AMB-T2-001	2024-07-05	1	2	3	3	2	1	0	2	4.6	2.3	0.9233	8328.0	109	1	0
20	AMB-T2-001	2024-07-12	1	1	2	10	1	1	0	1	0.5	0.5	0.9917	8328.0	109	1	0
21	AMB-T2-001	2024-07-19	1	1	2	17	1	1	0	1	0.5	0.5	0.9917	8328.0	109	1	0
22	AMB-T2-001	2024-07-26	1	1	2	24	1	1	0	1	0.5	0.5	0.9917	8328.0	109	1	0
23	AMB-T2-001	2024-08-02	1	1	2	31	1	1	0	1	0.5	0.5	0.9917	8646.0	113	1	0
24	AMB-T2-001	2024-08-09	1	0	1	38	1	0	0	0	0.0	0.0	1.0	8646.0	113	1	0
25	AMB-T2-001	2024-08-16	1	0	1	45	1	0	0	0	0.0	0.0	1.0	8646.0	113	1	0

Figura 22. Estructura del dataset de observaciones.

Nota. Primeras filas del archivo 'dataset_observaciones.csv' con las 14 variables explicativas agrupadas en cuatro categorías y la variable objetivo 'y_inoperativa_14'. n = 3 102 observaciones. Período 2024-2025.

Fuente. Elaboración propia.

La partición temporal para la validación retrospectiva (*backtesting*) dividió el *dataset* en dos subconjuntos sin superposición cronológica: el conjunto de entrenamiento, conformado por 1 452 observaciones correspondientes al período 2024, y el conjunto de validación, conformado por 1 650 observaciones del período 2025. Esta separación garantizó la coherencia temporal del proceso de validación, respetando estrictamente el orden cronológico de los registros conforme a los principios establecidos por Si et al. (2011). De este modo, el modelo fue entrenado con información histórica previa y evaluado sobre datos posteriores, reproduciendo la lógica de aplicación prospectiva del enfoque predictivo.

La variable objetivo **y_inoperativa_14** presentó una distribución desbalanceada. De las 3 102 observaciones del dataset completo, 2 515 corresponden a la clase 0, es decir, observaciones sin inoperatividad en el horizonte de predicción, y 587 corresponden a la clase 1, es decir, observaciones con al menos un evento de inoperatividad registrado en los 14 días posteriores al corte temporal. Esta distribución representa una prevalencia positiva total de 18,9%, evidenciando un desbalance moderado de clases. En el conjunto de entrenamiento 2024, la clase positiva representó el 16,7% de las observaciones, mientras que en el conjunto de validación 2025 representó el 20,8%, lo que confirma una mayor concentración relativa de eventos de inoperatividad en el período de validación.

Este desbalance no es un artefacto metodológico sino un reflejo del fenómeno operativo que se busca modelar: en flotas de ambulancias con mantenimiento activo, los períodos de operatividad continua son significativamente más frecuentes que los episodios de inoperatividad. Este patrón es coherente con la tasa de inoperatividad base del 21,2% mensual y con el diseño del horizonte de predicción de 14 días, que limita la ventana en la que puede ocurrir un evento positivo.

Para gestionar este desbalance durante el entrenamiento, se aplicó una estrategia de *oversampling* aleatorio con reemplazo sobre la clase minoritaria, con el objetivo de igualar la representación de ambas clases en

el conjunto de entrenamiento. Esta estrategia afecta exclusivamente al proceso de entrenamiento; el conjunto de validación (período 2025, $n = 1\,650$) mantiene la distribución original desbalanceada, lo que garantiza que las métricas de evaluación reflejen el desempeño del modelo bajo condiciones representativas del fenómeno real. La Figura 23 presenta la distribución de clases de la variable objetivo **y_inoperativa_14**.

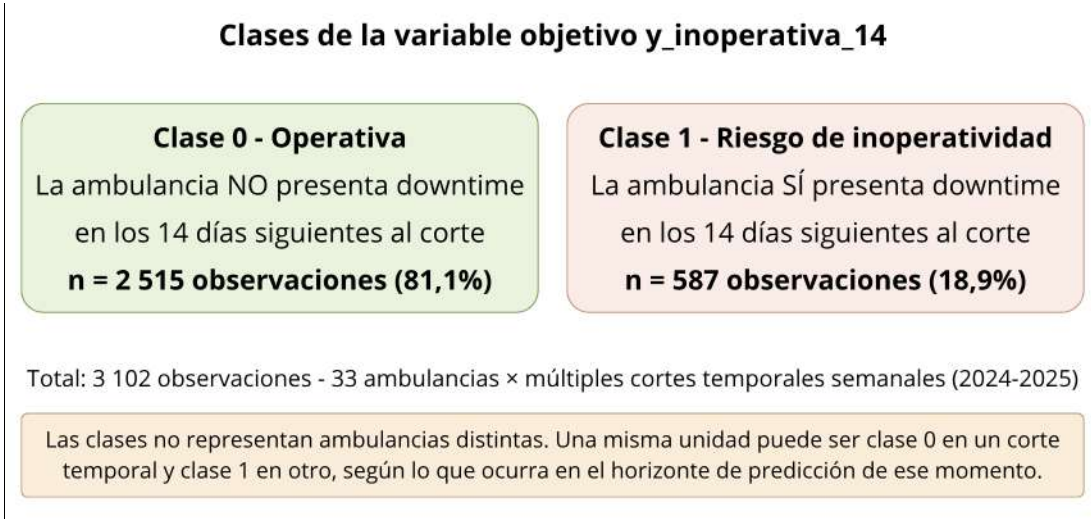


Figura 23. Distribución de clases de la variable objetivo.
Nota. Las clases representan observaciones clasificadas según la ocurrencia de inoperatividad en el horizonte de predicción $T = 14$ días, no ambulancias únicas.
Fuente. Elaboración propia.

La Figura 24 muestra la distribución de la variable objetivo por subconjunto de datos.

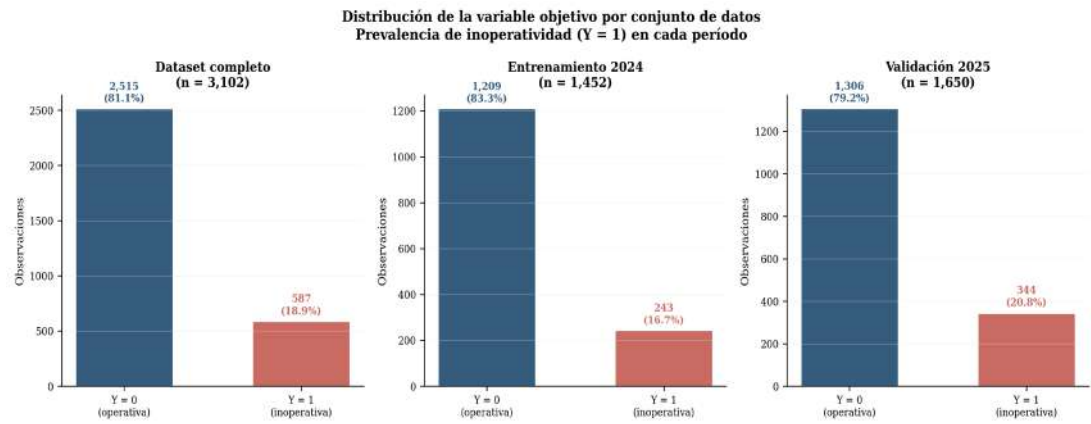


Figura 24. Distribución de la variable objetivo por conjunto de datos. Prevalencia de inoperatividad ($Y = 1$) en cada período.
Fuente. Elaboración propia.

V.1.2. *Análisis exploratorio de variables explicativas*

El análisis descriptivo de las variables explicativas reveló patrones operativos heterogéneos entre unidades, coherentes con la variabilidad esperada en una flota urbana de emergencias. Las variables de mantenimiento mostraron diferencias entre unidades según la frecuencia de intervenciones preventivas, correctivas y el tiempo transcurrido desde la última intervención, mientras que los indicadores de inoperatividad histórica reflejaron diferencias relevantes en la acumulación de *downtime* y disponibilidad reciente.

Los indicadores de inoperatividad mostraron que el 10,9% de las observaciones presentaron al menos un episodio de *downtime* en la ventana histórica, con una duración promedio de 3,5 días por episodio. La disponibilidad operativa promedio en la ventana W fue de 0,94, lo que indica que en promedio las unidades estuvieron operativas el 94% del tiempo de los 60 días previos al corte de evaluación.

Las variables de uso operativo, representadas por el kilometraje acumulado y los servicios prestados en la ventana histórica W, mostraron una dispersión amplia entre observaciones, reflejando la heterogeneidad de la carga operativa entre ambulancias de la flota simulada. Este análisis exploratorio permitió verificar que las variables construidas capturan dimensiones complementarias del comportamiento operativo de las ambulancias. En particular, la combinación de mantenimiento, inoperatividad histórica y uso operativo proporciona una representación más completa del estado reciente de cada unidad. Esta caracterización preliminar resulta necesaria para interpretar posteriormente el desempeño del modelo y la importancia relativa de los predictores. Asimismo, permitió identificar asociaciones preliminares entre las variables. La Figura 25 presenta el mapa de calor de correlaciones de Pearson entre las 14 variables explicativas, calculado sobre el *dataset* completo de 3 102 observaciones, revelando los principales patrones de asociación entre categorías y la presencia de multicolinealidad dentro del grupo de indicadores de inoperatividad histórica.

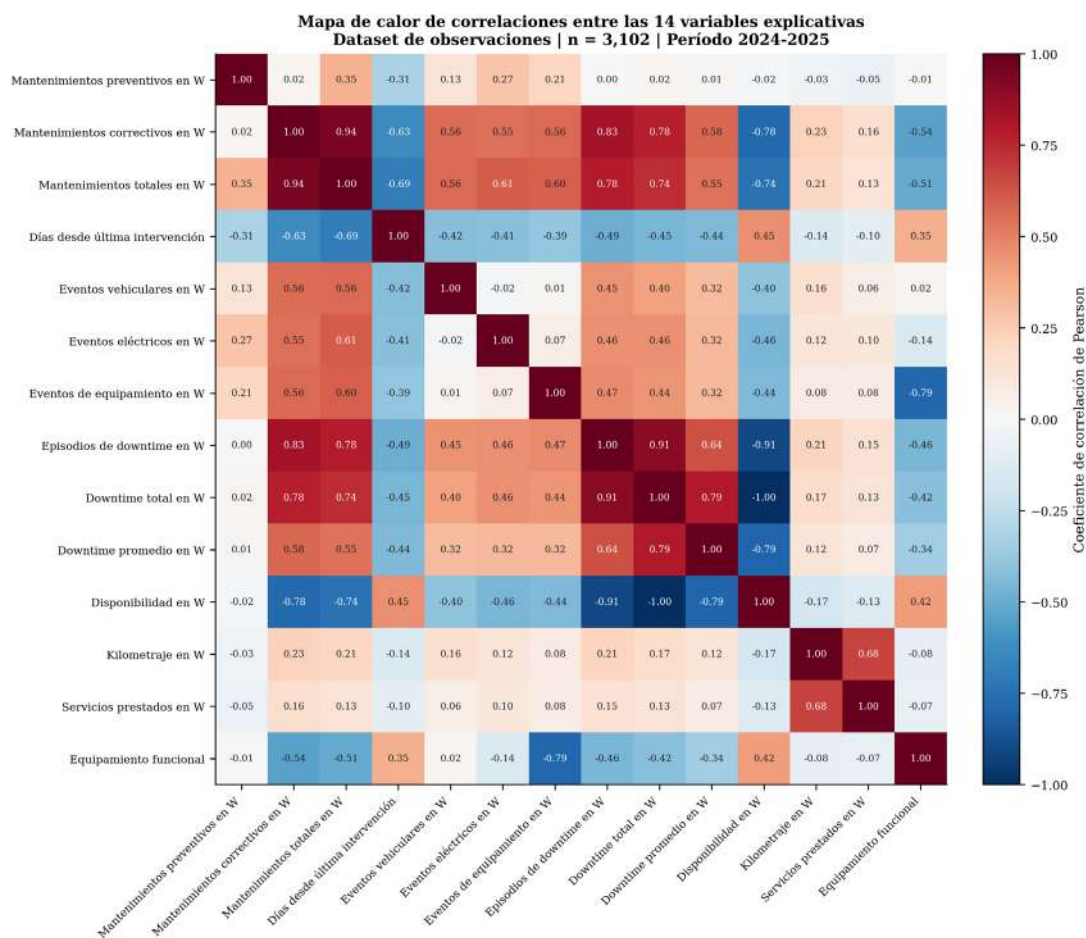


Figura 25. Mapa de calor de correlaciones entre las 14 variables explicativas. Dataset de observaciones. n = 3 102. Período 2024-2025.
Fuente. Elaboración propia.

La Figura 25 presenta el mapa de calor de correlaciones de Pearson entre las 14 variables explicativas, calculado sobre el dataset completo de 3 102 observaciones. Se observa alta correlación positiva entre “downtime_total_dias_w”, “downtime_promedio_dias_w” y “n_episodios_downtime_w”, patrón esperado al tratarse de indicadores derivados del mismo fenómeno de inoperatividad. Esta multicolinealidad interna al grupo afecta principalmente a la Regresión Logística, cuyo coeficiente de estimación se vuelve inestable, pero no compromete el desempeño del *Random Forest*, que maneja naturalmente predictores correlacionados. Se observaron correlaciones altas dentro del bloque de variables asociadas a inoperatividad histórica, especialmente entre episodios de *downtime*, *downtime* total, *downtime* promedio y disponibilidad en W, lo

que era esperable por tratarse de indicadores derivados del mismo fenómeno. También se identificaron asociaciones relevantes entre mantenimientos correctivos, eventos por macro-subsistema y variables de *downtime*, lo que sugiere que las categorías de mantenimiento e inoperatividad no son completamente independientes. En cambio, las variables de uso operativo: kilometraje y servicios prestados, aportan una dimensión complementaria vinculada a la carga de trabajo de la unidad.

Esta lectura exploratoria permitió reconocer que algunas variables contienen información parcialmente redundante, mientras que otras aportan dimensiones distintas del comportamiento operativo de las ambulancias. En particular, la presencia de correlaciones dentro del bloque de inoperatividad histórica confirma la coherencia interna de los indicadores derivados, pero también justifica la comparación entre modelos con distinta sensibilidad a la multicolinealidad. Por ello, el análisis de correlaciones no se utilizó para eliminar variables, sino como insumo interpretativo previo al modelado supervisado.

Complementariamente al análisis de correlaciones, se exploró la relación visual entre variables explicativas relevantes y la ocurrencia de inoperatividad futura. Para ello, se elaboraron diagramas de dispersión y distribuciones comparativas por clase, con el propósito de evaluar si la clase positiva podía distinguirse mediante patrones univariados o bivariados evidentes.

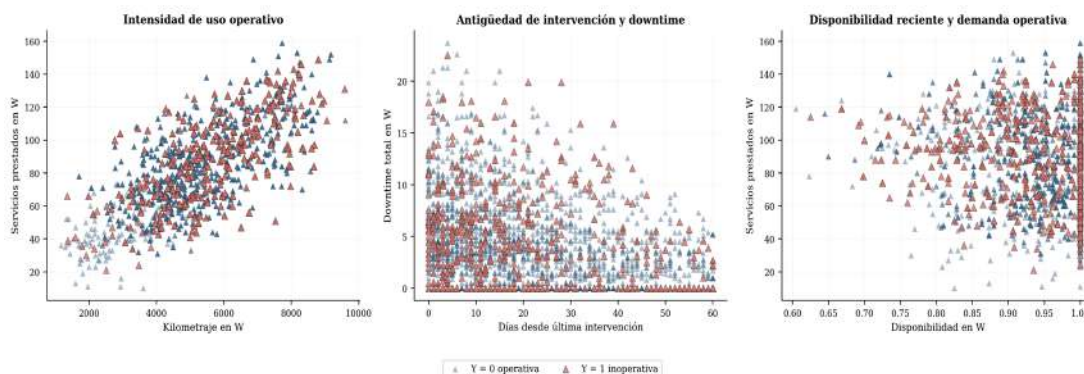


Figura 26. Diagramas de dispersión entre variables explicativas relevantes y eventos de inoperatividad.

Fuente. Elaboración propia.

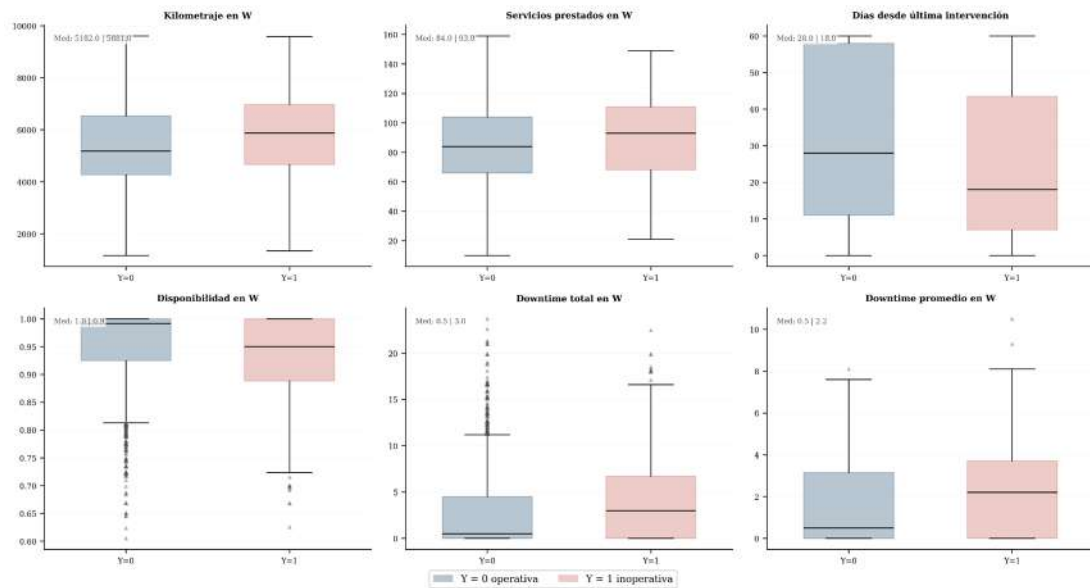


Figura 27. Distribución de variables explicativas principales según la clase objetivo.

Fuente. Elaboración propia.

Las Figuras 26 y 27 muestran una superposición considerable entre las observaciones operativas e inoperativas, tanto en las relaciones bivariadas como en las distribuciones univariadas de las variables principales. Este comportamiento sugiere que la inoperatividad futura no se explica adecuadamente por una sola variable considerada de forma aislada, sino por la interacción conjunta de múltiples factores operativos, de mantenimiento e inoperatividad histórica.

Esta superposición es coherente con la naturaleza del problema, dado que los eventos de inoperatividad no dependen únicamente del valor extremo de una variable, sino de combinaciones acumuladas de uso, historial de mantenimiento y antecedentes recientes de downtime. Por ello, la separación entre clases no resulta evidente en representaciones simples, lo que limita la utilidad de criterios univariados para anticipar el evento objetivo.

En consecuencia, estos resultados exploratorios respaldan la evaluación de modelos no lineales capaces de capturar interacciones multivariadas, como los árboles de decisión y los métodos de ensamble evaluados en la sección siguiente.

V.1.3. Desempeño del modelo computacional - Métricas de validación

El desempeño de los cuatro modelos predictivos fue evaluado sobre el conjunto de validación del período 2025, comparando sus resultados frente al esquema de mantenimiento preventivo programado como línea base. La Tabla 16 presenta las métricas de evaluación correspondientes a las declaradas en el OE4 de la investigación. Esta comparación permitió identificar la capacidad predictiva y el equilibrio entre detección de eventos y falsas alertas.

Tabla 16 Métricas de desempeño comparativas - Validación retrospectiva (período 2025)

Modelo	Precisión	Sensibilidad	Especificidad	Exactitud	F1-Score	AUC-ROC	MAE	Disp. Proyectada
Preventivo tradicional (línea base)	16,23%	30,52%	58,50%	52,67%	21,19%	42,47%	47,33%	60,79%
Regresión Logística	20,85%	100,00%	0,00%	20,85%	34,50%	53,09%	50,12%	0,00%
Árbol de Decisión	20,21%	55,52%	42,27%	45,03%	29,64%	50,46%	41,10%	42,73%
Random Forest	21,18%	91,86%	9,95%	27,03%	34,42%	56,03%	45,39%	9,58%
Gradient Boosting	22,17%	65,41%	39,51%	44,91%	33,11%	52,58%	42,05%	38,48%

Nota. Las métricas se expresan en valores porcentuales para facilitar su interpretación. El MAE se presenta como error medio absoluto equivalente sobre una escala normalizada de 0 a 100%. Umbral de clasificación: $\tau = 0,30$.

Fuente. Elaboración propia.

Los resultados completos de la validación retrospectiva para todos los modelos evaluados, incluyendo la matriz de confusión detallada, se presentan en el **Anexo F**. La Figura 28 presenta las matrices de confusión de los cinco enfoques evaluados de forma comparativa, permitiendo visualizar la distribución de verdaderos positivos, falsos negativos, falsos positivos y verdaderos negativos sobre el conjunto de validación.

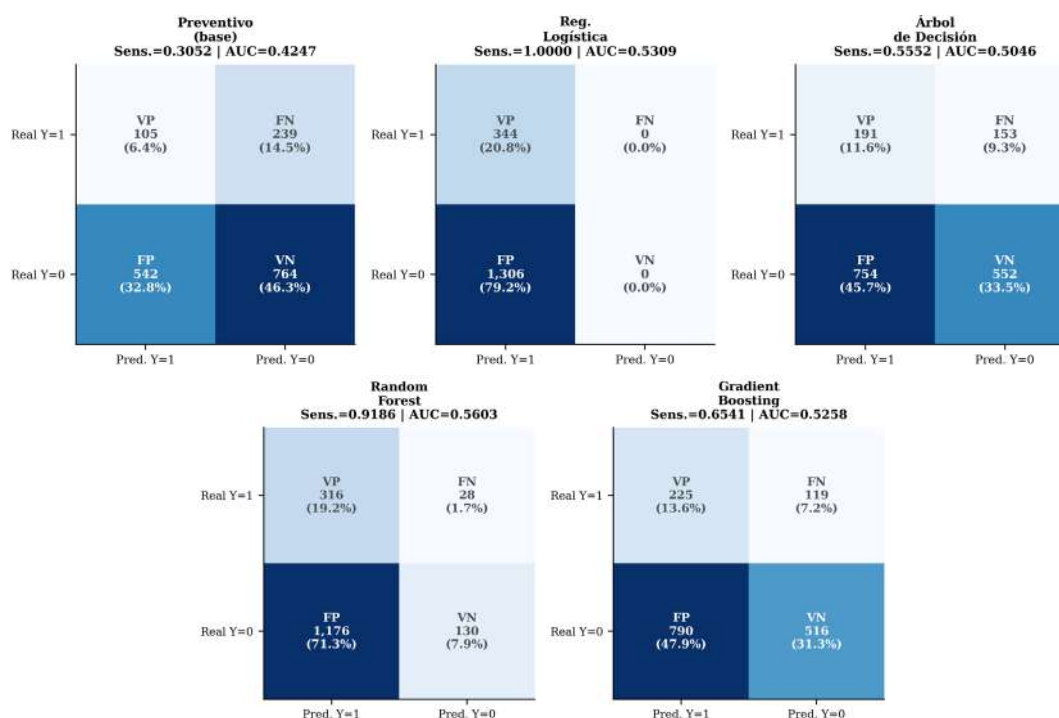


Figura 28. Matrices de confusión comparativas de los cinco enfoques evaluados. Período de validación 2025. $n = 1\,650$. Umbral $\tau = 0,30$.
Fuente. Elaboración propia.

Los cuatro modelos predictivos superaron la línea base preventiva en la métrica AUC-ROC, aunque con mejoras moderadas en términos absolutos. Este resultado indica que la incorporación de información histórica de operación, mantenimiento e inoperatividad aporta capacidad discriminativa adicional frente al criterio preventivo programado, pero no constituye evidencia de alto desempeño predictivo absoluto. El esquema preventivo tradicional obtuvo un AUC-ROC de 42,47%, inferior al valor de referencia de 0,50 asociado a un clasificador aleatorio, lo que evidencia una capacidad discriminativa limitada bajo la formulación retrospectiva adoptada. En contraste, los modelos basados en aprendizaje supervisado alcanzaron valores superiores, evidenciando que la incorporación de información histórica de operación y mantenimiento aporta capacidad predictiva adicional frente al criterio preventivo programado.

Desde la perspectiva operativa, el hallazgo más relevante fue la mejora relativa en sensibilidad. El modelo Random Forest identificó correctamente el 91,86% de los eventos de inoperatividad reales, frente al 30,52% del

esquema preventivo tradicional. Esto implica que, por cada 100 eventos de inoperatividad reales en el conjunto de validación, el modelo Random Forest anticiparía aproximadamente 92, mientras que la línea base preventiva anticiparía aproximadamente 31.

La Regresión Logística presentó sensibilidad máxima, pero con especificidad y disponibilidad proyectada extremadamente bajas, lo que indica un comportamiento operacionalmente degenerado al clasificar prácticamente todas las observaciones como positivas. Por su parte, el *Gradient Boosting* mostró un perfil más equilibrado en términos de exactitud y especificidad, aunque con menor sensibilidad que el *Random Forest*.

Considerando que la clase positiva corresponde a eventos de inoperatividad funcional de ambulancias Tipo II, la sensibilidad fue interpretada como la métrica de mayor relevancia operativa para la selección del modelo. En problemas de mantenimiento predictivo, la detección temprana de eventos críticos resulta prioritaria, debido a que los casos positivos suelen representar fallas potenciales cuya omisión puede comprometer la disponibilidad del sistema. Bajo este criterio, el *Random Forest* fue seleccionado como modelo predictivo principal, no por maximizar simultáneamente todas las métricas evaluadas, sino por mostrar el mejor desempeño relativo para soporte preventivo, combinando la mayor sensibilidad entre los modelos operacionalmente viables con el mayor AUC-ROC del conjunto evaluado. Esta selección debe interpretarse considerando su precisión limitada y su baja especificidad, asociadas al desbalance de clases y al umbral conservador $\tau = 0,30$. En consecuencia, la selección del modelo privilegió la reducción de falsos negativos sobre la maximización aislada de métricas globales, coherentemente con los objetivos del mantenimiento predictivo aplicado al contexto prehospitalario.

La Figura 29 ilustra la comparativa de las seis métricas de evaluación principales entre los modelos, consolidando en un único panel el comportamiento diferencial de cada algoritmo respecto a la línea base preventiva.

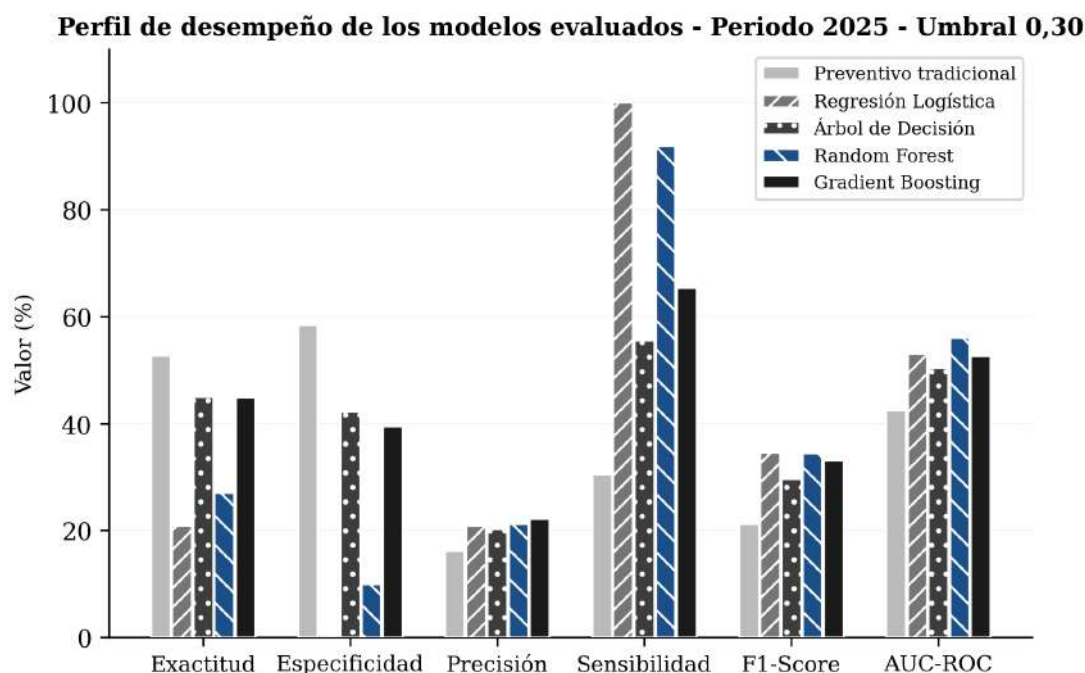


Figura 29. Perfil de desempeño de los modelos evaluados. Período de validación 2025. Umbral $\tau = 0,30$.

Fuente. Elaboración propia.

La Figura 30 presenta las curvas ROC de los cinco enfoques evaluados, ilustrando la capacidad discriminativa de cada uno respecto a la clasificación aleatoria. La curva ROC permite observar el comportamiento del clasificador a través de distintos umbrales de decisión, por lo que complementa la interpretación de las métricas calculadas con el umbral operativo $\tau = 0,30$. En este análisis, los modelos de aprendizaje supervisado se ubicaron por encima de la línea base preventiva, aunque sin alcanzar una separación amplia respecto al comportamiento aleatorio. Este resultado confirma que existe una mejora discriminativa relativa, pero también evidencia que la predicción de inoperatividad continúa siendo un problema complejo debido al desbalance de clases y a la superposición entre observaciones operativas e inoperativas. En particular, la cercanía entre las curvas refleja que la información disponible permite cierta ordenación del riesgo, pero no una discriminación perfecta entre ambas clases. Por ello, la interpretación del AUC-ROC debe realizarse junto con la sensibilidad, la especificidad y la matriz de confusión.

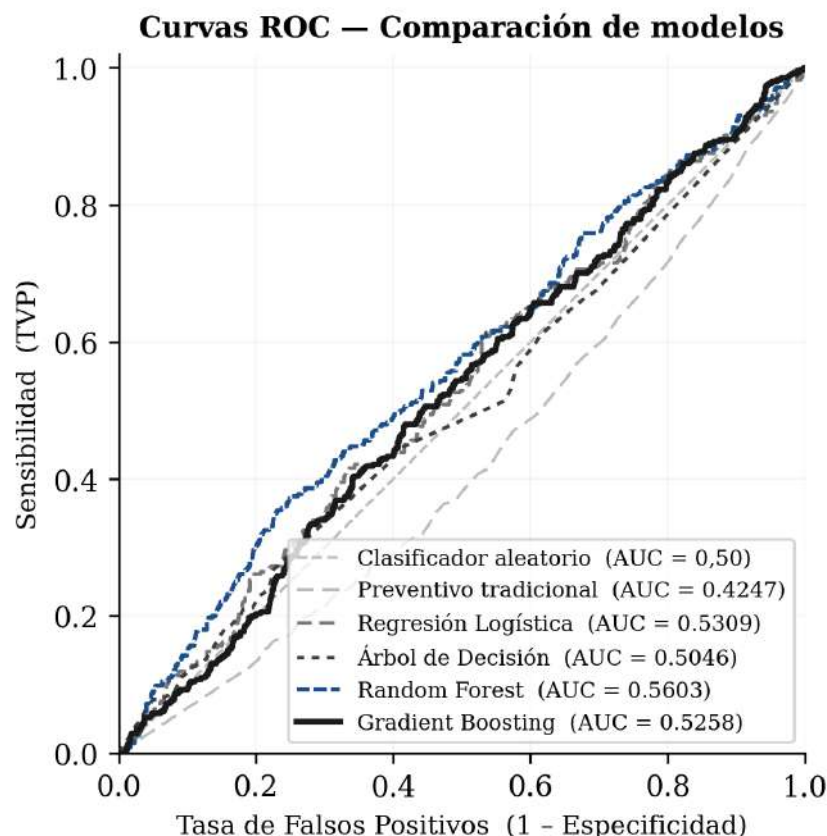


Figura 30. Curvas ROC comparativas - Validación retrospectiva período 2025.

Nota. La línea discontinua gris representa el clasificador aleatorio ($AUC = 0,50$). Los modelos supervisados evaluados presentan valores de AUC-ROC superiores a la línea base preventiva; Random Forest alcanzó el mayor AUC-ROC, aunque con una capacidad discriminativa moderada.

Fuente. Elaboración propia.

La Figura 31 muestra la distribución de probabilidades predichas por el modelo *Random Forest* para las clases $Y=0$ e $Y=1$, evidenciando un solapamiento importante entre ambas distribuciones. Este comportamiento indica que el modelo logra asignar probabilidades relativamente mayores a una parte de los eventos de inoperatividad, pero no separa de manera perfecta las observaciones positivas y negativas. Por ello, el desempeño debe interpretarse como una mejora relativa para priorización preventiva, más que como una clasificación determinística de alta certeza. La selección del umbral conservador $\tau = 0,30$ responde a esta condición, ya que permite aumentar la detección de eventos positivos a costa de generar un mayor número de falsas alertas. En consecuencia, la distribución de probabilidades

respalda la decisión metodológica de priorizar la sensibilidad sobre la especificidad en el contexto operativo de ambulancias Tipo II.

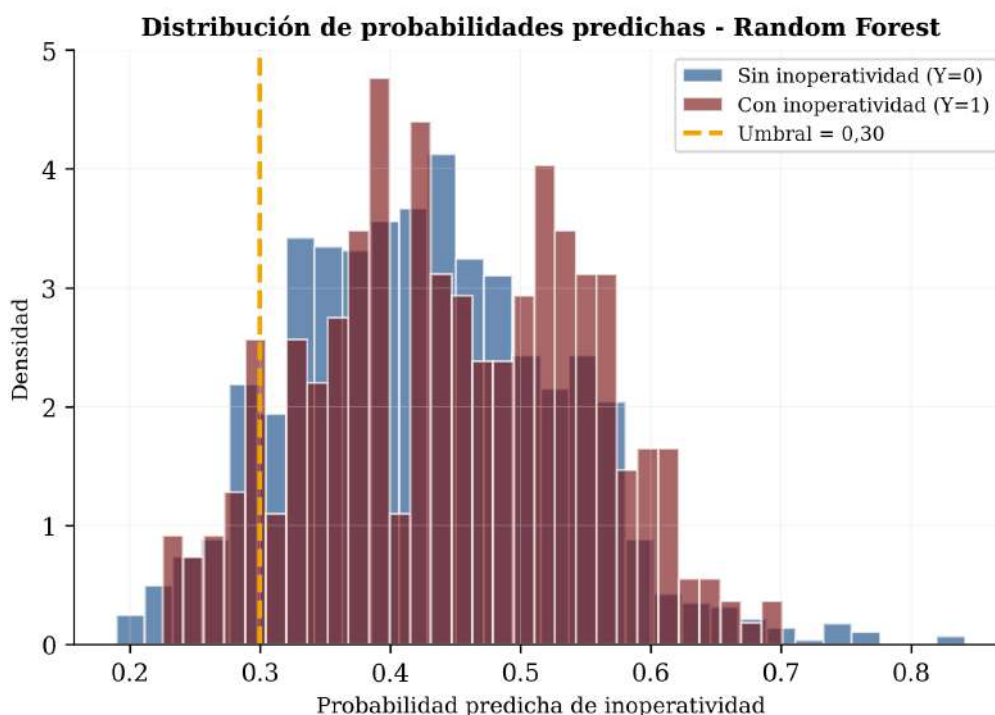


Figura 31. Distribución de probabilidades predichas por clase - Modelo Random Forest.

Nota. La línea discontinua naranja indica el umbral de clasificación adoptado (0,30). El solapamiento entre distribuciones $Y=0$ e $Y=1$ es consistente con los valores de AUC-ROC reportados en la Tabla 16.

Fuente. Elaboración propia.

V.1.3.1. Comparación estadística formal de clasificadores

Las métricas presentadas en la Tabla 16 permiten ordenar a los modelos evaluados según su desempeño en cada dimensión, pero no establecen por sí mismas si las diferencias observadas son estadísticamente distinguibles de la variabilidad propia del conjunto de validación. Con el fin de complementar dicha comparación descriptiva con un criterio inferencial, esta sección presenta la aplicación de la prueba de McNemar, la prueba de DeLong y el cálculo de la precisión media, desarrollados para los pares de modelos de mayor interés comparativo, tomando como referencia constante el modelo Random Forest, seleccionado como modelo predictivo principal. Este análisis permite valorar si las diferencias de desempeño observadas

responden a cambios consistentes en las predicciones y no solo a variaciones descriptivas entre métricas.

a) Prueba de McNemar

La prueba de McNemar (McNemar, 1947) evalúa si dos clasificadores correlacionados, es decir, evaluados sobre las mismas observaciones, presentan tasas de error estadísticamente diferentes. Dietterich (1998) demostró que, en problemas de clasificación supervisada donde dos modelos se comparan sobre un único conjunto de prueba, los procedimientos basados en la distribución t de Student violan el supuesto de independencia entre observaciones y subestiman sistemáticamente la varianza del estimador, generando una tasa de error de tipo I superior a la nominal. La prueba de McNemar evita este problema al operar directamente sobre la tabla de contingencia de discordancias entre los dos clasificadores.

Dado un par de clasificadores A y B evaluados sobre el mismo conjunto de n observaciones, se construye una tabla de contingencia 2×2 en la que n_{11} representa las observaciones clasificadas correctamente por ambos modelos, n_{00} las clasificadas incorrectamente por ambos, n_{10} las clasificadas correctamente por A e incorrectamente por B, y n_{01} las clasificadas incorrectamente por A y correctamente por B. El estadístico de McNemar con corrección de continuidad de Yates se define en la Ecuación 39:

$$\chi^2 = \frac{(|n_{10} - n_{01}| - 1)^2}{(n_{10} + n_{01})} \quad (39)$$

Donde:

- χ^2 : estadístico de la prueba de McNemar con corrección de continuidad de Yates.
- n_{10} : observaciones clasificadas correctamente por *Random Forest* e incorrectamente por el modelo comparado.
- n_{01} : observaciones clasificadas incorrectamente por *Random Forest* y correctamente por el modelo comparado.

Bajo la hipótesis nula de igualdad de tasas de error entre ambos clasificadores, el estadístico χ^2 se distribuye aproximadamente como una chi-cuadrado con 1 grado de libertad cuando $n_{10} + n_{01} \geq 25$ (Dietterich, 1998). Dado que en todas las comparaciones de la presente investigación se cumple esta condición (Tabla 18), se aplicó la aproximación con corrección de continuidad.

A modo ilustrativo del procedimiento, se desarrolla a continuación el cálculo correspondiente al par Random Forest frente a Gradient Boosting. A partir de las predicciones observación por observación sobre las 1 650 unidades de validación, se obtuvo $n_{10} = 129$ (observaciones clasificadas correctamente por *Random Forest* e incorrectamente por *Gradient Boosting*) y $n_{01} = 424$ (el caso inverso). Sustituyendo en la Ecuación 39:

$$\chi^2 = \frac{(|129 - 424| - 1)^2}{(129 + 424)}$$

$$\chi^2 = \frac{(294)^2}{553}$$

$$\chi^2 = \frac{86\,436}{553}$$

$$\chi^2 = 156,30$$

El valor p asociado, obtenido de la distribución χ^2 con 1 grado de libertad, es $p < 0,0001$. El valor obtenido supera ampliamente el valor crítico de 3,841 correspondiente a $\alpha = 0,05$; por lo que se rechaza la hipótesis nula de igualdad de tasas de error entre *Random Forest* y *Gradient Boosting*. La dirección de la discordancia es informativa: $n_{01} = 424$ es sustancialmente mayor que $n_{10} = 129$, lo que refleja el patrón ya documentado en la Tabla 16: *Random Forest*, al operar con una sensibilidad mucho mayor (91,86% frente a 65,41%) bajo el umbral $\tau = 0,30$; clasifica como positivas un número considerablemente mayor de observaciones, lo que incrementa tanto sus verdaderos positivos como sus falsos positivos respecto de *Gradient Boosting*. La prueba de McNemar confirma que esta diferencia no es atribuible a la variabilidad muestral,

sino que constituye una diferencia sistemática y estadísticamente significativa entre ambos clasificadores bajo el umbral operativo adoptado.

La Tabla 17 presenta los resultados de la prueba de McNemar para los cuatro pares de modelos de mayor interés comparativo, calculados siguiendo el mismo procedimiento ilustrado.

Tabla 17 Comparación estadística de clasificadores mediante la prueba de McNemar - Validación retrospectiva 2025.

Comparación	n_{10}	n_{01}	χ^2 (Ecuación 39)	Valor p
Random Forest vs. Preventivo tradicional	322	745	166,90	< 0,0001
Random Forest vs. Regresión Logística	130	28	64,56	< 0,0001
Random Forest vs. Árbol de Decisión	188	485	130,19	< 0,0001
Random Forest vs. Gradient Boosting	129	424	156,30	< 0,0001

Nota. n_{10} = observaciones clasificadas correctamente por Random Forest e incorrectamente por el modelo de comparación; n_{01} = el caso inverso. $n = 1\,650$. χ^2 calculado mediante la Ecuación 39, con corrección de continuidad de Yates (McNemar, 1947; Dietterich, 1998).

Fuente. Elaboración propia.

Las cuatro comparaciones resultan estadísticamente significativas ($p < 0,0001$ en todos los casos), lo que indica que el patrón de aciertos y errores de Random Forest bajo el umbral $\tau = 0,30$ difiere de manera sistemática del de cada uno de los enfoques alternativos evaluados, y que dicha diferencia no puede atribuirse a la variabilidad muestral del conjunto de validación.

b) Prueba de DeLong

El AUC-ROC reportado en la Tabla 16 para cada modelo es un estimador puntual de la probabilidad de que el clasificador asigne una probabilidad mayor a una observación positiva que a una observación negativa elegida al azar. DeLong, DeLong y Clarke-Pearson (1988) propusieron un método no paramétrico para estimar la varianza de este estimador y, a partir de ella,

comparar estadísticamente los AUC-ROC de dos clasificadores evaluados sobre el mismo conjunto de datos, considerando explícitamente la correlación entre ambos estimadores que surge de compartir las mismas observaciones. Sun y Xu (2014) propusieron un algoritmo de cálculo computacionalmente eficiente, matemáticamente equivalente al método original.

Para dos clasificadores A y B evaluados sobre las mismas n observaciones, la diferencia de AUC se evalúa mediante el estadístico z definido en la Ecuación 40:

$$z = \frac{(AUC_A - AUC_B)}{\sqrt{S_{AA} + S_{BB} - 2 \cdot S_{AB}}} \quad (40)$$

donde:

- z : estadístico de prueba bajo la hipótesis nula $AUC_A = AUC_B$, distribuido aproximadamente como una normal estándar.
- AUC_A : área bajo la curva ROC de Random Forest.
- AUC_B : área bajo la curva ROC del modelo comparado.
- S_{AA}, S_{BB} : varianzas de los estimadores AUC_A y AUC_B .
- S_{AB} : covarianza entre AUC_A y AUC_B , derivada de la estructura de los *midranks* (rangos promedio para observaciones empatadas) de las probabilidades predichas por ambos modelos.

Bajo la hipótesis nula $AUC_A = AUC_B$, z se distribuye aproximadamente como una normal estándar, y el valor p bilateral se obtiene como $p = 2 \cdot [1 - \Phi(|z|)]$, donde Φ es la función de distribución acumulada normal estándar. Las varianzas y la covarianza se estimaron mediante el algoritmo de Sun y Xu (2014), siguiendo el método de DeLong et al. (1988). Esta comparación resulta pertinente porque el AUC-ROC evalúa la capacidad discriminativa global del modelo, independientemente del umbral de clasificación adoptado. La Tabla 18 presenta los resultados obtenidos para los cuatro pares de modelos de interés.

Tabla 18 Comparación del AUC-ROC mediante la prueba de DeLong para clasificadores correlacionados - Validación retrospectiva 2025.

Comparación	AUC _A (RF)	AUC _B	Diferencia	z (Ecuación 40)	Valor p
Random Forest vs. Gradient Boosting	0,5603	0,5254	0,0349	2,178	0,029
Random Forest vs. Árbol de Decisión	0,5603	0,5046	0,0557	2,789	0,005
Random Forest vs. Regresión Logística	0,5603	0,5309	0,0294	1,508	0,132
Random Forest vs. Preventivo tradicional	0,5603	0,4247	0,1355	4,579	< 0,0001

Nota. $n = 1\,650$ ($m = 344$ positivas, $n - m = 1\,306$ negativas). z calculado mediante la Ecuación 40 (DeLong et al., 1988; Sun & Xu, 2014).

Fuente. *Elaboración propia.*

Random Forest presentó un AUC-ROC significativamente mayor que *Gradient Boosting* ($z = 2,178$; $p = 0,029$), Árbol de Decisión ($z = 2,789$; $p = 0,005$) y el esquema preventivo tradicional ($z = 4,579$; $p < 0,0001$). La comparación con Regresión Logística no resultó estadísticamente significativa ($z = 1,508$; $p = 0,132$), lo que indica que ambos modelos presentan una capacidad discriminativa global, medida de forma independiente del umbral de decisión, que no es distinguible bajo este criterio. La interpretación operativa conjunta de este resultado con el reportado en la Tabla 17 se desarrolla en la Sección VI.2.

c) Curva Precision-Recall y precisión media

Saito y Rehmsmeier (2015) demostraron formalmente que, en conjuntos de datos con clases desbalanceadas, el AUC-ROC puede sobreestimar la calidad discriminativa de un clasificador, debido a que la tasa de falsos positivos se calcula como $FP / (FP + VN)$, y el denominador está dominado

por el número de verdaderos negativos cuando la clase negativa es mayoritaria. En el conjunto de validación de la presente investigación, la clase negativa representa el 79,15% de las observaciones. La curva *Precision-Recall*, en cambio, relaciona directamente la precisión con la sensibilidad, ambas calculadas sobre cantidades directamente vinculadas a la clase positiva, lo que la hace más informativa cuando, como en el contexto de soporte preventivo de esta investigación, el costo operativo está asociado a cada alerta generada y a cada evento detectado.

La precisión media (*Average Precision*, AP) se define como el área bajo la curva *Precision-Recall* y se calcula mediante la Ecuación 41:

$$AP = \sum_k (R_k - R_{k-1}) \cdot P_k \quad (41)$$

Donde:

- AP : Precisión media, área bajo la curva *Precision-Recall*, $AP \in [0, 1]$.
- P_k : Precisión evaluada en el k-ésimo umbral de decisión,

$$P_k = \frac{VP_k}{(VP_k + FP_k)}$$
- R_k : Sensibilidad evaluada en el k-ésimo umbral de decisión,

$$R_k = \frac{VP_k}{(VP_k + FN_k)} ; \text{ con } R_0 = 0 \text{ y } R_k \text{ monótonamente creciente.}$$

A diferencia del AUC-ROC, que se calcula sobre el espacio (tasa de falsos positivos, sensibilidad), AP se calcula sobre el espacio (sensibilidad, precisión), y resume el desempeño del clasificador a lo largo de todos los umbrales posibles, sin fijarse en el umbral operativo $\tau = 0,30$ adoptado para las métricas puntuales de la Tabla 16. El cómputo de AP se realiza programáticamente sobre el conjunto de umbrales definidos por los valores distintos de probabilidad predicha por cada modelo. En los extremos de la curva, su comportamiento es interpretable de manera directa: cuando $R_k \rightarrow 0$ (umbral de decisión muy alto), la precisión tiende a su valor máximo, dado que las pocas observaciones clasificadas como positivas corresponden a las de mayor probabilidad predicha; cuando $R_k \rightarrow 1$ (umbral

de decisión muy bajo), la precisión converge a la prevalencia observada de la clase positiva ($\pi = 344 / 1\,650 = 0,2085$).

La Figura 30.1 presenta las curvas *Precision-Recall* comparativas de los cinco enfoques evaluados, junto con la línea de referencia correspondiente a un clasificador no informativo bajo la prevalencia observada.

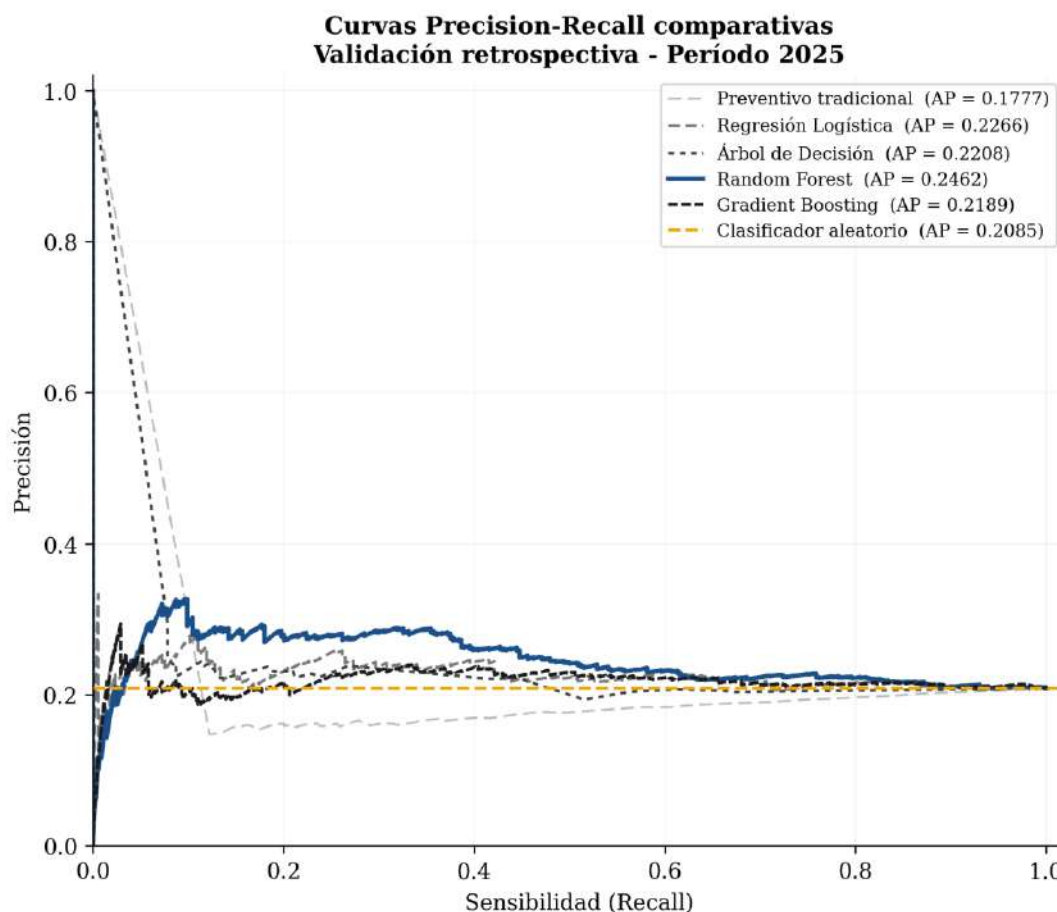


Figura 32. Curvas *Precision-Recall* comparativas - Validación retrospectiva período 2025.

Nota. La línea discontinua naranja representa el desempeño esperado de un clasificador no informativo bajo la prevalencia observada ($AP = 0,2085$). $AP = average\ precision$ (precisión media), área bajo la curva *Precision-Recall* calculada mediante la Ecuación 41.

Fuente. *Elaboración propia.*

Random Forest obtuvo la mayor precisión media ($AP = 0,2462$) entre los cinco enfoques evaluados, superando a *Gradient Boosting* ($AP = 0,2187$), *Árbol de Decisión* ($AP = 0,2208$), *Regresión Logística* ($AP = 0,2266$) y al esquema preventivo tradicional ($AP = 0,1777$), así como al valor de

referencia de un clasificador no informativo bajo la prevalencia observada ($AP = 0,2085$). Este resultado es particularmente relevante porque la Tabla 16 reporta que la precisión puntual de *Random Forest* bajo $\tau = 0,30$ (21,18%) es apenas marginalmente superior a la prevalencia, lo que podría sugerir, de forma aislada, que el modelo aporta poca capacidad discriminativa adicional frente a un clasificador trivial. Sin embargo, el valor de AP, al integrar la relación precisión-sensibilidad a lo largo de todos los umbrales posibles, evidencia que *Random Forest* mantiene una relación precisión-sensibilidad sistemáticamente por encima de la prevalencia en la mayor parte del rango de sensibilidad, lo cual no ocurre, en la misma magnitud, para los demás modelos evaluados. En conjunto con los resultados de las pruebas de McNemar y DeLong, este hallazgo constituye una tercera línea de evidencia, obtenida bajo un criterio metodológicamente independiente de las dos primeras, consistente con la selección de *Random Forest* como modelo predictivo principal.

Este resultado refuerza que la evaluación del modelo no debe limitarse a una única métrica calculada bajo un umbral fijo. En problemas con clases desbalanceadas, la precisión puntual puede subestimar el comportamiento global del clasificador, especialmente cuando el interés operativo se centra en detectar la mayor cantidad posible de eventos positivos. Por ello, la precisión media aporta una lectura complementaria sobre la capacidad del modelo para sostener un desempeño útil en distintos umbrales de decisión.

V.1.4. Importancia de variables explicativas

El análisis de importancia de variables se realizó sobre el modelo *Random Forest*, seleccionado como modelo predictivo principal de la investigación. Esta elección permite interpretar la contribución relativa de las variables explicativas al proceso de clasificación, dado que el algoritmo estima la importancia de cada predictor a partir de la reducción media de impureza generada por sus divisiones en los árboles del bosque. Los resultados permitieron identificar los factores con mayor capacidad

predictiva sobre la inoperatividad de las ambulancias médicas urbanas Tipo II. La Tabla 19 presenta las diez variables con mayor importancia relativa en el modelo Random Forest. La tabla completa con las 14 variables explicativas evaluadas se presenta en el **Anexo F** (Tabla F.2.).

Tabla 19 Importancia de variables explicativas - Modelo Random Forest, Top 10

Variable	Importancia (%)	Categoría
Kilometraje en W	18,52%	Uso operativo
Servicios prestados en W	16,91%	Uso operativo
Días desde última intervención	12,21%	Mantenimiento
<i>Downtime</i> promedio en W	9,99%	Inoperatividad
Disponibilidad histórica en W	8,71%	Inoperatividad
<i>Downtime</i> total en W	8,10%	Inoperatividad
Total de eventos de mantenimiento	6,04%	Mantenimiento
Mantenimientos correctivos en W	4,77%	Mantenimiento
Eventos en subsistema eléctrico/electrónico	3,38%	Mantenimiento
Eventos en subsistema vehicular	2,68%	Mantenimiento

Fuente. Elaboración propia.

La Figura 33 presenta la representación gráfica de los valores de importancia de la Tabla 19, evidenciando la concentración del poder predictivo en las primeras variables respecto al conjunto completo. Se observa que las variables de uso operativo ocupan las posiciones principales, lo que sugiere que la carga de trabajo reciente de la ambulancia constituye un factor relevante para anticipar eventos de inoperatividad. Asimismo, las variables asociadas a mantenimiento correctivo y *downtime* histórico aportan información complementaria sobre el estado técnico acumulado de la unidad. Esta distribución de importancias respalda la lógica del modelo, en tanto combina señales de uso, mantenimiento e inoperatividad previa para estimar el riesgo operativo futuro. En conjunto, los resultados muestran que la capacidad predictiva del modelo se sustenta en

múltiples dimensiones operativas y no en una única variable.

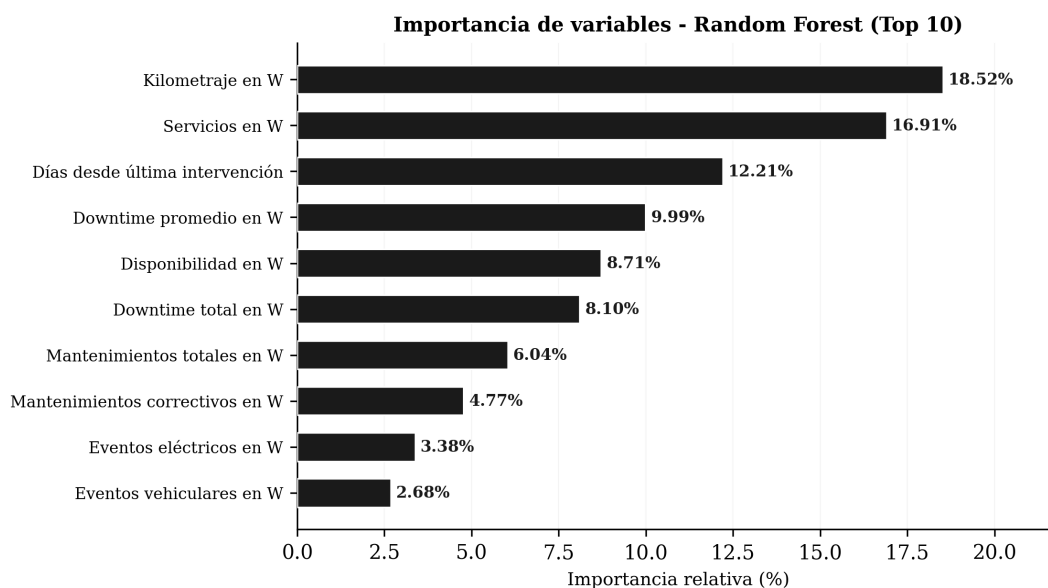


Figura 33. Importancia de variables explicativas - Modelo *Random Forest*. Top 10 variables.

Nota. Los valores representan la reducción media de impureza normalizada al 100%. Se muestran las diez variables con mayor importancia relativa del modelo *Random Forest*. Las variables de uso operativo, representadas por kilometraje acumulado en W y servicios prestados en W, concentran aproximadamente el 35,4% de la importancia relativa total. La tabla completa de las 14 variables explicativas se presenta en el Anexo F.

Fuente. Elaboración propia.

Las variables de uso operativo, representadas por el kilometraje acumulado en W y los servicios prestados en W, concentran aproximadamente el 35,4% de la importancia relativa total del modelo *Random Forest*, constituyendo los predictores más importantes de inoperatividad futura. Este hallazgo es coherente con los planteamientos de Prytz (2014), quien documentó que la intensidad de uso acumulada es uno de los factores más determinantes en la predicción de fallas vehiculares mediante modelos *data-driven*. La combinación de alta carga de servicios y elevado kilometraje en una misma ventana temporal es indicador de desgaste acelerado que el esquema preventivo tradicional no captura adecuadamente al basarse únicamente en intervalos fijos.

Los indicadores de inoperatividad histórica, representados por la disponibilidad en la ventana W, el *downtime* total y el *downtime* promedio,

aportan conjuntamente una proporción relevante de la capacidad predictiva, evidenciando que el comportamiento pasado de la unidad en términos de indisponibilidad es un predictor significativo de su comportamiento futuro. Esto respalda el principio fundamental del enfoque *data-driven*: los patrones históricos contienen información predictiva valiosa para la gestión del mantenimiento (Taoufyq et al., 2025).

La variable de días desde la última intervención confirma la relevancia del intervalo de mantenimiento como factor de riesgo, en línea con los fundamentos del mantenimiento centrado en confiabilidad descritos por Moubray (1997). Sin embargo, su importancia relativa muestra que el intervalo de tiempo por sí solo es un predictor insuficiente cuando no se complementa con información sobre la carga operativa real de la unidad.

En conjunto, la importancia de variables muestra que el modelo combina uso operativo, historial de inoperatividad e información de mantenimiento para estimar el riesgo futuro. Esta estructura es coherente con la naturaleza multifactorial del problema y respalda su uso como herramienta de priorización preventiva basada en evidencia.

V.1.5. *Prototipo de soporte preventivo*

A partir de las probabilidades estimadas por el modelo *Random Forest*, se construyó el prototipo de soporte preventivo correspondiente al Entregable 2 de la investigación. Cada ambulancia fue clasificada en uno de tres niveles de riesgo operativo: alto (probabilidad $\geq 0,50$), medio (probabilidad entre 0,25 y 0,49) y bajo (probabilidad $< 0,25$). El código fuente que implementa la Fase 5 de la metodología, generando los tres componentes del soporte preventivo, se encuentra disponible en el **Anexo D**.

La Figura 34 presenta las primeras filas del archivo de salida del prototipo, ilustrando el formato estructurado de las predicciones generadas por el modelo *Random Forest* para cada ambulancia, con la probabilidad estimada de inoperatividad y el nivel de riesgo operativo asignado.

1	id_ambulancia	t0	modelo_base	prob_inoperatividad	nivel_riesgo	dias_desde_ultima_interv	downtime_total_dias_w	n_cm_w	disponibilidad_w	equipamiento_funcional	km_en_w	servicios_en_w
2	AMB-T2-001	2025-06-13	Random Forest	0.8414	ALTO	3	1.5	4	0.975	1	7740.0	130
3	AMB-T2-001	2025-06-20	Random Forest	0.8392	ALTO	10	1.5	3	0.975	1	7740.0	130
4	AMB-T2-021	2025-06-13	Random Forest	0.7694	ALTO	10	1.6	1	0.9733	1	6976.0	134
5	AMB-T2-001	2025-05-09	Random Forest	0.7665	ALTO	12	3.0	4	0.95	1	8008.0	121
6	AMB-T2-021	2025-06-20	Random Forest	0.7663	ALTO	17	1.6	1	0.9733	1	6976.0	134
7	AMB-T2-001	2025-05-16	Random Forest	0.7508	ALTO	19	1.5	3	0.975	1	8008.0	121
8	AMB-T2-021	2025-06-06	Random Forest	0.7475	ALTO	3	1.6	1	0.9733	1	6976.0	134
9	AMB-T2-005	2025-11-07	Random Forest	0.7438	ALTO	8	2.8	3	0.9533	1	7260.0	141
10	AMB-T2-031	2025-06-27	Random Forest	0.7416	ALTO	5	3.3	2	0.945	0	7932.0	85
11	AMB-T2-001	2025-05-02	Random Forest	0.7363	ALTO	5	3.0	4	0.95	1	8008.0	121
12	AMB-T2-031	2025-02-28	Random Forest	0.7236	ALTO	0	2.8	3	0.9533	0	6621.0	111
13	AMB-T2-001	2025-01-24	Random Forest	0.7068	ALTO	11	6.9	3	0.885	0	8671.0	112
14	AMB-T2-001	2025-06-06	Random Forest	0.7033	ALTO	40	1.5	3	0.975	1	7740.0	130
15	AMB-T2-001	2025-10-17	Random Forest	0.701	ALTO	11	0.5	1	0.9917	0	8230.0	140
16	AMB-T2-001	2025-05-23	Random Forest	0.6991	ALTO	26	1.5	3	0.975	1	8008.0	121
17	AMB-T2-021	2025-04-11	Random Forest	0.6915	ALTO	4	3.2	3	0.9467	0	6988.0	120
18	AMB-T2-001	2025-10-10	Random Forest	0.6862	ALTO	4	0.5	1	0.9917	0	8230.0	140
19	AMB-T2-001	2025-06-27	Random Forest	0.684	ALTO	17	0.0	1	1.0	1	7740.0	130
20	AMB-T2-012	2025-08-29	Random Forest	0.6787	ALTO	8	16.6	4	0.7233	0	5830.0	118

Figura 34. Estructura del archivo de salida del prototipo de soporte preventivo. Primeras filas con la probabilidad estimada y el nivel de riesgo por unidad.

Fuente. Elaboración propia.

En el corte temporal del 12 de diciembre de 2025, correspondiente al último punto de evaluación del período de validación, el modelo clasificó 11 ambulancias con nivel de riesgo alto y 22 con nivel de riesgo medio, sin unidades clasificadas en riesgo bajo. Este resultado implica que el 33,3% de la flota presentó una probabilidad estimada de inoperatividad igual o superior a 0,50, requiriendo priorización preventiva dentro de los tres días hábiles siguientes.

La Figura 35 presenta los resultados del prototipo de soporte preventivo: la evolución temporal del nivel de riesgo de la flota durante el período de validación 2025, la distribución por categoría de riesgo en el corte del 12 de diciembre de 2025 y el ranking de probabilidad estimada de inoperatividad por unidad para dicho corte.

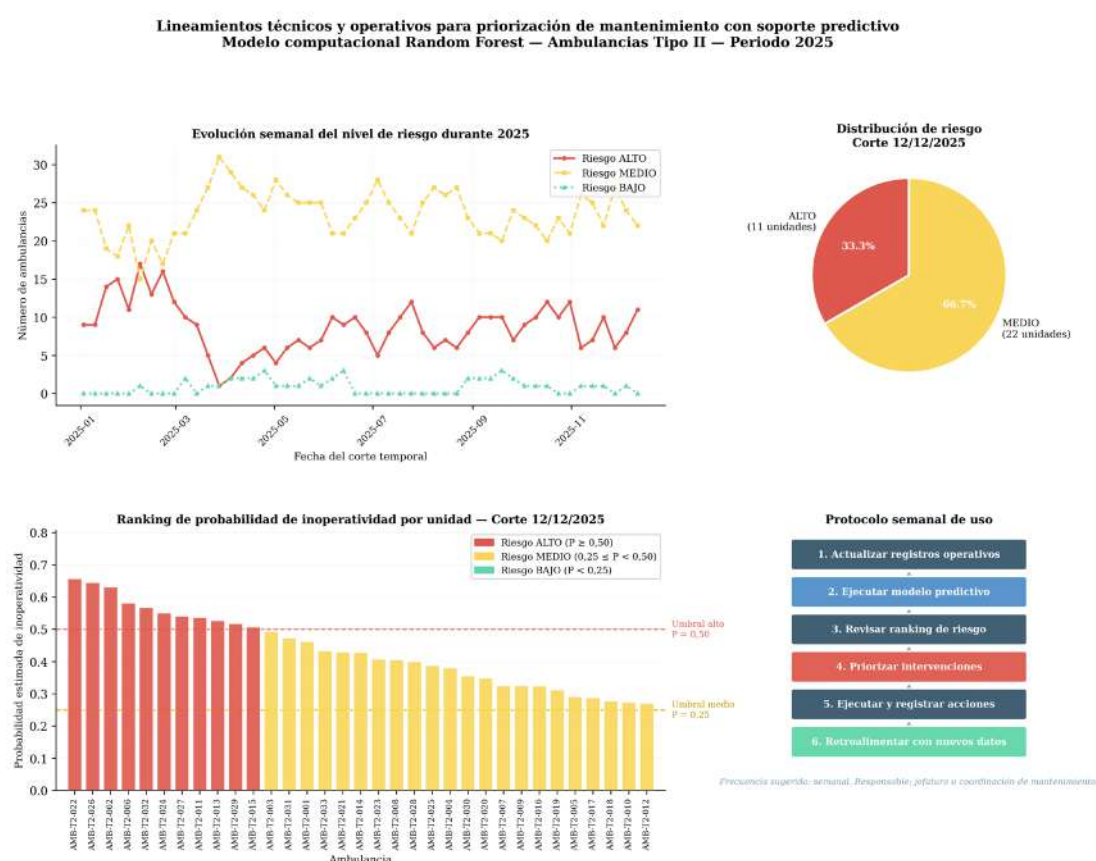


Figura 35. Lineamientos técnicos y operativos para la priorización de mantenimiento con soporte predictivo en ambulancias Tipo II durante el período de validación 2025.

Nota. El panel superior izquierdo muestra la evolución semanal del nivel de riesgo por categoría; el panel superior derecho presenta la distribución porcentual por nivel de riesgo en el corte del 12 de diciembre de 2025; el panel inferior izquierdo

presenta el ranking de probabilidad promedio de inoperatividad por unidad con los umbrales de clasificación; el panel inferior derecho presenta el esquema del protocolo semanal de uso del modelo, cuyo detalle completo se encuentra en el Anexo G.

Fuente. Elaboración propia.

La disponibilidad operativa proyectada por el modelo *Random Forest* para el período de validación fue menor que la disponibilidad proyectada por el esquema preventivo tradicional, debido al carácter conservador del umbral de clasificación adoptado. Esta diferencia no debe interpretarse como un error de pronóstico de disponibilidad real, sino como un reflejo de la estrategia de priorización preventiva: al clasificar como “en riesgo” a un mayor número de unidades, el modelo incrementa la sensibilidad y reduce la probabilidad de omitir eventos reales de inoperatividad.

Es importante precisar que la disponibilidad proyectada, en el contexto del presente modelo, no constituye un pronóstico de la disponibilidad real de la flota, sino un indicador derivado que refleja la proporción de observaciones clasificadas como de bajo riesgo ($y = 0$) bajo el umbral de decisión adoptado. Con un umbral de clasificación $\tau = 0,30$; el modelo clasifica como “en riesgo” a toda unidad cuya probabilidad estimada de inoperatividad supere ese valor, resultando en una clasificación deliberadamente conservadora que maximiza la sensibilidad a costa de incrementar las intervenciones preventivas.

Este comportamiento es funcionalmente coherente con la lógica de sistemas críticos: en el contexto prehospitalario, el costo operativo y sanitario de una inoperatividad no anticipada supera ampliamente al costo de una intervención preventiva innecesaria. Si se elevara el umbral de clasificación a $\tau = 0,50$; la disponibilidad proyectada se aproximaría a la disponibilidad real observada, pero a expensas de reducir la sensibilidad y perder eventos de inoperatividad que el modelo habría podido anticipar. La selección de umbral constituye, por tanto, una decisión de gestión de riesgo que debe ajustarse a las prioridades operativas del SAMU en una implementación real.

V.1.6. *Lineamientos técnicos derivados del modelo*

Los resultados del modelo computacional permitieron derivar tres categorías de lineamientos técnicos y operativos orientados a fortalecer la gestión del mantenimiento de las ambulancias Tipo II, correspondientes al Entregable 3 de la investigación.

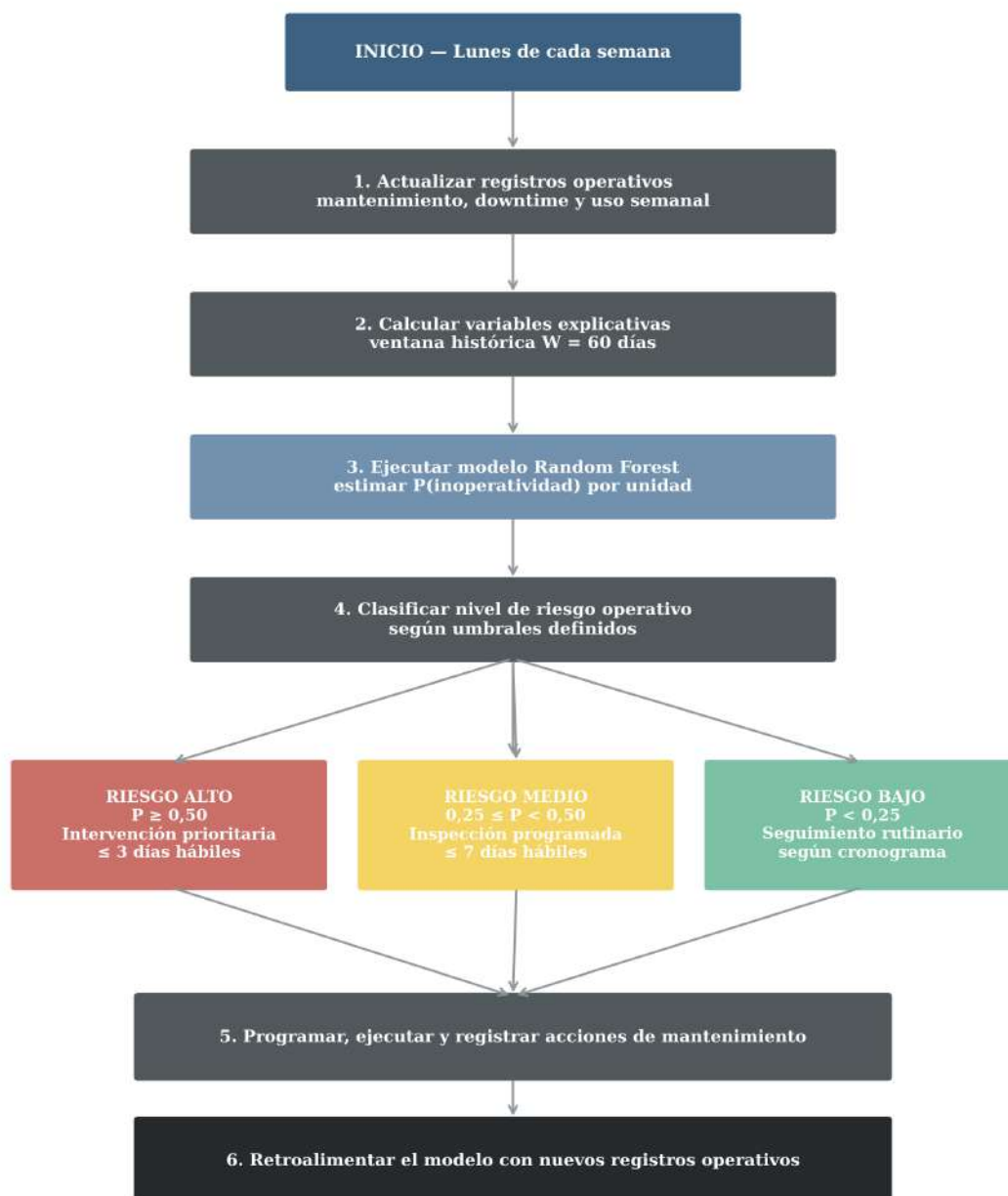
El primer conjunto de lineamientos establece criterios de priorización de intervenciones diferenciados por nivel de riesgo: las unidades con riesgo alto requieren intervención preventiva dentro de los tres días hábiles siguientes al corte de evaluación; las de riesgo medio, una inspección programada dentro de los siete días hábiles; y las de riesgo bajo, el cumplimiento del cronograma preventivo establecido según la institución y ficha técnica del fabricante.

El segundo conjunto define criterios de alerta temprana basados en las variables de mayor importancia predictiva. Se propusieron umbrales operativos específicos: más de 120 servicios prestados en 60 días, más de 8 000 km recorridos en la ventana histórica, más de 45 días sin intervención de mantenimiento, disponibilidad histórica inferior al 85% y más de 7 días acumulados de inoperatividad en la ventana W. La superación de cualquiera de estos umbrales debe activar una revisión técnica independientemente del nivel de riesgo asignado por el modelo. Los criterios de alerta temprana para cada variable crítica, incluyendo umbrales, interpretación y acción recomendada, se presentan en el **Anexo H**.

El tercer conjunto propone un protocolo operativo de seis pasos para la integración del modelo en el flujo de gestión del SAMU: actualización continua de registros, ejecución semanal del modelo, revisión del ranking de riesgo por el jefe de mantenimiento, programación de intervenciones priorizadas, ejecución y registro de intervenciones, y retroalimentación al modelo con los nuevos registros. Este protocolo no requiere modificaciones estructurales al sistema de gestión existente ni infraestructura tecnológica adicional, siendo ejecutable en cualquier entorno Python estándar.

La Figura 36 presenta el diagrama de flujo del protocolo semanal de uso del modelo computacional predictivo, sintetizando los seis pasos operativos y los tres niveles de acción diferenciados por nivel de riesgo.

Protocolo semanal de uso del modelo computacional predictivo para soporte al mantenimiento preventivo de ambulancias Tipo II



Frecuencia sugerida: semanal. Responsable: jefatura o coordinación de mantenimiento.

Figura 36. Protocolo semanal de uso del modelo computacional predictivo para soporte al mantenimiento preventivo de ambulancias médicas urbanas Tipo II.

Fuente. Elaboración propia.

El protocolo operativo completo de uso del modelo, con la descripción detallada de cada paso, responsables y frecuencia de ejecución, se presenta en el **Anexo G**. Este protocolo establece la secuencia mínima para convertir las probabilidades estimadas por el modelo en acciones de soporte preventivo, manteniendo una lógica trazable entre el resultado computacional y la decisión técnica de mantenimiento.

El protocolo operativo descrito se materializa en una plantilla de programación semanal en formato editable, generada automáticamente por el script “generar_plantilla_semanal.py” a partir del ranking de riesgo producido por el modelo. Esta plantilla permite organizar las unidades evaluadas según su nivel de riesgo operativo y facilita la revisión periódica de las ambulancias priorizadas.

Como se muestra en la Figura 37, la primera hoja funciona como un panel de control que integra la fecha de corte, el horizonte predictivo, el modelo utilizado, el resumen de priorización y la guía de interpretación por nivel de riesgo. Esta estructura permite identificar rápidamente las unidades que requieren revisión técnica inmediata, inspección programada o seguimiento rutinario.

La sección de programación semanal incorpora el listado priorizado de ambulancias según el riesgo estimado, junto con el número de alertas activas, la acción recomendada, el plazo sugerido y las observaciones correspondientes. Asimismo, la Figura 38 presenta la segunda hoja de la plantilla, donde se consigna el detalle completo de las unidades evaluadas, incluyendo la probabilidad estimada de inoperatividad, variables operativas relevantes, alertas activas e instrucción para el taller.

De este modo, la salida del modelo computacional se traduce en un instrumento operativo para apoyar la planificación semanal del mantenimiento con soporte preventivo, manteniendo la posibilidad de complementar la priorización automática con el criterio técnico del personal responsable. Para una visualización más amplia del formato completo propuesto, véase el **Anexo I**.

	A	B	C	D	E	F	G	H
1	PROGRAMACIÓN SEMANAL DE MANTENIMIENTO CON SOPORTE PREDICTIVO							
2	Ambulancias médicas urbanas Tipo II Semana 50 / 2025 Corte: 12/12/2025							
3								
4	1. DATOS GENERALES DEL CORTE							
5	Fecha de corte	12/12/2025	Flota evaluada	33 unidades				
6	Horizonte predictivo	14 días	Modelo predictivo	Random Forest				
7	Umbral alto	$P \geq 0.50$	Umbral medio	$0.25 \leq P < 0.50$				
8	Responsable	Jefatura / Coordinación de mantenimiento	Frecuencia sugerida	Semanal				
9								
10	2. RESUMEN DE PRIORIZACIÓN							
11	RIESGO ALTO		RIESGO MEDIO		RIESGO BAJO		CON ALERTAS ACTIVAS	
12	11		22		0		22	
13	$P \geq 0.50$		$0.25 \leq P < 0.50$		$P < 0.25$		Alertas técnicas registradas	
14								
15	3. GUÍA DE INTERPRETACIÓN							
16	Nivel / criterio	Rango o fuente	Interpretación operativa					
17	Riesgo alto	$P \geq 0.50$	Intervención prioritaria. Revisar la unidad en el plazo definido y validar alertas técnicas.					
18	Riesgo medio	$0.25 \leq P < 0.50$	Inspección programada. Confirmar condición técnica y priorizar según disponibilidad operativa.					
19	Riesgo bajo	$P < 0.25$	Seguimiento rutinario. Mantener programación preventiva institucional según corresponda.					
20	Alertas activas	Reglas técnicas	Orientan el tipo de revisión: uso operativo, historial de inoperatividad, disponibilidad o equipamiento biomédico.					
21								
22	4. PROGRAMACIÓN SEMANAL PRIORIZADA							
23	Unidad	P(inop.)	Riesgo	N.º alertas	Alertas activas	Acción recomendada	Plazo sugerido	Observaciones

Programación semanal

Datos completos

Figura 37. Primera hoja de la plantilla semanal de programación de mantenimiento con soporte predictivo.

Nota. La primera hoja presenta los datos generales del corte, el resumen de priorización, la guía de interpretación y la programación semanal de unidades según nivel de riesgo operativo.

Fuente. Elaboración propia.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	id_ambulancia	ti	prob_inoperatividad	nivel_risco	dias_desde_ultima_inspeccion	downtime_total_dias_w	n_cm_w	disponibilidad_w	equipamiento_funcional	km_cm_w	servicios_cm_w	alertas_activas	instruccion_taller
2	AMB-T2-022	2025-12-12	0.6576	ALTO	13	2.9	3	0.9517	1	7397	101	—	Inspección técnica general
3	AMB-T2-026	2025-12-12	0.6456	ALTO	2	0	1	1	1	9583	131	Servicios >120 (131) Km >9 000 (9583 km)	Inspección técnica general + revisión tren motriz
4	AMB-T2-002	2025-12-12	0.6314	ALTO	13	3.9	3	0.935	1	7471	94	—	Inspección técnica general
5	AMB-T2-006	2025-12-12	0.5823	ALTO	9	10.2	4	0.83	0	5023	33	Disp. <85% (83%) Downtime >7.0d (10.2d) Equipo biomédico no funcional	Inspección técnica general + análisis causa raíz inoperatividad + reemplazo/represión equipo biomédico
6	AMB-T2-032	2025-12-12	0.5677	ALTO	18	16	4	0.7333	0	5755	103	Disp. <85% (73%) Downtime >7.0d (16.0d) Equipo biomédico no funcional	Inspección técnica general + análisis causa raíz inoperatividad + reemplazo/represión equipo biomédico
7	AMB-T2-024	2025-12-12	0.5505	ALTO	4	4.3	3	0.9233	0	7219	115	Equipo biomédico no funcional	Inspección técnica general + reemplazo/represión equipo biomédico
8	AMB-T2-027	2025-12-12	0.5415	ALTO	11	2.7	1	0.955	1	3836	40	—	Inspección técnica general
9	AMB-T2-011	2025-12-12	0.5367	ALTO	10	2.2	3	0.9633	1	4715	81	—	Inspección técnica general
10	AMB-T2-013	2025-12-12	0.5273	ALTO	12	5.6	3	0.9057	0	5247	55	Equipo biomédico no funcional	Inspección técnica general + reemplazo/represión equipo biomédico
11	AMB-T2-029	2025-12-12	0.5175	ALTO	17	0	1	1	1	5284	91	—	Inspección técnica general
12	AMB-T2-015	2025-12-12	0.5075	ALTO	6	3.6	3	0.94	0	4881	95	Equipo biomédico no funcional	Inspección técnica general + reemplazo/represión equipo biomédico
13	AMB-T2-003	2025-12-12	0.4932	MEDIO	16	0	2	1	1	3958	68	—	Inspección dirigida
14	AMB-T2-031	2025-12-12	0.4733	MEDIO	4	21.3	5	0.545	0	5835	116	Disp. <85% (64%) Downtime >7.0d (21.3d) Equipo biomédico no funcional	Inspección dirigida: historial inoperatividad

Figura 38. Segunda hoja de la plantilla semanal de programación de mantenimiento con soporte predictivo.

Nota. La segunda hoja presenta el detalle tabulado de las unidades evaluadas, incluyendo variables operativas, alertas activas y criterios de apoyo para la toma de decisiones de mantenimiento.

Fuente. Elaboración propia.

V.1.7. Valoración técnica complementaria mediante juicio de especialistas

De manera complementaria a la evaluación cuantitativa del modelo computacional, se realizó una valoración técnica de los principales componentes de la propuesta mediante juicio de especialistas. Participaron siete profesionales, seleccionados considerando su formación y experiencia en ámbitos relacionados con ingeniería biomédica, ingeniería clínica, mantenimiento, equipamiento médico, ambulancias y transporte sanitario, gestión de activos, ingeniería electromecánica y proyectos del sector salud.

La valoración comprendió cinco dimensiones: pertinencia de las variables e información utilizada, coherencia del enfoque predictivo-preventivo, parámetros temporales y criterios operativos, priorización y toma de decisiones, y aplicabilidad del protocolo operativo. Adicionalmente, se incluyó una valoración global de la propuesta.

En conjunto, se registraron 252 respuestas correspondientes a los 36 ítems cerrados evaluados por los siete profesionales, de las cuales 251 constituyeron valoraciones válidas y una correspondió a la opción N/A. De las valoraciones válidas, 199 (79,28 %) correspondieron a la categoría “muy pertinente/adecuado” y 50 (19,92 %) a la categoría “pertinente/adecuado con ajustes menores”. Solo dos valoraciones (0,80 %) correspondieron a la categoría “poco pertinente/requiere cambios importantes”, mientras que no se registraron respuestas en la categoría “no pertinente/no adecuado”. En consecuencia, el 99,20 % de las valoraciones válidas correspondió a las categorías favorables 3 y 4, con una mediana global de 4 puntos sobre 4.

Por dimensiones, la coherencia del enfoque predictivo-preventivo, la priorización y toma de decisiones y la aplicabilidad del protocolo operativo alcanzaron un 100 % de valoraciones favorables. La pertinencia de las variables e información utilizada alcanzó 98,21 %, mientras que los parámetros temporales y criterios operativos presentaron 97,62 % de valoraciones favorables, tal como se indica en la Tabla 20.

Tabla 20. Resultados de la valoración técnica complementaria mediante juicio de especialistas

Dimensión evaluada	Mediana	Valoraciones favorables (3–4)	Interpretación
A. Pertinencia de las variables e información utilizada	4	98,21 %	Muy favorable
B. Coherencia del enfoque predictivo-preventivo	4	100,00 %	Muy favorable
C. Parámetros temporales y criterios operativos	4	97,62 %	Muy favorable
D. Priorización y toma de decisiones	4	100,00 %	Muy favorable
E. Aplicabilidad del protocolo operativo	4	100,00 %	Muy favorable
G. Valoración global	4	100,00 %	Muy favorable

Nota. Se consideraron valoraciones favorables las categorías 3 = pertinente/adequado con ajustes menores y 4 = muy pertinente/adequado. Una respuesta registrada como N/A en la dimensión de valoración global fue excluida del cálculo de los indicadores correspondientes.

La dimensión correspondiente a los parámetros temporales y criterios operativos presentó una mayor proporción relativa de valoraciones de nivel 3 respecto de las demás dimensiones. Este comportamiento resulta coherente con el carácter inicial de parámetros como la ventana histórica de 60 días, el horizonte predictivo de 14 días y la frecuencia semanal de actualización, los cuales se plantean como referencias metodológicas susceptibles de recalibración con registros institucionales reales.

En la valoración global, el 100 % de las respuestas válidas se concentró en las categorías pertinente y muy pertinente, con una mediana de 4 puntos. Estos resultados evidencian una apreciación profesional favorable respecto de la pertinencia técnica de las variables consideradas, la coherencia del mantenimiento predictivo como soporte complementario al mantenimiento preventivo y la potencial aplicabilidad del esquema de priorización y del protocolo operativo propuesto. El instrumento aplicado, el perfil de los profesionales participantes, la matriz consolidada de respuestas y el detalle de las observaciones cualitativas se presentan en el **Anexo K**.

CAPÍTULO VI : DISCUSIÓN

La presente investigación tuvo como propósito desarrollar un modelo computacional basado en datos simulados para el mantenimiento predictivo con soporte preventivo de ambulancias médicas urbanas Tipo II en Lima Metropolitana, evaluando su capacidad para anticipar eventos de inoperatividad mediante una estrategia de validación retrospectiva. Los resultados obtenidos se discuten a continuación en relación con los objetivos específicos planteados, la literatura especializada y las implicancias para el sistema prehospitalario peruano. Este enfoque permitió evaluar el potencial del modelo en condiciones controladas antes de una eventual validación con registros institucionales reales.

VI.1. Sobre el enfoque metodológico y la viabilidad del modelo *data-driven*

El primer hallazgo relevante de la investigación es de naturaleza metodológica: la viabilidad de construir un modelo computacional predictivo para ambulancias Tipo II a partir de registros históricos de operación y mantenimiento, sin requerir instrumentación adicional en campo. Este resultado valida el planteamiento central de la investigación y es coherente con lo documentado por Prytz (2014), quien demostró que los métodos de aprendizaje automático aplicados a datos históricos agregados de flotas vehiculares son capaces de generar predicciones útiles para la gestión del mantenimiento sin necesidad de sensores embarcados en tiempo real.

La construcción del dataset mediante ventanas temporales deslizantes con

W = 60 días, T = 14 días y paso semanal permitió generar 3 102 observaciones ambulancia-corte temporal a partir de una flota simulada de 33 unidades. Este resultado evidencia que incluso una flota de tamaño moderado puede generar un volumen de información estructurada suficiente para demostrar la viabilidad metodológica de modelos de clasificación binaria supervisada, siempre que existan registros consistentes de operación, mantenimiento e inoperatividad. Este aspecto es especialmente relevante en el contexto del sistema prehospitalario peruano, donde la disponibilidad de infraestructura tecnológica avanzada es limitada pero los registros operativos existen y se generan de forma continua (MINSA, 2006; Defensoría del Pueblo, 2023a).

VI.2. Sobre el desempeño del modelo y su comparación con la línea base

La comparación de los cinco enfoques evaluados evidencia distintos perfiles de desempeño operativo. La Regresión Logística alcanza una sensibilidad máxima, pero lo hace a costa de una especificidad prácticamente nula y una disponibilidad proyectada extremadamente baja, lo que indica que clasifica casi todas las observaciones como positivas. En consecuencia, aunque detecta todos los eventos de inoperatividad, su utilidad práctica es limitada debido al volumen excesivo de falsas alarmas.

El Árbol de Decisión presentó un comportamiento intermedio, con una sensibilidad de 55,52%, especificidad de 42,27% y AUC-ROC de 50,46%. Este resultado indica que un modelo individual basado en reglas jerárquicas mejora la sensibilidad frente al esquema preventivo tradicional, aunque su capacidad discriminativa global fue apenas superior al azar y menor que la observada en los modelos de ensamble. Esta diferencia respalda la utilidad de *Random Forest* y *Gradient Boosting* como enfoques basados en múltiples árboles para mejorar la detección de eventos de inoperatividad respecto a un árbol individual.

El *Random Forest* presentó el mejor desempeño relativo para la finalidad operativa de la investigación, al alcanzar una sensibilidad de 91,86% y el mayor AUC-ROC entre los modelos evaluados, con 56,03%. Sin embargo, este resultado debe interpretarse con cautela, debido a que la precisión fue limitada 21,18% y la especificidad reducida 9,95%, lo que indica una alta proporción de falsos positivos. Este comportamiento es consistente con el objetivo metodológico de priorizar la detección temprana de eventos de inoperatividad mediante un umbral conservador $\tau = 0,30$. Aunque su especificidad y exactitud son menores que las del esquema preventivo tradicional, esta reducción responde al umbral adoptado y al objetivo operativo de priorizar la detección temprana de inoperatividades.

Por su parte, el *Gradient Boosting* mostró un comportamiento más equilibrado en términos de exactitud y especificidad, pero con menor sensibilidad que el *Random Forest*. Esta diferencia es relevante porque, en un sistema crítico como el prehospitalario, el error de mayor consecuencia operativa corresponde a la omisión de una unidad que posteriormente quedará inoperativa. Por esta razón, la selección del *Random Forest* como modelo principal se sustenta en un criterio de utilidad operativa: maximizar la capacidad de anticipación de eventos de inoperatividad, aun cuando ello implique aceptar un mayor número de falsos positivos.

Esta interpretación cualitativa de la asimetría entre sensibilidad y especificidad, central en la justificación de *Random Forest* como modelo principal, encuentra sustento cuantitativo en el análisis estadístico formal desarrollado en la Sección V.1.3.1.

En primer lugar, la prueba de McNemar (Ecuación 39) indica que la diferencia entre el patrón de aciertos y errores de *Random Forest* y el de cada uno de los enfoques de comparación, es decir, el esquema preventivo tradicional, la Regresión Logística, el Árbol de Decisión y *Gradient Boosting*, es estadísticamente significativa en los cuatro casos ($p < 0,0001$; Tabla 17). En particular, frente a *Gradient Boosting*, el modelo de ensamble más cercano en desempeño global según la Tabla 16, el estadístico $\chi^2 = 156,30$ evidencia que la mayor sensibilidad de *Random Forest* (91,86% frente a

65,41%) no constituye una diferencia atribuible a la variabilidad muestral del conjunto de validación, sino un comportamiento sistemáticamente distinto bajo el umbral operativo $\tau = 0,30$.

En segundo lugar, la prueba de DeLong (Ecuación 40) indica que la capacidad discriminativa global de *Random Forest*, medida mediante el AUC-ROC e independiente del umbral de decisión, es también estadísticamente superior a la de *Gradient Boosting* ($z = 2,178$; $p = 0,029$), del Árbol de Decisión ($z = 2,789$; $p = 0,005$) y del esquema preventivo tradicional ($z = 4,579$; $p < 0,0001$; Tabla 18). Esto confirma que la selección de *Random Forest* no se sustenta exclusivamente en su comportamiento bajo un umbral específico, sino en una capacidad de ordenamiento de riesgo superior de manera consistente frente a los modelos de ensamble y los enfoques basados en árboles.

La única comparación en la que la prueba de DeLong no resultó significativa fue *Random Forest* frente a Regresión Logística ($z = 1,508$; $p = 0,132$), lo que indica que ambos modelos poseen una capacidad discriminativa global estadísticamente equivalente. No obstante, esta equivalencia en AUC-ROC contrasta con el resultado de la prueba de McNemar para el mismo par de modelos ($\chi^2 = 64,56$; $p < 0,0001$), que sí detecta una diferencia sistemática en la tasa de clasificación correcta bajo $\tau = 0,30$. Esta aparente disociación entre ambos resultados no constituye una contradicción, sino que refleja precisamente el comportamiento operacionalmente degenerado de la Regresión Logística descrito al inicio de esta sección. Su ordenamiento de las observaciones por probabilidad de riesgo es, en términos globales, comparable al de *Random Forest* (de allí la equivalencia en AUC-ROC), pero su decisión bajo el umbral adoptado, al clasificar la totalidad de las observaciones como positivas, constituye un comportamiento cualitativamente distinto, operativamente inviable para un esquema de soporte preventivo, y que la prueba de McNemar identifica como una diferencia sistemática.

Finalmente, el cálculo de la precisión media (Ecuación 41) proporciona una tercera línea de evidencia, obtenida bajo un criterio metodológico

independiente del umbral fijo y del espacio ROC. *Random Forest* obtuvo el valor de AP más alto entre los cinco enfoques evaluados (AP = 0,2462), superando a *Gradient Boosting* (AP = 0,2187) y al valor esperado bajo un clasificador no informativo dada la prevalencia observada (AP = 0,2085). Este resultado indica que, integrando la relación entre precisión y sensibilidad a lo largo de todos los umbrales posibles, y no únicamente en $\tau = 0,30$; *Random Forest* mantiene una ventaja relativa sobre los demás enfoques que es consistente con los resultados de las pruebas de McNemar y DeLong.

En conjunto, estas tres líneas de evidencia, es decir, la significancia de McNemar frente a los cuatro enfoques de comparación, la significancia de DeLong frente a tres de los cuatro enfoques con la excepción interpretable de la Regresión Logística, y el valor más alto de precisión media entre los cinco modelos evaluados, proporcionan un sustento matemático e inferencial a la decisión de seleccionar *Random Forest* como modelo predictivo principal, complementando la justificación cualitativa basada en la priorización de la sensibilidad operativa frente al costo de los falsos negativos en un sistema prehospitalario crítico.

La exactitud, la especificidad y la precisión no fueron descartadas, sino interpretadas como métricas complementarias. La exactitud permitió evaluar el desempeño global del clasificador; sin embargo, en escenarios desbalanceados puede ofrecer una visión optimista del rendimiento al favorecer modelos que clasifican correctamente la clase mayoritaria sin garantizar una detección adecuada de la clase positiva. La especificidad permitió cuantificar la capacidad de reconocer unidades que permanecerían operativas, mientras que la precisión permitió estimar qué proporción de alertas correspondía efectivamente a eventos reales de inoperatividad. En conjunto, estas métricas permitieron medir el costo operativo asociado al incremento de sensibilidad, especialmente en términos de falsos positivos. Esta interpretación es coherente con Hand et al. (2024), quienes sostienen que la medida de desempeño debe seleccionarse en función del propósito de la aplicación.

Estos resultados son consistentes con los hallazgos de Molęda et al. (2023), quienes documentaron que los enfoques predictivos basados en datos pueden superar a los esquemas preventivos programados en capacidad de anticipación de fallas, especialmente cuando se trabaja con datos operativos heterogéneos. Asimismo, Taoufyq et al. (2025) señalaron que los modelos *data-driven* presentan ventajas particulares en entornos con restricciones técnicas y presupuestales, precisamente el contexto del sistema prehospitalario público peruano.

Los valores de AUC-ROC obtenidos, cercanos a 0,55, deben interpretarse como una capacidad discriminativa moderada. Este resultado es esperable considerando que el estudio no emplea telemetría ni variables de condición en tiempo real, sino registros históricos simulados de operación, mantenimiento e inoperatividad. En aplicaciones vehiculares con monitoreo continuo de condición, la disponibilidad de señales más ricas, como: vibración, temperatura, presión o telemetría, puede incrementar la capacidad predictiva de los modelos, como se reporta en revisiones del sector transporte y automotriz (Ersöz et al., 2022).

Los valores de las métricas obtenidas deben interpretarse de forma integrada y en el contexto de los condicionantes del estudio, no como indicadores aislados de calidad predictiva. En particular, la baja precisión del *Random Forest* refleja el número elevado de alertas generadas como consecuencia del umbral conservador de clasificación. Este comportamiento es metodológicamente coherente con sistemas de tamizaje o alerta temprana: cuando el costo de no detectar una falla es mayor que el costo de una inspección preventiva innecesaria, se acepta una precisión baja a cambio de una sensibilidad alta. En el contexto prehospitalario, donde una ambulancia fuera de servicio puede comprometer la respuesta a una emergencia médica, esta lógica es operativamente justificable.

En este mismo sentido, la disponibilidad proyectada por el modelo no debe interpretarse como un pronóstico directo de la disponibilidad real de la flota, sino como un indicador derivado del umbral de decisión adoptado. Con un umbral de $\tau = 0,30$; el modelo clasifica como “en riesgo” a toda

observación cuya probabilidad estimada de inoperatividad supere ese valor, resultando en una clasificación deliberadamente conservadora. Si se elevara el umbral de clasificación, la disponibilidad proyectada aumentaría y se reducirían las falsas alarmas, pero a costa de disminuir la sensibilidad y perder eventos de inoperatividad que el modelo habría podido anticipar.

Con relación al tratamiento del desbalance de clases, cabe precisar que la estrategia de compensación empleada fue el oversampling aleatorio con reemplazo sobre la clase minoritaria durante el entrenamiento, complementado con el parámetro *class_weight='balanced'* en los modelos de Regresión Logística y *Random Forest* como medida de robustez. Esta decisión se tomó con el objetivo de mantener un pipeline de entrenamiento simple y reproducible, coherente con el carácter metodológico y exploratorio de la investigación. En estudios futuros con datos reales, la evaluación de técnicas más sofisticadas de balanceo, como SMOTE, *undersampling* controlado o ajuste sistemático de umbrales, podría mejorar las métricas de precisión y F1-Score del modelo sin comprometer su capacidad de anticipación.

Debe precisarse que los resultados obtenidos no constituyen una estimación del desempeño real del SAMU, sino una demostración metodológica sobre un dataset simulado, reproducible y calibrado con parámetros de referencia. Por tanto, la contribución principal del estudio radica en la formulación, implementación y validación retrospectiva del marco computacional, mientras que la validación externa con datos institucionales reales constituye el paso necesario para determinar su desempeño operativo definitivo.

VI.3. Sobre la importancia de variables y sus implicancias para la gestión

El análisis de importancia de variables del modelo *Random Forest* reveló que las variables de uso operativo, representadas por el kilometraje

acumulado y los servicios prestados en la ventana W, concentran aproximadamente el 35,4% de la importancia relativa total del modelo, constituyendo los factores más determinantes de inoperatividad futura. Este hallazgo tiene una implicancia práctica directa para la gestión del mantenimiento del SAMU: el criterio de programación de intervenciones no debería basarse exclusivamente en el tiempo calendario o el kilometraje acumulado desde la última intervención, sino incorporar la intensidad de uso reciente de cada unidad como variable de priorización.

Este resultado es coherente con los fundamentos de la gestión centrada en la confiabilidad descritos por Moubray (1997), quien argumentó que los programas de mantenimiento basados exclusivamente en intervalos fijos asumen erróneamente que el deterioro de los activos sigue patrones uniformes independientemente de las condiciones reales de operación. En el caso de ambulancias urbanas, donde la variabilidad de uso entre unidades puede ser considerable dependiendo de la zona de cobertura, el turno de operación y la demanda de emergencias, esta heterogeneidad no puede ser capturada por un esquema de mantenimiento uniforme.

La variable de días desde la última intervención, que ocupa el tercer lugar en importancia con 12,21%, confirma la relevancia del intervalo de mantenimiento como factor de riesgo, pero relativiza su peso frente al uso operativo. Esto sugiere que una ambulancia con alto uso reciente pero intervención reciente tiene mayor probabilidad de inoperatividad que una con uso moderado e intervención más antigua, lo que contradice la lógica del mantenimiento preventivo tradicional y respalda la necesidad de un enfoque adaptativo que integre ambos criterios.

VI.4. Sobre el soporte preventivo y su integración al sistema prehospitalario

El prototipo de soporte preventivo desarrollado como Entregable 2 de la investigación demuestra que los resultados del modelo pueden traducirse en

información accionable para la toma de decisiones técnicas sin requerir conocimientos especializados en ciencia de datos por parte del personal operativo. La clasificación de ambulancias en tres niveles de riesgo (alto, medio y bajo) con criterios de acción específicos y plazos definidos configura una herramienta de gestión operativa directamente aplicable al flujo de trabajo del SAMU.

Este enfoque es consistente con el planteamiento de Taoufyq et al. (2025), quienes señalaron que el valor del mantenimiento predictivo reside en la capacidad de transformar datos históricos en información útil para la toma de decisiones, y con los principios de la ISO 17359:2018, que encuadra el pronóstico como una actividad orientada a respaldar decisiones de mantenimiento a partir de información disponible. La integración del modelo como soporte al mantenimiento preventivo existente, sin pretender sustituirlo, es coherente con la perspectiva de Moubray (1997) sobre la gestión de activos críticos en sistemas de salud, donde la continuidad del plan preventivo es un requisito normativo establecido en la NTS N.º 051-MINSA/OGDN-V.01 (MINSA, 2006).

El protocolo operativo de seis pasos propuesto en los lineamientos técnicos responde directamente a las restricciones institucionales del sistema prehospitalario público: no requiere inversión en infraestructura, puede ejecutarse con recursos computacionales básicos y se integra al ciclo semanal de gestión del mantenimiento sin alterar su estructura organizativa. Esta característica distingue la presente propuesta de enfoques basados en monitoreo continuo con sensores o telemetría embarcada, cuya implementación suele requerir inversión adicional en infraestructura tecnológica, conectividad, mantenimiento de dispositivos y gestión de datos en tiempo real. En contraste, el enfoque propuesto aprovecha registros históricos estructurados y puede implementarse progresivamente con recursos computacionales básicos, lo que lo hace más viable para una primera etapa de adopción en el contexto del sistema público de salud.

En este sentido, la propuesta trasciende el ámbito estrictamente computacional y plantea una forma de incorporar los resultados del modelo

al proceso habitual de planificación del mantenimiento. La utilidad del enfoque depende, por tanto, no solo de su capacidad predictiva, sino también de su integración práctica en la gestión operativa de la flota.

La contribución del modelo adquiere una dimensión adicional cuando se analiza desde la perspectiva de la capacidad referencial de cobertura del sistema prehospitalario. En el planteamiento de la investigación se señaló que, para países de ingresos medios, se recomienda disponer al menos de una ambulancia por cada 50 000 habitantes como referencia mínima de dotación (Mishra et al., 2020). Bajo este criterio, la inoperatividad de una ambulancia no debe interpretarse de manera directa como la ausencia absoluta de cobertura para una población determinada, ya que la respuesta prehospitalaria depende también de la distribución territorial de las bases, la disponibilidad de personal, la demanda horaria, las condiciones de tráfico y los tiempos de desplazamiento. Sin embargo, sí representa una reducción de la capacidad referencial de respuesta del sistema.

Desde esta perspectiva, el modelo propuesto no estima cobertura poblacional ni tiempos de respuesta, pero contribuye indirectamente a preservar la capacidad operativa disponible al anticipar unidades con mayor probabilidad de inoperatividad. En contextos donde la dotación de ambulancias es limitada y la disponibilidad real se ve afectada por fallas técnicas documentadas, la identificación temprana de unidades en riesgo permite priorizar intervenciones antes de que ocurran salidas de servicio no planificadas. Por tanto, el valor sanitario del modelo no radica únicamente en clasificar observaciones como operativas o inoperativas, sino en ofrecer una herramienta de apoyo para reducir la pérdida evitable de capacidad de respuesta prehospitalaria.

En el *dataset* de validación del período 2025, el conjunto de validación contiene 344 observaciones con inoperatividad ($Y = 1$), correspondientes al 20,8% del total de 1 650 observaciones. El modelo *Random Forest* anticipó correctamente 316 de esos 344 eventos, frente a 105 identificados por el esquema preventivo tradicional bajo las mismas condiciones. Esta diferencia de 211 eventos adicionales anticipados ilustra la capacidad diferencial del

enfoque predictivo para adelantar intervenciones antes de que se produzca la salida de servicio, aunque debe interpretarse considerando el incremento simultáneo de falsos positivos asociado al umbral conservador de clasificación.

De manera complementaria, la valoración realizada por profesionales con experiencia relacionada mostró una apreciación favorable respecto de la pertinencia de las variables consideradas, la coherencia del enfoque predictivo-preventivo y la aplicabilidad potencial del esquema de priorización y del protocolo operativo propuesto. Este resultado aporta una perspectiva profesional adicional a los lineamientos derivados computacionalmente y respalda su plausibilidad técnica como herramientas de soporte a la toma de decisiones. No obstante, dicha valoración no constituye evidencia de desempeño operacional del modelo, por lo que se mantiene la necesidad de una etapa posterior de calibración y validación mediante registros institucionales reales.

VI.5. Sobre la contrastación de hipótesis

La hipótesis general se sustenta dentro del alcance metodológico del estudio: el modelo computacional desarrollado mostró mejor desempeño relativo que el esquema preventivo tradicional en la anticipación de eventos de inoperatividad, especialmente en sensibilidad. Este resultado constituye evidencia de utilidad potencial para apoyar la gestión técnica; no obstante, la optimización efectiva de la disponibilidad operativa real requiere una implementación piloto con registros institucionales del SAMU. La optimización definitiva de la disponibilidad operativa medida sobre indicadores MTBF, MTTR y disponibilidad real de flota requiere implementación con datos reales, condición que delimita el alcance del presente estudio y define la agenda de investigación inmediata, conforme a lo declarado en las limitaciones.

La H1 encuentra sustento en la identificación de 14 variables técnicas y

operativas con capacidad predictiva sobre la inoperatividad, siendo las de uso operativo las más determinantes. La H2 se sustenta en la construcción exitosa del *dataset* estructurado que permitió el entrenamiento y validación del modelo. La H3 se sustenta en la demostración de que el modelo de clasificación supervisada estima la probabilidad de inoperatividad con capacidad de anticipación superior al esquema preventivo. La H4 se sustenta parcialmente en los resultados cuantitativos obtenidos: el modelo *Random Forest* superó a la línea base preventiva en sensibilidad y AUC-ROC, mostrando mejor desempeño relativo para la anticipación de eventos de inoperatividad en las condiciones declaradas del estudio. Sin embargo, la precisión y la especificidad limitadas evidencian que el desempeño no debe interpretarse como alto en términos absolutos, sino como útil para un esquema de alerta temprana y priorización preventiva. La H5 encuentra sustento en los lineamientos técnicos derivados del modelo, que ofrecen criterios concretos para mejorar la priorización de intervenciones y la planificación del mantenimiento.

VI.6. Limitaciones y perspectivas de mejora

La limitación principal de la investigación es el origen simulado de los datos. Si bien la simulación fue calibrada con parámetros documentados en la normativa y la literatura técnica, los datos simulados no capturan la complejidad operativa real del SAMU en Lima Metropolitana: variabilidad geográfica y operacional, antigüedad de la flota y patrones de demanda estacional, entre otros factores. La replicación del estudio con datos reales del SAMU constituye la prioridad de investigación inmediata y permitiría evaluar el desempeño del modelo bajo condiciones operativas auténticas, así como recalibrar variables, umbrales y criterios de priorización.

Una segunda limitación es la ausencia de variables de condición técnica en tiempo real. La incorporación de variables adicionales, como kilometraje actualizado mediante GPS, telemetría vehicular, historial de reparaciones por componente específico, antigüedad de cada subsistema, condiciones de

operación, ubicación de bases y demanda horaria, permitiría enriquecer el *dataset* y evaluar si mejora la capacidad predictiva del modelo, acercándolo a los resultados obtenidos por Ersöz et al. (2022) en el sector automotriz con datos de monitoreo continuo.

Finalmente, el alcance de la investigación se limitó a ambulancias Tipo II del SAMU en Lima Metropolitana durante el período 2024-2025. La extensión de la metodología a otras regiones del país, a otros tipos de ambulancias o a otros vehículos críticos del sistema de salud representa una línea de trabajo futuro con potencial de impacto directo en la salud pública nacional, alineada con los Objetivos de Desarrollo Sostenible de las Naciones Unidas (2015) en materia de salud y bienestar e innovación en infraestructura. Asimismo, la valoración profesional complementaria se realizó mediante un panel intencional y de tamaño acotado, por lo que sus resultados representan juicio técnico especializado y no una inferencia generalizable a la totalidad de profesionales o instituciones vinculadas al mantenimiento de ambulancias.

CAPÍTULO VII : CONCLUSIONES

Las conclusiones se presentan en correspondencia directa con cada uno de los cinco objetivos específicos de la investigación, seguidas de una conclusión general que sintetiza el aporte de la tesis al campo de la ingeniería biomédica y la gestión del sistema prehospitalario. Esta organización permite mantener la trazabilidad entre los objetivos planteados, los resultados obtenidos y las conclusiones derivadas del análisis. Asimismo, facilita verificar el cumplimiento progresivo del propósito general de la investigación.

VII.1. Conclusión 1 - Sobre el objetivo específico 1: Análisis de factores técnicos y operativos

El análisis del *dataset* de las ambulancias médicas urbanas Tipo II permitió identificar 14 variables técnicas y operativas con capacidad predictiva sobre la inoperatividad futura de las unidades. Las variables de uso operativo, representadas por el kilometraje acumulado y los servicios prestados en la ventana histórica de 60 días, constituyen los factores de mayor importancia predictiva según el análisis del modelo *Random Forest*, concentrando aproximadamente el 35,4% de la importancia relativa total del modelo.

Este hallazgo evidencia que la intensidad de uso reciente es el predictor más determinante de inoperatividad futura, por encima del intervalo de tiempo transcurrido desde la última intervención de mantenimiento, lo que cuestiona la suficiencia del esquema preventivo tradicional basado exclusivamente en criterios temporales o de kilometraje acumulado total, el cual es expuesto como línea base de comparación.

VII.2. Conclusión 2 - Sobre el objetivo específico 2: Diseño de la base de datos

Se diseñó y construyó una base de datos sintética estructurada mediante simulación paramétrica, compuesta por tres tablas fuente (registros de mantenimiento, registros de inoperatividad y registros de uso operativo), cuya integración mediante ventanas temporales deslizantes generó un dataset de 3 102 observaciones ambulancia-corte temporal con 14 variables explicativas y una variable objetivo binaria. La simulación empleó distribuciones de probabilidad calibradas con parámetros de referencia a partir de la NTS N.º 051-MINSA y la literatura técnica, con semilla fija de reproducibilidad, lo que garantiza la trazabilidad y replicabilidad del proceso. La arquitectura del dataset es transferible a datos reales del SAMU, requiriendo únicamente la estandarización de identificadores, fechas y codificación de eventos.

VII.3. Conclusión 3 - Sobre el objetivo específico 3: Implementación del modelo computacional

Se implementó un modelo computacional de clasificación binaria supervisada mediante la evaluación comparativa de cuatro algoritmos: Regresión Logística, Árbol de Decisión, *Random Forest* y *Gradient Boosting*. El modelo *Random Forest* fue seleccionado como modelo predictivo principal, entrenado con registros del período 2024 y validado con datos del período 2025 bajo una estrategia de *backtesting* temporal. El modelo demostró capacidad para estimar la probabilidad de inoperatividad funcional en un horizonte de 14 días para cada ambulancia de la flota simulada. La implementación se realizó íntegramente en entorno Python estándar, sin requerir infraestructura especializada ni sensores en campo, lo que confirma la viabilidad técnica del enfoque en el contexto del sistema prehospitalario público peruano. Este resultado debe interpretarse como una

demostración de viabilidad metodológica a escala de simulación, pendiente de validación con datos institucionales reales.

VII.4. Conclusión 4 - Sobre el objetivo específico 4: Validación del modelo

El desempeño del modelo fue evaluado mediante un conjunto ampliado de métricas cuantitativas: exactitud, especificidad, precisión, sensibilidad, F1-Score, AUC-ROC, MAE y disponibilidad operativa proyectada. Frente a la línea base preventiva, *Random Forest* incrementó la sensibilidad en 61,34 puntos porcentuales, pasando de 30,52% a 91,86%. Esta diferencia evidencia una mayor capacidad de detección temprana de eventos de inoperatividad, aunque acompañada de una precisión limitada de 21,18%, una especificidad reducida de 9,95% y una menor disponibilidad proyectada debido al umbral conservador de clasificación $\tau = 0,30$. Aunque su exactitud y especificidad fueron menores que las del esquema preventivo tradicional, este comportamiento responde al umbral conservador adoptado y a la prioridad operativa de reducir falsos negativos en un sistema crítico. Dado el origen sintético de los datos, estos resultados constituyen una evaluación de la validez interna del enfoque metodológico y no deben interpretarse como indicadores de desempeño operativo definitivo.

VII.5. Conclusión 5 - Sobre el objetivo específico 5: Lineamientos técnicos y operativos

Se propusieron lineamientos técnicos y operativos estructurados en tres componentes derivados del modelo computacional. El primero establece criterios de priorización de intervenciones diferenciados por nivel de riesgo estimado, con plazos de acción específicos para cada categoría. El segundo define seis umbrales de alerta temprana basados en las variables de mayor importancia predictiva, orientado a activar revisiones

técnicas independientemente del ciclo preventivo programado. El tercero propone un protocolo operativo semanal de seis pasos para la integración del modelo en el flujo de gestión del SAMU, sin requerir modificaciones estructurales al sistema existente. Estos lineamientos constituyen una contribución metodológica concreta para fortalecer la gestión técnica, orientar la priorización de intervenciones y contribuir potencialmente a la mejora de la disponibilidad operativa de las ambulancias Tipo II en Lima Metropolitana, pendiente de validación con datos institucionales reales.

De manera complementaria, el juicio de expertos realizado con siete profesionales mostró una valoración altamente favorable de los principales componentes de la propuesta, con 99,20 % de las valoraciones válidas concentradas en las categorías pertinente y muy pertinente y una mediana global de 4 puntos sobre 4. La valoración respaldó especialmente la pertinencia de las variables consideradas, la coherencia de integrar el enfoque predictivo como soporte al mantenimiento preventivo y la utilidad potencial del esquema de priorización y del protocolo operativo.

Estos resultados aportan una valoración profesional complementaria sobre la pertinencia técnica de los lineamientos propuestos, manteniéndose como etapa posterior necesaria su calibración y validación mediante registros institucionales reales.

VII.6. Conclusión general

La presente investigación desarrolló y evaluó un modelo computacional de mantenimiento predictivo con soporte preventivo para ambulancias médicas urbanas Tipo II en Lima Metropolitana, construido sobre un dataset sintético generado mediante simulación paramétrica reproducible y calibrado con parámetros de referencia a partir de la NTS N.º 051-MINSA y la literatura técnica. El modelo *Random Forest* mostró el mejor desempeño relativo para la anticipación de inoperatividades, al alcanzar una sensibilidad de 91,86%

frente al 30,52% de la línea base preventiva durante el período de validación 2025. Asimismo, obtuvo el mayor AUC-ROC entre los modelos evaluados, con 56,03%, aunque con una precisión limitada de 21,18% y una especificidad reducida de 9,95%, lo que confirma que su utilidad principal corresponde a un esquema de alerta temprana y priorización preventiva, no a una clasificación definitiva de fallas.

El resultado demuestra que el enfoque *data-driven* es metodológicamente viable para este tipo de sistema, ya sea a partir de registros históricos institucionales de operación y mantenimiento o, en una etapa preliminar, mediante simulaciones calibradas y reproducibles. La contribución principal de la investigación es de naturaleza metodológica: se documenta un pipeline computacional reproducible, una arquitectura de *dataset* transferible y un conjunto de lineamientos técnicos operacionalizables. Los resultados obtenidos son preliminares dado el origen simulado de los datos y deben considerarse como punto de partida para una etapa de validación con registros históricos reales del SAMU, condición necesaria para cualquier implementación operativa. El marco metodológico desarrollado es extensible a otros tipos de unidades médicas móviles y otras regiones del país, contribuyendo al fortalecimiento de la gestión técnica de activos críticos en el sistema prehospitario peruano.

CAPÍTULO VIII : RECOMENDACIONES

Las recomendaciones de esta investigación se dirigen a tres audiencias: el SAMU como institución operadora de ambulancias urbanas Tipo II en Lima Metropolitana, la comunidad académica y la Escuela Profesional de Ingeniería Biomédica de la UNMSM.

VIII.1. Recomendación 1 - Al SAMU: gestión y estandarización de registros

Se recomienda al Sistema de Atención Móvil de Urgencias implementar un sistema de registro estandarizado de eventos de mantenimiento, inoperatividad y uso operativo para cada ambulancia de la flota, con identificadores únicos por unidad, trazabilidad temporal precisa y codificación uniforme de tipos de evento. Esta estandarización podría implementarse inicialmente con herramientas de bajo costo y constituye la condición necesaria para la aplicación del modelo predictivo con datos reales, incrementando progresivamente su capacidad predictiva con cada ciclo de operación.

VIII.2. Recomendación 2 - Al SAMU: implementación piloto del modelo

Se recomienda realizar una implementación piloto del modelo computacional con los registros históricos disponibles de la flota de Lima Metropolitana, comenzando con un subconjunto de unidades para evaluar su desempeño

en condiciones reales antes de una adopción a escala. Esta implementación permitiría ajustar los parámetros del modelo: ventana histórica, horizonte de predicción y umbrales de riesgo, a los patrones operativos auténticos del sistema, evaluar su capacidad predictiva real y validar los lineamientos técnicos propuestos en condiciones reales de operación.

VIII.3. Recomendación 3 - A la comunidad académica: replicación con datos reales

Se recomienda replicar la metodología propuesta con datos reales del SAMU obtenidos mediante los mecanismos de acceso a la información pública establecidos en la Ley N.º 27806, a fin de establecer el desempeño definitivo del modelo en condiciones operativas auténticas y comparar los resultados con los obtenidos en la presente investigación con datos simulados. Esta replicación constituye la prioridad de investigación más inmediata para consolidar el aporte científico de la tesis.

VIII.4. Recomendación 4 - A la comunidad académica: enriquecimiento del modelo

Se recomienda explorar el enriquecimiento del modelo mediante la incorporación de variables adicionales no consideradas en la presente investigación, incluyendo la antigüedad de cada subsistema, el historial de repuestos utilizados por componente, las condiciones de tráfico urbano por zona de cobertura y los patrones de demanda estacional de emergencias.

La incorporación de estas variables adicionales, en caso de encontrarse disponibles o poder ser estandarizadas en los registros institucionales del SAMU o del MINSA, permitiría capturar dimensiones del deterioro operativo que el modelo actual no considera y evaluar el incremento de la capacidad predictiva en condiciones reales con flotas de ambulancias urbanas Tipo II, hasta incluso considerar otros tipos de ambulancias como las rurales..

VIII.5. Recomendación 5 - A la Escuela Profesional de Ingeniería Biomédica

Se recomienda a la EPIBM de la UNMSM incorporar la gestión técnica de activos críticos en el sistema prehospitalario como línea de investigación activa, promoviendo el desarrollo de herramientas computacionales orientadas a fortalecer la disponibilidad operativa de las ambulancias y otros equipos médicos críticos del sistema de salud público peruano. La presente investigación ofrece un marco metodológico replicable que puede extenderse a ambulancias Tipo I y Tipo III, a otras regiones del país y a otros activos del sistema prehospitalario, contribuyendo al fortalecimiento del perfil investigativo de la ingeniería biomédica en el contexto de la salud pública nacional.

CAPÍTULO IX : BIBLIOGRAFÍA

Adem, M. A., Tezera, Z. B., & Agegnehu, C. D. (2024). The practice and determinants of ambulance service utilization in pre-hospital settings, Jimma City, Ethiopia. *BMC Emergency Medicine*, 24(81). <https://doi.org/10.1186/s12873-024-00999-8>

Aguilar Sota, L. M. (2024). *Gestión de mantenimiento basado en la confiabilidad (RCM), para incrementar la disponibilidad de la flota vehicular de la red de servicios de salud Cusco Norte-2022* [Tesis de pregrado, Universidad Nacional de San Antonio Abad del Cusco]. Repositorio UNSAAC. <http://hdl.handle.net/20.500.12918/9862>

Arena, S., Collotta, M., Luca, L., Pagano, M., & Tona, G. L. (2022). Predictive maintenance in the automotive sector: A literature review. *Mathematical and Computational Applications*, 27(1), 2. <https://doi.org/10.3390/mca27010002>

Barlow, R. E., & Proschan, F. (1965). *Mathematical theory of reliability*. John Wiley & Sons.

Bhattarai, H. K., Bhusal, S., Barone-Adesi, F., & Hubloue, I. (2023). Prehospital emergency care in low- and middle-income countries: A systematic review. *Prehospital and Disaster Medicine*, 38(4), 495–512. <https://doi.org/10.1017/S1049023X23006088>

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Blackwell, T. H., & Kaufman, J. S. (2002). Response time effectiveness: Comparison of response time and survival in an urban emergency medical services system. *Academic Emergency Medicine*, 9(4), 288–295. <https://doi.org/10.1197/aemj.9.4.288>

Boppana, V. R. (2023). *Data analytics for predictive maintenance in healthcare equipment*. SSRN. <https://doi.org/10.2139/ssrn.5005057>

Branco, P., Torgo, L., & Ribeiro, R. P. (2016). A survey of predictive modelling under imbalanced distributions. *ACM Computing Surveys*, 49(2), Article 31. <https://doi.org/10.1145/2907070>

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and regression trees*. Wadsworth.

Bürkner, P.-C., Gabry, J., & Vehtari, A. (2020). Approximate leave-future-out cross-validation for Bayesian time series models. *Journal of Statistical Computation and Simulation*, 90(14), 2499–2523. <https://doi.org/10.1080/00949655.2020.1783262>

Carrasco, J., López, D., Aguilera-Martos, I., García-Gil, D., Markova, I., García-Barzana, M., Arias-Rodil, M., Luengo, J., & Herrera, F. (2021). Anomaly detection in predictive maintenance: A new evaluation framework for temporal unsupervised anomaly detection algorithms. *Neurocomputing*, 462, 440–452. <https://doi.org/10.1016/j.neucom.2021.07.095>

Central de Compras Públicas – PERÚ COMPRAS. (2024, 2 de diciembre). *Ambulancias urbanas con equipamiento médico tipo I, II y III* [Prepublicación de proyectos de fichas de homologación]. Plataforma del Estado Peruano. <https://www.gob.pe/institucion/perucompras/informes-publicaciones/6247922-ambulancias-urbanas-con-equipamiento-medico-tipo-i-ii-y-iii>

Cerqueira, V., Torgo, L., & Mozetič, I. (2020). Evaluating time series forecasting models: An empirical study on performance estimation methods. *Machine Learning*, 109, 1997–2028. <https://doi.org/10.1007/s10994-020-05910-7>

Cimino, J., & Braun, C. (2023). Clinical research in prehospital care: Current and future challenges. *Clinical Practice*, 13, 1266–1285. <https://doi.org/10.3390/clinpract13050114>

Comité Europeo de Normalización. (2020). *UNE-EN 1789:2020 - Vehículos de transporte sanitario y su equipamiento - Ambulancias de carretera*. AENOR.

Commission on Accreditation of Ambulance Services. (2016). *Ground vehicle standard for ambulances: v.1.0 edition*. <https://www.groundvehiclestandard.org>

Congreso de la República del Perú. (2002). *Ley N.º 27806, Ley de Transparencia y Acceso a la Información Pública*. Diario Oficial El Peruano. <https://www.gob.pe/institucion/congreso-de-la-republica/normas-legales/118374-27806>

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2), 187–220.

Defensoría del Pueblo. (2023a, junio 7). *Defensoría del Pueblo supervisó funcionamiento del Sistema de Atención Móvil de Urgencias (SAMU)*. <https://www.defensoria.gob.pe/defensoria-superviso-funcionamiento-del-sistema-de-atencion-movil-de-urgencias-samu/>

Defensoría del Pueblo. (2023b, julio 31). *Defensoría del Pueblo advierte demora en arreglo de ambulancia y desabastecimiento de medicinas en centro de salud de Chavín Pariarca en Huánuco*. <https://www.defensoria.gob.pe/defensoria-del-pueblo-advierte-demora-en-arreglo-de-ambulancia-y-desabastecimiento-de-medicinas-en-centro-de-salud-de-chavin-pariarca-en-huanuco/>

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44(3), 837–845. <https://doi.org/10.2307/2531595>

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7, 1–30.

Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, 10(7), 1895–1923. <https://doi.org/10.1162/089976698300017197>

Ersöz, O. Ö., İnal, A. F., Aktepe, A., Türker, A. K., & Ersöz, S. (2022). A systematic literature review of the predictive maintenance from transportation systems aspect. *Sustainability*, 14(21), 14536. <https://doi.org/10.3390/su142114536>

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>

González Otárola, V. S. (2020). *Implementación de un plan de mantenimiento preventivo para mejorar la confiabilidad de las ambulancias del Hospital VRG, Huaraz-2020* [Tesis de pregrado, Universidad César Vallejo]. Repositorio UCV. <https://repositorio.ucv.edu.pe/handle/20.500.12692/65364>

Hand, D. J., Christen, P., & Ziyad, S. (2024). *Selecting a classification performance measure: Matching the measure to the problem* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2409.12391>

Instituto Colombiano de Normas Técnicas y Certificación. (2007). *NTC 3729: Tipología vehicular. Ambulancias de transporte terrestre*.

Instituto Ecuatoriano de Normalización. (2009). *NTE INEN 2512: Ambulancias “Estrella de la Vida”. Requisitos*.

International Organization for Standardization. (2018). *ISO 17359:2018: Condition monitoring and diagnostics of machines—General guidelines*.

Liao, L., & Köttig, F. (2014). Review of hybrid prognostics approaches for

remaining useful life prediction of engineered systems, and an application to battery life prediction. *IEEE Transactions on Reliability*, 63(1), 191–207. <https://doi.org/10.1109/TR.2014.2299152>

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2), 153–157. <https://doi.org/10.1007/BF02295996>

Ministerio de Salud del Perú. (2006). *Norma Técnica de Salud N.º 051-MINSA/OGDN-V.01: Atención prehospitalaria y transporte asistido de pacientes*. <https://www.gob.pe/institucion/minsa/normas-legales/250870-953-2006-minsa>

Ministerio de Salud del Perú. (2025, enero 18). *SAMU brindó atención médica prehospitalaria a 37 107 pacientes durante el 2024*. <https://www.gob.pe/institucion/minsa/noticias/1092841-samu-brindo-atencion-medica-prehospitalaria-a-37-107-pacientes-durante-el-2024>

Mishra, V., Ahuja, R., Nezamuddin, N., Tiwari, G., & Bhalla, K. (2020). Strengthening the capacity of emergency medical services in low and middle income countries using dispatcher-coordinated taxis. *Transportation Research Record*, 2674(6), 563–572. <https://doi.org/10.1177/0361198120929024>

Molęda, M., Małysiak-Mrozek, B., Ding, W., Sunderam, V., & Mrozek, D. (2023). From corrective to predictive maintenance—A review of maintenance approaches for the power industry. *Sensors*, 23(13), 5970. <https://doi.org/10.3390/s23135970>

Monks, T., Harper, A., & Anagnostou, A. (2023). Forecasting the daily demand for emergency medical ambulances in England and Wales: A benchmark analysis. *BMC Medical Informatics and Decision Making*, 23(1), 117. <https://doi.org/10.1186/s12911-023-02218-z>

Montáns, F. J., Chinesta, F., Gómez-Bombarelli, R., & Kutz, J. N. (2019). *Data-driven modeling and learning in science and engineering*. *Comptes*

Rendus Mécanique, 347(11), 845–855.
<https://doi.org/10.1016/j.crme.2019.11.009>

Moubray, J. (1997). *Reliability-centred maintenance* (2nd ed.). Butterworth-Heinemann.

Naciones Unidas. (2015). *Objetivos de Desarrollo Sostenible*.
<https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>

National Academies of Sciences, Engineering, and Medicine. (2020). *Crisis standards of care: Ten years of successes and challenges*. The National Academies Press. <https://doi.org/10.17226/25767>

Pons, P. T., & Markovchick, V. J. (2002). Eight minutes or less: Does the ambulance response time guideline impact trauma patient outcome? *Journal of Emergency Medicine*, 23(1), 43–48. [https://doi.org/10.1016/S0736-4679\(02\)00460-2](https://doi.org/10.1016/S0736-4679(02)00460-2)

Prytz, R. (2014). *Machine learning methods for vehicle predictive maintenance using off-board and on-board data* [Licentiate thesis, Halmstad University]. ResearchGate.
https://www.researchgate.net/publication/273452541_Machine_learning_methods_for_vehicle_predictive_maintenance_using_off-board_and_on-board_data

Rausand, M., & Høyland, A. (2004). *System reliability theory: Models, statistical methods, and applications* (2nd ed.). John Wiley & Sons.

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. <https://doi.org/10.1371/journal.pone.0118432>

Secretaría de Salud de México. (2013). *Norma Oficial Mexicana NOM-034-SSA3-2013, Regulación de los servicios de salud. Atención médica*

prehospitalaria. Diario Oficial de la Federación.

Si, X.-S., Wang, W., Hu, C.-H., & Zhou, D.-H. (2011). Remaining useful life estimation: A review on the statistical *data-driven* approaches. *European Journal of Operational Research*, 213(1), 1–14.
<https://doi.org/10.1016/j.ejor.2010.11.018>

Sun, X., & Xu, W. (2014). Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signal Processing Letters*, 21(11), 1389–1393.
<https://doi.org/10.1109/LSP.2014.2337313>

Taoufyq, H., El Guemmat, K., Mansouri, K., & Akef, F. (2025). Predictive maintenance approaches: A systematic literature review. *Journal of Industrial Engineering and Management*, 18(3), 427–458.
<https://doi.org/10.3926/jiem.8537>

World Health Organization. (2005). *Prehospital trauma care systems*.
<https://iris.who.int/handle/10665/43167>

World Health Organization. (2011). *Medical equipment maintenance programme overview*. <https://iris.who.int/handle/10665/44587>

CAPÍTULO X : ANEXOS

ANEXO A - Estructura del *dataset* simulado y diccionario de variables

A.1. Descripción general del dataset

El *dataset* de observaciones empleado en la presente investigación fue construido a partir de datos simulados calibrados con los parámetros operativos documentados en la NTS N.º 051-MINSA/OGDN-V.01 (MINSA, 2006) y en la literatura técnica de referencia, dado que los registros históricos institucionales del SAMU no estuvieron disponibles durante el período de ejecución del proyecto.

La simulación fue declarada explícitamente en la metodología de la investigación, precisando que el dataset no corresponde a registros institucionales reales del SAMU, sino a un escenario computacional reproducible diseñado para demostrar la viabilidad metodológica del modelo propuesto.

Esta precisión delimita el alcance de los resultados y evita atribuir al modelo una validación institucional aún no realizada. Por tanto, los resultados deben interpretarse como una validación computacional preliminar y no como una evaluación del desempeño operativo del SAMU.

El *dataset* final está compuesto por tres tablas fuente y un *dataset* de observaciones construido mediante ventanas temporales deslizantes, cuyos parámetros de generación se detallan en la Tabla A.1.

Tabla A.1 Parámetros de simulación del dataset

Parámetro	Valor	Fundamento
Número de ambulancias simuladas	33 unidades	Flota SAMU Lima Metropolitana reportada por Defensoría del Pueblo
Período de análisis	2024-01-01 a 2025-12-31	Definido en el diseño metodológico
Ventana histórica (W)	60 días	Decisión metodológica basada en enfoque de ventanas temporales deslizantes
Horizonte de predicción (T)	14 días	Horizonte operativo de corto plazo para priorización preventiva
Periodicidad de cortes temporales	7 días (semanal)	Frecuencia de reevaluación periódica del riesgo
Intervalo de mantenimiento preventivo	180 ± 15 días	Ficha de homologación prepublicada por PERÚ COMPRAS: mantenimiento preventivo semestral o según manual del fabricante
Referencia de inoperatividad base	21,2%	7 ambulancias inoperativas de 33 unidades reportadas
Macro-subsistemas analíticos	Vehicular; eléctrico/electrónico ; biomédico-asistencial	Agrupación funcional basada en la configuración técnica de ambulancia Tipo II

Fuente. Elaboración propia.

A.2. Diccionario de variables del dataset de observaciones

Tabla A.2 Diccionario de variables - Dataset de observaciones

Variable	Tipo	Descripción	Unidad
id_ambulancia	Categórica	Identificador único de la unidad	—
t0	Fecha	Corte temporal de evaluación	YYYY-MM-DD
n_pm_w	Numérica	Número de eventos preventivos en W	Eventos
n_cm_w	Numérica	Número de eventos correctivos en W	Eventos
n_total_mant_w	Numérica	Total de eventos de mantenimiento en W	Eventos
dias_desde_ultima_interv	Numérica	Días desde la última intervención técnica	Días

n_eventos_vehicular_w	Numérica	Eventos en subsistema vehicular en W	Eventos
n_eventos_electrico_w	Numérica	Eventos en subsistema eléctrico en W	Eventos
n_eventos_equipamiento_w	Numérica	Eventos en equipamiento médico en W	Eventos
n_episodios_downtime_w	Numérica	Número de episodios de inoperatividad en W	Episodios
downtime_total_dias_w	Numérica	Días acumulados de inoperatividad en W	Días
downtime_promedio_dias_w	Numérica	Duración promedio por episodio en W	Días
disponibilidad_w	Numérica	Proporción del tiempo operativo en W	0–1
km_en_w	Numérica	Kilometraje recorrido en W	Km
servicios_en_w	Numérica	Servicios prestados en W	Servicios
equipamiento_funcional	Binaria	Disponibilidad del equipamiento obligatorio	0/1
y_inoperativa_14	Binaria	Variable objetivo: inoperatividad en T	0/1

Nota. W = ventana histórica de 60 días. T = horizonte de predicción de 14 días.
Fuente. Elaboración propia.

ANEXO B - Código fuente del generador de *dataset* simulado

El presente anexo contiene el código fuente completo del script Python empleado para la generación del *dataset* simulado. El código está documentado con comentarios que describen cada función, sus parámetros de entrada y salida, y las referencias bibliográficas que fundamentan los valores empleados.

Este anexo tiene como finalidad facilitar la revisión técnica del proceso de generación de datos y permitir que el procedimiento pueda ser replicado bajo las mismas condiciones computacionales. La inclusión del código fuente permite verificar la trazabilidad entre los supuestos de simulación, las distribuciones probabilísticas empleadas y los archivos CSV generados como insumo para el modelado predictivo.

Archivo: [generar_dataset_ambulancias.py](#)

Versión de Python: 3.12.10

Dependencias:

- pandas: Manipulación de datos tabulares.
- numpy: Operaciones numéricas y generación aleatoria.

Instrucciones de ejecución:

```
pip install pandas numpy  
  
python generar_dataset_ambulancias.py
```

Salidas generadas:

- [datos_simulados/tabla_mantenimiento.csv](#)
- [datos_simulados/tabla_downtime.csv](#)
- [datos_simulados/tabla_uso.csv](#)
- [datos_simulados/dataset_observaciones.csv](#)
- [datos_simulados/dataset_entrenamiento_2024.csv](#)
- [datos_simulados/dataset_validacion_2025.csv](#)

Acceso al repositorio digital

El código fuente completo está disponible en el repositorio digital de acceso abierto de la presente investigación, que garantiza la trazabilidad y reproducibilidad de los resultados:

URL: <https://github.com/DJOA5799/tesis-mantenimiento-predictivo-ambulancias>

La Figura B.1 presenta el código QR de acceso al repositorio digital donde se encuentra disponible el script [generar_dataset_ambulancias.py](#).



Figura B.1. Código QR de acceso al repositorio digital de la investigación: generador del dataset simulado.

Nota. El código QR dirige al repositorio digital de la investigación, donde se encuentran disponibles los scripts de generación del dataset, modelamiento predictivo, lineamientos técnicos, plantilla semanal y análisis inferencial complementario.

Fuente. Elaboración propia.

ANEXO C - Código fuente del modelo computacional predictivo

El presente anexo contiene el código fuente completo del script Python que implementa las Fases 3 y 4 de la metodología: el modelado computacional y la validación retrospectiva.

Archivo: [modelo_predictivo_ambulancias.py](#)

Versión de Python: 3.12.10

Dependencias:

- pandas: Manipulación del dataset.
- numpy: Operaciones numéricas.
- scikit-learn: Modelos de clasificación y métricas.

- matplotlib: Generación de visualizaciones.

Modelos implementados:

- Regresión Logística ([LogisticRegression](#) - línea base interpretable)
- Árbol de decisión ([DecisionTreeClassifier](#) - modelo base no lineal interpretable)
- *Random Forest* ([RandomForestClassifier](#) - modelo predictivo principal seleccionado)
- *Gradient Boosting* ([GradientBoostingClassifier](#) - modelo comparativo de ensamble)

Instrucciones de ejecución:

```
pip install pandas numpy scikit-learn matplotlib  
  
python modelo_predictivo_ambulancias.py
```

Salidas generadas:

- [resultados/tabla_metricas_comparativas.csv](#)
- [resultados/importancia_variables.csv](#)
- [resultados/soporte_preventivo_completo.csv](#)
- [resultados_modelo.png](#)

Acceso al repositorio digital

El código fuente completo está disponible en el repositorio digital de acceso abierto de la presente investigación, que garantiza la trazabilidad y reproducibilidad de los resultados:

URL: <https://github.com/DJOA5799/tesis-mantenimiento-predictivo-ambulancias>

La Figura C.1 presenta el código QR de acceso al repositorio digital donde se encuentra disponible el script [modelo_predictivo_ambulancias.py](#).



Figura C.1. Código QR de acceso al repositorio digital de la investigación: modelo computacional predictivo y validación retrospectiva.

Nota. El código QR dirige al repositorio digital de la investigación, donde se encuentran disponibles los scripts de generación del dataset, modelamiento predictivo, lineamientos técnicos, plantilla semanal y análisis inferencial complementario.

Fuente. Elaboración propia.

ANEXO D - Código fuente de lineamientos técnicos, soporte preventivo y plantilla semanal

D.1. Script de lineamientos técnicos y soporte preventivo

El presente apartado contiene la descripción del script Python empleado para implementar la Fase 5 de la metodología, correspondiente a la generación de lineamientos técnicos y salidas operativas del modelo computacional predictivo.

Archivo: [lineamientos_tecnicos.py](#)

Versión de Python: 3.12.10

Descripción: Implementa la Fase 5 de la metodología, generando los componentes operativos derivados del modelo predictivo: lineamientos de priorización, protocolo de uso, criterios de alerta temprana, recomendaciones de registro y reporte ejecutivo.

Este script permite traducir las probabilidades estimadas por el modelo en productos operativos orientados a la toma de decisiones técnicas. Su inclusión en el anexo facilita verificar la trazabilidad entre la salida predictiva, los criterios de priorización preventiva y los archivos generados para apoyar la gestión semanal del mantenimiento.

Instrucciones de ejecución:

```
python lineamientos_tecnicos.py
```

Salidas generadas:

- [resultados/lineamiento_priorizacion.csv](#)
- [resultados/protocolo_uso_modelo.csv](#)
- [resultados/criterios_alerta_temprana.csv](#)
- [resultados/recomendaciones_datos.csv](#)
- [resultados/reporte_ejecutivo.txt](#)
- [lineamientos_visualizacion.png](#)

Acceso al repositorio digital

El código fuente completo está disponible en el repositorio digital de acceso abierto de la presente investigación, que garantiza la trazabilidad y reproducibilidad de los resultados:

URL: <https://github.com/DJOA5799/tesis-mantenimiento-predictivo-ambulancias>

La Figura D.1 presenta el código QR de acceso al repositorio digital donde se encuentra disponible el script [lineamientos_tecnicos.py](#).



Figura D.1. Código QR de acceso al repositorio digital de la investigación: lineamientos técnicos y soporte preventivo.

Nota. El código QR dirige al repositorio digital de la investigación, donde se encuentran disponibles los scripts de generación del dataset, modelamiento predictivo, lineamientos técnicos, plantilla semanal y análisis inferencial complementario.

Fuente. Elaboración propia.

D.2. Script de generación de plantilla semanal

El presente apartado contiene la descripción del script Python empleado para transformar las salidas del modelo predictivo y los lineamientos técnicos en una plantilla semanal editable de programación de mantenimiento. Este script no entrena modelos ni modifica las probabilidades generadas por el *Random Forest*; su función es organizar los resultados en un formato operativo, comprensible y utilizable por el responsable de mantenimiento. Este script cumple una función de conversión operativa, al transformar los resultados técnicos del modelo en una herramienta editable para la programación semanal. De este modo, facilita que la información predictiva sea utilizada sin modificar el modelo ni requerir conocimientos especializados de programación.

Archivo: [generar_plantilla_semanal.py](#)

Versión de Python: 3.12.10

Descripción: Genera una plantilla semanal en formato Excel a partir de las salidas del modelo predictivo, incorporando la clasificación por nivel de riesgo, las alertas activas, la acción recomendada, el plazo sugerido y los campos de observación y validación. La plantilla constituye el producto operativo final del flujo computacional propuesto.

Dependencias:

- pandas: lectura y procesamiento de archivos CSV.
- numpy: apoyo a operaciones numéricas.
- openpyxl: generación y formato del archivo Excel.
- datetime: gestión de fechas de corte.
- os: gestión de rutas y directorios.

Entrada esperada:

- resultados/soporte_preventivo_completo.csv

Salida generada:

- resultados/plantilla_semanal_YYYYMMDD.xlsx

Instrucciones de ejecución:

```
pip install pandas numpy openpyxl  
  
python generar_plantilla_semanal.py
```

Acceso al repositorio digital

El código fuente completo está disponible en el repositorio digital de acceso abierto de la presente investigación, que garantiza la trazabilidad y reproducibilidad de los resultados:

URL: <https://github.com/DJOA5799/tesis-mantenimiento-predictivo->

ambulancias

La Figura D.2 presenta el código QR de acceso al repositorio digital donde se encuentra disponible el script `generar_plantilla_semanal.py`.



Figura D.2. Código QR de acceso al repositorio digital de la investigación: generador de la plantilla semanal de programación de mantenimiento.

Nota. El código QR dirige al repositorio digital de la investigación, donde se encuentran disponibles los scripts de generación del dataset, modelamiento predictivo, lineamientos técnicos, plantilla semanal y análisis inferencial complementario.

Fuente. Elaboración propia.

ANEXO E - Código fuente del análisis inferencial complementario de clasificadores

El presente anexo contiene el código fuente completo del script Python empleado para el análisis inferencial complementario de la comparación entre clasificadores, descrito en la Sección V.1.3.1. El script no reentrena modelos, no modifica hiperparámetros, no altera la semilla de reproducibilidad fijada en `np.random.seed(42)` y no modifica el pipeline principal de la investigación. Reutiliza las funciones de carga de datos, balanceo, entrenamiento y evaluación del script “`modelo_predictivo_ambulancias.py`” (Anexo C) sin alteraciones, y aplica sobre las predicciones generadas en la Fase 4 los tres análisis inferenciales adicionales: la prueba de McNemar, la prueba de DeLong y el cálculo de la

precisión media (Average Precision).

Este anexo tiene como finalidad documentar íntegramente el procedimiento utilizado para complementar la comparación descriptiva entre modelos mediante herramientas de inferencia estadística. Su incorporación permite verificar que las conclusiones obtenidas sobre la selección del modelo principal no dependen exclusivamente de métricas descriptivas, sino que también se sustentan en pruebas estadísticas aplicadas sobre las predicciones generadas durante la validación retrospectiva. De este modo, el anexo refuerza la transparencia metodológica del proceso de comparación entre clasificadores.

Archivo: [analisis_inferencial_comparacion_modelos.py](#)

Versión de Python: 3.12.10

Dependencias:

- pandas: Manipulación de tablas de resultados.
- numpy: Operaciones numéricas y manejo de vectores de predicciones.
- scipy: Cálculo del valor p de la prueba de DeLong a partir de la distribución normal estándar.
- statsmodels: Implementación de la prueba de McNemar con corrección de continuidad de Yates.
- scikit-learn: Cálculo de la curva Precision-Recall y de la precisión media (Average Precision).
- matplotlib: Generación de la curva Precision-Recall comparativa (Figura 32).

Archivos de entrada esperados:

- [datos_simulados/dataset_entrenamiento_2024.csv](#)
- [datos_simulados/dataset_validacion_2025.csv](#)
- [modelo_predictivo_ambulancias.py](#) (en el mismo directorio o en el

PYTHONPATH)

Salidas generadas:

- [resultados/tabla_mcnemar.csv](#) (datos de la Tabla 17)
- [resultados/tabla_delong.csv](#) (datos de la Tabla 18)
- [resultados/figura_32_curva_precision_recall.png](#) (Figura 32)
- [resultados/resumen_analisis_inferencial.txt](#)

Instrucciones de ejecución:

```
pip install pandas numpy scipy statsmodels scikit-learn matplotlib
```

```
python analisis_inferencial_comparacion_modelos.py
```

Explicación metodológica:

El script reproduce, en su primer bloque, la ejecución de las Fases 3 y 4 del pipeline principal mediante la llamada directa a las funciones del Anexo C (cargar_datos, balancear_clases, entrenar_modelos, evaluar_modelo, calcular_linea_base), garantizando que los modelos generados, sus probabilidades predichas y sus predicciones binarias bajo el umbral $\tau = 0,30$ sean idénticos a los reportados en la Tabla 16. Sobre esta reproducción se aplican secuencialmente las tres pruebas inferenciales.

La prueba de McNemar (Ecuación 39) se calcula mediante la construcción explícita de las tablas de contingencia 2×2 entre los pares de modelos y el uso de la función `statsmodels.stats.contingency_tables.mcnemar` con corrección de Yates. La prueba de DeLong (Ecuación 40) se implementa siguiendo el algoritmo de cálculo eficiente basado en midranks propuesto por Sun y Xu (2014), matemáticamente equivalente al método original de DeLong et al. (1988). La precisión media (Ecuación 41) se calcula mediante la función `sklearn.metrics.average_precision_score` sobre las probabilidades predichas por cada modelo.

La implementación independiente de estos análisis garantiza que cualquier

investigador pueda reproducir los resultados inferenciales a partir del mismo conjunto de datos y bajo las mismas condiciones computacionales declaradas en la metodología. Asimismo, la separación del código respecto del pipeline principal preserva la trazabilidad del proceso y facilita futuras ampliaciones o incorporaciones de nuevas pruebas de comparación entre clasificadores sin modificar el modelo predictivo originalmente desarrollado.

Acceso al repositorio digital:

El código fuente completo está disponible en el repositorio digital de acceso abierto de la presente investigación, que garantiza la trazabilidad y reproducibilidad de los resultados:

URL: <https://github.com/DJOA5799/tesis-mantenimiento-predictivo-ambulancias>

La Figura E.1 presenta el código QR de acceso al repositorio digital donde se encuentra disponible el script `analisis_inferencial_comparacion_modelos.py`.



Figura E.1. Código QR de acceso al repositorio digital de la investigación: análisis inferencial complementario de clasificadores.

Nota. El código QR dirige al repositorio digital de la investigación, donde se encuentran disponibles los scripts de generación del dataset, modelamiento predictivo, lineamientos técnicos, plantilla semanal y análisis inferencial complementario.

Fuente. Elaboración propia.

ANEXO F - Resultados completos de la validación retrospectiva

F.1. Tabla completa de métricas por modelo

Tabla F.1 Resultados completos de la validación retrospectiva - Período 2025

Modelo	VP	FP	VN	FN	Precisión	Sensibilidad	Especificidad	Exactitud	F1-Score	AUC-ROC	MAE	Disp. Proyectada
Preventivo tradicional (línea base)	105	542	764	239	16,23%	30,52%	58,50%	52,67%	21,19%	42,47%	47,33%	60,79%
Regresión Logística	344	1 306	0	0	20,85%	100,00%	0,00%	20,85%	34,50%	53,09%	50,12%	0,00%
Árbol de Decisión	191	754	552	153	20,21%	55,52%	42,27%	45,03%	29,64%	50,46%	41,10%	42,73%
Random Forest	316	1 176	130	28	21,18%	91,86%	9,95%	27,03%	34,42%	56,03%	45,39%	9,58%
Gradient Boosting	225	790	516	119	22,17%	65,41%	39,51%	44,91%	33,11%	52,58%	42,05%	38,48%

Nota. VP = verdaderos positivos; FP = falsos positivos; VN = verdaderos negativos; FN = falsos negativos. Umbral de clasificación = 0,30.
Fuente. Elaboración propia.

F.2. Importancia de variables completa

Tabla F.2 Importancia de variables completa - Modelo *Random Forest*

Posición	Variable	Importancia	Importancia (%)
1	Kilometraje en W	0,185238	18,52%
2	Servicios prestados en W	0,169145	16,91%
3	Días desde última intervención	0,122138	12,21%
4	<i>Downtime</i> promedio en W	0,099889	9,99%
5	Disponibilidad histórica en W	0,087099	8,71%
6	<i>Downtime</i> total en W	0,081020	8,10%
7	Total de eventos de mantenimiento	0,060386	6,04%
8	Mantenimientos correctivos en W	0,047705	4,77%
9	Eventos en subsistema eléctrico/electrónico	0,033815	3,38%
10	Eventos en subsistema vehicular	0,026783	2,68%
11	Eventos en equipamiento biomédico-asistencial	0,023939	2,39%
12	Episodios de <i>downtime</i> en W	0,023723	2,37%
13	Mantenimientos preventivos en W	0,022811	2,28%
14	Equipamiento funcional	0,016311	1,63%

Fuente. Elaboración propia.

ANEXO G - Protocolo operativo completo de uso del modelo

El presente anexo desarrolla el protocolo operativo completo para la utilización del modelo computacional en un ciclo semanal de mantenimiento. Cada etapa define la actividad a realizar, su propósito, la frecuencia recomendada y el responsable de ejecución, con el fin de asegurar una aplicación sistemática, trazable y consistente del soporte preventivo. Su propósito es facilitar una aplicación uniforme del modelo en las actividades rutinarias de planificación y priorización técnica de las ambulancias.

Tabla G.1 Protocolo semanal de uso del modelo computacional predictivo

Paso	Actividad	Descripción	Frecuencia	Responsable
1	Actualización de registros	Registro de todos los eventos de mantenimiento, <i>downtime</i> y uso operativo en el sistema de gestión	Continua	Personal responsable del registro técnico Responsable de gestión de datos o mantenimiento Jefatura o coordinación de mantenimiento
2	Ejecución del modelo	Procesamiento de los registros de los últimos 60 días para cada unidad y generación de probabilidades de inoperatividad	Semanal - cada lunes	Encargado de gestión de datos
3	Revisión del ranking de riesgo	Revisión del tablero de soporte preventivo con el ranking de ambulancias por nivel de riesgo	Semanal - cada lunes	Jefe de mantenimiento
4	Programación de intervenciones	Programación de intervenciones priorizando unidades de riesgo ALTO y MEDIO	Semanal	Jefe de mantenimiento
5	Ejecución y registro	Ejecución de la intervención y registro de hallazgos, repuestos y tiempos	Según programación	Técnico de mantenimiento
6	Retroalimentación	Actualización del modelo con los nuevos registros para el siguiente ciclo	Continua	Encargado de gestión de datos

Fuente. Elaboración propia a partir de Taoufyq et al. (2025) e ISO 17359:2018.

ANEXO H - Criterios de alerta temprana por variable crítica

El presente anexo reúne los criterios de alerta temprana propuestos a partir de las variables de mayor importancia identificadas por el modelo Random Forest. Su finalidad es orientar la interpretación de las principales señales de riesgo y apoyar la priorización preventiva. Los umbrales presentados deberán recalibrarse posteriormente con registros reales del SAMU.

Tabla H.1 Criterios de alerta temprana basados en importancia de variables del modelo

Variable crítica	Importancia	Umbral de alerta	Interpretación	Acción
Kilometraje en W	18,52%	> 8 000 km en 60 días	Uso intensivo del vehículo asociado a falla mecánica	Verificar sistema vehicular y tren motriz
Servicios prestados en W	16,91%	> 120 servicios en 60 días	Alta carga operativa asociada a desgaste acelerado	Adelantar inspección técnica general
Días desde última intervención	12,21%	> 45 días sin intervención	Intervalo prolongado sin mantenimiento	Programar mantenimiento preventivo inmediato
Disponibilidad histórica en W	8,71%	< 0,85 (menos del 85%)	Historial reciente de inoperatividad	Revisión técnica integral prioritaria
<i>Downtime</i> total en W	8,10%	> 7 días acumulados en 60 días	Patrón de inoperatividad recurrente	Análisis de causa raíz
Equipamiento funcional	1,63%	Estado = 0 (no funcional)	Indisponibilidad del equipamiento obligatorio	Reemplazo o reparación inmediata

Nota. Los umbrales de alerta temprana son criterios técnicos propuestos a partir de la importancia relativa de variables del modelo y del comportamiento del dataset simulado. Su adopción operativa requiere calibración posterior con registros institucionales reales del SAMU.

Fuente. Elaboración propia a partir de NTS N.º 051-MINSA/OGDN-V.01 (MINSA, 2006) y resultados del modelo.

ANEXO I - Plantilla semanal de programación de mantenimiento con soporte predictivo.

En este anexo se presentan capturas ampliadas de la plantilla semanal generada por el sistema propuesto, con el fin de mostrar de manera más legible su estructura, contenido y utilidad operativa.

La Figura I.1 muestra la vista general de la primera hoja de la plantilla semanal; la Figura I.2 presenta el detalle operativo de las unidades priorizadas; la Figura I.3 muestra la sección de control y seguimiento semanal; y la Figura I.4 presenta la segunda hoja correspondiente a los datos completos generados por el modelo.

PROGRAMACIÓN SEMANAL DE MANTENIMIENTO CON SOPORTE PREDICTIVO			
Ambulancias médicas urbanas Tipo II Semana 50 / 2025 Corte: 12/12/2025			
1. DATOS GENERALES DEL CORTE			
Fecha de corte	12/12/2025	Flota evaluada	35 unidades
Horizonte predictivo	14 días	Modelo predictivo	Random Forest
Umbral alto	$P \geq 0.50$	Umbral medio	$0.25 \leq P < 0.50$
Responsable	Jefatura / Coordinación de mantenimiento	Frecuencia sugerida	Semanal
2. RESUMEN DE PRIORIZACIÓN			
RIESGO ALTO	RIESGO MEDIO	RIESGO BAJO	CON ALERTAS ACTIVAS
13	18	4	12
$P \geq 0.50$	$0.25 \leq P < 0.50$	$P < 0.25$	Alertas técnicas registradas
3. GUÍA DE INTERPRETACIÓN			
Nivel / criterio	Rango o fuente	Interpretación operativa	
Riesgo alto	$P \geq 0.50$	Intervención prioritaria. Revisar la unidad en el plazo definido y validar alertas técnicas.	
Riesgo medio	$0.25 \leq P < 0.50$	Inspección programada. Confirmar condición técnica y priorizar según disponibilidad operativa.	
Riesgo bajo	$P < 0.25$	Seguimiento rutinario. Mantener programación preventiva institucional según corresponda.	
Alertas activas	Reglas técnicas	Orientan el tipo de revisión: uso operativo, historial de inoperatividad, disponibilidad o equipamiento biomédico.	

Figura I.1. Vista ampliada de la primera hoja de la plantilla semanal.

Nota. La captura muestra los datos generales del corte, el resumen de priorización, y la guía de interpretación.

Fuente. Elaboración propia.

4. PROGRAMACIÓN SEMANAL PRIORIZADA							
Unidad	P(inop.)	Riesgo	N.º alertas	Alertas activas	Acción recomendada	Plazo sugerido	Observaciones
AMB-T2-022	0.6576	ALTO	0	—	Inspección técnica general	≤ 3 días hábiles	
AMB-T2-026	0.6456	ALTO	2	Servicios >120 (131) Km >8 000 (9583 km)	Inspección técnica general + revisión tren motriz	≤ 3 días hábiles	
AMB-T2-002	0.6514	ALTO	0	—	Inspección técnica general	≤ 3 días hábiles	
AMB-T2-006	0.5823	ALTO	3	Disp. <85% (83%) Downtime >7.0d (10.2d) Equipo biomédico no funcional	Inspección técnica general + análisis causa raíz inoperatividad + reemplazo/repación equipo biomédico	≤ 3 días hábiles	
AMB-T2-032	0.5677	ALTO	3	Disp. <85% (73%) Downtime >7.0d (16.0d) Equipo biomédico no funcional	Inspección técnica general + análisis causa raíz inoperatividad + reemplazo/repación equipo biomédico	≤ 3 días hábiles	
AMB-T2-024	0.5505	ALTO	1	Equipo biomédico no funcional	Inspección técnica general + reemplazo/repación equipo biomédico	≤ 3 días hábiles	
AMB-T2-027	0.5415	ALTO	0	—	Inspección técnica general	≤ 3 días hábiles	
AMB-T2-011	0.5367	ALTO	0	—	Inspección técnica general	≤ 3 días hábiles	
AMB-T2-013	0.5273	ALTO	1	Equipo biomédico no funcional	Inspección técnica general + reemplazo/repación equipo biomédico	≤ 3 días hábiles	
AMB-T2-029	0.5175	ALTO	0	—	Inspección técnica general	≤ 3 días hábiles	
AMB-T2-015	0.5075	ALTO	1	Equipo biomédico no funcional	Inspección técnica general + reemplazo/repación equipo biomédico	≤ 3 días hábiles	
AMB-T2-003	0.4932	MEDIO	0	—	Inspección dirigida	≤ 7 días hábiles	
AMB-T2-031	0.4733	MEDIO	3	Disp. <85% (64%) Downtime >7.0d (21.3d) Equipo biomédico no funcional Km >8 000 (9054 km)	Inspección dirigida: historial inoperatividad	≤ 7 días hábiles	
AMB-T2-001	0.4627	MEDIO	4	Disp. <85% (82%) Downtime >7.0d (10.9d)	Inspección dirigida: sistema vehicular + historial inoperatividad	≤ 7 días hábiles	

Figura I.2. Vista ampliada del detalle operativo de unidades priorizadas.

Nota. La captura muestra la tabla de programación semanal con probabilidad de inoperatividad, nivel de riesgo, alertas activas, acción recomendada, plazo sugerido y observaciones.

Fuente. Elaboración propia.

AMB-T2-027	0.4160	MEDIO	1	Servicios >120 (121)	Inspección dirigida: sistema vehicular	≤ 7 días hábiles	
AMB-T2-031	0.4125	MEDIO	1	Servicios >120 (121)	Inspección dirigida: sistema vehicular	≤ 7 días hábiles	
AMB-T2-012	0.4057	MEDIO	1	Servicios >120 (122)	Inspección dirigida: sistema vehicular	≤ 7 días hábiles	
AMB-T2-005	0.3960	MEDIO	1	Equipo biomédico no funcional	Inspección dirigida	≤ 7 días hábiles	
AMB-T2-023	0.3049	MEDIO	1	Servicios >120 (120)	Inspección dirigida: sistema vehicular	≤ 7 días hábiles	
AMB-T2-009	0.2730	MEDIO	1	Servicios >120 (122)	Inspección dirigida: sistema vehicular	≤ 7 días hábiles	
AMB-T2-011	0.2452	BAJO	0	—	Continuar operación normal. PM según cronograma.	Según cronograma	
AMB-T2-034	0.2275	BAJO	0	—	Continuar operación normal. PM según cronograma.	Según cronograma	
AMB-T2-008	0.2146	BAJO	0	—	Continuar operación normal. PM según cronograma.	Según cronograma	
AMB-T2-006	0.0424	BAJO	0	—	Continuar operación normal. PM según cronograma.	Según cronograma	

5. CONTROL Y SEGUIMIENTO SEMANAL							
Total alto	Total medio	Total bajo	Con alertas	Atendidas	Pendientes	Reprogramadas	Cierre
13	18	4	12				
Observaciones generales: Las acciones programadas deben validarse con la disponibilidad operativa, recursos técnicos, requisitos y criterio de la jefatura de mantenimiento.							

6. VALIDACIÓN DEL DOCUMENTO		
Elaborado por	Revisado por	Aprobado por
Nombre / firma: Fecha: 12/12/2025	Nombre / firma: Fecha: ____/____/____	Nombre / firma: Fecha: ____/____/____

Fuente: Elaboración propia. Plantilla generada a partir de probabilidades estimadas por el modelo predictivo Random Forest. Documento metodológico elaborado con datos simulados para fines de investigación. Generado: 27/05/2026 21:59.

Figura I.3. Vista ampliada del control y seguimiento semanal.

Nota. La captura muestra la tabulación planteada para el control y seguimiento semanal, además de la validación del documento para uso institucional.

Fuente. Elaboración propia.

Iden. de la unidad	id	prob. de pasar a la vida	clase de riesgo	clase de riesgo a la vida	clase de riesgo a la vida	clase de riesgo a la vida	clase de riesgo a la vida	clase de riesgo a la vida	clase de riesgo a la vida	clase de riesgo a la vida	clase de riesgo a la vida	clase de riesgo a la vida	clase de riesgo a la vida	clase de riesgo a la vida
AMB-1-001	2025-1-12	0.6536	AUTO	13	2.9	3	0.9817	1	7392	301	—	—	Inspección a la vida general	
AMB-1-006	2025-1-12	0.6090	AUTO	2	3	1	1	1	8383	131	—	—	Inspección a la vida general	
AMB-1-002	2025-1-12	0.6334	AUTO	13	2.9	3	0.9325	1	7471	94	—	—	Inspección a la vida general	
AMB-1-006	2025-1-12	0.5802	AUTO	9	10.2	4	0.93	9	3023	23	—	—	Inspección a la vida general	
AMB-1-007	2025-1-12	0.5677	AUTO	19	36	4	0.7382	8	5155	303	—	—	Inspección a la vida general	
AMB-1-008	2025-1-12	0.5375	AUTO	4	2.9	3	0.9283	8	7219	115	—	—	Inspección a la vida general	
AMB-1-007	2025-1-12	0.5415	AUTO	12	2.7	1	0.932	1	7026	96	—	—	Inspección a la vida general	
AMB-1-011	2025-1-12	0.5347	AUTO	10	2.3	3	0.9433	1	6715	81	—	—	Inspección a la vida general	
AMB-1-013	2025-1-12	0.5275	AUTO	19	5.6	3	0.9667	8	5187	55	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.5125	AUTO	17	8	1	1	1	5181	91	—	—	Inspección a la vida general	
AMB-1-015	2025-1-12	0.5015	AUTO	4	3.6	3	0.93	8	6831	83	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.5942	MEDIO	16	8	2	1	1	7819	69	—	—	Inspección a la vida general	
AMB-1-018	2025-1-12	0.4733	MEDIO	4	11.3	1	0.625	8	5535	116	—	—	Inspección a la vida general	
AMB-1-016	2025-1-12	0.5673	MEDIO	5	10.9	3	0.8383	8	3034	112	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.531	MEDIO	8	14.1	3	0.762	8	7915	96	—	—	Inspección a la vida general	
AMB-1-018	2025-1-12	0.4194	MEDIO	1	9	1	1	1	6642	119	—	—	Inspección a la vida general	
AMB-1-018	2025-1-12	0.4196	MEDIO	23	9	1	1	1	6695	112	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.5077	MEDIO	15	7.3	2	0.88	1	4340	70	—	—	Inspección a la vida general	
AMB-1-018	2025-1-12	0.4912	MEDIO	40	9	0	1	1	4244	42	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.4913	MEDIO	24	5	2	0.967	8	3566	63	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.3877	MEDIO	15	3.5	2	0.973	8	3186	25	—	—	Inspección a la vida general	
AMB-1-018	2025-1-12	0.3816	MEDIO	3	9	2	1	8	5016	55	—	—	Inspección a la vida general	
AMB-1-018	2025-1-12	0.296	MEDIO	31	9	1	1	1	4133	60	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.2400	MEDIO	60	9	0	1	1	3492	94	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.3151	MEDIO	7	9	0	1	1	4334	116	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.2150	MEDIO	60	9	0	1	1	3190	70	—	—	Inspección a la vida general	
AMB-1-016	2025-1-12	0.3104	MEDIO	33	9	1	1	1	4267	55	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.2113	MEDIO	23	4.3	2	0.923	8	3680	53	—	—	Inspección a la vida general	
AMB-1-016	2025-1-12	0.2913	MEDIO	43	2.8	1	0.9323	1	3469	117	—	—	Inspección a la vida general	
AMB-1-017	2025-1-12	0.2876	MEDIO	40	9	0	1	1	4990	127	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.279	MEDIO	51	3.9	1	0.9883	1	5115	56	—	—	Inspección a la vida general	
AMB-1-019	2025-1-12	0.2745	MEDIO	13	2.7	1	0.9383	8	6736	115	—	—	Inspección a la vida general	
AMB-1-012	2025-1-12	0.2705	MEDIO	1	57	3	0.7367	1	6175	206	—	—	Inspección a la vida general	

Figura I.4. Vista ampliada de la segunda hoja de datos completos.

Nota. La captura muestra el detalle tabulado de las unidades evaluadas y las variables operativas utilizadas para la toma de decisiones.
Fuente. Elaboración propia.

ANEXO J - Matriz de consistencia de la investigación

PROBLEMA		OBJETIVO	HIPÓTESIS	VARIABLES	DIMENSIONES	INDICADORES	METODOLOGÍA	ANTECEDENTES	
GENERAL		GENERAL	GENERAL	VARIABLE INDEPENDIENTE	DIMENSIONES VI	INDICADORES VI	DISEÑO GENERAL	ANTECEDENTES PRINCIPALES	
¿Cómo desarrollar un modelo computacional basado en datos para el mantenimiento predictivo con soporte preventivo que estime la probabilidad de inoperatividad y apoye la gestión técnica de las ambulancias médicas urbanas Tipo II en Lima Metropolitana durante 2024-2025?		Desarrollar un modelo computacional basado en datos para el mantenimiento predictivo con soporte preventivo de ambulancias médicas urbanas Tipo II en Lima Metropolitana durante el período 2024-2025, orientado a estimar la probabilidad de inoperatividad y apoyar la gestión técnica del sistema prehospitalario	El desarrollo de un modelo computacional basado en datos para el mantenimiento predictivo con soporte preventivo permite estimar la probabilidad de inoperatividad y apoyar la gestión técnica de las ambulancias médicas urbanas Tipo II en Lima Metropolitana durante el período 2024-2025	X: Modelo computacional basado en datos para el mantenimiento predictivo con soporte preventivo	X1: Historial de mantenimiento	X11: n_total_mant_w (total eventos en W)	Tipo de investigación: Aplicado - Descriptivo-Predictivo	AN1: González Otárola, V.S. (2020). Implementación de un plan de mantenimiento preventivo para mejorar la confiabilidad de las ambulancias del Hospital VRG, Huaraz-2020. Tesis de pregrado, UCV.	
						X12: dias_desde_ultima_interv (días desde última intervención)			
					X2: Inoperatividad histórica	X21: n_episodios_downtime_w (episodios de inoperatividad en W)	Método: Cuantitativo		
						X22: downtime_total_dias_w (días acumulados de inoperatividad)			Diseño: No experimental longitudinal
						X23: downtime_promedio_dias_w (duración promedio por episodio)			
					X3: Uso operativo	X31: servicios_en_w (servicios prestados en W)	Población: 35 ambulancias médicas urbanas Tipo II, Lima Metropolitana, período 2024-2025		
						X32: km_en_w (kilometraje recorrido en W)			
					X4: Condición del equipamiento	X41: n_eventos_vehicular_w	Técnica: Análisis de datos históricos - Aprendizaje supervisado (clasificación binaria) - Backtesting temporal		
						X42: n_eventos_electrico_w			
						X43: n_eventos_equipamiento_w			
						X44: disponibilidad_w			Instrumento: Dataset simulado calibrado con NTS N.º 051-MINSA/DGSP-V.01 y Defensoría del Pueblo (2023a) - Pipeline scikit-learn (Python 3.12.10)
						X45: equipamiento_funcional			
PROBLEMAS ESPECÍFICOS		OBJETIVOS ESPECÍFICOS	HIPÓTESIS ESPECÍFICAS	VARIABLE DEPENDIENTE	DIMENSIONES VD	INDICADORES VD		AN2: Aguilar Sota, L.M. (2024). Gestión de mantenimiento basado en la confiabilidad (RCM) para incrementar la disponibilidad de la flota vehicular de la red de salud Cusco Norte-2022. Tesis de pregrado, UNSAAC.	
X1-Y1	¿Cuáles son los principales factores técnicos y operativos que influyen en la ocurrencia de fallas, eventos de	Analizar los principales factores técnicos y operativos que influyen en la ocurrencia de fallas, eventos de inoperatividad y requerimientos de	H1: La identificación de los factores técnicos y operativos que influyen en las fallas, eventos de inoperatividad y	Y: Inoperatividad de las ambulancias médicas urbanas Tipo II	Y1: Capacidad predictiva	Y11: AUC-ROC			

	inoperatividad y requerimientos de mantenimiento de las ambulancias médicas urbanas Tipo II en Lima Metropolitana?	mantenimiento de las ambulancias médicas urbanas Tipo II, cuya configuración técnica está definida en la NTS N.º 051 (MINSA, 2006).	requerimientos de mantenimiento de las ambulancias médicas urbanas Tipo II permitirá determinar las variables críticas que deben incorporarse en el modelo predictivo.			Y12: Precisión		
X2-Y2	¿Cómo se puede estructurar una base de datos que integre variables operativas, técnicas, de mantenimiento e inoperatividad, de modo que permita el análisis predictivo del desempeño de las ambulancias médicas urbanas Tipo II?	Diseñar una base de datos estructurada que integre variables operativas, técnicas, de mantenimiento e inoperatividad representativas, permitiendo el almacenamiento y análisis de información para el entrenamiento del modelo computacional.	H2: La estructuración de una base de datos que integre variables operativas, técnicas, de mantenimiento e inoperatividad posibilitará el almacenamiento y análisis de la información necesaria para el entrenamiento del modelo computacional y la estimación predictiva del riesgo de inoperatividad de las ambulancias médicas urbanas Tipo II.			Y13: Sensibilidad		
						Y14: F1-Score		
						Y15: MAE (Mean Absolute Error)		
X3-Y3	¿Qué tipo de modelo computacional basado en datos permite estimar la probabilidad de inoperatividad de las ambulancias médicas urbanas Tipo II, integrando variables de uso operativo, mantenimiento, downtime y equipamiento funcional?	Implementar un modelo computacional basado en datos que combine componentes de mantenimiento predictivo y preventivo, mediante algoritmos de clasificación supervisada aplicados a registros históricos simulados de operación, mantenimiento e inoperatividad, para estimar la probabilidad de inoperatividad de las ambulancias médicas urbanas Tipo II.	H3: La implementación de un modelo computacional basado en datos, mediante algoritmos de clasificación supervisada, permitirá estimar la probabilidad de inoperatividad de las ambulancias médicas urbanas Tipo II en un horizonte de 14 días, integrando variables de uso operativo, mantenimiento, downtime y equipamiento funcional.		Y2: Disponibilidad operativa	Y21: Disponibilidad proyectada		AI1: Prytz, R. (2014). Machine learning methods for vehicle predictive maintenance using off-board and on-board data. Licentiate thesis, Halmstad University (Suecia)

X4-Y4	¿Qué indicadores de desempeño: exactitud, especificidad, precisión, sensibilidad, F1-Score, error medio absoluto (MAE), AUC-ROC y disponibilidad operativa proyectada permiten validar el desempeño del modelo predictivo propuesto frente al esquema de mantenimiento preventivo programado como línea base de comparación?	Validar el desempeño del modelo computacional mediante métricas cuantitativas de clasificación: exactitud, especificidad, precisión, sensibilidad, F1-Score, error medio absoluto (MAE), AUC-ROC y disponibilidad operativa proyectada, comparando sus resultados frente al esquema de mantenimiento preventivo programado como línea base.	H4: El modelo computacional propuesto presentará un desempeño evaluable mediante indicadores cuantitativos de exactitud, especificidad, precisión, sensibilidad, F1-Score, error medio absoluto (MAE), AUC-ROC y disponibilidad operativa proyectada, permitiendo comparar su capacidad de detección de eventos de inoperatividad frente al esquema de mantenimiento preventivo programado como línea base.		Y22: MTBF (Mean Time Between Failures)	AI2: Ersöz, O. Ö., İnal, A. F., Aktepe, A., Türker, A. K., & Ersöz, S. (2022). A systematic literature review of the predictive maintenance from transportation systems aspect. Sustainability, 14(21), 14536. Department of Industrial Engineering, Kırıkkale University, Kırıkkale 71450, Turkey
					Y23: MTTR: (Mean Time to Repair)	
X5-Y5	¿Qué lineamientos técnicos y operativos pueden derivarse del modelo computacional propuesto para apoyar la priorización preventiva, fortalecer la gestión del mantenimiento y orientar la mejora de la disponibilidad operativa de las ambulancias médicas urbanas en Lima Metropolitana?	Proponer lineamientos técnicos y operativos para la adopción del modelo computacional en la gestión del mantenimiento de ambulancias médicas urbanas Tipo II, orientados a apoyar la priorización preventiva, reducir eventos de inoperatividad y fortalecer la toma de decisiones técnicas del sistema prehospitalario en Lima Metropolitana.	H5: Los lineamientos técnicos y operativos derivados del modelo computacional permitirán apoyar la priorización preventiva, orientar los procedimientos de mantenimiento y fortalecer la toma de decisiones en la gestión técnica de las ambulancias médicas urbanas Tipo II.	Y3: Gestión del riesgo	Y31: Nivel de riesgo por unidad: ALTO/MEDIO/BAJO	
					Y32: Probabilidad estimada P (y=1)	
					Y33: Umbral de clasificación: 0,30	
VARIABLES DE CONTROL: (1) Período de análisis: 2024-2025 (2) Tipo de ambulancia: Tipo II - NTS N.º 051-MINSA/DGSP-V.01 (3) Ventana histórica W = 60 días (4) Horizonte de predicción T = 14 días (5) Registros históricos disponibles sobre fallas, mantenimiento y operación de las unidades						

ANEXO K - Valoración técnica complementaria mediante juicio de especialistas

K.1. Cuestionario aplicado

El cuestionario fue aplicado mediante la plataforma Google Forms y se estructuró en cuatro secciones. La primera para la contextualización y registro de datos:

Cuestionario de valoración técnica sobre gestión del mantenimiento de ambulancias médicas Tipo II

Estimado(a) profesional:

El presente cuestionario forma parte de la investigación titulada "**Modelo computacional basado en datos para el mantenimiento predictivo con soporte preventivo de ambulancias médicas urbanas Tipo II en Lima Metropolitana**", desarrollada para optar por el título profesional de Ingeniero Biomédico en la Universidad Nacional Mayor de San Marcos.

La investigación desarrolla una arquitectura computacional basada en información histórica de operación, mantenimiento e inoperatividad para estimar el riesgo de inoperatividad de ambulancias médicas Tipo II y generar información que apoye la priorización de acciones preventivas.

El objetivo de este instrumento es **recoger una valoración profesional sobre determinados criterios técnicos y operativos relacionados con la gestión del mantenimiento de ambulancias médicas Tipo II, en el marco de la investigación señalada.**

Esta valoración tiene carácter **complementario**. No tiene como finalidad validar estadísticamente el desempeño del algoritmo predictivo, cuya evaluación fue realizada mediante procedimientos cuantitativos propios de la investigación, sino valorar desde la experiencia profesional la pertinencia de sus componentes técnicos y operativos.

La metodología fue desarrollada inicialmente en un escenario computacional paramétrico-estocástico. Por ello, los parámetros y criterios operativos presentados deben entenderse como una propuesta preliminar susceptible de calibración posterior mediante registros institucionales reales.

Las respuestas proporcionadas serán utilizadas exclusivamente con fines académicos y representan el criterio profesional individual del participante, **sin constituir necesariamente una posición o aprobación institucional de la organización en la que desarrolla sus funciones.**

Tiempo estimado de respuesta: 8–10 minutos.

diego.osores@unmsm.edu.pe [Cambiar de cuenta](#)



* Indica que la pregunta es obligatoria

¿Acepta participar voluntariamente en el presente juicio de expertos y autoriza el uso académico de sus respuestas? *

- ☐ Sí, acepto.
- ☐ No acepto.

Nombres y apellidos *

Tu respuesta

Profesión *

Tu respuesta

Máximo grado académico alcanzado *

- ☐ Título profesional / Licenciado
- ☐ Maestría
- ☐ Doctorado
- ☐ Otro:

Años de experiencia profesional *

- ☐ 3–5 años
- ☐ 6–10 años
- ☐ 11–15 años
- ☐ 16–20 años
- ☐ Más de 20 años

Áreas de experiencia profesional *

- ☐ Ingeniería biomédica
- ☐ Ingeniería clínica
- ☐ Equipamiento médico
- ☐ Ambulancias y transporte sanitario
- ☐ Atención prehospitalaria
- ☐ Mantenimiento
- ☐ Gestión de activos
- ☐ Ingeniería electromecánica
- ☐ Gestión de proyectos
- ☐ Proyectos del sector salud
- ☐ Adquisición de equipamiento o tecnología sanitaria
- ☐ Analítica de datos
- ☐ Inteligencia artificial / Machine Learning
- ☐ Otro:

Describa brevemente su experiencia profesional relacionada con el objeto de esta evaluación. *

Tu respuesta

Información técnica previa

La propuesta considera a la ambulancia médica Tipo II como un sistema integrado por componentes **vehiculares, eléctricos/electrónicos y biomédico-asistenciales**.

El modelo utiliza información histórica relacionada con operación, mantenimiento e inoperatividad, incluyendo kilometraje, servicios prestados, intervenciones de mantenimiento, tiempo desde la última intervención, períodos de inoperatividad y disponibilidad histórica.

La metodología estima una probabilidad de inoperatividad futura que posteriormente puede transformarse en niveles de riesgo y en un ranking de unidades para apoyar la priorización preventiva.

Como parámetros iniciales, la investigación utiliza una **ventana histórica de 60 días**, un **horizonte predictivo de 14 días** y una **actualización semanal**.

Los resultados del modelo se plantean como **soporte para la toma de decisiones del profesional responsable del mantenimiento**, sin sustituir el mantenimiento preventivo ni el criterio técnico.

Los parámetros y umbrales operativos se consideran referencias iniciales y pueden requerir recalibración mediante datos institucionales reales antes de una implementación definitiva.

La segunda para la valoración técnica de los especialistas, acerca de cada dimensión propuesta de la investigación:

A. Pertinencia de las variables e información utilizada *

	No pertinente / no adecuado	Poco pertinente / requiere cambios importantes	Pertinente / adecuado con ajustes menores	Muy pertinente / adecuado	N/A – No cuento con elementos suficientes para valorarlo
A1. El kilometraje recorrido constituye información pertinente para representar la intensidad de uso de una ambulancia.	☐	☐	☐	☐	☐
A2. El número de servicios prestados constituye información pertinente para representar la carga operativa de una ambulancia.	☐	☐	☐	☐	☐
A3. El tiempo transcurrido desde la última intervención de mantenimiento constituye información pertinente para apoyar la evaluación técnica de la unidad.	☐	☐	☐	☐	☐
A4. El historial de mantenimiento preventivo y correctivo constituye información relevante para caracterizar el comportamiento técnico de una ambulancia.	☐	☐	☐	☐	☐

A5. La frecuencia y duración de los períodos de inoperatividad constituyen antecedentes relevantes para caracterizar el comportamiento histórico de la unidad.

☐ ☐ ☐ ☐ ☐

A6. La disponibilidad histórica constituye un indicador pertinente para apoyar el análisis del comportamiento operativo de una ambulancia.

☐ ☐ ☐ ☐ ☐

A7. La condición funcional del equipamiento biomédico debe considerarse dentro de la evaluación integral de la operatividad de una ambulancia Tipo II.

☐ ☐ ☐ ☐ ☐

A8. La integración de información de operación, mantenimiento, inoperatividad y condición funcional resulta técnicamente pertinente para un enfoque predictivo basado en datos.

☐ ☐ ☐ ☐ ☐

B. Coherencia del enfoque predictivo-preventivo *

	No pertinente / no adecuado	Poco pertinente / requiere cambios importantes	Pertinente / adecuado con ajustes menores	Muy pertinente / adecuado	N/A – No cuento con elementos suficientes para valorarlo
B1. Utilizar registros históricos para estimar el riesgo futuro de inoperatividad constituye un enfoque técnicamente coherente para complementar el mantenimiento preventivo.	☐	☐	☐	☐	☐
B2. La estimación de una probabilidad de inoperatividad futura puede aportar información útil para priorizar la revisión técnica de las unidades.	☐	☐	☐	☐	☐
B3. La clasificación de las unidades según niveles de riesgo facilita la interpretación operacional de los resultados del modelo.	☐	☐	☐	☐	☐
B4. La generación de un ranking de unidades según el riesgo estimado constituye un mecanismo pertinente para apoyar la	☐	☐	☐	☐	☐

priorización
preventiva.

B5. La
utilización del
modelo como
herramienta de
soporte, sin
sustituir el
criterio del
responsable de
mantenimiento,
resulta
coherente con
la gestión de
activos críticos.



B6. El uso de
registros
históricos
constituye una
alternativa
técnicamente
viable para
iniciar
estrategias
predictivas
cuando no se
dispone de
monitoreo
continuo o
sensorización
avanzada.



C. Parámetros temporales y criterios operativos *

	No pertinente / no adecuado	Poco pertinente / requiere cambios importantes	Pertinente / adecuado con ajustes menores	Muy pertinente / adecuado	N/A – No cuento con elementos suficientes para valorarlo
C1. Una ventana histórica de 60 días constituye un período razonable como punto de partida para representar el comportamiento reciente de operación y mantenimiento.	☐	☐	☐	☐	☐
C2. Un horizonte predictivo de 14 días constituye un período razonable para apoyar la planificación preventiva de corto plazo.	☐	☐	☐	☐	☐
C3. Una actualización semanal del riesgo constituye una periodicidad razonable para apoyar la planificación del mantenimiento.	☐	☐	☐	☐	☐
C4. Los parámetros temporales anteriores deberían recalibrarse posteriormente de acuerdo con el comportamiento de los datos reales de la institución donde se	☐	☐	☐	☐	☐

implemente la metodología.

C5. El uso de umbrales de alerta asociados a variables operativas y de mantenimiento resulta pertinente para identificar condiciones que requieran mayor atención técnica.



C6. Los valores numéricos específicos utilizados como umbrales de alerta deberían considerarse referencias iniciales y recalibrarse mediante información operacional real antes de su adopción definitiva.



D. Priorización y toma de decisiones *

	No pertinente / no adecuado	Poco pertinente / requiere cambios importantes	Pertinente / adecuado con ajustes menores	Muy pertinente / adecuado	N/A – No cuento con elementos suficientes para valorarlo
D1. Ante una clasificación de riesgo alto, resulta pertinente priorizar una evaluación técnica de la unidad en un período corto.	☐	☐	☐	☐	☐
D2. Ante una clasificación de riesgo medio, resulta pertinente programar una inspección técnica dentro del ciclo preventivo próximo.	☐	☐	☐	☐	☐
D3. Las unidades clasificadas con riesgo bajo pueden continuar bajo el cronograma preventivo establecido, manteniendo seguimiento periódico.	☐	☐	☐	☐	☐
D4. La prioridad propuesta por el modelo debería ser confirmada mediante criterio técnico antes de ejecutar una intervención específica.	☐	☐	☐	☐	☐

Us. La capacidad operativa disponible y el impacto de posibles falsas alertas deberían considerarse al establecer las prioridades finales de mantenimiento.



E. Aplicabilidad del protocolo operativo *

	No pertinente / no adecuado	Poco pertinente / requiere cambios importantes	Pertinente / adecuado con ajustes menores	Muy pertinente / adecuado	N/A – No cuento con elementos suficientes para valorarlo
E1. La actualización sistemática de los registros técnicos constituye una condición necesaria para mantener vigente una herramienta de soporte predictivo.	☐	☐	☐	☐	☐
E2. La ejecución periódica del modelo y la revisión posterior del ranking de riesgo resultan compatibles con un proceso de planificación preventiva.	☐	☐	☐	☐	☐
E3. La programación de intervenciones debería considerar conjuntamente el riesgo estimado y la evaluación del responsable técnico.	☐	☐	☐	☐	☐
E4. Los hallazgos, intervenciones realizadas y tiempos de inoperatividad deberían registrarse sistemáticamente para retroalimentar posteriores evaluaciones.					

E5. El ciclo de actualización, análisis, priorización, intervención y retroalimentación constituye una estructura operacional coherente para integrar progresivamente una herramienta predictiva.



E6. La estandarización y trazabilidad de los registros constituyen condiciones fundamentales para implementar la metodología con datos institucionales reales.



E6. La estandarización y trazabilidad de los registros constituyen condiciones fundamentales para implementar la metodología con datos institucionales reales.



La tercera sección de la encuesta presenta una consulta de valoración general de los principales puntos y variables de la investigación. Adicionalmente, se realizaron consultas abiertas de manera cualitativa para que el especialista pueda complementar su valoración de considerarlo necesario, estas preguntas son opcionales.

Valoración global *

	No pertinente / no adecuado	Poco pertinente / requiere cambios importantes	Pertinente / adecuado con ajustes menores	Muy pertinente / adecuado	N/A – No cuento con elementos suficientes para valorarlo
G1. En términos generales, las variables consideradas son técnicamente pertinentes para el propósito de la propuesta.	☐	☐	☐	☐	☐
G2. El enfoque de mantenimiento predictivo como soporte al mantenimiento preventivo presenta coherencia técnica.	☐	☐	☐	☐	☐
G3. Los criterios de priorización propuestos presentan utilidad potencial para apoyar la gestión del mantenimiento de ambulancias.	☐	☐	☐	☐	☐
G4. El protocolo operativo planteado presenta potencial de aplicabilidad dentro de un proceso institucional de mantenimiento.	☐	☐	☐	☐	☐
G5. Resulta pertinente continuar evaluando la metodología mediante una implementación piloto con registros institucionales reales.	☐	☐	☐	☐	☐

Preguntas abiertas

¿Considera que debería incorporarse alguna variable técnica u operacional adicional en una futura implementación de la metodología? En caso afirmativo, indíquela brevemente.

Tu respuesta

¿Cuál considera que sería la principal consideración o dificultad técnica para implementar una metodología de este tipo en una flota real de ambulancias?

Tu respuesta

¿Qué recomendación realizaría para fortalecer la aplicabilidad técnica u operacional de la propuesta?

Tu respuesta

La cuarta y última sección tuvo carácter declarativo y estuvo destinada a confirmar las respuestas proporcionadas en el cuestionario.

Declaración final

Declaro que las respuestas proporcionadas corresponden a mi criterio profesional independiente, sustentado en mi formación y experiencia, y que son emitidas exclusivamente con fines académicos. *

☐ Sí, confirmo.

K.2. Perfil de los profesionales participantes

Especialista	Profesión	Grado	Experiencia profesional	Áreas relevantes
ESP1	Ingeniero Biomédico	Título profesional	3–5 años	Ingeniería biomédica, equipamiento médico, ambulancias, proyectos de salud e IA/ML
ESP2	Ingeniera Biomédica	Bachiller	3–5 años	Ingeniería clínica, equipamiento médico, ambulancias, mantenimiento y proyectos de salud
ESP3	Gerente de Proyecto	Maestría	Más de 20 años	Ambulancias, equipamiento médico, gestión de activos, mantenimiento y proyectos de salud
ESP4	Ingeniero Biomédico	Título profesional	11–15 años	Ingeniería biomédica, ambulancias, mantenimiento, gestión de activos y equipamiento médico
ESP5	Ingeniera Biomédica	Título profesional	6–10 años	Ingeniería clínica, mantenimiento, equipamiento médico y proyectos de salud
ESP6	Ingeniero Electrónico	Maestría	16–20 años	Ambulancias, adquisiciones y coordinación de proyectos de salud
ESP7	Ingeniero Biomédico	Maestría	11–15 años	Ingeniería clínica, gestión de activos, mantenimiento y equipamiento médico

K.3. Matriz consolidada de respuestas

Codificación: 1 = No pertinente/no adecuado; 2 = Poco pertinente/requiere cambios importantes; 3 = Pertinente/adecuado con ajustes menores; 4 = Muy pertinente/adecuado.

A. Pertinencia de las variables e información utilizada

Especialista	A1	A2	A3	A4	A5	A6	A7	A8
ESP1	4	4	3	3	3	2	3	4
ESP2	4	3	4	4	4	4	4	3
ESP3	4	4	4	4	4	4	4	3

ESP4	4	4	3	4	4	4	4	4
ESP5	4	4	4	4	3	4	4	4
ESP6	4	4	3	4	4	4	3	4
ESP7	4	4	4	4	4	4	4	4

B. Coherencia del enfoque predictivo-preventivo

Especialista	B1	B2	B3	B4	B5	B6
ESP1	4	4	4	4	4	4
ESP2	4	4	4	3	3	4
ESP3	4	4	3	3	4	4
ESP4	3	4	4	3	4	4
ESP5	4	4	4	4	4	3
ESP6	4	4	4	3	4	4
ESP7	4	4	3	4	4	4

C. Parámetros temporales y criterios operativos

Especialista	C1	C2	C3	C4	C5	C6
ESP1	4	4	4	4	4	4
ESP2	4	4	4	4	4	4
ESP3	3	4	3	3	4	4
ESP4	3	3	4	4	4	4
ESP5	3	3	4	4	4	4
ESP6	3	3	4	4	3	4
ESP7	3	2	3	4	4	4

D. Priorización y toma de decisiones

Especialista	D1	D2	D3	D4	D5
ESP1	4	4	4	4	4
ESP2	4	3	4	4	3
ESP3	4	4	3	3	3
ESP4	4	3	4	4	4
ESP5	4	4	4	4	3
ESP6	4	3	4	4	4
ESP7	4	4	3	4	4

E. Aplicabilidad del protocolo operativo

Especialista	E1	E2	E3	E4	E5	E6
ESP1	4	3	4	4	3	4
ESP2	4	4	3	4	4	4
ESP3	3	4	4	4	4	4
ESP4	4	3	4	4	4	4
ESP5	4	4	4	4	4	4
ESP6	4	3	4	4	4	4
ESP7	4	4	4	4	3	4

F. Valoración global

Especialista	G1	G2	G3	G4	G5
ESP1	4	4	4	4	4
ESP2	3	4	4	4	4
ESP3	N/A	4	4	4	4
ESP4	4	4	4	4	4
ESP5	4	4	4	4	4
ESP6	4	4	4	3	4
ESP7	4	4	4	4	4

K.4. Resultados consolidados por dimensión

Dimensión	n respuestas válidas	Media	Mediana	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Favorables
Variables	56	3,77	4	0,0 %	1,79 %	19,64 %	78,57 %	98,21 %
Coherencia	42	3,79	4	0,0 %	0,00 %	21,43 %	78,57 %	100,00 %
Parámetros	42	3,67	4	0,0 %	2,38 %	28,57 %	69,05 %	97,62 %
Priorización	35	3,74	4	0,0 %	0,00 %	25,71 %	74,29 %	100,00 %
Protocolo	42	3,83	4	0,0 %	0,00 %	16,67 %	83,33 %	100,00 %
Valoración global	34	3,94	4	0,0 %	0,00 %	5,88 %	94,12 %	100,00 %
Total	251	3,78	4	0,0 %	0,80 %	19,92 %	79,28 %	99,20 %

K.5. Observaciones y recomendaciones profesionales

Especialista	Variable adicional sugerida	Principal consideración/dificultad	Recomendación
ESP1	—	—	—
ESP2	—	Disponibilidad de toda la información necesaria	Filtrar y depurar información no confiable
ESP3	Costos, telemetría y vida útil del vehículo considerando las condiciones geográficas de operación	Disponer de mecanismos obligatorios o automatizados para el registro de información; alternatively, incorporar esta obligación al servicio de mantenimiento del proveedor	Realizar una implementación piloto
ESP4	Rotación entre unidades y personal, con el fin de considerar posibles efectos asociados a la conducción	Factor conductor	Reforzar las buenas prácticas de uso y manejo, así como el reporte y documentación de eventos
ESP5	Antigüedad de la ambulancia	Disponibilidad de personal técnico	Incorporar la capacidad disponible de mantenimiento dentro de la priorización
ESP6	Intensidad de operación según zona	Disponibilidad e interoperabilidad de la información	Implementación piloto progresiva
ESP7	Condiciones de tráfico y entorno	Calidad y trazabilidad de los registros	Reentrenar periódicamente el modelo

Las observaciones recogidas identificaron principalmente la disponibilidad, calidad, estandarización y trazabilidad de los registros como condiciones relevantes para una futura implementación institucional. Asimismo, se señalaron potenciales variables y factores adicionales relacionados con los costos, la telemetría, la vida útil de las unidades, las condiciones geográficas de operación, el factor conductor y la rotación entre unidades y personal.

Entre las recomendaciones formuladas destacaron la implementación de mecanismos sistemáticos para el registro de información, el fortalecimiento de las buenas prácticas de uso y documentación de eventos, la depuración de información no confiable y el desarrollo de una

implementación piloto. En conjunto, estas observaciones complementan los lineamientos propuestos y refuerzan la necesidad de adaptar y recalibrar la metodología de acuerdo con las condiciones operativas y la disponibilidad de registros institucionales reales.