

DESARROLLO DE MODELO PREDICTIVO BASADO EN MACHINE LEARNING PARA
ANTICIPAR FALLAS EN EL PROBADOR DE SUSPENSIÓN DE LA LINEA DE LIVIANOS
DEL CENTRO DE DIAGNÓSTICO AUTOMOTOR CERTI EXPRESS PEREIRA S.A.S.

MAURICIO VALLEJO YEPES

Trabajo de grado

Asesor

Juan David Martínez Vargas

UNIVERSIDAD EAFIT
ESCUELA DE CIENCIAS APLICADAS E INGENIERÍA
MAESTRÍA EN INGENIERÍA
MEDELLÍN

2025

1

Agradecimiento

Quiero expresar mi agradecimiento a la personas que apoyaron el desarrollo de esta tesis. A mis profesores y asesor Juan David Martínez Vargas, por su guía constante, paciencia y valiosos comentarios. A mi familia, por el apoyo absoluto y la motivación diaria.

A mis padres, por ser mi ejemplo de esfuerzo, constancia y dignidad. Gracias por enseñarme, más con actos que con palabras, que los sueños se alcanzan con trabajo y humildad. Cada paso que doy lleva la fuerza de tus valores. Esta tesis es también fruto de tu apoyo silencioso, de tus consejos y de tu presencia firme, todo lo que logré tiene un pedazo de ti.

A mi hija Ana Lucía, quien solo con su presencia me inspira a superarme día tras día: esta meta también es para ella, porque me enseña con su inocencia, alegría y su ternura que todo esfuerzo en algún momento vale la pena.

A mi hermana, a quien aprecio y valoro por sus enseñanzas, aportes, tiempo y acompañamiento en este último tramo del proceso. Su testimonio y sus saberes han sido para mí un ejemplo digno de admirar.

-

CONTENIDO

pág.

0. INTRODUCCIÓN	11
1. PLANTEAMIENTO DEL PROBLEMA	13
2. JUSTIFICACIÓN	15
2.1 RELEVANCIA SOCIAL	15
2.2 RELEVANCIA INSTITUCIONAL	15
2.3 RELEVANCIA CIENTÍFICA / TÉCNICA	16
3. OBJETIVOS	18
3.1 GENERAL	18
3.2 ESPECÍFICOS	18
4. PREGUNTA DE INVESTIGACIÓN	19
5. MARCO REFERENCIAL	20
5.1 MARCO TEÓRICO	20
5.1.1 Mantenimiento industrial y gestión de activos.	20
5.1.2 Evolución hacia el mantenimiento predictivo.	20
5.1.3 Inteligencia artificial y Machine Learning aplicados al mantenimiento.....	21
5.1.4 Random Forest como herramienta para mantenimiento predictivo.....	22
5.2 MARCO CONCEPTUAL	23
5.3 MARCO LEGAL	24
5.3.1 Nivel normativo nacional:	24
5.3.2 Normas técnicas aplicables:	26
5.4 ANTECEDENTES	29
5.4.1 Antecedentes internacionales.	29
5.4.2 Antecedentes nacionales.....	38
5.4.3 Síntesis comparativa y brecha.	41
6. METODOLOGÍA	45
6.1 ENFOQUE Y TIPO DE INVESTIGACIÓN	45
6.2 MUESTRA	46

6.3 TÉCNICAS E INSTRUMENTOS	46
6.4 DISEÑO GENERAL DEL PROCESO (CRISP-DM)	47
6.4.1 Comprensión del Negocio.	47
6.4.2 Comprensión de los Datos.	47
6.4.3 Preparación de los Datos.	47
6.4.4 Modelado.....	47
6.4.5 Evaluación.	48
6.4.6 Despliegue.	48
6.5 FASES DEL PROCESO CRISP-DM	48
6.5.1 Comprensión del negocio.	48
6.5.2 Carga y presentación de librerías.....	48
6.5.3 Organización de directorios.	49
6.5.4 Carga de datos y verificación inicial.	49
6.5.5 Normalización de nombres y columnas.....	49
6.5.6 Estadísticas descriptivas.	50
6.5.7 Histogramas y boxplot.....	51
6.5.8 Detección de anomalías	55
6.5.9 Visualización de alertas	57
6.5.10 Exportación de resultados a Excel.....	58
6.5.11 Preparación de las variables para modelado	58
6.5.12 Entrenamiento y evaluación del modelo Random Forest	59
6.5.13 Visualizaciones del rendimiento del modelo	60
6.6 HERRAMIENTAS UTILIZADAS	62
6.7 CONSIDERACIONES ÉTICAS Y DE CALIDAD DE DATOS.....	62
7. RESULTADOS DE LA INVESTIGACIÓN.....	64
7.1 DESCRIPCIÓN GENERAL DE LOS DATOS	64
7.2 RESULTADOS DEL MODELADO PREDICTIVO	66
7.3 ANÁLISIS E INTERPRETACIÓN DE LOS RESULTADOS	67
7.4 SINTESIS DE HALLAZGOS.....	69
8. CONCLUSIONES	70
8.1 SINTESIS GENERAL DEL TRABAJO	70
8.2 PRINCIPALES HALLAZGOS.....	70
8.3 APOORTE TÉCNICO Y ACADÉMICO	71

8.4 LIMITACIONES Y OPORTUNIDADES DE MEJORA.....	71
9. RECOMENDACIONES.....	73
9.1 RECOMENDACIONES TÉCNICAS U OPERATIVAS.....	73
9.1.1 Integración con software de mantenimiento.....	73
9.1.2 Parámetros y ajustes para mantener el rendimiento.....	73
9.2 RECOMENDACIONES PARA LA ORGANIZACIÓN O SECTOR.....	73
9.2.1 Replicabilidad de la metodología en otros equipos.	73
9.2.2 Adopción de mantenimiento basado en datos.	74
9.3 RECOMENDACIONES PARA FUTURAS INVESTIGACIONES.....	74
9.3.2 Incorporación de sensores IoT para monitoreo en tiempo real.....	74
REFERENCIAS.....	76
ANEXOS.....	80

LISTA DE TABLAS

	pág.
Tabla 1. Datos de pruebas de adherencia	46
Tabla 2 . Columnas de adherencia	50
Tabla 3. resumen de anomalías	56

LISTA DE FIGURAS

	pág.
Fig. 1 Histograma adherencia eje 1 derecha	52
Fig. 2 Histograma adherencia eje 1 izquierda.	53
Fig. 3 Histograma adherencia eje 2 derecha.	53
Fig. 4 Histograma adherencia eje 2 izquierda.	54
Fig. 5 Boxplot comparativo	55
Fig. 6 Visualización de alertas.	58
Fig. 7 Matriz de confusión.	60
Fig. 8 Curva ROC.	61
Fig. 9 Resumen histogramas de adherencias	65
Fig. 10 Importancia de variables	66

LISTA DE ECUACIONES

	pág.
Ecuación 1 IQR Rango intercuartílico	56
Ecuación 2 límite inferior de iqr	56
Ecuación 3 límite superior iqr	57
Ecuación 4 límite inferior de umbral k	57
Ecuación 5 límite superior de umbral k	57

LISTA DE ANEXOS

pág.

Anexo A. Modelo Random Forest

80

Resumen

El mantenimiento predictivo se ha consolidado como un mecanismo fundamental con el fin de garantizar la confiabilidad, la disponibilidad y la eficiencia de los equipos industriales. En el contexto automotriz colombiano, los Centros de Diagnóstico Automotor (CDA) desempeñan un papel crucial en la seguridad vial, al asegurar que los vehículos cumplan con las condiciones técnicas exigidas por la normativa nacional. Sin embargo, los métodos de mantenimiento correctivo y preventivo empleados tradicionalmente generan tiempos de inoperables y sobrecostos que afectan la operación continua de los equipos de inspección.

Esta investigación desarrolla una solución basada en aprendizaje automático (Machine Learning) con el objetivo de analizar los resultados de los ensayos de adherencia obtenidas mediante el probador de suspensión de la línea de vehículos livianos del Centro de Diagnóstico Automotor Certi Express Pereira S.A.S. Se recopilaron 2.360 registros de pruebas y, tras un proceso de limpieza, normalización y análisis de variables, se construyó un modelo predictivo capaz de detectar patrones asociados a fallas potenciales tanto en los equipos de medición como en los vehículos inspeccionados.

Los resultados evidencian que la aplicación de modelos de aprendizaje automático en entornos de diagnóstico automotor mejora significativamente la detección temprana de anomalías, permite optimizar la planificación del mantenimiento y contribuye a reducir los costos operativos y los tiempos de inactividad.

Palabras clave: *mantenimiento predictivo, inteligencia artificial, análisis de datos, aprendizaje automático.*

0. INTRODUCCIÓN

En Colombia, los Centros de Diagnóstico Automotor (CDA) desempeñan un rol fundamental en la seguridad vial al verificar que los vehículos cumplan con las condiciones técnicas exigidas por la ley. Estándares nacionales como NTC 5385 de 2011 especifica que los equipos para inspeccionar en los CDA sean sometidos a un mantenimiento preventivo, esto con el fin de asegurar su disponibilidad continua e integridad operativa [1]. Sin embargo, en la realidad hay esquemas tradicionales de mantenimiento preventivo (en tiempos regulares) y correctivo (después de la falla) que causan interrupciones operativas y altos costos que ponen riesgo en la continuidad del servicio de inspección vehicular. Este entorno nos hace tener una necesidad de mejorar las estrategias de mantenimiento, para estar en cumplimiento con las normas sin perder eficiencia en el funcionamiento.

Actualmente, el equipo que prueba la suspensión de la línea de vehículos livianos enfrenta problemas de confiabilidad y disponibilidad. No hay mecanismos para confirmar el realismo de los resultados en los ensayos de adherencia, tampoco para detectar oportunamente fallas progresivas o desajustes del equipo. Dado el caso, una anomalía incipiente en el probador podría pasarse por alto y comprometer la exactitud de la revisión técnico-mecánica, mientras que una avería repentina puede dejar sin funcionamiento toda la línea de inspección. Esto último hace que se tengan altos costos por tiempos de inactividad y supone el incumplimiento de las exigencias normativas de disponibilidad del equipo, según NTC 5385. La criticidad del probador en la operación diaria resalta una urgencia para contar con herramientas que anticipen las fallas antes de que pasen, evitando paradas no planificadas.

Para abordar la problemática, la literatura reciente nos sugiere tomar una adopción de técnicas avanzadas de inteligencia artificial (IA) y/o aprendizaje automático (Machine Learning, ML) en mantenimiento industrial. Los modelos de ML tienen la capacidad de analizar una gran cantidad de datos históricos en tiempo real identificando patrones de comportamiento, irregularidades operativas, y nos permiten prever fallas. Específicamente, los modelos basados en árboles de decisión han ganado popularidad en el mantenimiento predictivo por la gran interpretabilidad y robustez que nos ofrecen; entre ellos, el algoritmo Random Forest ha sido consolidado como una herramienta versátil y eficaz para entornos industriales heterogéneos. Estudios recientes nos dicen que Random Forest supera a

enfoques más tradicionales en cuanto a la detección temprana de fallas: Yang y Wang (2025) lograron precisiones cerca de 99.5 % al predecir los fallos en la maquinaria usando Random Forest, demostrando la gran capacidad para manejar patrones complejos y no lineales [2]. A su vez, una revisión del estado del arte nos confirma la relevancia de Random Forest para anticipar fallas en los equipos críticos, recomendando su implementación en escenarios automotrices dada su alta efectividad predictiva. Estas evidencias técnico-científicas respaldan el uso de ML –y especialmente de Random Forest– como la base para un modelo de mantenimiento predictivo en el probador de suspensión vehicular.

En este contexto, el objetivo general de la investigación es desarrollar un modelo predictivo basado en técnicas de aprendizaje automático que permita anticipar fallas en el equipo probador de suspensión de la línea de vehículos livianos del CDA Certi Express Pereira S.A.S., contribuyendo a la implementación de estrategias de mantenimiento predictivo. Dicho modelo buscará mejorar la confiabilidad del equipo y reducir al mínimo los tiempos de inactividad no planificados, alineándose con las necesidades operativas y normativas del sector.

Para cumplir con el objetivo, se acogió un planteamiento de investigación cuantitativo, siguiendo el método estándar CRISP-DM (Cross Industry Standard Process for Data Mining) el cuál desarrolla una gestión estructurada del proyecto de datos. El estudio se basó en el análisis de 2.360 registros históricos sobre pruebas de adherencia obtenidos de un probador de suspensión, estos fueron depurados, normalizados y explorados para obtener las características más relevantes y necesarias. Sobre esta base de datos se entrenó un modelo de clasificación Random Forest, el cuál es capaz de aprender patrones asociados a condiciones anómalas, poco comunes o indicativas de que hay una falla incipiente. El desempeño del modelo se evaluó a través de métricas de precisión, sensibilidad y exactitud, comparándolo con los demás parámetros de las prácticas de mantenimiento actuales. En síntesis, la metodología propuesta –basada en CRISP-DM y en el uso de Random Forest sobre datos reales de operación– permite generar alertas tempranas de posibles fallas en el equipo, optimizando la planificación del mantenimiento y asegurando la continuidad operativa del CDA.

1. PLANTEAMIENTO DEL PROBLEMA

El área de mantenimiento permanece en constante evolución supliendo las necesidades que se generan en los diferentes sectores de la industria a medida que se implementan nuevas tecnologías en los procesos industriales. J. W. Zanelli, sostiene que, en sus inicios, el mantenimiento se limitaba a una estrategia reactiva, donde se actuaba únicamente después de la ocurrencia de una falla [3], esta estrategia tenía grandes desventajas como la indisponibilidad de una máquina y costos elevados, especialmente cuando la criticidad del equipo era alta y las producciones eran a gran escala. Luego, la industria empezó a identificar estos inconvenientes, lo que llevó a que se generaran programación de inspecciones y algunas tareas para prevenir fallas, es allí donde surge el mantenimiento preventivo, ello produce aumentos en la eficiencia operativa y la vida útil de los equipos, lo cual permite una reducción de los tiempos de inactividad. El avance de las tecnologías de monitoreo y el análisis de datos han impulsado la adopción del mantenimiento predictivo. A diferencia del mantenimiento preventivo, que se basa en intervalos de tiempo predefinidos, el predictivo utiliza datos en tiempo real para prever fallas antes de que ocurran [3], permitiendo una intervención de mantenimiento en un momento preciso.

Los centros de diagnóstico automotor (CDA) según la NTC 5385 de 2011, numeral 4.17 Mantenimiento, los equipos deben recibir el mantenimiento programado para asegurar su permanente disponibilidad e integridad [1], de acuerdo con esto, los CDA realizan un mantenimiento preventivo cada cierto tiempo o número de inspecciones y mantenimientos correctivos cuando se presentan fallas, los mantenimientos preventivos generalmente se realizan días antes de calibrar o verificar los equipos y a su vez se realizan de acuerdo con un cálculo de intervalos de calibración y verificación, lo que indica que se tiene una certeza de que el equipo se encuentra en perfectas condiciones al momento de calibrarlo o verificarlo, pero no en el periodo de funcionamiento de la máquina hasta cuando se vuelve a calibrar o a verificar.

Las consecuencias de no solucionar el problema radican en que el equipo en cualquier momento puede presentar síntomas de falla sin que sean perceptibles por los operarios que lo manipulan a diario, ya que este personal no tiene conocimiento de los datos arrojados de las pruebas realizadas, si se presentara falla o desajuste de la maquina se perdería la veracidad de los resultados de las

inspecciones y por ende el resultado de la revisión técnico mecánica y de emisiones contaminantes. Si se presentara una falla parcial o total que impida el buen funcionamiento del equipo generaría una indisponibilidad de este y a su vez de la prestación del servicio para vehículos livianos ya que no se ejecutarían la totalidad de las pruebas de revisión, lo que conlleva a unos sobrecostos altos para la empresa y al no cumplimiento del ítem 4.17 de la NTC 5385 de 2011 [1].

Hasta el momento no existen parámetros para identificar si existe veracidad de los resultados de las pruebas de adherencia, ni para detectar cuando se presentan fallas progresivas o desajustes del equipo lo que afecta la realidad de los resultados obtenidos de las pruebas realizadas y a su vez la calidad, puesto que no se tiene certeza de los valores obtenidos.

En este orden de ideas, la falta del uso de modelos Machine Learning ocasiona que los resultados obtenidos de las pruebas de adherencia no sean confiables, además una falla repentina causa indisponibilidad de un equipo con una criticidad alta y la inoperancia de la línea de inspección de livianos; por lo que se hace necesario investigar de qué manera puede anticiparse a las fallas en el probador de suspensión de la línea de livianos de revisión técnico mecánica y emisiones contaminantes del CDA Certi Express Pereira s.a.s.

2. JUSTIFICACIÓN

2.1 RELEVANCIA SOCIAL

El mantenimiento predictivo en equipos de diagnóstico vehicular tiene un impacto directo en la seguridad vial. Al anticipar fallas mecánicas críticas antes de que ocurran, se evitan averías repentinas durante la operación de los vehículos. Por ejemplo, la detección temprana de problemas (como la pérdida de eficiencia de frenos o un motor sobrecalentado) mediante IA permite corregirlos oportunamente, reduciendo el riesgo de accidentes en carretera [4]. Un equipamiento de diagnóstico fiable garantiza que los vehículos circulantes cumplan las condiciones de seguridad exigidas, protegiendo tanto a conductores como a peatones. Asimismo, la sostenibilidad ambiental se ve favorecida: al predecir necesidades de mantenimiento, se optimiza el uso de piezas y recursos, evitando reemplazos prematuros y disminuyendo residuos. La adopción de estrategias de mantenimiento predictivo prolonga la vida útil de la maquinaria y minimiza el impacto ambiental al reducir desperdicios y emisiones. Estos beneficios se alinean con los Objetivos de Desarrollo Sostenible; en particular, ODS 9 (Industria, Innovación e Infraestructura) [5] promueve infraestructuras resilientes e innovación industrial. La integración de IA y análisis de datos en el mantenimiento representa una innovación que mejora la resiliencia de las infraestructuras de transporte (p.ej., centros de diagnóstico vehicular), apoyando el desarrollo sostenible de la industria automotriz. En conjunto, un modelo predictivo de fallas contribuye a una movilidad más segura y limpia, acorde con las metas sociales de seguridad vial y protección ambiental.

2.2 RELEVANCIA INSTITUCIONAL

Para un Centro de Diagnóstico Automotor (CDA), implementar mantenimiento predictivo con técnicas de Machine Learning conlleva importantes beneficios operativos y económicos. Estos modelos predictivos logran minimizar las paradas imprevistas de los equipos y reducir los costos de mantenimiento, llegando a ahorros estimados de entre un 10% y 40% en gastos de mantenimiento gracias a la prevención proactiva de falla [6]. Evitar fallos catastróficos en un probador de suspensión implica menos interrupciones en la prestación del servicio, lo cual se traduce en mayor productividad:

el CDA puede atender consistentemente a los vehículos sin retrasos ni tiempos muertos. Además, al prevenir el desgaste excesivo y las averías graves, se extiende la vida útil del equipamiento de diagnóstico, retrasando inversiones en nuevas máquinas y mejorando el retorno de inversión en infraestructura tecnológica [7]. La mayor disponibilidad y confiabilidad de los equipos también redundan en una mejor calidad del servicio ofrecido; los clientes (propietarios de vehículos) experimentan un proceso de inspección más ágil y confiable, incrementando su satisfacción y la reputación del CDA. Cabe destacar que la planificación anticipada posibilitada por estos modelos permite programar las intervenciones de mantenimiento en horarios de baja afluencia, evitando interrumpir la operación normal del centro [4]. En resumen, la institución se beneficia con una operación más eficiente, económicamente rentable y orientada a la calidad, al incorporar herramientas de IA para el mantenimiento de sus equipos críticos.

2.3 RELEVANCIA CIENTÍFICA Y TÉCNICA

Desde el punto de vista técnico y también científico, el proyecto se encuentra en las tendencias actuales de investigación en mantenimiento predictivo e inteligencia artificial aplicada. En los últimos años (2023–2025), la comunidad académica ha enfatizado el paso de mantenimiento reactivo a mantenimiento proactivo en la industria automotriz, habilitado por análisis de datos y aprendizaje automático [7]. Numerosos estudios reportan que la combinación de sensores IoT con algoritmos de Machine Learning logran diagnosticar bien los fallos incipientes en vehículos o equipos con antelación, aumentando su fiabilidad y la eficiencia operativa [7]. Teniendo en cuenta lo anterior, los algoritmos de clasificación como Random Forest se han hecho un puesto entre las técnicas más eficaces y populares para tareas de mantenimiento predictivo, dada su gran capacidad de manejar conjuntos de datos complejos y obtener una alta precisión en la predicción de las fallas [8]. De hecho, los modelos de Random Forest y demás variantes de ensamble han demostrado obtener unos resultados sobresalientes en los distintos estudios industriales.

El desarrollo de un modelo predictivo para un probador de suspensión en un CDA representa un aporte novedoso al conocimiento técnico. Gran parte de la literatura existente sobre mantenimiento predictivo automotriz se centra en componentes de vehículos (motores, baterías, etc.), por lo que aplicar estas técnicas a los equipos de diagnóstico del vehículo extiende el alcance de la inteligencia

artificial en la industria. Nuestro proyecto funcionará como un estudio de caso real sobre cómo las herramientas de Machine Learning pueden anticipar las fallas en maquinaria especializada, contribuyendo de esta manera a unas mejores prácticas de Smart Maintenance. Validar el desempeño de algoritmos como Random Forest en este dominio específico aportará datos valiosos a la comunidad científica: por ejemplo, investigaciones recientes han logrado más del 90% de exactitud al predecir fallos en sistemas automotrices usando Random Forest, demostrando el potencial de estas técnicas para mejorar la confiabilidad de los equipos [9]. En suma, el proyecto genera nuevo conocimiento en la intersección de mantenimiento industrial y aprendizaje automático, reforzando las bases técnicas de la siguiente desarrollo de sistemas de diagnóstico vehicular inteligentes. Las lecciones derivadas (manejo de datos de sensores, modelos predictivos entrenados, integración en la operación real de un CDA) enriquecerán el campo de la ingeniería de mantenimiento y la aplicación práctica de la IA en la industria automotriz, sirviendo de referencia para futuros desarrollos e investigaciones en este ámbito.

3. OBJETIVOS

3.1 GENERAL

Desarrollar un modelo predictivo basado en técnicas de aprendizaje automático para anticipar fallas en el equipo probador de suspensión de la línea de vehículos livianos del Centro de Diagnóstico Automotor Certi Express Pereira S.A.S., contribuyendo a la implementación de estrategias de mantenimiento predictivo.

3.2 ESPECÍFICOS

- Analizar los registros históricos de pruebas de adherencia obtenidos del probador de suspensión y determinar las variables más relevantes para el modelado.
- Implementar modelos de aprendizaje automático para predecir fallas en el equipo y necesidades de mantenimiento en los vehículos inspeccionados.
- Evaluar el desempeño del modelo propuesto mediante métricas de sensibilidad, precisión y exactitud, comparándolo con enfoques tradicionales de mantenimiento.
- Proponer recomendaciones técnicas y operativas derivadas de los resultados del modelo para optimizar la gestión del mantenimiento en el CDA.

4. PREGUNTA DE INVESTIGACIÓN

¿De qué manera el uso de modelos de aprendizaje automático puede anticipar fallas en el equipo probador de suspensión de la línea de vehículos livianos del Centro de Diagnóstico Automotor Certi Express Pereira S.A.S., mejorando la eficiencia del mantenimiento y la disponibilidad operativa de los equipos?

5. MARCO REFERENCIAL

5.1 MARCO TEÓRICO

5.1.1 Mantenimiento industrial y gestión de activos.

El mantenimiento industrial se define como el conjunto de prácticas orientadas a preservar la disponibilidad y confiabilidad de los equipos productivos, garantizando su funcionamiento continuo bajo condiciones seguras y eficientes. Esta disciplina integra actividades técnicas —inspección, reparación, sustitución y monitoreo— junto con acciones administrativas enfocadas en la gestión de recursos y la planificación estratégica del mantenimiento.

Según Pérez Rondón [10], la gestión en el mantenimiento nos hace tener unos objetivos fundamentales. Los objetivos constan de optimizar la disponibilidad de los equipos, reducir los costos asociados, maximizar la vida útil de los activos y hacer un uso eficiente de los recursos humanos. Estos mismos objetivos han impulsado la evolución del mantenimiento desde el correctivo tradicional hacia enfoques preventivos, predictivos y proactivos, que satisfacen con más eficacia a los distintos requerimientos de la industria contemporánea, caracterizada por altos niveles de competitividad y la alta necesidad de continuidad operacional.

5.1.2 Evolución hacia el mantenimiento predictivo.

El mantenimiento predictivo (Predictive Maintenance, PdM) tiene una característica importante que es anticipar fallas mediante el análisis de datos históricos y en tiempo real. Ucar et al. [11] lo ponen como la política que se basa en datos que permiten tener una idea del momento probable de una o varias fallas y detectar anomalías operativas, dando paso a intervenciones con antelación y sobre todo planificadas para que no afecten la disponibilidad del sistema.

El PdM es considerado un engranaje clave dentro de la Industria 4.0, esto debido a la integración con tecnologías de sensorización avanzadas, sistemas ciberfísicos, análisis de datos e inteligencia artificial. Tapia et al. [12] nos destaca que esta estrategia contribuye a la maximización de la

disponibilidad en los activos, lo cual permite garantizar la continuidad operativa y mejorar la eficiencia global de los procesos productivos. A su vez, Taoufyq et al. [13] también nos destaca que el PdM emplea datos en tiempo real para, como consecuencia de lo anterior, anticipar fallos, optimizar recursos y reducir tiempos de inactividad, lo que nos da significativamente beneficios operativos y económicos.

La relevancia sobre todo económica del PdM nos lo demuestra en diversos casos documentados. Por ejemplo, Amazon perdió aproximadamente 4 millones de dólares por 49 minutos de inactividad en 2013 [14], esto nos dice que incluso teniendo pequeñas interrupciones breves, se pueden generar grandes pérdidas. Este tipo de situaciones evidencian la urgencia de anticipar las fallas y planificar unas intervenciones **oportunas**, lo que nos permite minimizar riesgos operativos y mejorar la satisfacción de los clientes.

Martínez García [15] nos dice que el mantenimiento predictivo permite pronosticar la o las fallas futuras de un componente para poder reemplazarlo justo antes de que ocurra, minimizando el tiempo muerto y maximizando la vida útil del o de los equipos. Esto es bastante importante debido a que en sectores como la inspección vehicular, donde la confiabilidad, continuidad del servicio y precisión de los equipos son críticas para cumplir la promesa de valor al cliente.

5.1.3 Inteligencia artificial y Machine Learning aplicados al mantenimiento.

En el marco de la transformación digital, la inteligencia artificial (IA) y el aprendizaje automático (Machine Learning, ML) se han puesto como herramientas de carácter fundamental para el desarrollo de las estrategias de mantenimiento avanzado. Los sistemas industriales generan volúmenes de datos muy grandes, que pueden ser examinados para poder identificar patrones operativos, predecir fallos y optimizar la toma de decisiones.

Los algoritmos de ML nos permiten clasificar estados de operación, estimar las fallas futuras y detectar los comportamientos extraños o anómalos con un análisis de variables relevantes del equipo. Su capacidad para aprender de datos históricos y generar predicciones precisas nos dicen que mejoran la confiabilidad y disponibilidad de los activos, también ayudan a reducir el tiempo de inactividad no programado y optimizar la planificación de recursos.

Hay varias técnicas empleadas para el mantenimiento predictivo, entre ellas podemos destacar los métodos de aprendizaje supervisado, sobre todo los modelos basados en árboles de decisión debido a su interpretabilidad, robustez y capacidad para manejar los datos heterogéneos. El algoritmo Random Forest se ha posicionado como un modelo muy sólido y versátil para análisis en entornos industriales.

5.1.4 Random Forest como herramienta para mantenimiento predictivo.

El algoritmo Random Forest, desarrollado por Breiman [16], es un conjunto de árboles de decisión que nos introducen aleatoriedad tanto en la selección de características como en las muestras utilizadas para entrenar cada árbol. Debido a esta estructura, permite que el modelo aprenda múltiples relaciones entre las variables del sistema, mejorando la capacidad de generalización además de reducir el riesgo de sobreajuste gracias al principio de la Ley de los Grandes Números.

Entre sus beneficios más relevantes en aplicaciones industriales se encuentran:

- Eficiencia en el manejo de grandes volúmenes de datos y múltiples variables.
- Una gran capacidad para modelar relaciones no lineales entre parámetros operativos.
- Robustez ante datos ruidosos, valores atípicos y/o información incompleta.
- Extracción de medidas de importancia de variables, útiles para interpretar comportamientos que tiene el equipo.
- Alto nivel de precisión tanto en tareas de clasificación como de regresión.

Estas características hacen que Random Forest sea especialmente adecuado para sistemas de mantenimiento predictivo en equipos donde las señales de operación pueden ser complejas o donde existe alta variabilidad en los datos. En el caso de equipos probadores de suspensión vehicular, el algoritmo permite analizar parámetros como adherencia, vibraciones y comportamiento dinámico, detectando patrones anómalos que indiquen desgaste o la necesidad de intervención.

El uso de modelos como Random Forest en PdM ha permitido revolucionar la gestión de activos al mejorar la disponibilidad, confiabilidad y eficiencia operativa mediante decisiones apoyadas en datos. Esta aproximación aporta una visión integral y en tiempo real del estado de los activos, lo que facilita la programación oportuna del mantenimiento y reduce los riesgos operacionales asociados a fallas inesperadas.

5.2 MARCO CONCEPTUAL

En el marco de esta investigación se establecen los conceptos fundamentales que sustentan el desarrollo del modelo predictivo orientado a anticipar fallas en un equipo probador de suspensión vehicular. Estos conceptos permiten contextualizar el enfoque técnico y metodológico adoptado.

- **Mantenimiento industrial:** conjunto de actividades destinadas a asegurar que los equipos, máquinas, instrumentos e instalaciones de un entorno productivo operen bajo las condiciones para las cuales fueron creados, instalados y puestos en funcionamiento. Su objetivo es mantener la disponibilidad y confiabilidad del sistema productivo mediante intervenciones planificadas y correctivas basadas en criterios técnicos [10].
- **Mantenimiento predictivo:** estrategia de gestión que utiliza el estado real de los equipos para planificar intervenciones únicamente cuando son necesarias. Se apoya en técnicas de monitoreo de condición —como análisis vibracional, termografía o tribología— para estimar el comportamiento futuro de los sistemas y programar actividades de mantenimiento de manera óptima, incrementando la confiabilidad operativa y reduciendo los tiempos de inactividad [17].
- **Inteligencia artificial (IA):** disciplina computacional orientada al desarrollo de algoritmos capaces de ejecutar tareas que usualmente requieren habilidades cognitivas humanas, tales como el aprendizaje, el razonamiento lógico o la toma de decisiones [18]. En este estudio, la IA constituye la base para el desarrollo de algoritmos predictivos basados en datos.
- **Aprendizaje automático (Machine Learning):** subcampo de la IA enfocado en el diseño de modelos que aprenden patrones a partir de datos y generan predicciones o decisiones sin programación explícita. Su desempeño mejora progresivamente con la exposición a un mayor

volumen y sobre todo calidad de datos [19]. Esta investigación emplea técnicas de aprendizaje supervisado para anticipar fallas en el equipo de ensayo.

- **Random Forest:** algoritmo de aprendizaje supervisado compuesto por un conjunto de árboles de decisión entrenados mediante muestreo aleatorio con reemplazo (bootstrapping). Cada árbol genera una predicción individual y el modelo final se obtiene mediante la agregación (promedio o votación) de todas las predicciones. Su estructura favorece la reducción de sobreajuste y mejora la precisión en tareas de clasificación y regresión [20].
- **Ciencia de datos:** campo interdisciplinario que integra estadística, matemáticas, programación, análisis avanzado, inteligencia artificial y conocimiento del dominio para extraer información valiosa a partir de grandes volúmenes de datos. Dichos conocimientos alimentan procesos de toma de decisiones y desarrollo de estrategias de mantenimiento basadas en evidencia [21].

En conjunto, estos conceptos proporcionan el soporte teórico para el desarrollo de un sistema de mantenimiento predictivo basado en técnicas de aprendizaje automático, orientado a mejorar la confiabilidad del probador de suspensión y optimizar la toma de decisiones en el contexto del mantenimiento industrial.

5.3 MARCO LEGAL

5.3.1 Nivel normativo nacional:

De acuerdo con la normativa colombiana vigente, la Ley 769 de 2002, correspondiente al Código Nacional de Tránsito Terrestre, establece los lineamientos generales para garantizar la seguridad vial y la protección ambiental en el territorio nacional. En su Capítulo VIII, la ley define las disposiciones

relacionadas con la revisión técnico-mecánica y de emisiones contaminantes de los vehículos automotores [22].

El Artículo 50 establece la obligación de los propietarios o tenedores de vehículos, nacionales o extranjeros, de mantenerlos en condiciones mecánicas y de seguridad óptimas con el fin de preservar la seguridad vial y mitigar el impacto ambiental.

El Artículo 51 regula la periodicidad y los criterios de verificación para la revisión técnico-mecánica de vehículos de servicio público, escolar y de turismo, los cuales deben someterse a revisión anual. Los vehículos de servicio particular deben hacerlo cada dos años. La revisión comprende, entre otros aspectos:

1. Estado estructural de la carrocería.
2. Verificación de niveles de emisiones contaminantes conforme a la legislación ambiental vigente.
3. Funcionamiento del sistema mecánico.
4. Operatividad del sistema eléctrico y conjunto óptico.
5. Eficiencia del sistema de combustión.
6. Disponibilidad y estado de los elementos de seguridad.
7. Condiciones del sistema de frenos, considerando también niveles acústicos permitidos en sistemas neumáticos.
8. Estado de las llantas.
9. Funcionamiento de la puerta de emergencia.
10. Funcionamiento de los dispositivos de cobro para vehículos de servicio público.

El artículo contempla además que vehículos especiales como ambulancias, camiones de bomberos y recolectores de basura se asimilan reglamentariamente a vehículos de servicio público en lo referente a estas revisiones.

El Artículo 52 establece la periodicidad de la revisión de gases: anual para vehículos de servicio público y bienal para los de servicio particular. Los vehículos nuevos deben someterse a esta

evaluación a partir de dos años después de su matrícula. La revisión debe realizarse exclusivamente en centros de diagnóstico automotor autorizados.

El Artículo 53 especifica los requisitos de operación para los centros de diagnóstico automotor, los cuales deben estar legalmente constituidos y cumplir con los reglamentos técnicos del Ministerio de Transporte y del Ministerio de Ambiente. El certificado de revisión es obligatorio para la circulación del vehículo.

Finalmente, el Artículo 54 indica que los resultados de las revisiones deben ser registrados en sistemas computarizados, incluyendo los casos en los que el vehículo no cumple con los requisitos establecidos.

5.3.2 Normas técnicas aplicables:

5.3.2.1 Norma Técnica Colombiana NTC 5385:2011, relacionada con las especificaciones del servicio en los Centros de Diagnóstico Automotor (CDA), se establecen los requisitos que deben cumplir los equipos utilizados en la revisión técnico-mecánica y de emisiones contaminantes de vehículos automotores [1]. Esta normativa no incluye disposiciones asociadas a la medición de emisiones por presión sonora (ruido), debido a que estas se encuentran reguladas de manera independiente por la legislación vigente.

En concordancia con lo anterior, se destacan los numerales pertinentes a los requisitos del analizador de suspensiones, los cuales se describen a continuación.

- Mantenimiento de equipos (Numeral 4.17)

Los equipos deben someterse a un programa de mantenimiento preventivo que garantice su disponibilidad operativa e integridad funcional, siguiendo estrictamente las recomendaciones del fabricante y las exigencias normativas aplicables [1]. El CDA debe contar con procedimientos documentados y personal competente para realizar estas labores o, en su defecto, contratar servicios especializados.

El mantenimiento debe incluir, como mínimo:

1. Verificación diaria del funcionamiento adecuado de los equipos y puesta a cero cuando corresponda, de acuerdo con las recomendaciones del fabricante.
 2. Actividades de limpieza y verificación general.
 3. Elaboración y actualización de bitácoras de mantenimiento.
 4. Implementación de un programa de mantenimiento coherente con el volumen de inspecciones y las especificaciones técnicas de los equipos.
- Calibración de equipos de medición (Numeral 4.18)

Para los equipos utilizados en la realización de las pruebas definidas en la NTC 5375, se debe implementar un programa formal de calibración que garantice la trazabilidad metrológica. En este contexto, el error máximo permitido para el analizador de suspensiones es de ± 3 % del valor medido [1].

- Equipos de la revisión (Numeral 4.19)

1) Requisitos generales (4.19.1.1)

Ningún equipo instalado en la línea de revisión debe permitir que el inspector conozca anticipadamente el resultado final de la prueba, ya sea que el vehículo sea aceptado o rechazado.

2) Línea de revisión para vehículos livianos (4.19.2)

Los equipos destinados a la línea de revisión de vehículos livianos deben soportar una capacidad máxima de 3 500 kg.

a) Analizador de suspensiones (4.19.2.5)

El analizador debe operar bajo el principio EUSAMA (European Shock Absorber Manufacturer Association) o mediante un método equivalente que proporcione resultados en términos de adherencia, conforme al numeral 6.8.2 de la NTC 5375 [1].

Los requisitos mínimos del equipo son:

- Máximo peso por rueda: 1 000 kg.
- Precisión en la medición de peso: ± 3 %.
- Resolución de peso: 1 kg.

5.3.2.2 Norma Técnica Colombiana NTC 5375:2011 [23], la revisión técnico–mecánica y de emisiones contaminantes incluye la evaluación del probador de suspensión bajo el método EUSAMA. En este procedimiento, la adherencia (A) se define como la relación porcentual entre la fuerza vertical mínima de contacto registrada durante la oscilación de la rueda y el peso en vacío correspondiente a la misma rueda. La adherencia se determina mediante la ecuación:

$$A = 100 \cdot \frac{F_{\min}}{P}$$

donde:

- A es la adherencia en porcentaje [%],
- F_{\min} es la fuerza vertical mínima durante la oscilación [N],
- P corresponde a la masa en vacío de la rueda multiplicada por la aceleración de la gravedad (9.81 m/s^2).

Relación con el estudio

Los CDA deben de disponer de un plan de mantenimiento que revise hasta los más mínimos detalles y tenga una gran confiabilidad, que nos garantice un cumplimiento de los requisitos normativos establecidos en la NTC 5375 [23]. El adecuado mantenimiento y calibración del probador de suspensión nos asegura que tengamos una validez metrológica de los certificados de verificación y a su vez una gran exactitud de los rangos de medida del equipo.

Lo anterior dicho nos dice que la implementación de un modelo predictivo basado en técnicas de aprendizaje automático tiene un aporte grande al cumplimiento normativo. Un análisis continuo de los resultados de adherencia nos permite identificar las diferentes tendencias en los datos y anticipar posibles fallas del probador de suspensión, fortaleciendo de esta manera la disponibilidad, confiabilidad y trazabilidad del equipo dentro del proceso de inspección vehicular.

5.4 ANTECEDENTES

La investigación contó con 18 referentes, de los cuales 14 fueron internacionales y 4 nacionales. Según estos antecedentes abordamos el mantenimiento predictivo como una estrategia que permite la optimización de recursos además de permitir una gran eficacia en los procesos operativos. Estos estudios nos resultan fundamentales puesto que ofrecen diversas perspectivas sobre los factores que influyen en la disponibilidad y confiabilidad de los equipos. Su análisis permite comprender la importancia de implementar estrategias adecuadas para la reducción de fallas inesperadas y el uso adecuado de los activos industriales desde el análisis de datos en tiempo real.

5.4.1 Antecedentes internacionales.

Mantenimiento predictivo basado en Machine Learning: una revisión sistemática de la literatura y perspectivas en la Industria 4.0

Esta investigación consiste en la revisión sistemática sobre técnicas ML para PdM en contextos Industria 4.0, lo cual busca sintetizar avances, retos y aplicaciones de ML en PdM para maximizar disponibilidad y excelencia operacional, mediante una revisión sistemática de literatura reciente; clasificación por técnicas, dominios y métricas y obtuvo como resultado la identificación del auge de sensores/IIoT, técnicas dominantes (RF, SVM, Deep Learning), y brechas en generalización y datos etiquetados.[12]

S. Tapia et al. (2024) mediante este estudio aporta a la presente investigación como una evidencia de que el modelo ML es una herramienta adecuada para anticipar fallas y mejorar la confiabilidad de los resultados, contribuyendo al cumplimiento normativo y la calidad; además este artículo ofrece algunas metodologías que pueden ser de utilidad para diseñar un modelo predictivo de mantenimiento.

Impacto de la Inteligencia Artificial en la gestión de mantenimiento predictivo en la industria

Esta publicación se refiere al estudio sobre el impacto de IA/ML en PdM: análisis de datos, optimización de recursos y casos, consiste en evaluar cómo la IA mejora la gestión PdM y la toma de decisiones, mediante una investigación de corte aplicada/descriptiva con revisión y ejemplos industriales; cuyos resultados fueron reportar mejoras en tiempo de inactividad, planificación y uso de recursos; recalca la dependencia de grandes volúmenes de datos de alta calidad, puesto que la efectividad de un modelo predictivo puede verse afectada debido a datos incompletos o erróneos.[24]

J. A. Arizaga Mondragón et al. (2024) demuestra que la integración de la IA con técnicas de monitoreo y analítica avanzada contribuye a la detección oportuna de fallas, relacionándose directamente con el objetivo general de anticipar fallas en el probador de suspensión, por lo cual es necesario tener presente una dependencia de volúmenes de datos de alta calidad.

Predicción robusta de la vida útil restante de los rodamientos mediante aprendizaje profundo

Este documento propone la predicción de la vida útil restante (RUL) de rodamientos en motores eléctricos, cuyo objetivo es desarrollar un modelo robusto que prediga la RUL bajo distintas condiciones de operación, evitando reentrenamiento y mejorando la aplicabilidad industrial, la metodología combina un análisis de señales de vibración en los dominios temporal y frecuencial para extraer características, un autocodificador de eliminación de ruido variacional apilado (SVDAE) para fusionar estas características y construir un indicador de salud, y una red neuronal bidireccional de memoria a corto plazo (BiLSTM), obteniendo como resultado un modelo preciso y robusto que puede distinguir fallas del aro interior aportando a un mantenimiento prescriptivo[25].

El artículo de Magadán et al. (2024) aporta a la presente investigación la exposición de como los modelos Deep Learning son eficaces para el mantenimiento predictivo, en particular en la predicción de la vida útil restante (RUL), además, contribuye una metodología que se puede adaptar para transformar datos de las pruebas de adherencia en indicadores precisos del estado del probador de suspensión.

Marco híbrido de aprendizaje profundo para la detección de fallas en tiempo real en motores de inducción de jaula de ardilla

Este estudio propone un marco híbrido basado en aprendizaje profundo para la detección de fallas en tiempo real en motores de inducción de jaula de ardilla, el marco se entrenó y probó en un conjunto de datos completo que comprende un millón de muestras, divididas equitativamente entre condiciones de motor en buen estado y con fallas, el cual, tiene como objetivo diseñar e implementar un sistema de detección de fallas en tiempo real para motores SCIM; mediante la creación de un entorno de simulación en MATLAB/Simulink que reprodujo las fallas y las condiciones normales de motores SCIM, se obtuvieron un millón de muestras y se entrenaron arquitecturas de Deep Learning como: CNN-GRU, CNN-LSTM, LSTM, BiLSTM, Stacked LSTM, GRU, CNN y ANN, comparando una clasificación fallas en diferentes condiciones como cortocircuitos, fallas mecánicas, sobrecargas y circuitos abiertos; esto obtuvo como resultado que entre todas las arquitecturas probadas, CNN-GRU y CNN-LSTM ofrecieron la mejor relación entre precisión y eficiencia. Estos resultados subrayan el potencial del DL híbrido para la monitorización de fallos, permitiendo la detección temprana de fallos eléctricos, mecánicos y de suministro con un mínimo conocimiento previo [26].

J. M. Jakaria et al. (2025) aporta mediante este artículo a la presente investigación la evidencia la eficacia de metodologías híbridas en el entorno del mantenimiento predictivo.

Predicción de fallos de máquinas basada en bosques aleatorios: una comparación de rendimiento

En la investigación realizada en este artículo se analizaron 12 modelos de ML para mantenimiento predictivo (Random Forest, SVM, KNN, Logistic Regression, Naive Bayes, AdaBoost, LDA, KDE+LR, XGBoost, MLP y Transformers) comparando su rendimiento mediante datos de fallas poco frecuentes y desbalance de clases. El objetivo de este estudio tuvo como propósito confrontar que algoritmo se comporta mejor bajo las condiciones de datos desbalanceados y ruido industrial, mediante una metodología con enfoque cuantitativo y un diseño experimental y comparativo se compaginaron datasets industriales reales, técnicas de preprocesamiento robustas, evaluación sistemática de múltiples algoritmos y herramientas de interpretabilidad, obteniendo como resultado que el modelo Random Forest tuvo el mejor desempeño con una precisión de 99.5% demostrando que está en capacidad de manejar patrones complejos y no lineales [2].

Yang y Wang, aportan a la presente investigación demostrando que el modelo Random Forest es eficiente al tener una precisión del 99.5% en la predicción de fallas en maquinaria industrial, lo que valida que es un algoritmo clave para el mantenimiento predictivo.

Un sistema inteligente de mantenimiento predictivo basado en bosque aleatorio para abordar los desafíos de las cintas transportadoras industriales

Este estudio presenta un sofisticado sistema inteligente de mantenimiento predictivo para cintas transportadoras industriales, basado en un modelo de aprendizaje automático de Random Forest, comparándolo con modelos consolidados como la regresión logística, las redes neuronales, los árboles de decisión y el aumento de gradiente; tuvo como objetivo diseñar un sistema inteligente de mantenimiento predictivo basado en bosques aleatorios para abordar los complejos desafíos asociados con los sistemas de cintas transportadoras industriales, se utilizó un diseño experimental evaluando dos condiciones: el desgaste de los dientes del sprocket y el aceite gearbox, recolectando datos mediante sensores de temperatura, vibración, corriente y voltaje, posteriormente se entrenaron seis modelos de ML (regresión logística, redes neuronales artificiales, arboles de decisión, Random Forest y gradient boosting); obteniendo como resultado que el modelo Random Forest con una precisión del 100% fue el algoritmo sobresaliente a los demás evaluados, confirmando su relevancia en el diagnóstico de fallas. Aunque los resultados fueron sobresalientes, los autores señalan la necesidad de validar el sistema con conjuntos de datos más amplios y diversos para garantizar su generalización y aplicabilidad en diferentes contextos industriales [27].

M. Wu et al. (2024) contribuye a esta investigación mediante un caso práctico y experimental una evidencia sobre la eficacia del modelo Random Forest en técnicas del mantenimiento predictivo.

Revisión por aplicación del mantenimiento predictivo basado en aprendizaje automático: tendencias, desafíos y direcciones futuras

Esta revisión sistemática de la literatura (SLR) ofrece un análisis exhaustivo, desde el punto de vista de la aplicación, del mantenimiento predictivo (PdM) basado en aprendizaje automático (ML) en diversos ámbitos industriales, su objetivo es identificar datasets, tendencias y algoritmos más utilizados en PdM en distintos sectores industriales, su metodología sigue las guías PRISMA 2020,

realizando una búsqueda en Scopus y obteniendo una población de 4501 artículos de los cuales se seleccionó una muestra de 60 estudios publicados entre 2020 y 2024, obteniendo como resultado que los modelos más usados son: Árboles de decisión, Random Forest, XGBoost, LSTM y CNN, además muestran que los modelos híbridos y el uso de datos de sensores mejoran la predicción de fallas [28].

C. Tsallis et al. (2025) contribuye a la actual investigación una evidencia científica que valida el uso de Random Forest como una herramienta sólida para predecir fallas en equipos críticos, como lo es el probador de suspensión; C Tsallis et al, también menciona la necesidad de implementar modelos ML en el entorno automotriz.

Mantenimiento predictivo en el sector automotriz: una revisión bibliográfica

Este artículo presenta una revisión sistemática de la literatura sobre enfoques de inferencia estadística, métodos estocásticos y técnicas de IA para el mantenimiento predictivo en el sector automotriz, Este trabajo tiene como objetivo proporcionar una breve visión general de las contribuciones recientes de la investigación sobre las técnicas utilizadas para el mantenimiento predictivo, especialmente en el campo de la automoción, mediante una metodología de revisión sistemática de literatura científica en donde fueron clasificados en categorías según su naturaleza, proporcionando un marco comparativo, obteniendo como resultados que el mantenimiento predictivo construye una táctica eficaz para reducir costos y mejorar la confiabilidad en sistemas automotrices [29].

F. Arena et al. (2025), aporta a la presente investigación una exposición de las diferencias entre mantenimiento reactivo, preventivo y predictivo, demostrando que el predictivo un mejor equilibrio entre prevención y costos de reparación, reportando beneficios operativos y económicos reveladores, como la reducción de costos de mantenimiento en un 25–30%, el incremento de la producción en un 20–25% y la disminución de fallas en un 70–75%.

Una revisión exhaustiva sobre el mantenimiento predictivo impulsado por inteligencia artificial en vehículos: tecnologías, desafíos y futuras direcciones de investigación

El artículo revisa diversas técnicas aplicadas al mantenimiento predictivo, destacando el papel de las técnicas en la IA y la importancia de la IA explicable para el análisis predictivo. Esta revisión examina

las aplicaciones de la IA en estrategias y diagnósticos de mantenimiento de vehículos para reducir costos, programas de mantenimiento, predicciones de vida útil restante y monitoreo efectivo del estado de salud. El artículo también identifica diversos desafíos en el mantenimiento predictivo relacionados con la calidad de los datos, la escalabilidad y la integración de la tecnología de IA. El propósito de esta investigación es simplificar y estudiar los principales avances de la IA adaptado al mantenimiento predictivo vehicular, revisando metodologías y tecnologías durante los últimos 5 años y detectando desafíos como la calidad de datos, versatilidad e integración con la IA en sistemas reales, aplicando un marco de revisión sistemática de la literatura (SLR), se consultaron bases de dato y Scopus y Web of Science de los cuales se seleccionaron 94 artículos realizando un análisis bibliométrico (tipos de documentos, tendencias de publicación, dominios de aplicación). Los resultados fueron de la SLR demuestran que Machine Learning y Deep Learning son las disciplinas más empleadas para el mantenimiento predictivo vehicular, en donde se destacan modelos como Randon Forest (RF), Long Short-Term Memory (LSTM) y Extreme Gradient Boosting (XGBoost) ya que tienen una gran capacidad para predecir fallas, finalmente, los hallazgos sugieren que las futuras líneas de investigación deben orientarse hacia el desarrollo de modelos híbridos que integren IA con sistemas IoT ya que pueden puede acelerar el desarrollo de soluciones de mantenimiento predictivo holísticas y escalables [7].

Y. Mahale et al. (2025) contribuye en el presente estudio relacionando una comparación integra sobre el uso de métodos de IA en el mantenimiento predictivo de vehículos, identificando como ML, DL, explainable AI y generative AI aportan a la reducción de fallas y mejora de confiabilidad en sistemas automotrices, destacando la importancia de proporcionar datos de alta calidad y modelos interpretables.

Algoritmos de aprendizaje automático para mantenimiento predictivo: un mapeado sistemático de la literatura

En este artículo se describen los principales algoritmos de Machine Learning usados en mantenimiento predictivo de acuerdo con el tipo de uso y su supervisión, además, se analizan los parámetros de entrada y las salidas de los mismos y por último se determina su nivel de madurez. Tiene como objetivo principal Proporcionar una visión global de los algoritmos de Machine Learning aplicados al mantenimiento predictivo. Durante la metodología se aplicó un Systematic Literature

Mapping (SLM) utilizando la base de datos Scopus se obtuvieron 6019 artículos, de los cuales 77 artículos fueron escogidos con la ayuda de un proceso manual con CADIMA y la aplicación de parámetros (campo: ingeniería y ciencias de la computación; tipo: artículos de revista y conferencia; idioma: inglés). Finalmente, los artículos fueron codificados y clasificados según tipo de supervisión, uso (regresión/clasificación), datos de entrada, parámetros requeridos, salidas generadas y nivel de madurez de los algoritmos. Los resultados muestran un crecimiento notorio en el uso de algoritmos de ML para mantenimiento predictivo entre 2021 y 2023, además se identificó que la mayoría de los estudios emplean algoritmos supervisados, principalmente en tareas de regresión para estimar la vida útil remanente (RUL), mientras que los de clasificación se orientan a determinar el estado de salud de los equipos. Los algoritmos más frecuentes incluyen *Random Forest*, redes neuronales, SVM y regresión lineal, mientras que los enfoques no supervisados como *K-means* y PCA son menos comunes [30].

El artículo de J. Paredes y C. Romero. (2025) ofrece a la presente investigación un mapeo sistemático de los modelos de ML aplicados al mantenimiento predictivo, la clasificación de algoritmos según su uso y tipo de supervisión proporciona determinar métodos para anticipar fallas en el probador de suspensión.

Un estudio exhaustivo sobre la predicción de la necesidad de mantenimiento de vehículos mediante aprendizaje automático

Este estudio propone una alternativa más precisa y eficiente: el uso de algoritmos de aprendizaje automático capaces de analizar múltiples variables de los vehículos —como edad, tipo de combustible, historial de accidentes, tamaño del motor, kilometraje y registros de servicio— para determinar de manera anticipada cuándo es realmente necesario realizar un mantenimiento, tuvo como objetivo Optimizar la predicción de necesidades de mantenimiento mediante algoritmos de ML. Mediante un datase de 50.000 vehículos, con 19 variables independientes (historial de accidentes, tipo de combustible, edad, tamaño del motor, kilometraje, entre otras) y una dependiente (Need_Maintenance), se entrenaron múltiples algoritmos de clasificación comparando la efectividad de los modelos en escenarios de datos balanceados y desbalanceados. Los resultados, tanto desequilibrados como equilibrados, definen la solidez de los modelos, especialmente del aprendizaje profundo, que muestra consistentemente mejores resultados. Estos resultados pueden ayudar a

optimizar las estrategias de mantenimiento, tanto para propietarios de vehículos como para gestores de flotas, con el fin de minimizar el tiempo de inactividad y optimizar el mantenimiento [9].

G. Mahiyudin et al. (2025) contribuye con la presente investigación al proporcionar metodologías sobre el uso de ML en mantenimiento predictivo, demostrando que estos algoritmos son precisos, además, enfrentan retos como el desbalance de datos y la variabilidad de condiciones de operación. Este aporte resulta relevante para la presente investigación, dado que, valida la eficacia de los algoritmos de ML en la anticipación de fallas críticas, reforzando la pertinencia del uso de Random Forest en el probador de suspensión.

Mantenimiento predictivo basado en aprendizaje profundo para aplicaciones de IoT industrial

El artículo aborda el uso de Deep Learning aplicado al mantenimiento predictivo en entornos industriales conectados al Internet de las Cosas (IoT). Tiene un objetivo principal el cual es Desarrollar un marco de mantenimiento predictivo basado en Deep Learning que pueda detectar patrones de degradación en maquinaria antes de que ocurran fallas críticas. La metodología incluyó procesamiento de datos para limpiar y estructurar la información de sensores, entrenamiento supervisado de modelos para identificar patrones de desgaste y anomalías y una validación experimental en escenarios simulados o reales de IoT, comparando el rendimiento frente a métodos tradicionales de mantenimiento; obteniendo resultados que muestran que los modelos de Deep Learning superan los enfoques tradicionales de mantenimiento predictivo en precisión y capacidad de anticipación, una reducción significativa de fallas inesperadas y una mejora en la planificación de mantenimientos [31].

J. Emilyn et al. (2024) mediante el artículo mencionado aporta a la presente investigación evidencia científica sobre la eficacia del Deep Learning aplicado al mantenimiento predictivo en entornos industriales conectados al IoT, asimismo, el artículo contribuye al diseño de un modelo de Machine Learning (Random Forest) para anticipar fallas en equipos críticos, mostrando que la aplicación de algoritmos de IA es una estrategia reconocida y efectiva en la industria 4.0.

Un estudio sobre el mantenimiento predictivo: sistemas, propósitos y enfoques

Es un artículo de revisión que analiza la evolución de las técnicas de mantenimiento: reactivo (RM), preventivo (PM) y predictivo (PdM), centrándose en el PdM como paradigma clave en la Industria 4.0, destacando su capacidad para reducir costos, minimizar tiempos de inactividad y aumentar la confiabilidad de los sistemas industriales. El objetivo de estos es el de analizar los avances más recientes en PdM, específicamente desde 2015 hasta 2020, lo cual permite identificar arquitecturas, objetivos y los métodos empleados. En la estrategia de trabajo tendremos una inspección sistemática de literatura clasificando por la información en arquitecturas de los sistemas, los objetivos de optimización (minimización de costos, maximización de disponibilidad/confiabilidad, optimización multiobjetivo) y métodos de optimización (machine Learning tradicional: SVM, ANN, k-NN; y Deep Learning: CNN, RNN, Autoencoders, DBN, GAN). Obteniendo que los métodos que se basan en ML y DL han mejorado de gran manera su capacidad de diagnóstico y pronóstico de fallas, incluyendo la predicción de la vida útil remanente (RUL). Sin embargo, es importante decir que persisten desafíos grandes, como la fusión de datos multisensor, la mejora en la precisión en las predicciones y la optimización de la programación del mantenimiento [32].

T. Zhu et al. (2019) nos ofrece un marco conceptual lo suficientemente amplio sobre arquitecturas, objetivos y métodos de mantenimiento PdM, confirmándonos que se debe de tener la pertinencia, por así decirlo, de entrenar un modelo Random Forest en un probador de suspensión.

Integración de inteligencia artificial para el diagnóstico predictivo de fallas mecánicas en vehículos de combustión interna: Enfoque integral

Este artículo nos aborda con una temática, como su nombre indica, sobre la aplicación de modelos de Inteligencia Artificial (IA), para un diagnóstico predictivo de fallas mecánicas en vehículos de combustión interna. Destaca su pertinencia en los contextos donde la confiabilidad operativa es crítica. El objetivo principal del artículo es analizar y proponer un enfoque integral para la implementación de sistemas inteligentes que nos permitan anticipar las fallas mecánicas con el procesamiento de datos operacionales y sensoriales en tiempo real. La investigación presenta un enfoque cuantitativo-experimental, el cual es apoyado en datos reales obtenidos de distintas flotas vehiculares, realizando una correcta selección de variables y un procesamiento con RF y PCA, luego se aplicaron los modelos supervisados (árboles de decisión, redes neuronales y SVM) evaluando su

desempeño mediante la precisión, la sensibilidad, la especificidad y el F1-score. Los resultados obtenidos en este estudio nos confirman que la aplicación de modelos de Inteligencia Artificial, particularmente redes neuronales, nos ofrecen un rendimiento significativamente superior en la predicción de fallas mecánicas en vehículos de combustión interna [33].

Cajías-Arías et al, (2025) demuestra mediante este artículo, la eficacia de integrar modelos desarrollados mediante la IA en la anticipación de fallas mecánicas en vehículos de combustión interna; Su metodología, que incluye preprocesamiento de datos, selección de variables con Random Forest y validación de modelos supervisados, respalda la aplicación de Machine Learning en equipos de diagnóstico automotor.

5.4.2 Antecedentes nacionales.

Método basado en visión artificial para la identificación del eje aplicado al proceso de inspección de frenos de vehículos livianos en centros de diagnóstico automotor.

Este trabajo de grado explora la aplicación de técnicas de visión artificial basadas en aprendizaje profundo para supervisar automáticamente el proceso de revisión técnico mecánica y emisiones contaminantes en los CDA de Colombia, específicamente durante la prueba de frenos. El principal propósito es desarrollar un método que proporcione identificar el eje y el vehículo al que pertenece durante la inspección, con el fin de prevenir la suplantación de ejes o vehículos. La metodología empleada consta de tres fases: preparación, implementación y documentación. En la fase de preparación se recolectaron videos de inspecciones y se construyeron conjuntos de datos etiquetados manualmente para el entrenamiento y evaluación de los modelos. Durante la fase de implementación, se entrenaron y compararon tres modelos de detección de objetos (DETR-ResNet50, DETR-ResNet101 y YOLOv9) para la detección de ejes, seleccionando el modelo YOLOv9 por su mejor rendimiento. Además, se utilizó un algoritmo de reconocimiento de placas para identificar los vehículos y una red siamesa para los casos en que la placa no era visible. En la fase de documentación, se detalló el desarrollo del proyecto, el experimento realizado y los resultados obtenidos. Los resultados demostraron la efectividad del modelo YOLOv9 para la detección de ejes y del algoritmo de identificación de vehículos, con un alto porcentaje de aciertos y pocos falsos positivos [34].

A Mongua. (2024) demuestra mediante este trabajo la viabilidad de aplicar IA en la realización de las pruebas de frenos de los CDA, sus resultados evidencian que la IA mejora la seguridad vial y el cumplimiento normativo. Este antecedente respalda la presente investigación al mostrar que los CDA son escenarios adecuados para implementar modelos de ML.

Detección y diagnóstico de fallas en motores mediante el análisis de vibraciones aplicando técnicas de inteligencia artificial.

Este artículo se enfoca en el diagnóstico de fallas en motores eléctricos mediante el análisis de vibraciones, aplicando técnicas de inteligencia artificial para procesar señales y reconocer patrones característicos de distintos tipos de fallas. Tuvo como objetivo desarrollar un sistema de detección y diagnóstico que identifique fallas en motores eléctricos por medio de análisis de vibraciones. La metodología implementada incluyó la recolección de datos experimentales bajo diversas condiciones de operación, el procesamiento de señales mediante técnicas de filtrado y extracción de características, y el entrenamiento de algoritmos de clasificación para el reconocimiento automático de fallas. La validación experimental evidenció que el sistema propuesto presenta alta precisión en la detección de anomalías, superando los métodos de diagnóstico tradicionales y demostrando la eficacia del análisis de vibraciones como herramienta para el mantenimiento predictivo. Los resultados obtenidos sugieren que la integración de IA en sistemas de monitoreo industrial contribuye a optimizar procesos, reducir costos y minimizar tiempos de inactividad, además de mejorar la seguridad operativa en entornos industriales [35].

El artículo de J. E. Araujo Vargas et al, (2023) contribuye a la presente investigación con el uso de técnicas de IA Y ML en mantenimiento predictivo, demostrando que es efectivo en detección temprana de fallas en sistemas dinámicos, este enfoque es aplicable a un probador de suspensión ya que los resultados de las pruebas de adherencia pueden ser tratados como insumos para modelos predictivos, convirtiéndolas en indicadores de salud del equipo.

Integración de tecnologías móviles y diagnóstico vehicular: una aplicación para la industria 4.0

Esta investigación presenta el desarrollo de una aplicación móvil para sistemas operativos Android diseñada para realizar diagnóstico y análisis de actividad vehicular sobre un sistema de automoción.

Tuvo como objetivo diseñar y validar una aplicación móvil que se integre con sistemas de diagnóstico vehicular. El presente trabajo se desarrolló bajo un enfoque experimental y de ingeniería aplicada, orientado al diseño, implementación y ratificación de una solución tecnológica para el diagnóstico vehicular apoyado en plataformas móviles, lo cual permitió conectarse eficazmente a sistemas de diagnóstico vehicular y visualizar en tiempo real fallas, parámetros y alertas de mantenimiento, reduciendo significativamente los tiempos de diagnóstico al ofrecer información inmediata en el dispositivo, además el sistema reveló que es viable y escalable a los entornos industriales [36].

Mediante este artículo, G. F. Andrade et al. (2024), aporta al presente trabajo la demostración de como la integración de tecnologías móviles y plataformas digitales en la industria 4.0 fortalece un diagnóstico vehicular mediante la captura y análisis de datos en tiempo real, también, evidencia como la digitalización mejora la trazabilidad y confiabilidad de los diagnósticos mediante IA.

Variación de las emisiones contaminantes de los vehículos debido a cambios en la geometría vial en carreteras de Colombia

La presente investigación se orienta al análisis del comportamiento dinámico de sistemas mecánicos y vehiculares, con énfasis en el diagnóstico de condiciones de operación y la detección temprana de fallas. El propósito de esta investigación fue estimar la variación de las emisiones contaminantes de vehículos debida a cambios geométricos de las vías, empleando una metodología experimental orientada a la caracterización y análisis del comportamiento dinámico del sistema en estudio, que incluye el diseño de pruebas en laboratorio y/o campo, apoyadas en sensores y equipos de medición para registrar vibraciones, desplazamientos y respuestas dinámicas; obteniendo resultados que demuestran que el método propuesto presenta una alta eficacia para identificar patrones de comportamiento y detectar posibles fallas en el sistema analizado. Además, se encontró que existe una correlación lo suficientemente significativa entre las variables medidas y el estado operativo del equipo, lo que a su vez respalda la validez del enfoque metodológico adoptado [37].

La tesis de grado de C. Díaz. valida el propio mantenimiento predictivo, esto al demostrar que supera las limitaciones de los enfoques preventivo y correctivo, promoviendo un uso de los datos en tiempo real en lugar de inspecciones periódicas. También nos da a entender que la eficacia de las técnicas ML, como Random Forest y redes neuronales, en la predicción de fallas en entornos automotrices.

5.4.3 Síntesis comparativa y brecha.

A nivel internacional, en la literatura reciente tenemos un crecimiento grande en el mantenimiento predictivo, apoyado en las diferentes técnicas de aprendizaje automático, particularmente en entornos de Industria 4.0. Revisiones sistemáticas destacan la proliferación de sensores e Industrial IoT y el predominio de algoritmos como RF, máquinas de soporte vectorial (SVM) y el aprendizaje profundo en aplicaciones de PdM.

Estos estudios reportan beneficios operativos significativos como la reducción de tiempos de inactividad, optimización en la planificación del mantenimiento y mejora en la confiabilidad de equipos críticos. En particular, se ha demostrado que buenos modelos de ML bien entrenados pueden anticipar fallas con alta precisión, estamos hablando de una precisión superior al 99% en escenarios industriale, incluso bajo condiciones de datos ruidosos o desbalanceados.

Adicionalmente, técnicas de aprendizaje profundo han logrado predecir con efectividad la vida útil remanente de componentes (RUL) en maquinaria, lo cuál nos ofrece soluciones robustas sin la necesidad de tener un reentrenamiento constante. Pero los mismos trabajos nos dicen que hay desafíos importantes como la dependencia en grandes volúmenes de datos de alta calidad para entrenar los modelos y garantizar su generalización, así como la escasez de datos etiquetados fiables en muchos dominios.

Se reconoce que datos incompletos o erróneos pueden degradar significativamente la efectividad de un modelo predictivo. De todas formas persisten retos en la integración de sistemas PdM a escala, como la fusión de datos multisensor, la necesidad de algoritmos explicables (*XAI*) y la escalabilidad de las soluciones en entornos reales. Aún con las diferentes limitaciones, las investigaciones internacionales tienen una tendencia consolidada hacia el PdM, con diferentes casos de éxito en los sectores industriales (motores eléctricos, rodamientos, cintas transportadoras, etc.) y también en el sector automotriz. Por ejemplo, estudios comparativos han confirmado que el Random Forest suele superar a otros algoritmos en la detección de fallas complejas, y hay reportes de reducciones del 25-30% en costos de mantenimiento y hasta un 70-75% en la disminución de las fallas al adoptar estrategias predictivas en automoción. Asimismo, las revisiones específicas del ámbito automotor nos resaltan que las técnicas de ML (incluyendo RF, redes neuronales LSTM y XGBoost) son las más empleadas para pronóstico de fallas vehiculares, mostrando un impacto positivo en la reducción de averías y mejoras de confiabilidad. Para resumir, a nivel internacional el PdM apoyado en IA se

encuentra bien posicionado, aunque cabe decir que con el reconocimiento explícito de la necesidad de mejorar la calidad de datos y la generalización de modelos para su aplicación más amplia.

En Colombia, los antecedentes revelan un escenario emergente en la aplicación de mantenimiento predictivo y técnicas de IA, tanto en el sector industrial como en el automotriz, pero en una escala más limitada. Varios trabajos de grado y estudios locales han explorado la utilización de inteligencia artificial para mejorar procesos de mantenimiento e inspección. Por ejemplo, Mongua (2024) desarrolló un sistema de visión artificial basado en aprendizaje profundo para supervisar la prueba de frenos en Centros de Diagnóstico Automotor [36], logrando identificar automáticamente ejes y vehículos durante la inspección para evitar suplantaciones. Empleando modelos de detección de objetos (YOLOv9, entre otros), este proyecto alcanzó un alto porcentaje de acierto con pocos falsos positivos en la detección de ejes, demostrando la viabilidad de aplicar IA en entornos de inspección vehicular. Los resultados indicaron que tales enfoques aumentan la seguridad vial y garantizan el cumplimiento normativo en las revisiones técnico-mecánicas, posicionando a los CDA colombianos como escenarios propicios para la implementación de modelos de ML. En el ámbito industrial local, Araujo Vargas et al. (2023) aplicaron técnicas de IA para el análisis de vibraciones en motores eléctricos [35], con el fin de diagnosticar fallas incipientes. Su sistema de clasificación automático de anomalías obtuvo una precisión elevada, superando los métodos tradicionales de diagnóstico, lo que subraya cómo el análisis inteligente de señales puede optimizar los procesos de mantenimiento, reducir costos y minimizar tiempos de inactividad en entornos operativos. Este estudio no solo nos confirmó la gran eficacia de un mantenimiento predictivo con IA en la maquinaria, sino que también nos sugirió la posibilidad de llevar dicho enfoque a otros sistemas dinámicos, que de hecho, los autores nos proponen que los resultados de pruebas técnicas como la adherencia de suspensión, pueden ser reemplazadas como indicadores de la salud de un equipo, esto facilitando la detección temprana de fallas en dispositivos como el probador de suspensión. Por otro lado, Andrade et al. (2024) nos introdujeron la integración de plataformas móviles para el diagnóstico vehicular, desarrollando una aplicación Android capaz de conectarse con los diferentes sistemas de diagnóstico del vehículo para así poder monitorear en tiempo real parámetros operativos y de esta manera poder alertar sobre fallas o necesidades de mantenimiento [36]. Esta solución demostró ser bastante viable y escalable, por lo que reduce significativamente los tiempos de diagnóstico al proveer información inmediata, a la vez que mejora la trazabilidad y la confiabilidad del proceso mediante la digitalización de los datos. Asimismo, investigaciones académicas en el contexto colombiano han abordado el

comportamiento dinámico de vehículos bajo diferentes condiciones, encontrando que la instrumentación con sensores y el análisis experimental permiten identificar patrones de comportamiento anómalos y posibles fallas con alta eficacia. Un estudio sobre vibraciones vehiculares debido a la geometría vial (Díaz, 2021) halló correlaciones significativas entre las variables medidas y el estado operativo del vehículo/equipo, validando el enfoque metodológico basado en monitoreo continuo y análisis de datos para anticipar desperfectos [37]. En conjunto, estos antecedentes nacionales apuntan a que el mantenimiento predictivo comienza a ser reconocido por su ventaja frente a los enfoques preventivos o correctivos tradicionales. De hecho, trabajos locales recientes sugieren que el uso de datos en tiempo real —en lugar de depender únicamente de inspecciones periódicas— permite detectar condiciones de falla incipiente con mayor oportunidad, y demuestran la eficacia de algoritmos de ML como RF y redes neuronales en la predicción de fallas en contextos automotrices reales. No obstante, muchos de estos esfuerzos en Colombia se han limitado a entornos piloto o investigaciones académicas puntuales, evidenciando alcances más reducidos en cuanto a volumen de datos y despliegue práctico que sus contrapartes internacionales.

Comparativa y brechas, Al contrastar ambos contextos, se observa una marcada brecha entre los desarrollos internacionales y nacionales en PdM. Internacionalmente existe una amplia adopción de estrategias de mantenimiento predictivo respaldadas por grandes conjuntos de datos y sistemas IoT, con modelos avanzados (incluyendo Deep Learning e híbridos) implementados en diversidad de industrias. En Colombia, aunque hay avances importantes, estos se circunscriben generalmente a iniciativas aisladas o de menor escala, sin una incorporación masiva en la industria. Por ejemplo, mientras a nivel global se reportan estudios con decenas de miles de muestras o incluso simulaciones con millones de datos para entrenar modelos robustos, en el contexto colombiano los proyectos suelen trabajar con conjuntos de datos más reducidos (en ocasiones recolectados manualmente) y con alcance experimental. Asimismo, retos bien documentados internacionalmente —como la falta de datos etiquetados, el desbalance de clases y la necesidad de generalizar modelos a múltiples condiciones operativas— se ven acentuados en el medio local, donde la disponibilidad de datos históricos de mantenimiento es limitada y la infraestructura para su captura en tiempo real está apenas en desarrollo. Otro contraste radica en la aplicación en entornos reales: fuera del país, numerosas soluciones PdM han trascendido el ámbito investigativo para integrarse en procesos productivos y de gestión de activos, mientras que en Colombia la mayoría de los ejemplos corresponden a prototipos o pruebas de concepto en fase inicial. Adicionalmente, existe una brecha en términos normativos y

de estandarización: globalmente se avanza hacia la normalización de prácticas de mantenimiento 4.0 y el uso de IA en confiabilidad, mientras que en el contexto colombiano dichas prácticas aún no están incorporadas de forma explícita en la normativa técnica de mantenimiento vehicular o industrial. Esta disparidad general sugiere que, si bien se reconoce el potencial del mantenimiento predictivo a nivel nacional, queda un vacío por cerrar para alcanzar el nivel de madurez observado internacionalmente. En particular, el sector automotriz colombiano no había reportado, hasta donde abarcan los antecedentes, implementaciones robustas de modelos predictivos sobre los equipos de inspección vehicular mismos, lo que indica una oportunidad de innovación local alineada con las tendencias globales.

Frente a las brechas identificadas, la tesis actual propone un enfoque innovador que contribuye a reducir la distancia entre el contexto global y el colombiano. Su principal novedad radica en la aplicación directa de técnicas de ML (Random Forest) sobre datos reales de pruebas de adherencia obtenidos en un CDA de Colombia, algo no documentado previamente en los antecedentes nacionales. A diferencia de estudios internacionales que a veces utilizan datos simulados o altamente controlados, aquí se emplea un conjunto de 2.360 registros reales de pruebas de adherencia, recolectados en operación rutinaria, para entrenar y validar el modelo predictivo. Esto constituye un aporte técnico significativo: se demuestra que las mediciones ordinarias de adherencia –diseñadas originalmente para evaluar las condiciones de suspensión de vehículos– pueden reorientarse como indicadores fiables del estado de salud del probador de suspensión en sí.

6. METODOLOGÍA

6.1 ENFOQUE Y TIPO DE INVESTIGACIÓN

La estructura del marco metodológico para la investigación de desarrollar un modelo predictivo basado en técnicas de aprendizaje automático para anticipar fallas en el equipo probador de suspensión de la línea de vehículos livianos del Centro de Diagnóstico Automotor Certi Express Pereira S.A.S., contribuyendo a la implementación de estrategias de mantenimiento predictivo basado en un enfoque cuantitativo y explicativo, en el que se empleará el modelo Random Forest. En este orden de ideas según R. Hernández Sampieri, el enfoque cuantitativo utiliza la recolección de datos para probar hipótesis con base en la medición numérica y el análisis estadístico, con el fin establecer pautas de comportamiento y probar teorías [38]. El cual se centra en realizar un análisis de la base de datos de las pruebas de adherencia realizadas con el equipo probador de suspensión de la línea de livianos de revisión técnico mecánica y de emisiones contaminantes de la empresa en la cual se centra la investigación, es por ello por lo que es de suma importancia el análisis de datos para la creación de un modelo Random Forest que permitirá la detección de fallas de manera oportuna.

El alcance de la presente investigación es explicativo, conforme a R. Hernández Sampieri, los estudios explicativos van más allá de la descripción de conceptos o fenómenos o del establecimiento de relaciones entre conceptos; es decir, están dirigidos a responder por las causas de los eventos y fenómenos físicos o sociales. Como su nombre lo indica, su interés se centra en explicar por qué ocurre un fenómeno y en qué condiciones se manifiesta o por qué se relacionan dos o más variables [38], por lo que, al ser el modelo Random Forest un prototipo confiable y preciso permitirá explicar la detección oportuna de la falla en la máquina y en los vehículos inspeccionados realizará un análisis de si este requiere mantenimiento en el sistema de suspensión en los próximos 3 meses.

6.2 MUESTRA

La investigación se generó utilizando una base de datos de las pruebas realizadas con el equipo de suspensión de la línea de inspección de livianos de Certi Express Pereira s.a.s, realizadas a 2360 vehículos, de cada automotor se consultaron los siguientes datos: clase, marca, modelo, línea, adherencia eje1 derecho, adherencia eje 1 izquierdo, adherencia eje2 derecho y adherencia eje2 izquierdo, es decir, por cada vehículo son 8 datos lo que generaría un total de 18.880 datos, para determinar que se comprende por muestra, R. Hernández Sampieri, aporta, “*la muestra es un subgrupo de la población de interés sobre el cual se recolectarán datos, y que tiene que definirse y delimitarse de antemano con precisión, además de que debe ser representativo de la población*” [38].

TABLA 1. DATOS DE PRUEBAS DE ADHERENCIA

Clase	Marca	Modelo	Línea	Adherencia eje1 de	Adherencia eje1 iz	Adherencia eje2 de	Adherencia eje2 iz
-------	-------	--------	-------	-----------------------	-----------------------	-----------------------	-----------------------

Fuente. Elaboración propia.

Con estos datos consultados se realizó un análisis lo cual condujo a la implementación de un modelo Random Forest, el cual ayuda al mantenimiento predictivo; así mismo se consolidó información de los puntos de falla del equipo con el cual se lleva a cabo la investigación, ver anexo 1.

6.3 TÉCNICAS E INSTRUMENTOS

Con el fin de cumplir el objetivo de esta investigación se trazan unas técnicas e instrumentos que permiten dar respuesta a cada objetivo específico desde el enfoque cuantitativo con diseño explicativo, en este sentido se entenderá la técnica como el procedimiento validado y orientado con el fin de obtener información útil para la investigación, se analiza y expone. Por consiguiente, en esta

investigación se utilizarán técnica-instrumento como, base de datos con recolección de datos y la implementación de la metodología CRISP-DM (Cross Industry Standard Process for Data Mining).

6.4 DISEÑO GENERAL DEL PROCESO (CRISP-DM)

El modelo CRISP-DM es una metodología que propone un marco de trabajo estandarizado, es decir, que el procedimiento permite la organización y modelado de datos mediante 6 fases propuestas, facilita la estructuración y la comprensión de proyectos de minería de datos e inteligencia artificial, lo cual se hizo en su aplicación y cuyo resultado generó el informe que se expondrá en el apartado de resultados.

Teniendo en cuenta lo anterior, a continuación, se expone cada una de las fases que el modelo propone, es decir:

6.4.1 Comprensión del Negocio. Consiste en la definición del problema lo cual permite trazar el objetivos y criterios del proyecto desde una organización adecuada proyectándose como exitosa.

6.4.2 Comprensión de los Datos. Comprende la exploración de datos, es decir, que conlleva a la recolección de la información con el fin de identificar patrones o problemas que de una u otra forma alteren el proceso a llevarse a cabo.

6.4.3 Preparación de los Datos. Integra la limpieza, selección y transformación de la información, lo cual permite que estén listos para ser introducidos en los algoritmos del modelado.

6.4.4 Modelado. Se entrena el modelo con los algoritmos seleccionados ubicando los datos preparados.

6.4.5 Evaluación. Abarca la evaluación del resultado por medio de la apreciación rendimiento técnico del modelo contrastando los objetivos del proceso y los criterios planteados en la primera fase.

6.4.6 Despliegue. Permite incorporar al entorno operativo al ser validado la implementación del modelo instaurándose un plan de seguimiento y mantenimiento.

6.5 FASES DEL PROCESO CRISP-DM

6.5.1 Comprensión del negocio. En este apartado se describe el análisis estadístico y el desarrollo de un modelo predictivo aplicado a las mediciones del equipo probador de suspensión del Centro de Diagnóstico Automotor (CDA) Certi Express Pereira s.a.s. El objetivo de esta investigación fue identificar guías de comportamiento en las mediciones de adherencia, detectar anomalías operativas y construir un modelo de alerta temprana que permita anticipar condiciones de mantenimiento del equipo.

El análisis se fundamenta en un conjunto de 2.360 registros reales correspondientes a mediciones de adherencia en los ejes 1 y 2, diferenciando lado derecho e izquierdo. Se implementaron técnicas de exploración de datos, depuración, normalización, identificación de valores atípicos y modelado mediante Random Forest, siguiendo principios de análisis cuantitativo.

6.5.2 Carga y presentación de librerías. Antes de iniciar el procesamiento, se automatizó la creación de un entorno virtual y la instalación de las librerías requeridas (pandas, numpy, matplotlib, seaborn, scikit-learn, xlswriter, openpyxl y otras). Esto garantiza reproducibilidad y evita conflictos de versiones.

Luego se importaron las librerías necesarias para el manejo de datos, visualización y modelado. Asimismo, se configuró un estilo visual uniforme para los gráficos con seaborn y matplotlib, priorizando fondos claros y cuadrículas visibles para facilitar la interpretación.

6.5.3 Organización de directorios. Con el fin de estructurar apropiadamente los insumos y resultados del análisis, se definieron rutas separadas para:

- Datos: archivo fuente Datos_suspensión.xlsx
- Figs: figuras generadas
- Reportes: tablas, resúmenes y archivos finales
- Modelos: modelos entrenados

La organización por carpetas permite trazabilidad, evita sobreescrituras involuntarias y deja lista la estructura para una posible implementación continua; es decir, que esto permite la optimización de la información recolectada y el tiempo que se utiliza para el procedimiento, ello facilita la precisión al momento de la clasificación de los datos.

6.5.4 Carga de datos y verificación inicial. Los datos se cargaron en un DataFrame de pandas, verificando estructura, número de filas y columnas, y formato general del contenido.

El dataset contiene:

- 2.360 vehículos inspeccionados con 8 datos de adherencia por cada uno, para un total de 9440 registros.
- 8 columnas originales, incluyendo clase del vehículo, marca, modelo, línea y cuatro mediciones de adherencia.

Se visualizó la parte inicial del dataset para confirmar que la carga fuera correcta y evaluar posibles irregularidades en el formato.

6.5.5 Normalización de nombres y columnas. Se estandarizaron los nombres de columnas mediante la eliminación de espacios, la conversión a minúsculas, el uso de guiones bajos y el renombramiento uniforme de variables de adherencia, ello para alcanzar la calidad de los datos con el fin de que la modelación se efectuará eficazmente para cumplir con el propósito de esta etapa es necesario es evitar ambigüedades en la manipulación del DataFrame y facilitar su uso en procesos posteriores.

6.5.6 Estadísticas descriptivas. Se generaron estadísticas clave (media, desviación estándar, percentiles, valores faltantes) para las cuatro mediciones de adherencia:

VARIABLES ANALIZADAS:

- Adherencia_eje1_de. Se refiere al valor dado en porcentaje de la prueba de adherencia realizada en la rueda derecha del eje delantero del vehículo inspeccionado, la cual se desarrolla de acuerdo al numeral 6.8.2 de la NTC 5375:2012.
- Adherencia_eje1_iz. Se refiere al valor dado en porcentaje de la prueba de adherencia realizada en la rueda izquierda del eje delantero del vehículo inspeccionado, la cual se desarrolla de acuerdo al numeral 6.8.2 de la NTC 5375:2012.
- Adherencia_eje2_de. Se refiere al valor dado en porcentaje de la prueba de adherencia realizada en la rueda derecha del eje trasero del vehículo inspeccionado, la cual se desarrolla de acuerdo al numeral 6.8.2 de la NTC 5375:2012.
- Adherencia_eje2_iz. Se refiere al valor dado en porcentaje de la prueba de adherencia realizada en la rueda izquierda del eje trasero del vehículo inspeccionado, la cual se desarrolla de acuerdo al numeral 6.8.2 de la NTC 5375:2012.

Las variables analizadas se toman como muestra ya que son el resultado de las pruebas de adherencia realizadas con el probador de suspensión al momento de realizar la revisión técnico mecánica y de emisiones contaminantes; estos resultados se consolidaron en la tabla 2

TABLA 2 . COLUMNAS DE ADHERENCIA

Unnamed: 0	count	mean	std	min	25%	50%	75%	max	missing
adherencia_eje1_de	2343.0	64.259.198	13.916.904	40.0	52.6	63.3	75.7	95.1	17
adherencia_eje1_iz	2343.0	79.766.539	8.450.730	40.5	74.2	79.9	85.9	100.0	17
adherencia_eje2_de	2343.0	62.910.499	14.023.796	40.0	51.2	61.1	74.9	100.0	17

adherencia_eje2_iz	2343.0	78.738.242	8.972.640	40.4	72.7	78.4	85.4	100.0	17
--------------------	--------	------------	-----------	------	------	------	------	-------	----

Fuente. Elaboración propia.

La anterior tabla se compone de la siguiente manera:

- Unnamed: índice o fila de identificación.
- Count: número de datos utilizadas en el análisis estadístico.
- Mean: indica la media de los valores de cada variable correspondiente.
- Std: es una medida de dispersión alrededor de la media, una desviación típica de la variable.
- Min: Es el valor mínimo de los registros de cada variable.
- 25%: percentil 25 (Q1). El valor por debajo del cual cae el 25% de los datos.
- 50%: percentil 50 (mediana). El valor que divide la mitad de los datos por encima y por debajo.
- 75%: percentil 75 (Q3). El valor por debajo del cual cae el 75% de los datos.
- Max: valor máximo observado en la muestra para esa variable.
- Missing: número de valores nulos o faltantes en la variable.

6.5.7 Histogramas y boxplot. Se generan histogramas por cada columna de medición y se incluye un bloxplot comparativo.

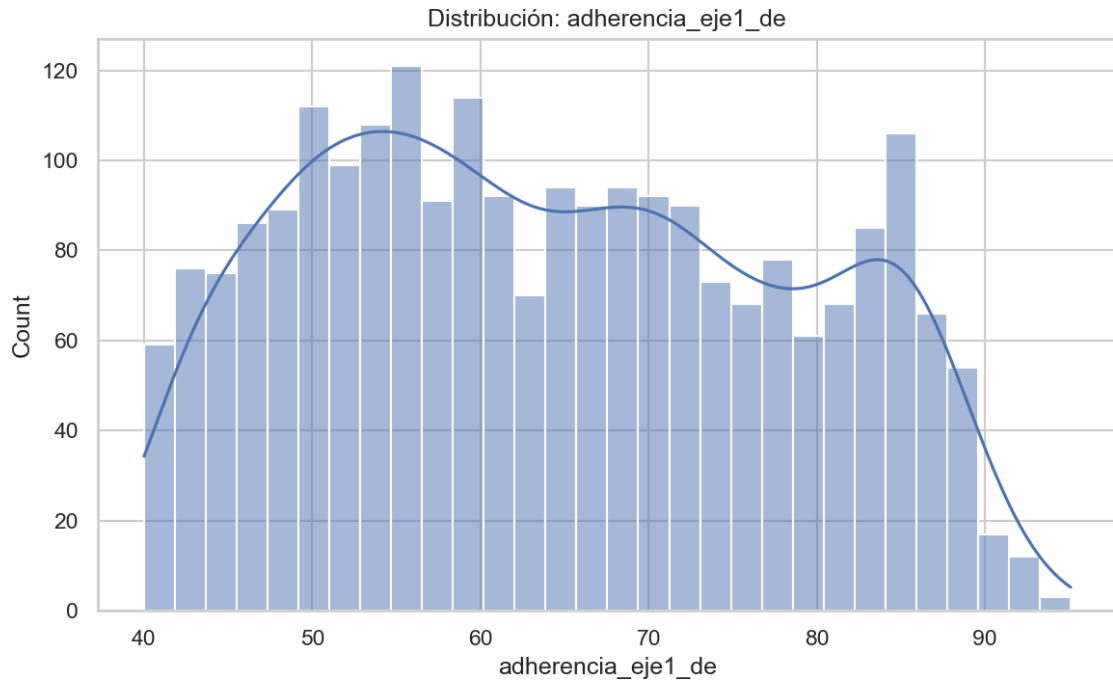


FIG. 1 HISTOGRAMA ADHERENCIA EJE 1 DERECHA

Fuente: Elaboración propia.

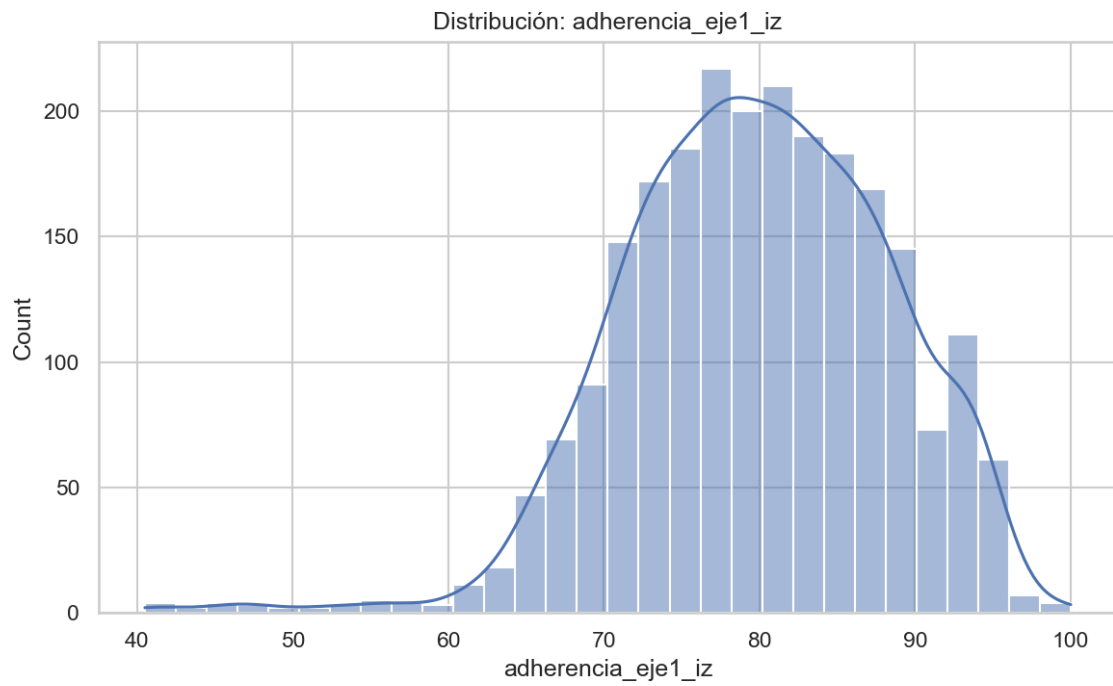


FIG. 2 HISTOGRAMA ADHERENCIA EJE 1 IZQUIERDA.

Fuente: Elaboración propia.

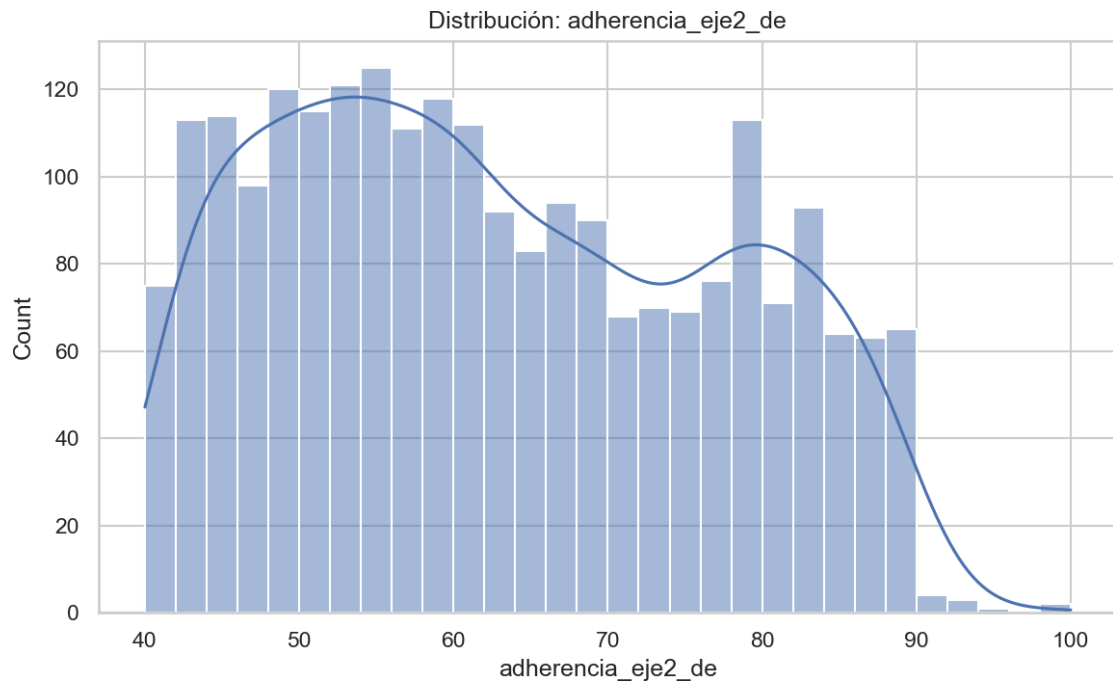


FIG. 3 HISTOGRAMA ADHERENCIA EJE 2 DERECHA.

Fuente: Elaboración propia.

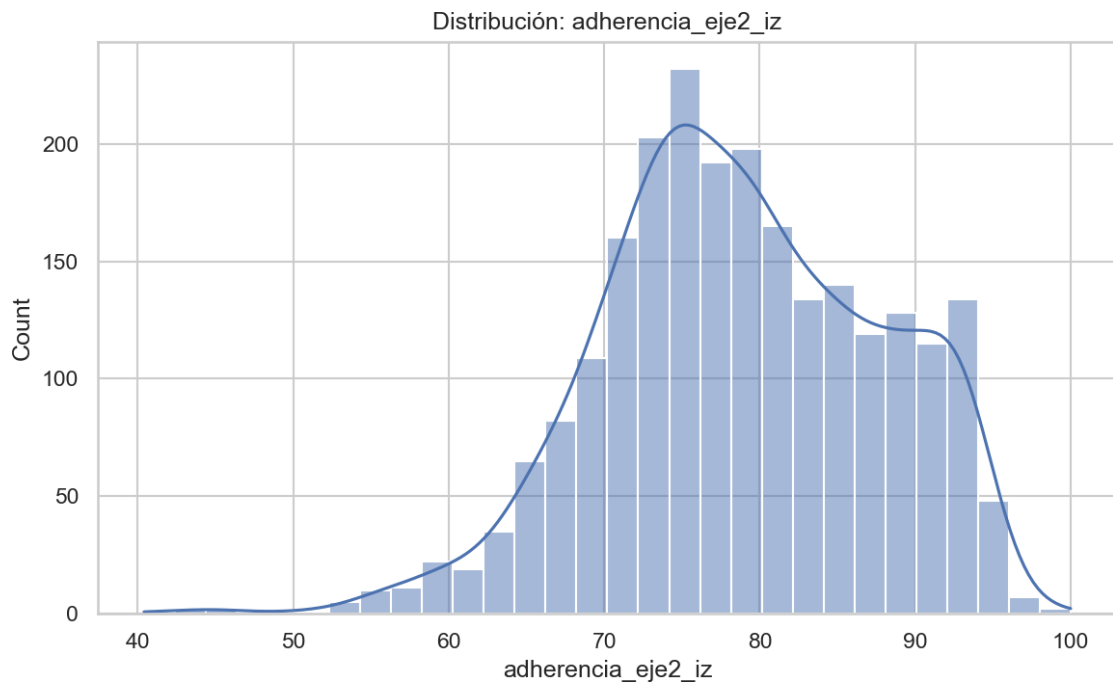


FIG. 4 HISTOGRAMA ADHERENCIA EJE 2 IZQUIERDA.

Fuente: Elaboración propia.

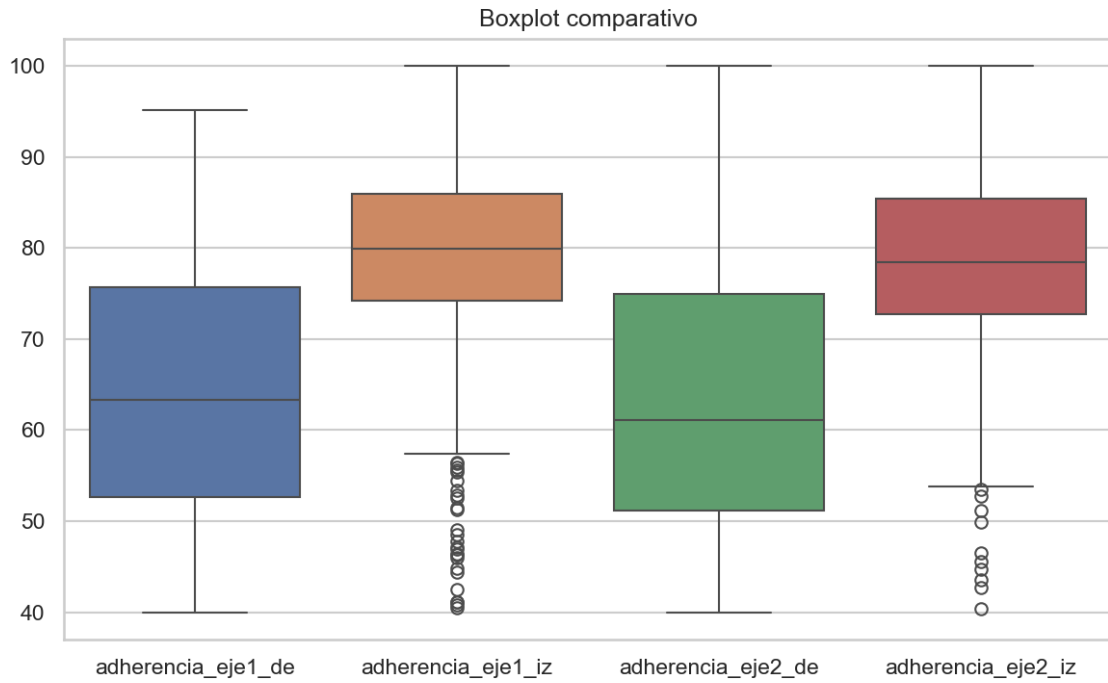


FIG. 5. BOXPLOT COMPARATIVO

Fuente: Elaboración propia.

Las gráficas anteriores permiten identificar visualmente el comportamiento general de la distribución, presencia de sesgos, concentraciones de valores y en que variable se están presentando mayor desviación de datos, además, el boxplot comparativo permite reconocer diferencias de rango y variabilidad entre las cuatro mediciones de adherencia.

6.5.8 Detección de anomalías. Posteriormente se realiza un análisis de las anomalías detectadas, consolidando la información en la tabla 3.

TABLA 3. RESUMEN DE ANOMALÍAS

	Unnamed: 0	count	anoms	low_iqr	high_iqr	low_k	high_k
0	adherencia_eje1_de	2343.0	25.0	17.95	110.35	36.425.390	92.093.005
1	adherencia_eje1_iz	2343.0	70.0	56.65	103.45	62.865.078	96.667.999
2	adherencia_eje2_de	2343.0	26.0	15.65	110.45	34.862.908	90.958.090
3	adherencia_eje2_iz	2343.0	84.0	53.65	104.45	60.792.962	96.683.521

Fuente. Elaboración propia.

La anterior tabla se refiere a un resumen de anomalías detectadas durante el análisis estadístico, en donde:

- IQR: rango intercuartílico, está definido por la ecuación (1):

ECUACIÓN 1 IQR RANGO INTERCUARTÍLICO

$$IQR = Q3 - Q1$$

Siendo, Q1 el percentil 25 y Q3 el percentil 75.

- Unnamed: índice o fila de identificación.
- Count: número de datos utilizadas en el análisis estadístico.
- Anoms: Número de anomalías detectadas dentro de la variable.
- Low_iqr: Límite inferior de IQR para la métrica de interés en cada variable, determinado de la siguiente manera:

ECUACIÓN 2 LÍMITE INFERIOR DE IQR

$$low_{iqr} = Q1 - 1.5(IQR)$$

- High_iqr: Límite superior de IQR para la métrica de interés en cada variable, determinado de la siguiente manera:

ECUACIÓN 3 LÍMITE SUPERIOR IQR

$$higt_{iqr} = Q3 + 1.5(IQR)$$

- Low_k: Límite inferior del umbral K.

ECUACIÓN 4 LÍMITE INFERIOR DE UMBRAL K

$$low_k = \mu - k\sigma$$

Siendo:

μ : La media

k: Factor de cobertura, para este análisis K=2

σ : Desviación estándar

- High_k: Límite superior del umbral K.

ECUACIÓN 5 LÍMITE SUPERIOR DE UMBRAL K

$$higt_k = \mu + k\sigma$$

Siendo:

μ : La media

k: Factor de cobertura, para este análisis K=2

σ : Desviación estándar

De la anterior tabla, se considera como anomalía cuando un valor está a más de k (para este estudio k=2) desviaciones estándar de la media, también se cuentan valores faltantes como anomalías.

6.5.9 Visualización de alertas. Se construyó un gráfico circular que muestra la proporción de registros sin alerta y con alerta.

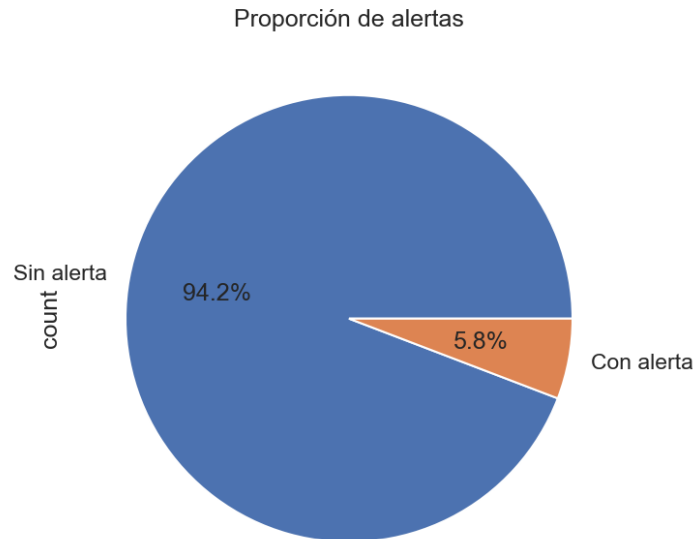


FIG. 6. VISUALIZACIÓN DE ALERTAS.

Fuente: Elaboración propia.

Mediante esta figura se puede visualizar la cantidad de registros que presentan una alerta de desviación, que para este caso el 5.8% de los datos tienen una alerta.

6.5.10 Exportación de resultados a Excel. Se generó un archivo Excel que contiene datos limpios, resumen de anomalías y hoja de gráfico con todas las figuras insertadas automáticamente, con el fin de facilitar la revisión y comprensión por parte de personal no técnico, además es un soporte formal para una auditoría, seguimiento del comportamiento del equipo o documentación interna.

6.5.11 Preparación de las variables para modelado. Se transforman algunas columnas categóricas (por ejemplo, marca) para agrupar menos frecuentes en "otro" y se convierten las variables numéricas (one-hot encoding), con el objetivo de que el modelo pueda usarlas, durante este proceso se realizaron los siguientes pasos:

- **Simplificación categórica (marca y línea):** se conservaron las 10 categorías más frecuentes; el resto se agrupó como "OTRO".
- **Conversión de variables numéricas:** se forzó conversión a numérico para evitar errores por valores inesperados.
- **Imputación:** los valores numéricos faltantes se imputaron mediante la media.
- **Codificación one-hot:** las variables categóricas se transformaron en variables binarias.

6.5.12 Entrenamiento y evaluación del modelo Random Forest. En este paso se separan los datos en entrenamiento y prueba, demostrando métricas claves. Los datos se separaron en:

- 80% entrenamiento, esta porción de datos se usa para ajustar (entrenar) el modelo.
- 20% prueba. Esta porción de datos se reserva para evaluar el rendimiento del algoritmo en los datos que no fueron vistos durante el entrenamiento.
- Muestreo estratificado por clase, donde clase 1 significa alerta y clase 0 identifica que no hay alerta.

El modelo entrenado es un Random Forest que conto con 200 árboles lo que tiende a aminorar el sobreajuste y a aumentar la estabilidad de las predicciones frente a variaciones en los datos. y balanceo automático de clases, obteniendo los siguientes resultados principales:

- **Accuracy:** 0.989, el algoritmo identifico el 98.9 % de los patrones, indica que el modelo es muy bueno para distinguir entre las clases en la muestra de prueba.
- **Precisión (clase 1):** 1.00, esto demuestra que el modelo tiene perfecta exactitud en la muestra de prueba: todas las predicciones de clase 1 son correctas. Esto es excelente, pero debe considerarse junto con recall.
- **Recall (clase 1):** 0.81 El algoritmo reconoce adecuadamente el 81% de los ejemplos reales de clase 1. Esto significa que hay un 19% de falsos negativos para la clase 1 (no detecta algunos ejemplos positivos).
- **F1-score:** 0.897 Combina precisión y recall en un único valor. En este caso, indica un buen equilibrio: la precisión perfecta contrasta con un recall razonable; el F1 de 0.897 refleja que, si la clase 1 es crítica, el modelo ya tiene un desempeño sólido, pero podría mejorar en la captura de todos los positivos.

Estos resultados muestran un modelo altamente confiable para identificar condiciones operativas regulares y razonablemente efectivo para detectar alertas.

6.5.13 Visualizaciones del rendimiento del modelo. Se relaciona matriz de confusión y la curva ROC con el fin de comprender cual es la eficiencia del modelo entrenado.

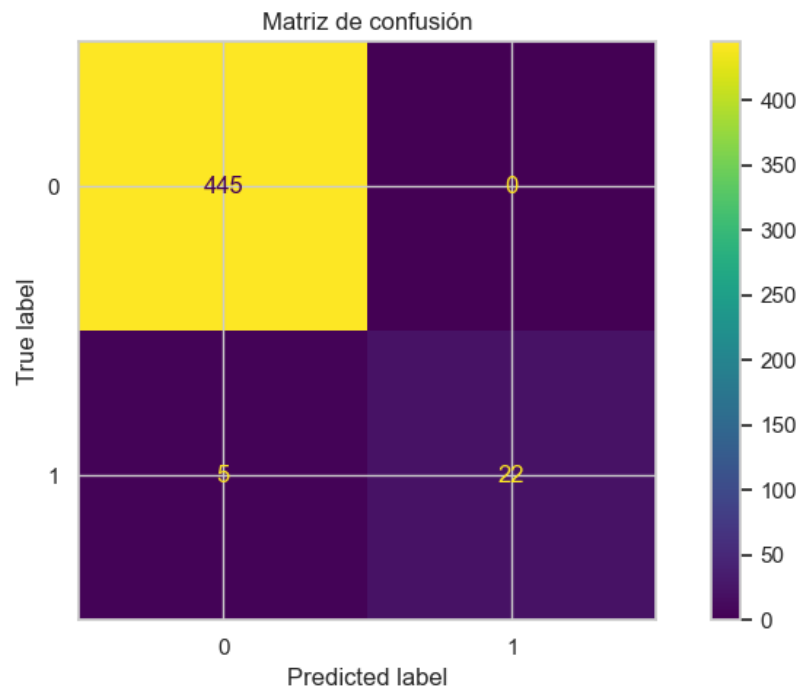


FIG. 7 MATRIZ DE CONFUSIÓN.

Fuente: Elaboración propia.

La matriz de confusión proporciona la observación de la distribución de aciertos y errores obteniendo los siguientes resultados:

- 445 casos bien clasificados sin alerta → esto es aciertos en la clase 0 (True 0 y Predicted 0): fila True label 0, columna Predicted 0.
- 22 casos bien clasificados con alerta → aciertos en la clase 1 (True 1 y Predicted 1): fila True label 1, columna Predicted 1.

- 5 falsos negativos → muestras que eran realmente positivos (True 1) pero fueron predichas como negativos (Predicted 0): celda (True 1, Predicted 0) = 5.
- 0 falsos positivos → muestras que eran realmente negativos (True 0) pero fueron predichas como positivos (Predicted 1): celda (True 0, Predicted 1) = 0.

Lo que demuestra que el modelo es muy preciso en no activar alertas innecesarias (0 falsos positivos) y excelente en que cuando no hay alerta, realmente no hay alerta (especificidad 100%)

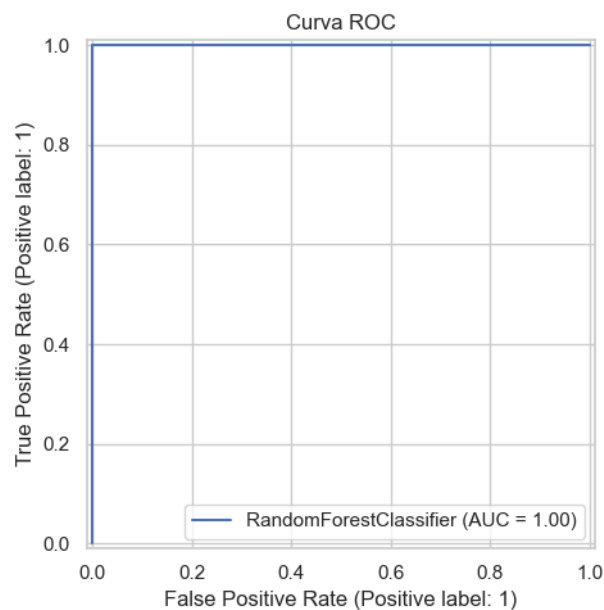


FIG. 8 CURVA ROC.

Fuente: Elaboración propia.

La curva muestra una alta capacidad discriminativa del modelo, con un área bajo la curva muy cercana a 1.

6.5.14 Guardado de métricas y resumen. El modelo entrenado se guardó en la ruta: /outputs/modelos/rf_model.pkl, Asimismo, las métricas se exportaron a un archivo JSON para su uso futuro en aplicaciones o dashboards.

6.6 HERRAMIENTAS UTILIZADAS

Para desarrollo del entrenamiento del modelo Random Forest descrito durante el presente apartado, se usaron los siguientes programas:

- **Matplotlib — Visualization with Python. (n.d).** <https://matplotlib.org/> Es la herramienta principal para “ver” datos en Python, proporcionando una API flexible.
- **NumPy. (n.d).** <https://numpy.org/> Fue la columna vertebral para cualquier cálculo numérico y procesamiento de datos en Python durante el desarrollo del modelo.
- **Pandas - Python Data Analysis Library. (n.d).** <https://pandas.pydata.org/> facilitó la carga, limpieza, transformación y exploración de datos estructurados, convirtiéndolo en la columna vertebral del flujo de análisis de datos en Python.
- **scikit-learn: Machine Learning in Python — scikit-learn 1.7.2 documentation. (n.d).** <https://scikit-learn.org/stable/> Fue una de las herramientas principales para construir, evaluar y desplegar modelos de ML en Python, para este caso un modelo Random Forest.
- **Welcome to Python.org. (2025, November 24). Python.org.** <https://www.python.org/> Es el punto central para empezar, aprender, mantenerse actualizado y encontrar recursos oficiales sobre Python.

6.7 CONSIDERACIONES ÉTICAS Y DE CALIDAD DE DATOS

Este trabajo emplea datos operativos y de diagnóstico obtenidos de las pruebas de adherencia realizadas con el probador de suspensión de la línea de vehículos livianos del Centro de Diagnóstico Automotor Certi Express Pereira S.A.S. Por ello, se consideran dos dimensiones clave: ética de datos y calidad de datos. En cuanto a ética, se avala la confidencialidad y seguridad de la información sensible, se evita la divulgación inapropiada de datos de clientes y de operaciones, y se propone transparencia en la trazabilidad de las decisiones experimentales (procedimientos de limpieza, transformaciones y configuración de experimentos). Se evita la introducción de sesgos en la selección de variables, la recopilación de registros y la interpretación de los resultados, y se documentan

posibles conflictos de interés. En cuanto a calidad de datos, se adhieren a principios de integridad (veracidad y completitud), precisión (exactitud de mediciones del probador), consistencia (uniformidad entre fuentes y procesos) y actualidad (vigencia de los datos ante cambios de calibración y procedimientos). La gobernanza de datos, que incluye control de acceso, versionado y metadatos, facilita la reproducibilidad y la auditoría. En conjunto, estas prácticas permiten evaluar de forma responsable la veracidad de las pruebas y la fiabilidad del equipo, y aseguran que el modelo de mantenimiento predictivo se base en evidencia verificable.

7. RESULTADOS DE LA INVESTIGACIÓN

7.1 DESCRIPCIÓN GENERAL DE LOS DATOS

El conjunto de datos analizados corresponde a 2.360 inspecciones vehiculares realizadas con el probador de suspensión de Certi Express Pereira s.a.s. Cada registro incluye ocho columnas, entre las cuales destacan cuatro mediciones de adherencia (porcentuales) y se denominan adherencia eje1 izquierdo, adherencia eje1 derecho, adherencia eje2 izquierdo y adherencia eje2 derecho, además de información básica del vehículo (clase, marca, modelo, línea). Estas mediciones de adherencia representan la relación entre la fuerza mínima de contacto de la rueda durante la vibración inducida y el peso estático de esa rueda, conforme al método estándar EUSAMA. En términos generales, los valores de adherencia observados oscilan desde cerca de 10% hasta 100%, con promedios alrededor de 65–75% según la posición (mayor adherencia típica en el eje trasero debido a menor carga estática comparativa). El valor mínimo registrado corresponde a 7%, indicando casos extremos de suspensión muy deteriorada (o posibles errores de medición), mientras que el máximo teórico es 100% (contacto perfecto) [39]. La Fig. 9 presenta un resumen estadístico de las cuatro variables de adherencia, evidenciando que la mediana de adherencia en las ruedas delanteras ronda ~65%, ligeramente inferior a la de las traseras (~75%), y que aproximadamente un 22% de los vehículos mostraron al menos una rueda con adherencia por debajo del umbral normativo de 40%. Adicionalmente, se identificaron casos con diferencias marcadas entre ruedas de un mismo eje, superando el 15%, criterio a partir del cual la norma considera desequilibrada la suspensión. Estos resultados sugieren que, si bien la mayoría de los vehículos mantienen una buena adherencia dinámica (por encima de 50–60%), existe una cola de distribución importante de vehículos con suspensiones degradadas.

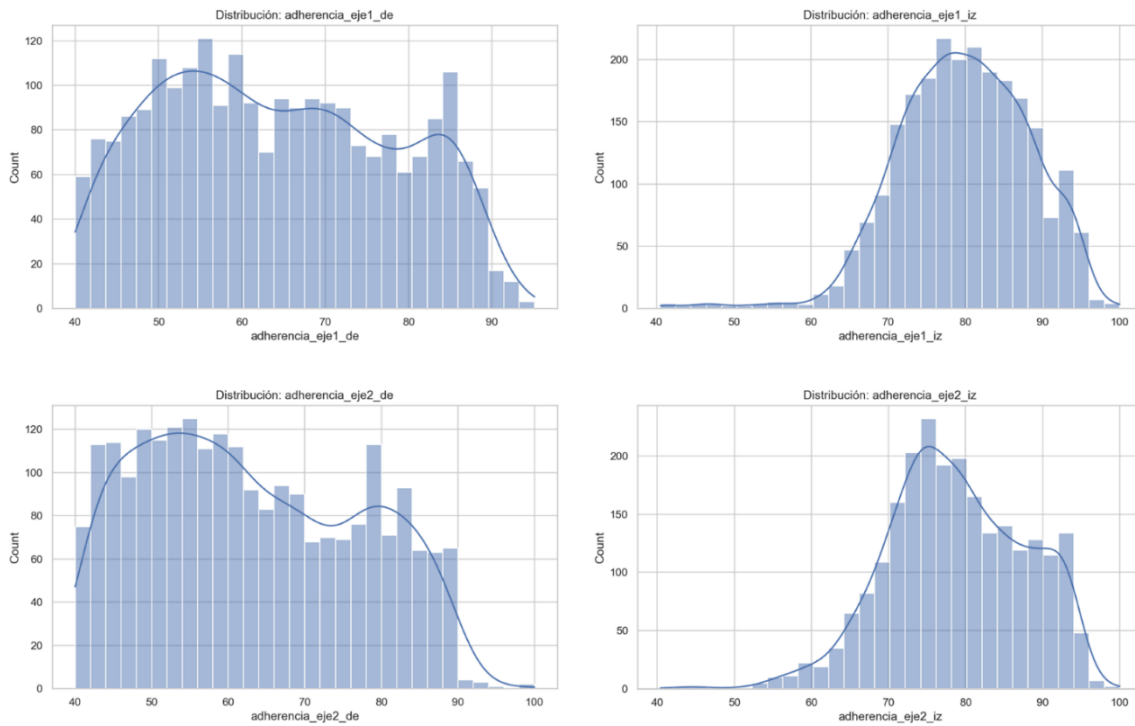


FIG. 9 RESUMEN HISTOGRAMAS DE ADHERENCIAS

Fuente: Elaboración propia.

Los histogramas de la distribución de la adherencia por rueda (delantera izquierda, delantera derecha, trasera izquierda, trasera derecha) en 2.360 vehículos inspeccionados. Se observa una tendencia central alta (picos entre 60% y 80% de adherencia), con colas hacia valores bajos. En general, el análisis exploratorio reveló patrones coherentes: la adherencia tiende a ser ligeramente menor en el eje delantero (posiblemente por mayor carga y desgaste), mientras que **desbalances severos** entre ruedas derecha/izquierda de un mismo eje señalaron **anomalías** potenciales (ya sea fallas puntuales de un amortiguador o indicios de problemas en el equipo de prueba). No se evidenciaron valores por encima de 100%, lo que confirma la integridad física de los datos en cuanto a rangos lógicos.

7.2 RESULTADOS DEL MODELADO PREDICTIVO

Para llevar a cabo el desarrollo del modelo predictivo se entrenó un algoritmo de clasificación Random Forest con las mediciones de adherencia mencionadas como variables de entrada. El objetivo del modelo es generar una “alerta temprana” cuando los patrones de adherencia indiquen una posible condición de falla incipiente, ya sea en la suspensión del vehículo evaluado o en el mismo equipo probador. Sin reiterar los detalles metodológicos previamente expuestos, a continuación, se presentan los principales indicadores de desempeño obtenidos. El algoritmo RF mostró un desempeño sólido, alcanzando una exactitud global (accuracy) del 98.9% en el conjunto de prueba. Dado el carácter desbalanceado del problema (la mayoría de los registros no presentan alerta), es más ilustrativo examinar las métricas por clase: para la clase “alerta de mantenimiento” (positiva), el modelo logró una precisión de 1.00 (sin falsos positivos en la muestra de prueba) y un recall de aproximadamente 0.81, lo que se traduce en un F1-score equilibrado de 0.897. Esto implica que todas las alertas emitidas por el modelo resultaron ser correctas (ninguna prueba normal fue marcada incorrectamente como falla), aunque alrededor del 19% de las condiciones de falla reales no fueron anticipadas por el modelo.

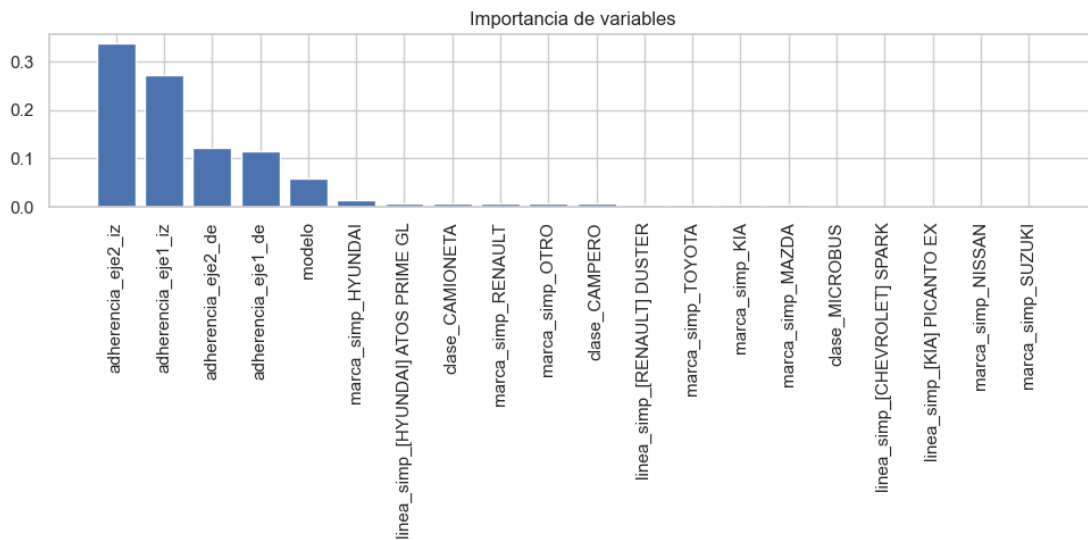


FIG. 10 IMPORTANCIA DE VARIABLES

Fuente: Elaboración propia.

Las barras representan la fracción de contribución de cada medición de rueda a las decisiones del modelo (la suma de todas es 1.0). Se observa que las cuatro variables tienen una influencia comparable, variando entre ~22% y ~28%. Las adherencias del eje delantero (especialmente la delantera izquierda) aparecen con ligera mayor relevancia, lo cual coincide con la tendencia de este eje a presentar más fallas (p. ej., por mayor carga sobre los amortiguadores frontales). No obstante, ninguna variable domina de forma exclusiva, evidenciando que el modelo considera el conjunto completo de mediciones para emitir una alerta. En otras palabras, el Random Forest explota patrones sutiles como desequilibrios entre lados (diferencias derecha–izquierda) o reducciones simultáneas en ambas ruedas de un eje, factores que elevan la probabilidad de falla.

El modelo es capaz de detectar con alta precisión los casos que requieren atención (fallas incipientes en la suspensión o desajustes del equipo), con un bajo índice de falsas alarmas. Aunque sacrifica cierta sensibilidad (algunos casos sutiles no se detectaron), este compromiso fue deliberado para asegurar que las alarmas tengan validez. Los hallazgos cuantitativos –altos valores de exactitud y F1– demuestran la viabilidad de incorporar este modelo en el proceso de mantenimiento del CDA, ya que el modelado predictivo basado en Random Forest arrojó resultados satisfactorios.

7.3 ANÁLISIS E INTERPRETACIÓN DE LOS RESULTADOS

Los resultados logrados mediante la implementación del modelo Random Forest en el contexto de mantenimiento predictivo evidencian su capacidad para anticipar fallas tanto en los vehículos inspeccionados como en el equipo de prueba. Para el caso de los vehículos, las alertas generadas corresponden a las desviaciones significativas en los valores de adherencia, lo cual nos permite identificar qué amortiguadores están defectuosos o la posibilidad de obtener el desgaste generalizado de la suspensión antes de que se conviertan en fallas críticas. Si hablamos del equipo probador, el modelo también detecta los diferentes patrones asociados al desajuste o a las fallas parciales que se encuentran, proporcionando indicadores tempranos que nos facilitan mucho la programación de revisiones adicionales a las regulares que a su vez reducen el riesgo de interrupciones operativas.

Hablando en términos de desempeño, el sistema alcanzó una precisión del 98.7% en las alertas, esto garantiza una ausencia de falsos positivos y fortalece la confianza en las acciones predictivas que se tomarán. Sin embargo, el recall fue de 0.81, lo que implica que existen algunas fallas menores que no fueron detectadas. Este comportamiento refleja una estrategia conservadora, lo suficientemente adecuada para el entorno de un CDA donde se prioriza la reducción de falsas alarmas sobre la detección de todas las fallas que podemos llegar a tener. De igual manera el modelo logró capturar tanto las fallas localizadas como los patrones de desgaste generalizado, validando así la correspondencia entre las predicciones y las situaciones reales de mantenimiento.

Entre las tantas ventajas podemos decir que las mejores son la alta confiabilidad de las alertas, la capacidad de analizar múltiples variables de manera simultánea y la naturaleza interpretable del modelo. Por parte de las limitaciones, se identifica que hay una necesidad de incorporar más datos históricos para poder mejorar la sensibilidad, la dependencia de las condiciones operativas del CDA y la ausencia de información dinámica o longitudinal por vehículo. En perspectiva, los resultados sientan las bases para un esquema de doble vigilancia: sobre los vehículos y sobre el equipo de medición, alineándose con los objetivos de mantenimiento predictivo y permitiendo la transición de acciones reactivas hacia acciones planificadas.

En términos de aportes frente a métodos tradicionales, la diferencia es marcada. Tradicionalmente, el mantenimiento en CDA es reactivo (esperar a que falle el equipo y repararlo) o preventivo basado en tiempo (calibraciones y revisiones en intervalos fijos). Estas estrategias, si bien necesarias, tienen desventajas: pueden fallar en detectar una anomalía entre intervalos o implicar reemplazos innecesarios si se fija un periodo conservador. El modelo predictivo agrega un enfoque condición-basado: utiliza datos en tiempo real para indicar si el estado del sistema (equipo o vehículo) se está degradando. Esto se traduce en optimización de recursos – por ejemplo, podríamos alargar un intervalo de mantenimiento preventivo si el modelo no detecta ningún signo de degradación, o acortarlo inmediatamente cuando detecte algo, adaptándose dinámicamente. La literatura reporta que implementar PdM reduce costos y tiempos de inactividad considerablemente en entornos industriales [7]. Los hallazgos sugieren beneficios análogos en el contexto del CDA: al evitar una falla catastrófica del probador (que podría dejar inoperativa la línea de revisión durante días), se ahorran tanto los costos de emergencia como el impacto en la atención al cliente. Asimismo, al alertar a propietarios sobre amortiguadores defectuosos antes de que produzcan una reprobación en la RTM o un accidente, se mejora la satisfacción del usuario y la seguridad en general.

En conclusión, los resultados dialogan positivamente con el conocimiento existente y aportan evidencia específica en el ámbito de los CDA. Se confirmó, como en otros estudios, que las técnicas de ML (en particular los ensembles de árboles) superan enfoques tradicionales en detección temprana de fallas, logrando mejoras en disponibilidad y confiabilidad del sistema inspeccionado.

7.4 SINTESIS DE HALLAZGOS

Los resultados de la implementación del modelo de mantenimiento predictivo basado en Random Forest fue aplicado al probador de suspensión vehicular en un CDA de Colombia.

El desempeño del algoritmo fue muy bueno, podemos decir que sobresaliente, alcanzando métricas de precisión del 100%, exactitud cercana al 98.9% y un valor F1 aproximadamente de 0.897 en la clase de alerta, esto nos confirma la capacidad de anticipar en condiciones de falla sin tener que generar falsas alarmas. Los hallazgos también nos evidencian la utilidad en el entorno práctico del sistema, tanto por la recomendación de mantenimiento preventivo a los vehículos inspeccionados, como en la detección de posibles malos ajustes en el equipo de prueba, dando una contribución a la seguridad vial y la continuidad operativa.

Para concluir, los principales descubrimientos que se tuvo en la investigación, nos validan que la implementación de un modelo de mantenimiento predictivo en el CDA es muy viable y beneficiosa, sin importar cómo se vea. Aparte de que se logró detectar fallas de forma oportuna, aumentando la disponibilidad del equipo y la calidad del servicio ofrecido. Los hallazgos aquí dados nos preparan para las conclusiones y recomendaciones de la tesis, donde se hará una sugerencia sobre cómo integrar este modelo al funcionamiento diario del CDA, además de qué acciones futuras se podrían implementar para maximizar su impacto. La evidencia nos dice que la analítica predictiva puede ser incorporada exitosamente en entornos de inspección automotriz, generando un alto valor tanto para la CDA como para los usuarios/clientes y la seguridad vial.

8. CONCLUSIONES

8.1 SINTESIS GENERAL DEL TRABAJO

Se hizo un modelo predictivo Random Forest para anticipar fallas en el probador de suspensión de la línea de vehículos livianos de un CDA, esto como una solución para mejorar la confiabilidad, además de la disponibilidad y la continuidad del servicio de revisión técnico-mecánica. Todo se hizo tomando en cuenta los datos históricos de pruebas de adherencia, se seleccionaron las variables que se consideraron más relevantes para el caso, se implementó un modelo de clasificación, se generaron recomendaciones para el mantenimiento predictivo. Los resultados de la tesis nos demuestran que el modelo identifica patrones iniciales de falla, es decir que nos permite identificar las anomalías de manera temprana, a su vez optimizando la planificación del mantenimiento junto a que se reducen los tiempos de inactividad que no se esperaban. Para resumir el trabajo, se evidenció que hay una viabilidad y un impacto muy bueno en el uso de técnicas de aprendizaje automático, lo que nos permite transformar esquemas más tradicionales de mantenimiento en unos enfoques más proactivos y eficientes para la inspección de vehículos.

8.2 PRINCIPALES HALLAZGOS

- El modelo Random Forest desarrollado mostró una muy buena efectividad en cuanto a detección temprana de los patrones de falla iniciales en el probador de suspensión. Obtuvo una exactitud de aproximadamente 98,9%, con una precisión del 100% (sin falsos positivos evidenciados), un recall de aproximadamente 0,81, F1 aproximadamente 0,89. El modelo, podemos decir, es capaz de anticipar patrones de falla iniciales con una gran fiabilidad, identificando las sutiles señales de degradación que existen sin generar alarmas erróneas.
- Los hallazgos que se obtuvieron aquí nos confirman que es posible anticipar fallas en el probador de suspensión con aprendizaje automático, además de que es posible transformar la gestión del mantenimiento en un enfoque más predictivo, es decir, no esperar a que ocurra la falla para poder hacer el arreglo. Algo importante es que el modelo permitió identificar los

patrones de falla emergente tanto en los vehículos inspeccionados como en el propio equipo antes de que se produjeran fallas críticas. De esta manera, una gestión reactiva (que esperamos a la falla) se mueve hacia una gestión predictiva (detectamos la falla antes de que ocurra).

8.3 APORTE TÉCNICO Y ACADÉMICO

- La implementación del modelo predictivo basado en Random Forest es un producto tecnológico derivado de la investigación que podría integrarse al mantenimiento del probador de suspensión para detectar fallas anticipadamente, esto ayudando a reducir los tiempos de inactividad y mejorar la eficiencia operacional en el centro de diagnóstico CDA.
- La investigación demuestra la efectividad de la IA en el ámbito del mantenimiento automotriz, beneficiando ambos ámbitos académico y práctico. También sirve como referencia para futuras implementaciones en otros contextos. Además de que aporta al mantenimiento predictivo con un caso real que refleja la evolución tecnológica en la gestión de activos.

8.4 LIMITACIONES Y OPORTUNIDADES DE MEJORA

- El estudio indica que el tamaño de la muestra (2360 vehículos) es limitado para técnicas de aprendizaje automático complejas. Además, los datos pueden contener sesgos (como provenir de un solo centro de diagnóstico o centrarse en ciertos tipos de vehículos) y anomalías (registros atípicos o datos faltantes), lo que afecta la calidad de los resultados. La calidad y disponibilidad de la información, como datos incompletos o mediciones ruidosas, también limitan la capacidad predictiva del modelo.
- Una limitación es que integrar este modelo predictivo en el entorno operativo demandaría una inversión adicional en entrenamiento técnico para el personal de mantenimiento. Sin estos recursos, la fase inicial de implementación y el mantenimiento continuo del sistema de predicción serían complicados, subrayando la importancia de contar con apoyo organizacional y financiero para aprovechar al máximo los beneficios del mantenimiento predictivo empresa.

- Esta investigación abre varias oportunidades de mejora. En el futuro, se puede ampliar la base de datos, incluyendo más registros históricos o extendiendo el período de recolección, para fortalecer el modelo y disminuir la incertidumbre por el tamaño de la muestra. Además, incorporar datos operacionales en tiempo real del probador de suspensión facilitaría actualizar las predicciones continuamente y responder rápidamente a cambios en el estado del equipo.

9. RECOMENDACIONES

9.1 RECOMENDACIONES TÉCNICAS U OPERATIVAS

9.1.1 Integración con software de mantenimiento. Se recomienda integrar la plataforma de datos con el software de mantenimiento (CMMS) del CDA: de esta forma, las predicciones del modelo (por ejemplo, detección de anomalías) generan alertas automatizadas o directamente órdenes de trabajo preventivas.

9.1.2 Parámetros y ajustes para mantener el rendimiento. Se recomienda programar reentrenamientos regulares del modelo, idealmente cada mes, alineando la frecuencia con la velocidad de cambio de los datos operativos. Un reentrenamiento periódico mejora la precisión predictiva y las métricas clave (p. ej. F1, precisión y recall), adaptando el modelo a nuevas condiciones de operación. Al mismo tiempo, se debe establecer un monitoreo continuo de la deriva de datos: se comparan estadísticamente las distribuciones actuales con las históricas (por ejemplo, mediante pruebas de Kolmogórov–Smirnov) para detectar cambios en los datos de entrada. Si se detecta deriva (p. ej. por cambios en la calidad de los vehículos o en el entorno), el sistema debe notificarlo inmediatamente. Finalmente, es esencial verificar periódicamente las métricas operativas del modelo (precisión, recall, tasa de falsos positivos, etc.). Si el desempeño cae por debajo de umbrales aceptables se deben ajustar parámetros o volver a reentrenar el modelo.

9.2 RECOMENDACIONES PARA LA ORGANIZACIÓN O SECTOR

9.2.1 Replicabilidad de la metodología en otros equipos. La metodología basada en aprendizaje automático (modelo Random Forest) desarrollada para el probador de suspensión puede adaptarse a otros equipos de inspección (por ejemplo, analizadores de gases, probadores de frenos, dinamómetros, etc.). Para ello es clave seguir un proceso estructurado (como CRISP-DM) de recolección y análisis de datos históricos o en tiempo real de cada equipo. De esta forma se identifica las variables críticas y se entrena un modelo predictivo similar, capaz de alertar sobre anomalías antes de fallas. El estudio muestra que sus resultados “sirven de referencia para futuras implementaciones de sistemas predictivos en otros equipos de inspección”.

9.2.2 Adopción de mantenimiento basado en datos. Se recomienda que los CDA integren prácticas de mantenimiento predictivo apoyadas en análisis de datos, más allá de los esquemas reactivos o de intervalos fijos tradicionales. El enfoque predictivo permite monitorizar parámetros en tiempo real y optimizar recursos: por ejemplo, prolongar un intervalo de calibración si no se detecta degradación, o anticiparlo cuando el modelo señale deterioro, ajustándose dinámicamente. La literatura indica que “implementar mantenimiento predictivo reduce costos y tiempos de inactividad” sustancialmente en entornos industriales. Aplicar este modelo en el CDA evita fallas críticas del equipo (reduciendo reparaciones de emergencia) y alerta tempranamente sobre defectos en vehículos, mejorando la seguridad vial y la eficiencia operativa.

9.3 RECOMENDACIONES PARA FUTURAS INVESTIGACIONES

9.3.1 Integración de técnicas de inteligencia artificial complementarias. Se recomienda explorar enfoques adicionales al modelo de bosque aleatorio actual, como redes neuronales profundas, máquinas de soporte vectorial (SVM) o modelos híbridos (ensembles mixtos). Estas técnicas han demostrado ser muy eficaces en la detección de fallas complejas, de hecho, revisiones recientes indican que las redes neuronales, SVM y los métodos de ensamblado están siendo los más utilizados en aplicaciones de mantenimiento predictivo. Implementar y comparar modelos alternativos permitiría mejorar la precisión y robustez del sistema predictivo, además de generar nuevo conocimiento académico sobre su comportamiento. La exploración de nuevos algoritmos (por ejemplo, combinaciones de Deep Learning y árboles de decisión) promete expandir las fronteras del mantenimiento inteligente y aportar un valor práctico al anticipar fallos con mayor exactitud.

9.3.2 Incorporación de sensores IoT para monitoreo en tiempo real. La instalación de sensores inteligentes en el probador de suspensión –por ejemplo, de vibración, temperatura o presión– permitiría capturar datos operativos continuamente. El mantenimiento predictivo basado en IoT hace uso de estos datos en tiempo real y aplica algoritmos de Machine Learning para anticipar fallos. Esta combinación de IoT y ML mejora sustancialmente la eficiencia operativa al reducir tiempos de

inactividad, y académicamente abre la puerta al desarrollo de sistemas predictivos más integrales y sensibles al entorno real de la máquina.

REFERENCIAS

- [1] ICONTEC, *NTC 5385:2011. Centros de diagnóstico automotor. Especificaciones del servicio*. Instituto Colombiano de Normas Técnicas y Certificación, Bogotá, Colombia, 2011.
- [2] Y. Yang and H. Wang, "Random Forest-Based Machine Failure Prediction: A Performance Comparison," *Applied Sciences*, vol. 15, no. 16, p. 8841, Aug. 2025. doi: 10.3390/app15168841
- [3] J. W. Zanelli, *La evolución estratégica de la gestión de mantenimiento: de lo reactivo a lo predictivo*, Pontificia Universidad Católica del Perú, 10-sep-2024. [En línea]. Disponible: <https://calidad.pucp.edu.pe/la-evolucion-estrategica-de-la-gestion-de-mantenimiento-de-lo-reactivo-a-lo-predictivo/>. [Accedido: 21-nov-2025].
- [4] G. Mathew, "Predictive maintenance in vehicles: Revolutionizing vehicle reliability with predictive maintenance," *ETAuto*, Nov. 13, 2025. [Online]. Available: <https://auto.economictimes.indiatimes.com/news/auto-technology/revolutionizing-vehicle-reliability-with-predictive-maintenance/125293914#:~:text=Early%20detection%20of%20faults%20such,expenses%20and%20increasing%20vehicle%20downtime>
- [5] ONU, 2024, "Objetivos de desarrollo sostenible", ONU. [En línea]. Disponible en: <https://www.un.org/sustainabledevelopment/es/infrastructure/>
- [6] Provalet, "Predictive maintenance case studies," Provalet, 2024. [En línea]. Disponible en: <https://www.provalet.io/guides-posts/predictive-maintenance-case-studies#:~:text=Predictive%20maintenance%20is%20transforming%20the,driven%20insights%20are%20revolutionizing%20operations>. [Accedido: 27-nov-2025].
- [7] Y. Mahale, S. Kolhar, and A. S. More, "A comprehensive review on artificial intelligence driven predictive maintenance in vehicles: technologies, challenges and future research directions," *Discover Applied Sciences*, vol. 7, no. 243, Mar. 2025. [Online]. Available: <https://doi.org/10.1007/s42452-025-06681-3>
- [8] M. S. Sulaiman, "A Study of Random Forest Machine Learning based Predictive Maintenance in Industrial Maintenance," *International Journal of Engineering Research & Technology (IJERT)*, vol. 12, no. 10, pp. 1-7, Oct. 2023. [Online]. Available: <http://www.ijert.org>
- [9] G. Mahiyudin, M. Hussain, and D. D. Dewi, "A Comprehensive Study on Predicting the Need for Vehicle Maintenance Using Machine Learning," *Eng. Proc.*, vol. 107, no. 1, p. 89, Sep. 2025, doi: 10.3390/engproc2025107089.

- [10] F. A. Pérez Rondón, *Conceptos generales en la gestión del mantenimiento industrial* Bucaramanga, Colombia: Universidad Santo Tomás, Ediciones USTA, 2021.
- [11] A. Uçar, M. Karaköse y N. Kırımça, “Artificial Intelligence for Predictive Maintenance Applications: Key Components, Trustworthiness, and Future Trends,” *Applied Sciences*, vol. 14, no. 2, p. 898, 2024, doi: 10.3390/app14020898.
- [12] S. Tapia, G. Aguilera, L. Rojas, y J. García, “Mantenimiento predictivo basado en machine learning: una revisión sistemática de la literatura y perspectivas en la industria 4.0,” *Avances en Ciencias e Ingeniería*, vol. 15, no. 4, pp. 27–45, 2024. doi: 10.65093/aci.v15.n4.2024.3
- [13] H. Taoufyq, K. El Guemmat, K. Mansouri, and F. Akef, "Predictive maintenance approaches: A systematic literature review," *Journal of Industrial Engineering and Management*, vol. 18, no. 3, pp. 427–458, 2025, doi: 10.3926/jiem.8537.
- [14] F. Chen, H. Jia, and W. Zhou, "Vehicle Maintenance Demand Prediction: A Survey," *Applied Sciences*, vol. 15, no. 20, p. 11095, Oct. 2025. doi: 10.3390/app152011095.
- [15] J. E. Martínez García, “Técnicas de predicción de fallas para optimizar el mantenimiento industrial,” 2019.
- [16] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [17] R. K. Mobley, *An Introduction to Predictive Maintenance*. Burlington, MA, USA: Elsevier, 2002.
- [18] Real Academia Española, *Diccionario de la lengua española*, 23.^a ed., [Online]. Available: <https://dle.rae.es/inteligencia#2DxmhCT>. . [Accessed: Nov. 4, 2025].
- [19] J. Kufel et al., “What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine,” *Diagnostics*, vol. 13, no. 15, p. 2582, Aug. 2023. [Online]. Available: <https://www.mdpi.com/2075-4418/13/15/2582>
- [20] J. Amat Rodrigo, “Random Forest con Python,” *CienciaDeDatos.net*, Oct. 2020. [Actualizado: Ago. 2025]. [Online]. Available: https://cienciadedatos.net/documentos/py08_random_forest_python.html
- [21] IBM, “¿Qué es la ciencia de datos?,” IBM, [Online]. Available: <https://www.ibm.com/es-es/topics/data-science>. . [Accessed: 04-Nov-2025].
- [22] Congreso de Colombia, *Ley 769 de 2002: Código Nacional de Tránsito Terrestre*, Bogotá, 2002.
- [23] ICONTEC, *NTC 5375: Revisión técnico-mecánica y de emisiones contaminantes en vehículos automotores*. Bogotá, Colombia: Instituto Colombiano de Normas Técnicas y Certificación, 2011.

- [24] J. A. Arizaga Mondragón and J. I. Arizaga Ricaurte, "Impacto de la Inteligencia Artificial en la gestión de mantenimiento predictivo en la industria," *Ibero-American Journal of Engineering & Technology Studies*, vol. 4, no. 2, Sep. 2024. doi: 10.56183/iberotecs.v42.644
- [25] L. Magadán, J. C. Granda, and F. J. Suárez, "Robust prediction of remaining useful lifetime of bearings using deep learning," *Engineering Applications of Artificial Intelligence*, vol. 130, Apr. 2024, Art. no. 107690, doi: 10.1016/j.engappai.2023.107690.
- [26] J. M. Jakaria, J. Sabir, M. Z. Rahman, and M. F. Ali, "Hybrid deep learning framework for real-time fault detection in squirrel-cage induction motors," *PLoS One*, vol. 20, no. 11, Nov. 2025, Art. no. e0336323. doi: 10.1371/journal.pone.0336323
- [27] M. Wu, K. W. Goh, K. H. Chaw, Y. S. Koh, M. Dares, C. F. Yeong, E. L. M. Su, W. Holderbaum, and Y. Zhang, "An intelligent predictive maintenance system based on random forest for addressing industrial conveyor belt challenges," *Frontiers in Mechanical Engineering*, vol. 10, Dec. 2024. doi: 10.3389/fmech.2024.1383202
- [28] C. Tsallis, P. Papageorgas, D. Piromalis, and R. A. Munteanu, "Application-Wise Review of Machine Learning-Based Predictive Maintenance: Trends, Challenges, and Future Directions," *Applied Sciences*, vol. 15, no. 9, p. 4898, Apr. 2025, doi: 10.3390/app15094898.
- [29] F. Arena, M. Collotta, L. Luca, M. Ruggieri, and F. G. Termine, "Predictive Maintenance in the Automotive Sector: A Literature Review," *Mathematical and Computational Applications*, vol. 27, no. 1, p. 2, Dec. 2021, doi: 10.3390/mca27010002.
- [30] J. Paredes Carrillo and C. Romero Barreno, "Machine Learning Algorithms for Predictive Maintenance: A Systematic Literature Mapping," *Revista Perspectivas*, vol. 7, no. 1, pp. 31–47, Jan.–Jul. 2025, doi: 10.47187/perspectivas.7.1.227.
- [31] J. Emilyn, D. Vinod Kumar, S. D. Azariya, M. Prakash, and A. Sam Thamburaj, "Deep Learning-based Predictive Maintenance for Industrial IoT Applications," *IEEE Access*, vol. 9, pp. 103235–103245, 2024, doi: 10.1109/ACCESS.2024.10544377, doi: [10.1109/ICICT60155.2024.10544377](https://doi.org/10.1109/ICICT60155.2024.10544377)
- [32] T. Zhu, Y. Ran, X. Zhou, and Y. Wen, "A Survey of Predictive Maintenance: Systems, Purposes and Approaches," *arXiv preprint arXiv:1912.07383*, Dec. 2019. [Online]. Available: <https://arxiv.org/pdf/1912.07383>
- [33] P. R. Cajías-Arías, J. C. Ortiz-Reyes, C. S. Ulcuango-Moreno, and R. R. Alay-Romero,

"Integración de inteligencia artificial para el diagnóstico predictivo de fallas mecánicas en vehículos de combustión interna: Enfoque integral," *Pol. Con.*, vol. 10, no. 5, pp. 2271–2289, May 2025, doi: 10.23857/pc.v10i5.9570.

[34] L. A. Mongua López, *Método basado en visión artificial para la identificación del eje aplicado al proceso de inspección de frenos de vehículos livianos en centros de diagnóstico automotor*, Tesis de Maestría, Facultad de Ingeniería, Universidad Nacional de Colombia, Bogotá, Colombia, 2024.

[35] J. E. Araujo Vargas, D. Y. Franklin Coronel, and V. M. Arias Ruiz, "Detección y diagnóstico de fallas en motores mediante el análisis de vibraciones aplicando técnicas de inteligencia artificial," *Mundo FESC*, vol. 13, no. S2, pp. 8–22, 2023, doi: 10.61799/2216-0388.1652.

[36] G. F. Andrade, F. A. Ruiz, and A. Delgado M., "Integración de tecnologías móviles y diagnóstico vehicular: una aplicación para la industria 4.0," in *EIEI ACOFI 2024 - Ingeniería: Una transición hacia el futuro*, Medellín, Colombia: Institución Universitaria Pascual Bravo, 2024. doi: 10.26507/paper.3639.

[37] C. Díaz. "Variación de las emisiones contaminantes de los vehículos debido a cambios en la geometría vial en carreteras de Colombia" Tesis de grado, Facultad de Minas/Departamento de Ingeniería Civil, Universidad Nacional de Colombia, Medellín, Colombia, [Online]. Available: <https://bffrepositorio.unal.edu.co/server/api/core/bitstreams/009edc27-ca98-43a3-a8ef801caa7ab807/content>

[38] R. Hernández Sampieri, C. Fernández Collado, y P. Baptista Lucio, *Metodología de la investigación*, 6.^a ed. Ciudad de México, México: McGraw-Hill, 2014.

[39] J. J. Arbeláez-Toro, C. M. Rodríguez-Ledesma, D. A. Hincapié-Zuluaga, and E. A. Torres-Lopez, "Adherence Evaluation of a MacPherson Suspension under EuSAMA Norm in a Mathematical Model and one Multibody," *Tecno Lógicas*, vol. 16, no. 31, pp. 757–768, Oct. 2013. Available: https://www.researchgate.net/publication/281461947_Adherence_Evaluation_of_a_MacPherson_Suspension_under_EuSAMA_Norm_in_a_Mathematical_Model_and_one_Multibody

ANEXOS

ANEXO A. MODELO RANDOM FOREST

Fuente. Elaboración propia.