



# POSGRADOS

## MAESTRÍA EN TRANSFORMACIÓN DIGITAL E INNOVACIÓN

RPC-SO-13-NO.734-2024

OPCIÓN DE TITULACIÓN:

ARTÍCULOS PROFESIONALES DE ALTO NIVEL

TEMA:

PREDICCIÓN DEL VOLUMEN DE VENTAS  
MEDIANTE LA APLICACIÓN DE RANDOM  
FOREST EN TALLERES DE LA CORPORACIÓN  
MQ

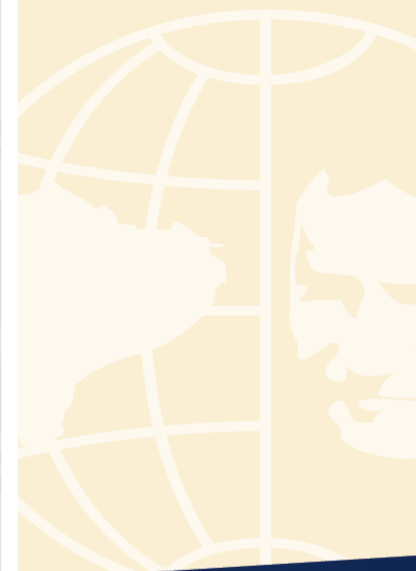
AUTOR:

MARCO VINICIO QUITO RIVERA

DIRECTOR:

IRENE GISEL BUELE NUGRA

CUENCA – ECUADOR  
2026



**Autor:**



**Marco Vinicio Quito Rivera**

Ingeniero en Sistemas informáticos.  
Candidato a Magister en Transformación Digital e  
Innovación por la Universidad Politécnica Salesiana  
sede Cuenca.  
mquitor1@est.ups.edi.ec

**Dirigido por:**



**Irene Gisel Buele Nugra**

Ingeniera en Administración de Empresas.  
Ingeniera en Contabilidad y Auditoría.  
Máster en Auditoría Integral.  
Máster en Dirección y Gestión Financiera.  
Máster en Planificación de Proyectos de Desarrollo  
Rural y Gestión Sostenible.  
Doctora en Planificación de Proyectos de Desarrollo  
Rural y Gestión Sostenible.  
ibuele@ups.edu.ec

Todos los derechos reservados.

Queda prohibida, salvo excepción prevista en la Ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra para fines comerciales, sin contar con autorización de los titulares de propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual. Se permite la libre difusión de este texto con fines académicos investigativos por cualquier medio, con la debida notificación a los autores.

DERECHOS RESERVADOS

2026 © Universidad Politécnica Salesiana.

CUENCA – ECUADOR – SUDAMÉRICA

MARCO VINICIO QUITO RIVERA

Predicción del volumen de ventas mediante la aplicación de Random Forest en talleres de la Corporación MQ

**Tabla de Contenidos**

1.	RESUMEN .....	iv
2.	ABSTRACT.....	v
1.	INTRODUCCIÓN .....	6
2.	MATERIALES Y MÉTODOS .....	9
3.	RESULTADOS.....	15
4.	DISCUSIÓN .....	25
5.	CONCLUSIONES .....	27
6.	AGRADECIMIENTOS .....	28
7.	REFERENCIAS BIBLIOGRÁFICAS.....	28

## Lista de tablas

<b>Tabla 1</b> Metadatos de Variables.....	12
<b>Tabla 2</b> Detalle de variables dependientes e independientes usados para la predicción .....	13
<b>Tabla 3</b> Muestra los datos cargados en Python.....	21
<b>Tabla 4</b> Configuración del modelo Random Forest Regressor y Hiperparámetros de entrenamiento .....	21
<b>Tabla 5</b> Importancia de las variables en el modelo Random Forest.....	23
<b>Tabla 6</b> Escenarios para comprobar aprendizaje del modelo Random Forest.....	24

## Lista de figuras

<b>Figura 1</b> Ventas por La línea de Negocio de Corporación MQ, numero de facturas y monto de ventas (valores representados en %) .....	15
<b>Figura 2</b> Ventas por canal de Corporación MQ, numero de facturas y monto de ventas (valores representados en %) .....	16
<b>Figura 3</b> Ventas por Bodega de Corporación MQ, numero de facturas y monto de ventas (valores representados en %) .....	18
<b>Figura 4</b> Ventas por Clase Cliente de Corporación MQ, numero de facturas y monto de ventas (valores representados en %) .....	19
<b>Figura 5</b> Ventas por zona de Corporación MQ, numero de facturas y monto de ventas (valores representados en %) .....	20
<b>Figura 6</b> Fragmento código Python ingreso de datos.....	22
<b>Figura 7</b> Código Python de procesamiento de datos.....	22
<b>Figura 8</b> Diagrama simplificado de un árbol del Random Forest .....	23
<b>Figura 9</b> Volumen de Ventas Pronosticados .....	24

PREDICCIÓN DEL  
VOLUMEN DE VENTAS  
MEDIANTE LA  
APLICACIÓN DE  
RANDOM FOREST EN  
TALLERES DE LA  
CORPORACIÓN MQ

AUTOR(ES):

MARCO VINICIO QUITO RIVERA

## 1. RESUMEN

En el marco de la transformación digital de la industria automotriz, la gestión eficiente de la demanda en los talleres de posventa representa un reto estratégico debido a la variabilidad del mercado, la heterogeneidad de los servicios, lo que exige el uso de herramientas analíticas avanzadas para apoyar la toma de decisiones basadas en datos. El objetivo de la investigación fue aplicar y entrenar un modelo predictivo que permita estimar con mayor precisión el volumen de ventas en los talleres de la Corporación MQ empresa automotriz ecuatoriana dedicada a la comercialización de vehículos, repuestos y servicios de taller (preventivo y correctivo), considerando variables clave como línea de negocio, clase de cliente, bodega, canal y zona geográfica. La metodología adoptó un enfoque cuantitativo y experimental, sustentado en el análisis de 114.937 registros históricos de ventas correspondientes al año 2024, de los cuales se depuró una muestra válida de 87.778 registros, procesados y modelados en Python mediante técnicas de aprendizaje automático. El modelo Random Forest fue entrenado y evaluado bajo criterios estadísticos de desempeño, permitiendo identificar la importancia relativa de las variables que influyen en el comportamiento de las ventas. Los resultados evidenciaron que el modelo alcanzó un coeficiente de determinación de  $R^2 = 0.7769$ , lo que demuestra una alta capacidad predictiva, aprendizaje consistente y su aplicabilidad para anticipar escenarios de ventas. La investigación contribuye al uso de la analítica de datos como herramienta estratégica para la planeación, eficiencia operativa y transformación digital en la Corporación MQ y en la industria automotriz ecuatoriana.

**Palabras clave:** Predicción de ventas, Inteligencia artificial, Machine learning, Random Forest, Transformación digital.

## 2. ABSTRACT

Within the context of the digital transformation of the automotive industry, efficient demand management in after-sales workshops represents a strategic challenge due to market variability and service heterogeneity, which requires the use of advanced analytical tools to support data-driven decision-making. The objective of this research was to apply and train a predictive model to more accurately estimate sales volume in the workshops of Corporación MQ, an Ecuadorian automotive company dedicated to the commercialization of vehicles, spare parts, and workshop services (preventive and corrective), considering key variables such as line of business, customer class, warehouse, channel, and geographic area. The methodology adopted a quantitative and experimental approach, based on the analysis of 114,937 historical sales records from the year 2024, from which a refined sample of 87,778 valid records was obtained, processed, and modeled in Python using machine learning techniques. The Random Forest model was trained and evaluated under statistical performance criteria, allowing the identification of the relative importance of variables influencing sales behavior. The results showed that the model achieved a coefficient of determination of  $R^2 = 0.7769$ , demonstrating high predictive capacity, consistent learning, and applicability for anticipating sales scenarios. This research contributes to the use of data analytics as a strategic tool for planning, operational efficiency, and digital transformation within Corporación MQ and the Ecuadorian automotive industry.

**Keywords:** Sales forecasting, Artificial intelligence, Machine learning, Random Forest, Digital transformation.

## 1. INTRODUCCIÓN

La industria automotriz se encuentra inmersa en un proceso de transformación profunda, impulsado por la digitalización, la innovación tecnológica y las nuevas demandas del mercado (Gutiérrez, 2024). Más allá de la venta de vehículos, los servicios de posventa, que incluyen la atención en talleres, la distribución de repuestos y la gestión de garantías, representan una de las principales fuentes de ingresos y fidelización para los concesionarios y fabricantes. En efecto, en la industria automotriz, las ventas de nuevos vehículos representan alrededor del 37 % de los ingresos, mientras que los servicios de posventa concentran cerca del 54 % de las utilidades totales, lo que evidencia su papel estratégico para la sostenibilidad del negocio automotor. (Eslava et al., 2021)

El sector automotriz mundial, según Ordóñez (2024) experimenta una transición tecnológico-productiva marcada por la convergencia entre la ingeniería mecánica, la electrónica y las telecomunicaciones, lo que ha originado un nuevo núcleo dinámico basado en la información y el conocimiento. Este desarrollo no solo ha transformado la producción de vehículos eléctricos, autónomos y conectados, sino que también ha afectado al servicio de posventa, que se consolida como un pilar estratégico para la rentabilidad de las organizaciones de la industria automotriz (Millar et al., 2020)

La causa del problema es la creciente demanda de repuestos y servicios en los talleres de reparación y mantenimiento (preventivo y correctivo) de automóviles, lo que dificulta la planificación precisa del inventario, la asignación eficiente de recursos humanos y la toma de decisiones operativas y estratégicas. En este escenario, la proyección de ventas se convierte en un elemento muy importante para una gestión organizada e inteligente de los talleres de servicio, especialmente en el ámbito de la posventa. Factores como el envejecimiento del parque vehicular, las variaciones en los ciclos de reemplazo de piezas, la estacionalidad de la demanda y los cambios en el comportamiento del consumidor generan una complejidad considerable en la planificación de inventarios y en la gestión de los talleres. Una predicción inadecuada puede provocar tanto en la falta de repuestos, que afecta directamente en la satisfacción del cliente como en un exceso de inventario que aumentan los costos operativos y reduce la rentabilidad (Cedeño, 2025).

La literatura académica ha destacado que la predicción de la demanda en la industria automotriz plantea desafíos adicionales en comparación con otras industrias. A diferencia de los productos de consumo masivo, los repuestos de vehículos tienen ciclos de vida heterogéneo: mientras algunos se reemplazan con frecuencia (por ejemplo, filtros o frenos), otros pueden tener una vida útil prolongada y su demanda es más difícil de anticipar (Weizhi et al., 2020). Del mismo modo, la fidelidad de los clientes a los talleres oficiales disminuye con la antigüedad del vehículo, generando una dispersión de la demanda hacia talleres independientes (García González, 2015). Todo eso refuerza la necesidad de contar con modelos predictivos que integren grandes cantidades de datos y sean capaces de identificar patrones en entornos complejos.

El aprendizaje automático (machine learning) se ha convertido en una poderosa herramienta para abordar problemas de la predicción de ventas y la planificación de la demanda según (Silva, 2024). A diferencia de los métodos estadísticos tradicionales, que se basan en gran medida en supuestos lineales y de series históricas estables, los algoritmos de aprendizaje automático permiten procesar grandes volúmenes de datos, identificar relaciones no lineales y adaptarse a entornos dinámicos y cambiantes (Zheng, 2025).

Dentro de los algoritmos está Random Forest que se destaca por su capacidad de generalizar, lo que lo hace especialmente útil en contextos con datos grandes, heterogéneos o incompletos. La investigación de (Malik et al., 2024) han demostrado su utilidad para pronosticar las ventas en un entorno minorista, obteniendo mejoras de precisión del 25 % en comparación con los modelos de regresión lineal. La eficacia del Random Forest en la predicción de ventas ha sido ampliamente documentada en diferentes industrias. Por ejemplo, en la industria alimentaria, se ha demostrado que este algoritmo supera a modelos como la regresión lineal o el soporte vectorial en precisión y robustez (Naik et al., 2022). En este sentido, el método Random Forest es ampliamente reconocido por su capacidad para manejar datos de alta dimensión, detectar interacciones no lineales entre variables y mitigar el sobreajuste. Investigaciones como las de Liang (2024) y Yadav (2023) destacan el uso de Random Forest en escenarios de predicción de ventas con mejores resultados en comparación a métodos tradicionales, tanto en términos de precisión como de resistencia al sesgo y al ruido en los datos.

En Ecuador, el análisis de la necesidad de repuestos y servicios en talleres automotrices es muy importante, particularmente en la Corporación MQ, empresa automotriz ecuatoriana fundada en

1926, que se dedica a la venta de vehículos nuevos y usados, venta de repuestos y servicio de taller de las marcas Chevrolet y Multimarca (todas las marcas). El parque vehicular nacional crece de manera sostenida y muestra una diversidad en antigüedad y origen de unidades, lo que genera necesidades de mantenimiento preventivo y reparaciones derivadas de desgaste y siniestros (colisiones). En este escenario, los talleres oficiales enfrentan una demanda variable, marcada por la competencia de servicios independientes y las crecientes expectativas de los clientes respecto a agilidad, transparencia y calidad, todos estos aspectos orientados a la experiencia del usuario.

El periodo de análisis de esta investigación comprende los datos históricos del año 2024, se aplica el modelo de predicción Random Forest en Corporación MQ para anticipar el volumen de ventas en servicios de posventa (Talleres y Repuestos) con mayor precisión a partir de datos de variables como línea de negocio, clase de cliente, entre otros, la aplicación de modelos de predicción basados en Random Forest ofrecen ventajas como anticipar el volumen de ventas de repuestos con mayor precisión, ajustando los inventarios y evitando pérdidas por desabastecimiento o sobre stock. Los resultados de la predicción proporcionan información estratégica para la planificación de recursos en los talleres, como la asignación de personal técnico, la gestión de citas y la optimización de procesos de atención al cliente. Todo ello contribuye no solo a mejorar la rentabilidad de la organización, sino también a fortalecer la fidelización mediante campañas y la experiencia del cliente.

El estudio está destinado a quienes toman decisiones como, jefes de taller, analistas de inventarios, analistas comerciales y directivos de la Corporación MQ, la investigación se justifica por la importancia económica del negocio de posventa y la necesidad de modelos de gestión innovadores que apoyan la sostenibilidad de la industria automotriz del Ecuador.

Una revisión de investigaciones recientes confirma que la predicción de la demanda de repuestos y servicios en la industria automotriz se ha beneficiado de enfoques híbridos que combinan algoritmos de machine learning. Por ejemplo, Zheng (2025) presenta un modelo que integra Random Forest con AdaBoost y ARMA, obteniendo mejoras significativas en la precisión de los pronósticos y en la detección temprana de riesgos en la cadena de suministro. De igual forma, Liao et al. (2020). plantean un modelo de aprendizaje automático diseñado para datos pequeños y ciclos de sustitución largos, como los que caracterizan a determinados repuestos de vehículos,

demostrando que el enfoque es capaz de producir predicciones confiables incluso en condiciones de información limitada o con sesgos.

La capacidad de predecir el comportamiento de la demanda, basándose en el análisis de datos históricos y en la identificación de patrones poco visibles, se convierte en un factor crítico para la competitividad y el desarrollo comercial. De esta manera, se pretende no solo optimizar los resultados de la Corporación MQ, sino también crear un referente de innovación que pueda ser aplicado a otras organizaciones de la sector automotriz en Ecuador y la región.

## **2. MATERIALES Y MÉTODOS**

La investigación se basa en un enfoque cuantitativo y experimental, respaldado por la recopilación y el análisis estructurado de los datos numéricos provenientes de registros de historial de ventas, para evaluar el desempeño de un modelo predictivo basado en aprendizaje automático. El diseño experimental fue elegido por su capacidad de controlar las condiciones de entrenamiento y prueba del modelo, garantizando objetividad, aplicabilidad y una comparación válida de los escenarios de predicción en corporación MQ.

El análisis descriptivo de los datos se realizó para comprender el comportamiento general de las ventas durante el año 2024 en Talleres de la Corporación MQ, este análisis permitió identificar la distribución de las ventas según variables independientes como línea de negocio, clase de cliente, bodega, canal y zona, así como identificar patrones, tendencias, desviaciones para el proceso de entrenamiento.

La metodología garantiza validez interna y confiabilidad a través de un proceso ordenado que incluye, limpieza rigurosa de datos mediante la preparación y depuración del conjunto de datos, eliminando registros inconsistentes, incompletos, duplicados, nulos, atípicos como ventas con valores irrelevantes como por ejemplo 0.01 centavo o ventas muy grandes y puntuales para reducir sesgos y errores que pueden afectar el rendimiento del modelo. Las variables categóricas se codificaron numéricamente para el procesamiento del modelo mediante el computador y el conjunto de datos se dividió aleatoriamente en un conjunto de entrenamiento (80 %) y un conjunto de pruebas (20 %).

Se utilizó el modelo de pronóstico de Random Forest, entrenado y preparado a partir de los datos históricos de ventas, para estimar el volumen de ventas en los talleres de posventa de corporación

MQ. El entrenamiento del modelo se realizó bajo un diseño de prueba controlada, que incluyó técnicas de validación cruzada para minimizar el sobreajuste y garantizar que el modelo generalice adecuadamente a datos no vistos.

El funcionamiento del modelo Random Forest fue evaluado mediante medidas estadísticas confiables, que permitió medir la precisión y efectividad del modelo en la predicción del volumen de ventas en los talleres de corporación MQ. Esta evaluación fue clave para comprobar que el modelo no solo se ajusta a los datos de entrenamiento, sino que también es capaz de generar estimaciones coherentes y cercanas a la realidad, fortaleciendo la validez interna del estudio, así como la confiabilidad y la posibilidad de replicar los resultados en ámbitos similares.

Para el procesamiento y análisis de los datos se aplicó el lenguaje de programación Python, debido a su solidez y adaptabilidad en el manejo de grandes volúmenes de información, se usa ampliamente a nivel mundial en proyectos de machine learning y existen bibliotecas especializadas disponibles para la predicción.

El entrenamiento del modelo de predicción Random Forest se realizó directamente con librerías:

**Pandas:** Librería utilizada para la limpieza, transformación y manipulación de Datos, permite cargar datos, eliminar duplicados, corregir valores nulos, blancos y estructurar las columnas para el análisis.

**NumPy:** La Librería se emplea para cálculos numéricos y manejo de matrices, fundamentales para la preparación de variables que se entregan al modelo Random Forest.

**Scikit-learn:** Es la librería más importante para machine learning.

Se utilizó para:

- Preprocesamiento (codificación de variables categóricas),
- División en entrenamiento y prueba,
- Creación del modelo Random Forest,
- Ajuste de hiperparámetros,
- Evaluación del rendimiento ( $R^2$ , MSE, RMSE).

Es el núcleo del proceso predictivo.

**Matplotlib / Seaborn:** Se usó para el análisis exploratorio y visualización de datos, generando gráficos como:

Importancia de las variables, comparación de ventas reales versus pronosticadas.

Estas visualizaciones permitieron identificar tendencias y evaluar el comportamiento del modelo.

**Power BI:** Para la visualización de datos, generando gráficos como: ventas por línea de negocio, por canal, por bodega, por zonas y clase de cliente.

Para el análisis de los datos históricos de ventas de talleres, se continuo con la recopilación de información importante de los datos de ventas correspondientes al año 2024. Estos datos incluirán una serie de variables clave como la línea de negocio, bodega, zona geográfica, clase de cliente, totales de ventas, entre otros, además del volumen de ventas asociado a cada variable antes mencionadas. Se obtuvieron un total de 114937 registros originales, una vez recopilada la información, se inició el preprocesamiento de los datos, paso fundamental para asegurar su calidad y la fiabilidad. En este proceso, se corrigieron posibles errores o sesgos, valores nulos, en blanco, bodegas que ya no existen en la actualidad en la corporación MQ, registros duplicados o inconsistencias que puedan interferir con el análisis. Como resultado de esta depuración, el conjunto de datos quedo conformado por 87778 registros válidos. Además, las variables categóricas como la línea de negocio, la zona o la clase de cliente fueron transformados en un formato numérico para su análisis, como la codificación de los datos aplicando técnicas procesamiento de datos (Limpieza, organización, transformación) para garantizar la limpieza y la calidad de la información. Con los datos limpios, se llevó a cabo un análisis exploratorio de los mismos.

Con los datos ya preprocesados y las variables seleccionadas en el objetivo anterior para el entrenamiento del modelo predictivo basado en el algoritmo Random Forest. En primer instancia, se dividieron los datos en dos conjuntos: uno de entrenamiento y otro de prueba. Este paso fue esencial para poder evaluar posteriormente el rendimiento del modelo y garantizar su capacidad de extender a nuevos datos. Normalmente, el 80% de los datos se utilizaba para el conjunto de entrenamiento y el 20% restante para las pruebas.

Antes del entrenamiento de los datos mediante el algoritmo Random Forest, se realizó la estructuración del conjunto de metadatos que detalla cada una de las variables utilizadas para pa predicción.

En la Tabla 1 se presentan la descripción de cada variable, detallando su tipo de dato y formato de dato.

**Tabla 1**

**Metadatos de Variables**

METADATOS	NOMINAL	NOMINAL	NOMINAL	NOMINAL	NOMINAL	ESCALAR
	TEXTO	TEXTO	TEXTO	TEXTO	TEXTO	NUMERICO
	1 - ACCESORIOS 2 - DISPOSITIVOS 3 - REPUESTOS 4 - T.O.T 5 - TALLER	1 - AGENDAMIENTO 2 - ASEGURADORA 3 - FLOTAS 4 - RETAIL	1 - 01.1.3 - CUE ESPAÑA TAL 2 - 01.2.3 - CUE QUINTA CHICA TAL 3 - 03.1.3 - AZO IGNACIO NEIRA TAL 4 - 07.1.3 - MCH AV 25 DE JUNIO TAL 5 - 07.2.3 - MCH TERM. TERRESTRE TAL 6 - 09.1.3 - GYE AV AMERICAS TAL 7 - 09.2.3 - GYE JUAN TANCA TAL 8 - 09.3.3 - GYE DAULE TAL MEC 9 - 09.4.3 - GYE CALLE TERCERA TAL 10 - 09.5.3 - GYE CALLE SEXTA TAL 11 - 11.1.3 - LOJ ISIDRO AYORA TAL 12 - 11.2.3 - LOJ SALVADOR BUST TAL 13 - 17.1.3 - QTO GRANADOS TAL 14 - 17.13.3 - QTO NARANJOS TAL 15 - 17.2.3 - QTO SIMON BOLIVAR TAL 16 - 17.4.3 - QTO CAYAMBE TAL 17 - 17.7.3 - QTO RUSIA TAL	1 - ZONA 1 2 - ZONA 2 3 - ZONA 3	1 - ASEGURADORA 2 - CLIENTE 3 - CLIENTE / PROVEEDOR 4 - CONCESIONARIO 5 - DEMANDA Y EXTRAJUDICIALES 6 - EMPLEADO 7 - FLOTISTA 8 - INSTITUCION FINANCIERA 9 - INSTITUCION PUBLICA 10 - MARCA / IMPORTADORA 11 - MAYORISTAS 12 - MECANICAS 13 - PATIOS 14 - RELACIONADAS	Valor total Factura
No	Línea Negocio	Canal	Bodega	Zona	Clase Cliente	Total
1	5	4	1	3	14	1.01
2	1	1	1	3	7	1.02
3	5	1	1	3	3	1.09
4	3	4	1	3	2	1.10
5	5	4	1	3	14	1.13
6	5	4	1	3	14	1.28
7	1	2	1	3	1	1.29
8	5	4	1	3	14	1.32
9	5	4	1	3	14	1.33
10	5	4	1	3	14	1.36

**Fuente:** Elaboración propia

**Tabla 2**

Detalle de variables dependientes e independientes usados para la predicción

<b>Tipo de variable</b>	<b>Nombre de la variable</b>	<b>Descripción</b>	<b>Relevancia en el estudio</b>
<b>Variable dependiente</b>	<b>Volumen de ventas</b>	Representa la cantidad de ventas realizadas en los talleres de la Corporación MQ en un periodo determinado.	Es el resultado principal que se busca predecir mediante el modelo Random Forest.
<b>Variables independientes (predictoras)</b>	<b>Línea de Negocio</b>	Segmento o categoría del negocio (por ejemplo, repuestos, Talleres, accesorios).	Permite identificar patrones específicos de comportamiento de ventas.
	<b>Bodega</b>	Agencia, sucursal o lugar físico parte de corporación MQ donde se prestan servicios de Posventa	Influye en la disponibilidad de inventario y los tiempos de entrega, factores que afectan la demanda.
	<b>Canal</b>	Medio o vía mediante la cual se gestiona y ejecuta la venta	Facilita analizar el desempeño comercial
	<b>Zona</b>	Región geográfica donde se realiza la venta o prestación del servicio.	Permite analizar las diferencias de comportamiento del mercado entre regiones.
	<b>Clase de cliente</b>	Tipo o categoría del cliente (Retail, flota, corporativo, etc.).	Facilita la segmentación del mercado.

**Fuente:** Elaboración propia

Durante el entrenamiento se utilizó el algoritmo Random Forest, que es un conjunto de árboles de decisión, para que aprenda a pronosticar las ventas basándose en variables clave como la línea empresa, bodega, zona geográfica y tipo de cliente. Random Forest es un algoritmo potente, permitiendo el manejo de variables tanto continuas como categóricas, lo que lo hace adecuado para este tipo de predicción. Además, se ajustaron los hiperparámetros del modelo, el número de árboles, la profundidad máxima de cada árbol y el número mínimo de muestras para realizar particiones utilizando métodos como la validación cruzada. Esta configuración de los parámetros están destinados para mejorar la precisión del modelo y evitar problemas como el sobreajuste, que garantiza que el modelo no sólo se ajuste bien a los datos de entrenamiento, sino que también sea capaz de predecir correctamente con los datos reales. Después de entrenar el modelo, se evaluó utilizando el conjunto de prueba. Las medidas utilizadas, como el error cuadrático medio (MSE), la raíz del error cuadrático medio (RMSE) y coeficiente de determinación ( $R^2$ ) para evaluar su desempeño. Estas medidas permitieron medir qué tan cerca están las predicciones del modelo de los valores de ventas reales, lo que proporciona una medida precisa del poder predictivo del modelo. Se analizó los resultados en detalle para identificar áreas potenciales de mejora, y realizar correcciones de hiperparámetros o preprocesamiento de datos si fuera necesario. Uno los objetivos se centran en entrenar y evaluar un modelo predictivo robusto y confiable para predecir las ventas en los talleres de posventa de Corporación MQ, para facilitar la toma de decisiones y la planificación estratégica.

Una vez entrenado y evaluado el modelo predictivo, el siguiente paso es aplicar el modelo a diferentes Escenarios para predecir ventas basándose en variables independientes se definieron tres escenarios basados en combinaciones de variables del giro de negocio como son, canal de ventas, bodega, zona y clase de cliente. Estos escenarios permitieron examinar cómo diferentes condiciones afectan las ventas, que son de gran utilidad para la empresa a la hora de considerar diferentes contextos comerciales.

Los escenarios seleccionados reflejan situaciones comunes dentro del negocio y permiten analizar el comportamiento del modelo en condiciones reales. Entre ellos se incluyen los siguientes escenarios:

- 1) La empresa aplica una campaña de marketing a Clientes enfocados en la línea de negocio de Accesorios, con clientes provenientes de aseguradoras (Zona 3, Bodega España TAL).

Este escenario permite evaluar la capacidad del modelo para estimar la demanda generada por siniestros y atenciones gestionadas a través de canales aseguradoras, con el objetivo de analizar relaciones comerciales y crear nuevas alianzas para mejorar ingresos en este rubro.

- 2) La empresa se orienta a Repuestos y clientes que ingresan por Agendamiento (Zona 3, Bodega España TAL).

Este caso representa una situación frecuente en la operación, donde el modelo debe predecir volúmenes asociados al mantenimiento programado y a clientes particulares del día a día que adquieren repuestos (cliente normal).

- 3) La empresa decide enfocarse a servicios de Taller con clientes normales que ingresan vía Agendamiento (Zona 3, Bodega España TAL).

Este escenario permite evaluar el desempeño del modelo en servicios de mano de obra, donde la variabilidad puede ser mayor debido a la naturaleza del trabajo en Taller.

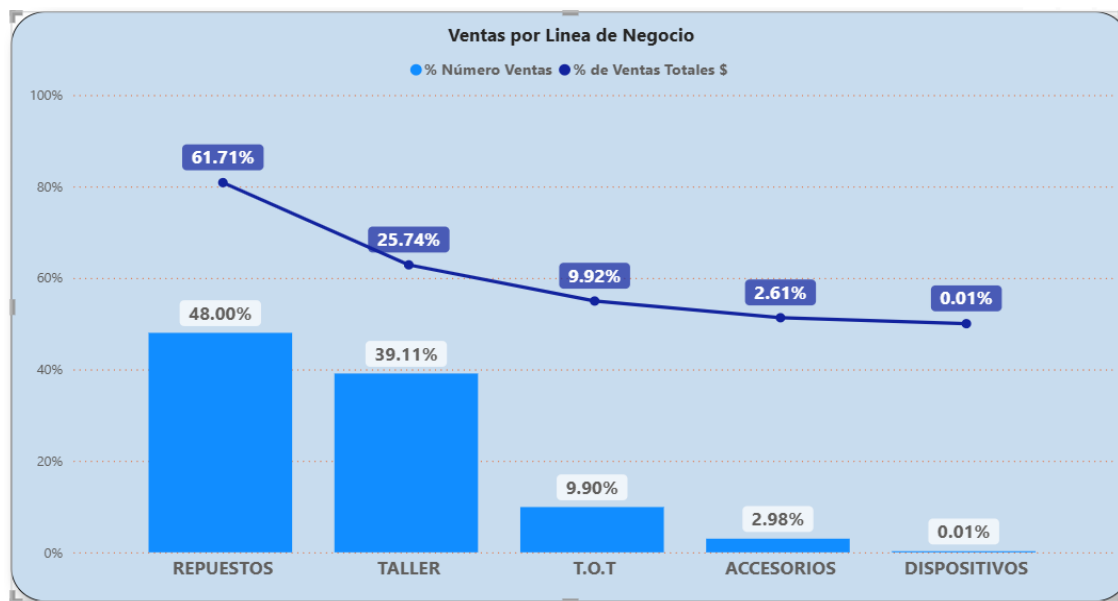
### 3. RESULTADOS

#### a) Análisis Descriptivo:

En esta sección se presentan gráficos descriptivos de la variables dependientes e independientes, lo que permite observar tendencias generales y posibles patrones en los datos.

#### Figura 1

Ventas por La línea de Negocio de Corporación MQ, número de facturas y monto de ventas (valores representados en %)



**Fuente:** Elaboración propia con base en la Base de Datos de Corporación MQ, año 2024

La figura 1 de ventas por la línea de negocio evidencia una clara concentración en Repuestos y Taller, mientras que las demás líneas (Trabajos otros Talleres, Accesorios y Dispositivos) contribuyen con muy poco margen tanto en número de facturas como en valor monetario. Este hallazgo refleja una dependencia significativa de dos segmentos principales, lo cual plantea oportunidades estratégicas para diversificar la oferta o, por el contrario, focalizar la innovación digital en los procesos de postventa y servicio técnico asociados a estas dos áreas críticas.

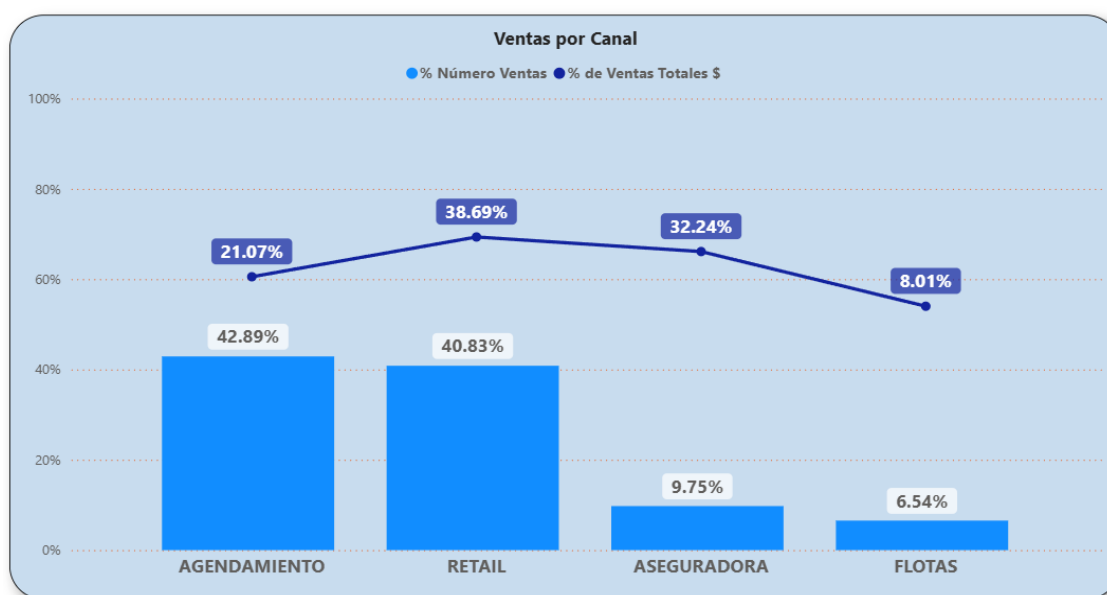
Además, la baja participación de líneas como Accesorios y Dispositivos revela un potencial de crecimiento que podría aprovecharse mediante campañas de marketing digital más segmentadas y el uso de plataformas digitales (Facebook, Instagram, entre otras).

La concentración de ingresos en pocos segmentos también implica un riesgo operativo: cualquier disrupción en la cadena de suministros de repuestos o en la operación de los talleres afectaría de manera directa y significativa los resultados globales.

Por ello, en un contexto de transformación digital, es indispensable explorar la automatización de inventarios, el uso de analítica predictiva y la integración de canales digitales, como medios para optimizar la gestión de estas líneas críticas y, al mismo tiempo, fortalecer las de menor participación.

## Figura 2

Ventas por canal de Corporación MQ, número de facturas y monto de ventas (valores representados en %)



**Fuente:** Elaboración propia con base en la Base de Datos de Corporación MQ, año 2024

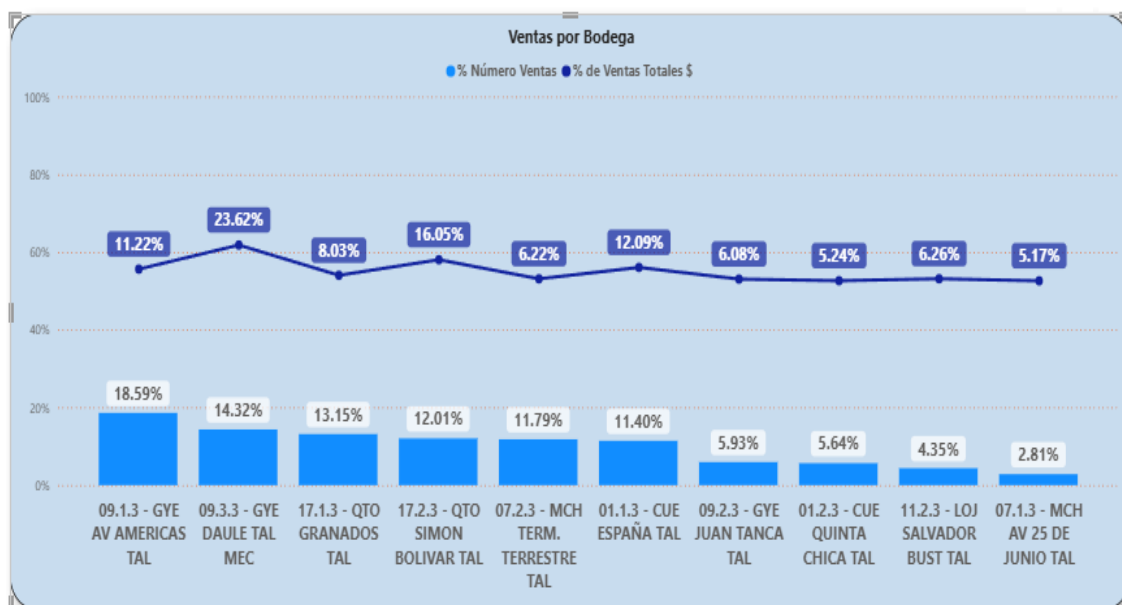
La figura 2 de ventas por canal para el año 2024 nos permite comprender, más allá de las cifras, la historia que hay detrás de los hábitos y preferencias de los clientes. Los datos reflejan que el Agendamiento y el Retail concentran juntos más del 80% de las ventas, lo que evidencia que los usuarios valoran tanto la comodidad de planificar sus compras como la confianza en la atención presencial. Este comportamiento muestra una transición interesante: los clientes han aprendido a combinar la practicidad de lo digital con la calidez del contacto humano.

No obstante, los canales de Aseguradora y Flotas muestran una participación considerablemente menor. Lejos de considerarse una debilidad, este escenario representa una valiosa oportunidad para impulsar la innovación. Potenciar estos canales mediante alianzas estratégicas con aseguradoras, programas de descuentos dirigidos a Flotas y la implementación de soluciones digitales que simplifiquen los procesos y mejoren la experiencia del cliente, no solo permitiría dinamizar su crecimiento, sino también diversificar las fuentes de ingreso y reducir la dependencia de los canales principales. Esta estrategia abriría el camino hacia un modelo de negocio más equilibrado, inclusivo y adaptable a las nuevas demandas del mercado.

La Transformación Digital no se trata únicamente de incorporar nuevas tecnologías; se trata de poner a las personas en el centro: entender cómo compran, qué valoran y qué barreras enfrentan en su interacción con los distintos canales. Este análisis nos recuerda que la digitalización debe ir acompañada de un enfoque humano, capaz de conectar con las necesidades y emociones de los usuarios.

**Figura 3**

Ventas por Bodega de Corporación MQ, numero de facturas y monto de ventas (valores representados en %)



**Fuente:** Elaboración propia con base en la Base de Datos de Corporación MQ, año 2024

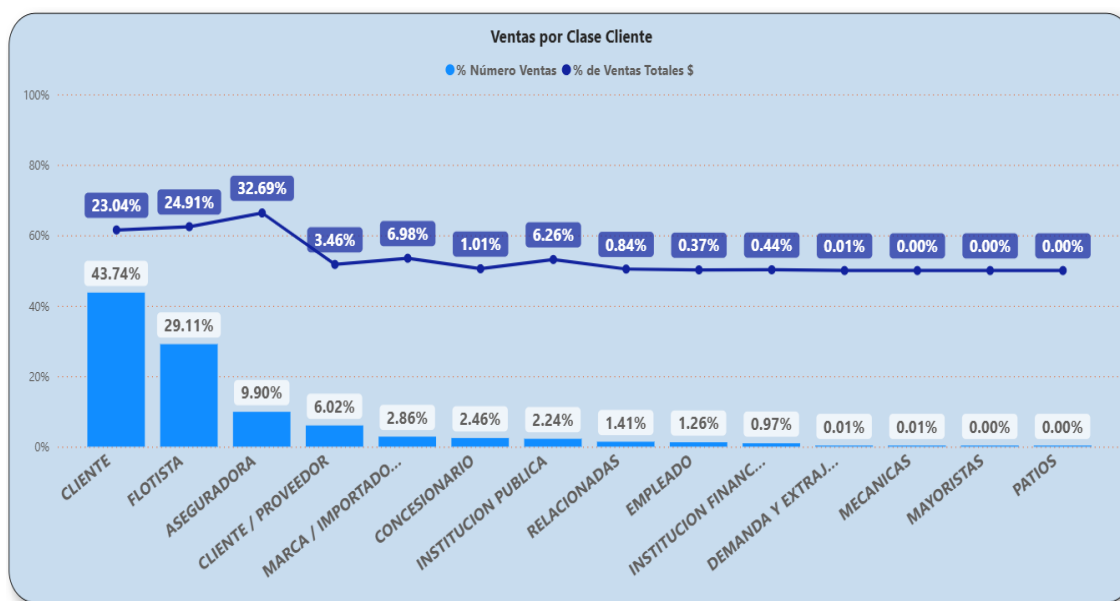
La figura 3 de ventas por bodega del año 2024 se visualiza la 10 principales con mayor porcentaje de ventas (numero, valor), evidenciando que el volumen de ventas está altamente concentrado en pocas sedes, siendo las seis principales responsables de cerca del 60% del total. Este patrón indica que algunas bodegas han logrado posicionarse como puntos estratégicos de alto rendimiento o a su vez son las bodegas Matrices de cada ciudad, mientras que otras mantienen una participación mucho menor, lo que genera un amplio margen para el crecimiento.

Esta brecha entre bodegas líderes y las de menor volumen resalta la necesidad de aprovechar la transformación digital como un motor de cambio. La implementación de herramientas tecnológicas como analítica avanzada de datos para entender la demanda local, plataformas digitales de atención al cliente y estrategias de marketing segmentadas puede ayudar a optimizar procesos, mejorar la eficiencia operativa y aumentar las ventas en las bodegas rezagadas.

Más que un desafío, esta distribución desigual representa una oportunidad estratégica: al fortalecer digitalmente las bodegas con menor participación, la empresa no solo equilibrará su red de ventas, sino que incrementará el volumen global, reducirá la dependencia de las sedes líderes y ofrecerá una experiencia más ágil y consistente a todos los clientes.

**Figura 4**

Ventas por Clase Cliente de Corporación MQ, numero de facturas y monto de ventas (valores representados en %)



**Fuente:** Elaboración propia con base en la Base de Datos de Corporación MQ, año 2024

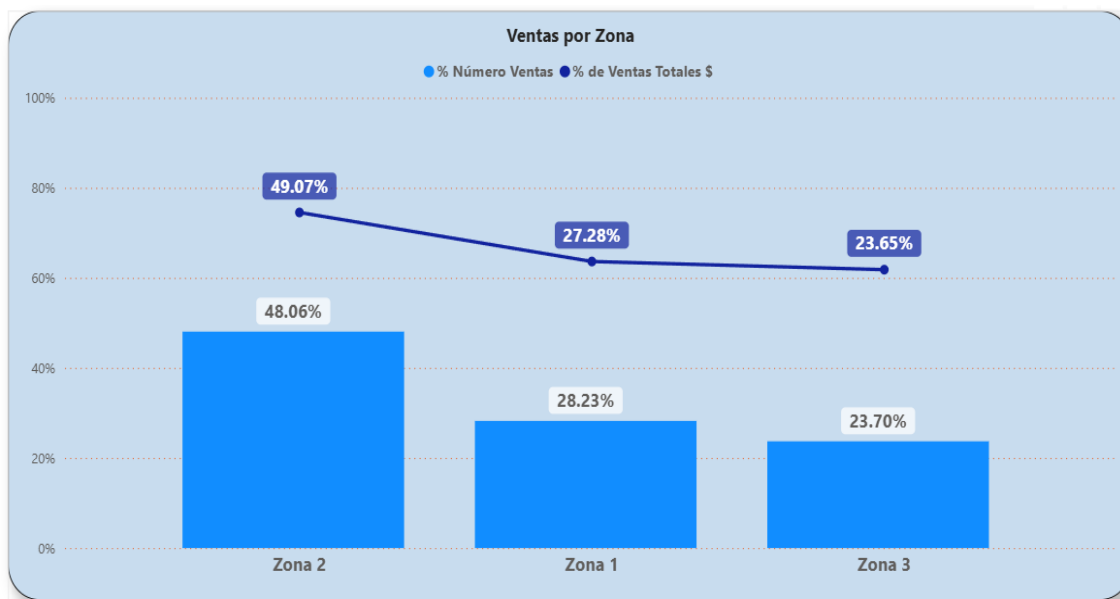
La figura 4 de ventas por clase de cliente para 2024 muestra una alta concentración del volumen de ventas en solo dos segmentos principales: Clientes individuales y Flotas, que juntos representan más del 70% de las transacciones. Esta concentración evidencia que el negocio depende en gran medida de estos dos tipos de clientes, mientras que otros segmentos presentan un volumen significativamente menor, y el resto mantiene participaciones marginales.

Este escenario, aunque destaca el buen rendimiento de los segmentos principales, revela oportunidades claras de crecimiento en los clientes con baja participación. La Transformación Digital juega un papel clave aquí: la adopción de soluciones tecnológicas como plataformas de autoservicio, integraciones B2B, sistemas de analítica avanzada para segmentar ofertas y estrategias de marketing digital personalizado puede dinamizar estos segmentos menos desarrollados, mejorar la experiencia de compra y fortalecer las relaciones comerciales.

Aprovechar estas herramientas permitirá diversificar el portafolio de clientes, disminuir el riesgo de depender de pocos segmentos y aumentar el volumen total de ventas. En otras palabras, la digitalización no solo optimizaría los canales actuales, sino que también abriría nuevas oportunidades de mercado para aquellos clientes que hoy representan una participación menor.

**Figura 5**

Ventas por zona de Corporación MQ, número de facturas y monto de ventas (valores representados en %)



**Fuente:** Elaboración propia con base en la Base de Datos de Corporación MQ, año 2024

La figura 5 de ventas por zonas para el año 2024 revela una clara concentración del volumen en la Zona 2, mientras que la Zona 1 y la Zona 3 está ligeramente distante. Este comportamiento indica que casi la mitad del negocio depende de una sola región, lo que puede generar riesgos si esta zona enfrenta cambios en la demanda o en las condiciones del mercado.

La diferencia en los volúmenes entre las zonas pone en evidencia áreas de oportunidad para impulsar el crecimiento en las regiones con menor participación. La Transformación Digital puede ser el motor que equilibre esta disparidad, mediante el uso de herramientas de analítica de datos georreferenciados, campañas de marketing digital localizadas, plataformas de ventas en línea y optimización de la logística y distribución para mejorar la cobertura y el acceso en las zonas menos fuertes.

Más allá de solo incrementar las cifras de ventas, estas estrategias permitirán reducir la dependencia de la Zona 2, aumentando la resiliencia del negocio y creando un ecosistema de ventas más balanceado, competitivo y sostenible.

En conclusión, el gráfico destaca que el verdadero desafío está en aprovechar la tecnología y los datos para potenciar el desempeño de las Zonas 1 y 3, cerrando las brechas actuales y maximizando el potencial de cada región para impulsar el crecimiento global de la organización.

### b) Pronóstico con Random Forest

A continuación, se visualiza la predicción del volumen de ventas mediante la aplicación de Random Forest, los datos son cargados en Python.

**Tabla 3**

Muestra los datos cargados en Python

Línea Negocio	Canal	Bodega	Zona	Clase Cliente	Total, venta
5	4	1	3	14	1.01
1	1	1	3	7	1.02
5	1	1	3	3	1.09
3	4	1	3	2	1.1

**Fuente:** Elaboración propia

Luego del análisis respectivo, se aplicó el algoritmo **Random Forest**, configurado con los hiperparámetros que permitieron alcanzar el mayor nivel de precisión en el proceso de aprendizaje del modelo, mismo que se muestran la tabla 4.

**Tabla 4**

Configuración del modelo Random Forest Regressor y Hiperparámetros de entrenamiento

Hiperparámetros	Valor asignado
Número de árboles ( <i>n_estimators</i> )	200
Profundidad máxima ( <i>max_depth</i> )	None (sin límite definido)
Mínimo de muestras para dividir un nodo ( <i>min_samples_split</i> )	2
Mínimo de muestras en una hoja ( <i>min_samples_leaf</i> )	1
Tamaño del conjunto de entrenamiento	(70 222, 5)
Tamaño del conjunto de prueba	(17 556, 5)
Modelo implementado	RandomForestRegressor( <i>n_estimators</i> =200, <i>random_state</i> =42)

**Fuente:** Elaboración propia a partir del proceso de modelado y ajuste de hiperparámetros

Como resultado, el modelo alcanzó un Coeficiente de Determinación ( $R^2$ ) de 0.7769, valor que evidencia un buen nivel de ajuste y capacidad explicativa del modelo sobre la variabilidad del volumen de ventas predicho.

También se muestra el fragmento del código desarrollado en Python para la captura de los datos de entrada del usuario, así como la ejecución del modelo Random Forest entrenado. En esta etapa, el sistema solicita los valores correspondientes a las variables independientes y genera automáticamente el valor de venta.

**Figura 6**

Fragmento código Python ingreso de datos

```

print("=" * 50)
print("PREDICCIÓN DE TOTAL - RANDOM FOREST")
print("=" * 50)

#DATOS DE ENTRADA DEL USUARIO
linea_negocio = float(input("Ingrese Línea de Negocio: "))
canal = float(input("Ingrese Canal: "))
bodega = float(input("Ingrese Bodega: "))
zona = float(input("Ingrese Zona: "))
clase_cliente = float(input("Ingrese Clase de Cliente: "))

```

**Fuente:** Elaboración Propia

Una vez ingresados los datos por el usuario a través del teclado, estos son procesados y organizados en una estructura denominada `nuevo_dato`, la cual agrupa de manera ordenada los valores correspondientes a las variables independientes del modelo. Posteriormente, dicha estructura es utilizada como entrada para el modelo entrenado Random Forest Regressor, mediante la función `model.predict(nuevo_dato)`, con el fin de generar la estimación del volumen de ventas. Este procedimiento permite aplicar el modelo predictivo de forma dinámica, simulando distintos escenarios de negocio y obteniendo resultados inmediatos a partir de los parámetros proporcionados por el usuario.

**Figura 7**

Código Python de procesamiento de datos

```

# Realizar la predicción
prediccion = model.predict(nuevo_dato)
# Mostrar el resultado
print("\n" + "=" * 50)
print("RESULTADO DE LA PREDICCIÓN")
print("=" * 50)
print(f"Línea de Negocio: {linea_negocio}")
print(f"Canal: {canal}")
print(f"Bodega: {bodega}")
print(f"Zona: {zona}")
print(f"Clase de Cliente: {clase_cliente}")
print(f"\n✓ Total Pronosticado: ${prediccion[0]:.2f}")
print("=" * 50)

```

**Fuente:** Elaboración Propia

A partir del modelo entrenado, se obtuvo el nivel de importancia de cada variable independiente dentro del proceso de predicción. Esta medida refleja la contribución relativa de cada atributo en

la estimación del volumen de ventas, es decir, cuánto influye cada variable independiente en el resultado final de predicción del modelo.

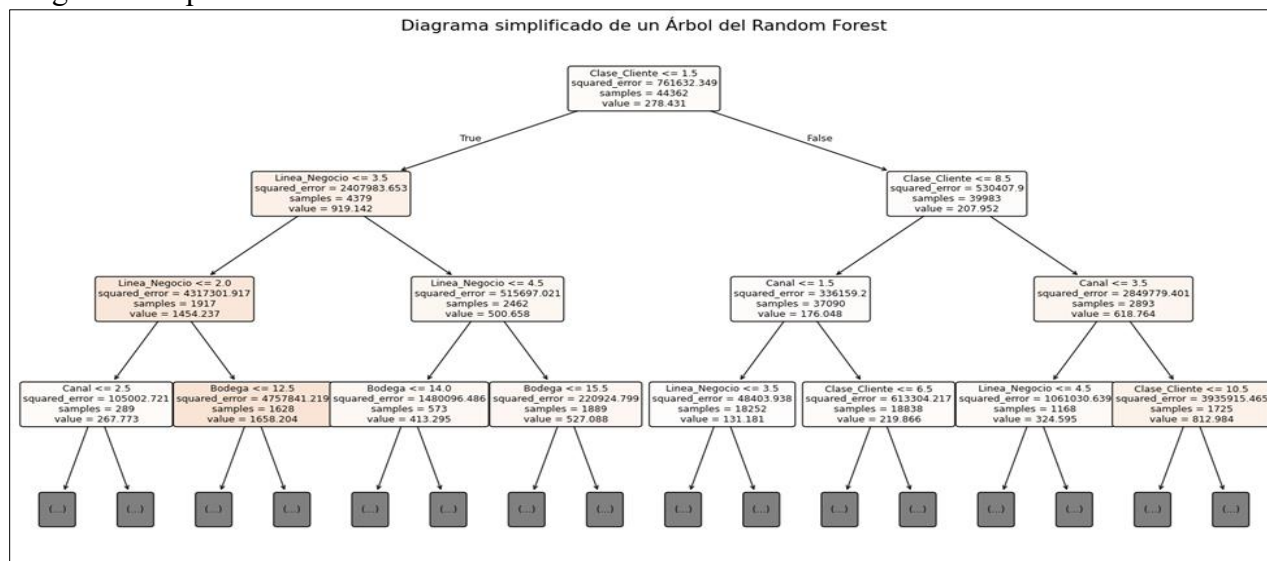
**Tabla 5**  
Importancia de las variables en el modelo Random Forest

Variable	Importancia (%)
Clase de cliente	38,44 %
Línea de negocio	29,53 %
Bodega	20,48 %
Canal	10,62 %
Zona	0,93 %

**Fuente:** Elaboración propia a partir de los resultados del modelo Random Forest

A continuación, la figura 8 se presenta una visualización parcial del diagrama del modelo Random Forest, en la cual se muestran únicamente los primeros niveles del árbol de decisión. Esto permite apreciar la estructura del modelado desde su raíz hasta las primeras divisiones, evidenciando la forma en que el algoritmo realiza las segmentaciones iniciales de los datos para establecer los criterios predictivos del volumen de ventas. En el árbol inicia con la información de las ventas por clase de cliente y va haciendo preguntas para ir dividiendo los datos en grupos más pequeños donde ya va involucrando a las demás variables con una profundidad de hasta 200 árboles.

**Figura 8**  
Diagrama simplificado de un árbol del Random Forest

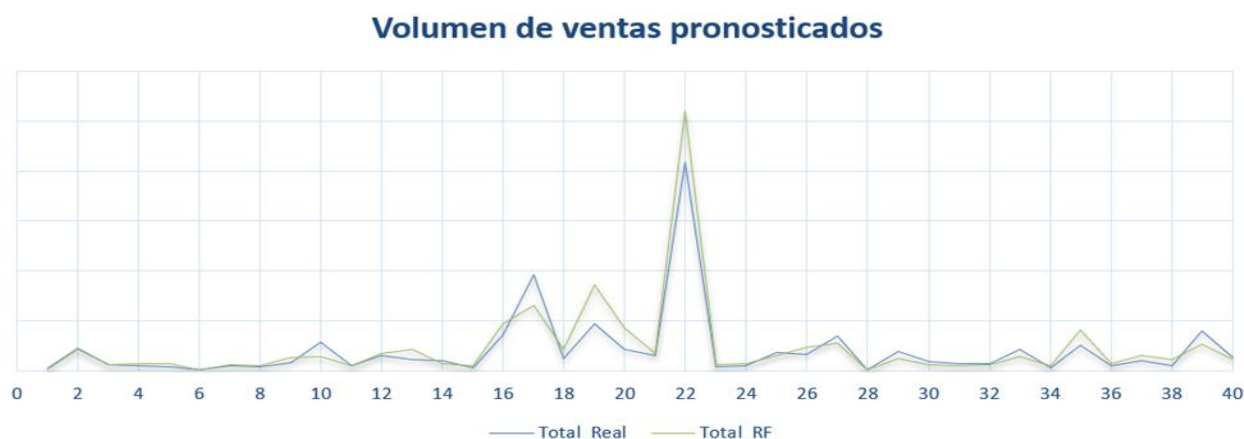


**Fuente:** Elaboración Propia

También, se realizó un gráfico figura 9 para poder comparar a lo largo del tiempo las ventas reales versus las predecidas en la mayoría de los periodos se mantienen muy cerca, lo que significa que el modelo asimila las ventas reales y reconoce cuando las ventas suben o bajan. La cercanía de ambas líneas evidencia la capacidad predictiva del modelo y su precisión en la estimación del volumen de ventas en los talleres de corporación MQ analizados.

**Figura 9**

Volumen de Ventas Pronosticados



**Fuente:** Elaboración Propia

Para finalizar, se presenta la tabla 6 donde se puede visualizar el total de ventas pronosticadas versus las reales realizadas por el modelo Random Forest aplicando varias combinaciones entre las variables independientes (Línea de negocio, canal, bodega, zona y tipo de cliente). Esta tabla demuestra que el modelo es una herramienta confiable por la cercanía de los valores predecidos con los reales, por lo tanto, ayuda enormemente en la planificación y toma de decisiones en los talleres de corporación MQ.

**Tabla 6**

Escenarios para comprobar aprendizaje del modelo Random Forest

1=Accesorios 3=Repuestos 5=Taller	1=Agendamiento 2=Aseguradora	1=01.1.3 Cue España Tal	3= Zona3	1=Aseguradora, 2=Cliente Normal		
Línea Negocio	Canal	Bodega	Zona	Clase Cliente	Total, Venta Real	Total, Venta RF
1	2	1	3	1	227	219
3	1	1	3	2	65	61
3	1	1	3	2	53	53
5	1	1	3	2	63	68

**Fuente:** Elaboración Propia

#### 4. DISCUSIÓN

El modelo Random Forest aplicado para la predicción del volumen de ventas en el área de posventa en corporación MQ alcanzó un coeficiente de determinación de  $R^2 = 0.7769$  (77.69%), el modelo logra predecir con precisión gran parte de los cambios observados en las ventas reales y un aprendizaje efectivo a partir de las variables como Clase de Cliente, Línea de Negocio, Bodega, Canal y Zona. Estos resultados reflejan un desempeño competitivo y superan o se alinean con los métodos tradicionales de pronóstico (como regresión lineal, promedios móviles o suavizamiento exponencial) y que coinciden con lo que muestran estudios recientes sobre la utilización de algoritmos de aprendizaje automático en la gestión de demanda, los cuales han comprobado que los algoritmos de inteligencia artificial como Random Forest, ofrecen mejores predicciones.

Este hallazgo respalda la postura de diversos autores que sostienen que los algoritmos de aprendizaje automático superan a los métodos tradicionales de pronóstico cuando se enfrentan a relaciones no lineales y patrones complejos. En este sentido, los resultados obtenidos coinciden con los planteamientos de Huang et al. (2018) desarrollaron un modelo de Incremental Random Forest para servicios O2O y demostraron que la selección de variables relevantes reduce el error medio absoluto porcentual (MAPE) en un 11.64 %, y que la actualización incremental del modelo mejora la precisión en un 3.1 % adicional, confirmando la superioridad del Random Forest sobre métodos convencionales para series temporales con alta fluctuación. Los resultados de nuestro modelo, con una efectividad del 77.69 %, son consistentes con esa mejora en exactitud atribuida a la capacidad del Random Forest puede trabajar eficazmente con datos que no siguen una relación lineal ni ordenada (por ejemplo, cuando las ventas no crecen o bajan de forma constante), además, al ser un conjunto de muchos árboles de decisión, tolera el “ruido” en los datos es decir, valores atípicos, errores o fluctuaciones inesperadas sin que su precisión se deteriore, porque promedia múltiples resultados y así reduce el impacto de las anomalías.

Asimismo, los resultados refuerzan lo señalado por Yadav et al. (2023) comprobaron que el algoritmo Random Forest superó a las redes neuronales y la regresión lineal en precisión y estabilidad para pronósticos de ventas minoristas, destacando su resistencia al sobreajuste y su facilidad de interpretación. De manera similar, en este estudio el Random Forest mostró una clara

capacidad para explicar la importancia de las variables predictoras, siendo Clase de Cliente la más influyente (38.44 %), lo que evidencia la utilidad del modelo para decisiones comerciales basadas en segmentación y comportamiento del consumidor final.

De igual forma, los resultados obtenidos guardan coherencia con los hallazgos de Izquierdo y San Martín (2024) aplicaron Random Forest y regresión Lasso para pronosticar la demanda de autopartes en una cadena de 4000 tiendas en Norteamérica, logrando reducir pérdidas por sobre stock y ventas no realizadas. Concluyeron que los modelos basados en machine learning son más eficientes para identificar productos que se comportan de forma inusual, es decir, aquellos cuya demanda sube o baja de manera inesperada o diferente al resto. Este hallazgo coincide con la capacidad de nuestro modelo para reconocer patrones de venta diferenciados según las variables independientes analizadas, se puede evidenciar que este modelo aprende patrones complejos en los datos históricos y pueden reconocer automáticamente cuándo un producto se desvía del comportamiento normal, ayudando a ajustar inventarios y evitar pérdidas.

A nivel teórico, los resultados también se alinean con la revisión sistemática de Douaioui et al. (2024) identificó que los modelos de inteligencia artificial, especialmente Random Forest y XGBoost, se encuentran entre los más precisos en la predicción de demanda en cadenas de suministro, al adaptarse dinámicamente a los cambios del mercado y reducir la incertidumbre logística. En esa línea, la robustez del Random Forest en este trabajo refleja su aplicabilidad en contextos industriales reales, como la Posventa de corporación MQ, donde la demanda es irregular y depende de múltiples factores operativos (Políticos, sociales, etc.).

No obstante, algunos autores como Hasani (2023) comparó nueve métodos de pronóstico de repuestos industriales y determinó que, aunque modelos de redes neuronales (LSTM o MLP) pueden mejorar la precisión en demandas muy intermitentes, el Random Forest mantiene una relación más equilibrada entre exactitud y control del inventario, lo que lo hace recomendable para entornos empresariales con gran volumen de datos y múltiples variables categóricas.

Todos los estudios revisados anteriormente de las diferentes citas bibliográficas, avalan que el modelo de predicción aplicada en esta investigación funciona de manera sólida y confiable. Los resultados obtenidos son muy parecidos a los alcanzados en investigaciones internacionales, lo que demuestra que el modelo Random Forest es una herramienta eficaz para mejorar la planificación

de ventas y la gestión posventa en los talleres Corporación MQ, ayudando a tomar decisiones más precisas mejorando la rentabilidad del negocio.

## 5. CONCLUSIONES

Con la implementación del modelo Random Forest que permite predecir las ventas, el resultado obtenido fue muy bueno se logró una precisión del 77,69 % en el volumen de ventas en los talleres de posventa de la Corporación MQ del año 2024, demostrando que la inteligencia artificial puede convertirse en una herramienta importante y estratégica para optimizar la planificación operativa y comercial dentro de la industria automotriz.

Entre los hallazgos clave mostraron que variables independientes como Clase de Cliente y Línea de Negocio influyen directamente en el comportamiento de las ventas analizado, lo que sugiere la necesidad de fortalecer la segmentación del mercado y las estrategias de atención personalizada a los diferentes clientes que predominan. Este análisis indica que la empresa concentra gran parte de sus ingresos en pocas líneas de negocio y zonas del país, lo si bien refleja eficiencia operativa, también aumenta la dependencia de segmentos específicos, una desventaja estructural que podría mitigarse mediante la diversificación de productos y canales digitales.

Como fortaleza destacada de la Corporación MQ es la disponibilidad de datos históricos estructurados de las ventas en talleres del año 2024, lo que permitió entrenar un modelo potente y confiable. Sin embargo, el estudio presentó ciertas limitaciones en los datos estadísticos facilitados por la empresa relacionadas con la ausencia de variables externas como factores económicos, ventas en ciertas épocas del año, promociones por temporada. La inclusión de estos elementos en investigaciones futuras podría aportar una visión más completa del fenómeno analizado y mejorar la precisión de las predicciones.

El modelo demostró importantes ventajas sobre métodos estadísticos tradicionales en el manejo de datos no lineales, con valores atípicos y múltiples variables categóricas sin perder estabilidad ni interpretabilidad. Esto demuestra que el modelo Random Forest puede ser una opción viable para empresas que buscan modernizar su gestión de posventa y adoptar soluciones basadas en datos utilizando Inteligencia artificial para toma de decisiones.

El proyecto también encontró que en Corporación MQ tiene una alta posibilidad para avanzar hacia un modelo de transformación digital inteligente ya que cuenta con una gran base de datos en la nube de Azure, donde el análisis predictivo puede apoyar decisiones importantes en las áreas de

inventario, gestión operativa de talleres y campañas comerciales de posventa. Sin embargo, el éxito de esta implementación dependerá de capacitar al personal de Tecnología y directores en herramientas analíticas y fortalecer una cultura organizacional enfocada en el uso de datos para la toma de decisiones. Las investigaciones muestran que el uso de modelos de aprendizaje automático como Random Forest no solo mejora la precisión de los pronósticos de ventas, sino que también contribuye a la eficiencia operativa, la satisfacción del cliente final y la rentabilidad empresarial.

## 6. AGRADECIMIENTOS

Agradezco a Dios, por la sabiduría y fortaleza, por guiarme en cada paso de este camino de estudios, su presencia constante me ha dado la inspiración, la paciencia y la claridad necesarias para culminar este trabajo con éxito, a mi familia, por su amor incondicional, apoyo y comprensión en todo momento, a mis seres queridos, por su compañía, sus palabras de aliento y su confianza en mí.

A los directivos de Corporación MQ, por facilitar el acceso a la información necesaria para el desarrollo de esta investigación. Su apertura y colaboración fueron fundamentales para aplicar el modelo de predicción y demostrar el valor de la analítica de datos en la industria automotriz. También, expreso mi más sincero agradecimiento a mi tutor y docentes de la Maestría en Transformación digital e innovación de la universidad Politécnica Salesiana, por compartir su conocimiento en todas las asignaturas, en cada una de ella he aprendido algo nuevo para mi crecimiento profesional.

## 7. REFERENCIAS BIBLIOGRÁFICAS

- Cedeño, G. (15 de septiembre de 2025). GESTIÓN DE INVENTARIOS Y SU INCIDENCIA EN LA EFICIENCIA OPERATIVA DE LUBRIREPUESTOS NAVARRETE.  
<http://repositorio.unesum.edu.ec/handle/53000/8194>
- Douaioui, K., Oucheikh, R., Benmoussa, O., y Mabrouki, C. (2024). Machine Learning and Deep Learning Models for Demand Forecasting in Supply Chain Management. *Applied System Innovation (ASI)*, 5, 7. <https://doi.org/https://doi.org/10.3390/asi7050093>
- Eslava, j., Caitlin, C., & Mark, B. (2021). Impacts of Decarbonization on the Automotive After Sales Sector: A Review of Evidence. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 235(6), 1516–1526.  
<https://doi.org/https://doi.org/10.1177/0954407020974793>

- Garcia Gonzalez, A. (2015). Service Quality and Repurchase Behaviour in the Spanish Automotive After Sales Business. *Journal of Relationship Marketing*, 14. <https://doi.org/10.1080/15332667.2015.1069456>
- Gutiérrez, S. (diciembre de 2024). Redes productivas globales y núcleos dinámicos en Norteamérica: la transición tecnológico-productiva de la industria automotriz. 19. <https://doi.org/https://doi.org/10.22201/cisan.24487228e.2024.657>
- Hasani, Y. (2023). Comparing spare parts demand forecasting methods. *Master Thesis Data Science & Marketing Analytics*. <https://doi.org/https://thesis.eur.nl/pub/70009/>
- Huang, e., Xiao, Q., Dai, H., y Yan, N. (2018). Sales Forecast for O2O Services - Based on Incremental Random Forest Method. *15th International Conference on Service Systems and Service Management*. Beijing, China: Business School Central University of Finance and Economics. <https://doi.org/10.1109/ICSSSM.2018.8465121>
- Izquierdo Munoz, L., y San Martin Galindo, J. (2024). Pesky Stock Keeping Unit (SKU) demand forecasting model for American Auto Parts Retailer. *Society and Development*, 9, 13. <https://doi.org/https://doi.org/10.33448/rsd-v13i9.46809>
- Liang, X.-x., Lin, Y., Deng, C., Mo, Y.-h., Lu, B., Yang, J., . . . Liu, M. (2024). Machine Learning-Based Sales Prediction Using Bayesian Optimized XGBoost Algorithms. *Modern Management based on Big Data V*. <https://doi.org/10.3233/FAIA240263>
- Malik, S., Khan, M., Abid, M., y Aslam, N. (2024). Sales Forecasting Using Machine Learning Algorithm in the Retail Sector. *Journal of Computing & Biomedical Informatics*, 6. <https://doi.org/hCps://doi.org/10.56979/602/2024>
- Millar, K., Oliver, B., y Hussain, A. (2020). Aftermarket services Transforming manufacturing in the wake of the COVID-19 pandemic. *Deloitte Insights*. <https://doi.org/https://www.deloitte.com/us/en/insights/industry/manufacturing-industrial-products/aftermarket-services-digital-differentiator-beyond-COVID-19.html>
- Naik , H., Yashwanth, K., P, S., y Jayapandian, N. (2022). Machine Learning based Food Sales Prediction using Random Forest Regression. *Proceedings of the Sixth International Conference on Electronics, Communication and Aerospace Technology*. <https://ieeexplore.ieee.org/abstract/document/10009277>
- Ordóñez, S. (2024). Global Production Networks and Dynamic Cores in North America: The Technological-Productive Transition of the Automotive Industry. *Revista Académica del Cisan-Unam*, 19. <https://doi.org/https://doi.org/10.22201/cisan.24487228e.2024.2.657>
- Silva, E. (2024). OPTIMIZACIÓN DE PROCESOS ORGANIZACIONALES A TRAVÉS DE LA IA. *Revista analisis Organizacional*, pág. 70. [https://www.remineo.org/remineoorg/files/RAO\\_No\\_16.pdf#page=70](https://www.remineo.org/remineoorg/files/RAO_No_16.pdf#page=70)
- Weizhi , L., Guanglei, Y., Yanchao, Y., Weijun, Y., Yaheng, M., y Dongzhou, Z. (2020). Auto Parts Sales Prediction Based on Machine Learning for small data and a long replacement cycle. *IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA)*. IEEE, 2020. China. <https://ieeexplore.ieee.org/abstract/document/9316540>

Weizhi , L., Guanglei, Y., Yanchao, Y., Weijun, Y., Yaheng, M., y Dongzhou, Z. (2020). Auto Parts Sales Prediction Based on Machine Learning for small data and a long replacement cycle. *IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA)*. IEEE, 2020. <https://ieeexplore.ieee.org/abstract/document/9316540>

Yadav, P. K., Kumar, V., Bhushan, R., y Kumar Singh, P. (2023). Analysis of Machine Learning Model for Predicting Sales Forecasting. *First International Conference on Advances in Electrical, Electronics and Computational Intelligence (ICAEECI)*. India: School of Engineering and Design, Alliance University, Bangalore. <https://doi.org/https://ieeexplore.ieee.org/abstract/document/10370914/>

Zheng, S. (2025). Research on market demand forecasting algorithm based on machine learning. *2025 IEEE 3rd International Conference on Image Processing and Computer Applications (ICIPCA)*. China. <https://ieeexplore.ieee.org/abstract/document/11138443>