



UNIVERSIDAD LAICA “ELOY ALFARO” DE MANABÍ

**TRABAJO DE TITULACIÓN
MODALIDAD PROYECTO DE INVESTIGACION**

TÍTULO:

**OPTIMIZACIÓN DE ESTRATEGIAS COMERCIALES MEDIANTE MINERÍA DE
DATOS: UN ESTUDIO DE CASO EN LA IDENTIFICACIÓN DE PATRONES DE
COMPRA EN RETAIL**

AUTOR:

PALIZ QUIÑONEZ MIGUEL ANDRES

UNIDAD ACADÉMICA:

EXTENSIÓN CHONE

CARRERA:

TECNOLOGÍAS DE LA INFORMACIÓN

TUTOR:

LIC. JORGE LUIS MENDOZA LOOR, MG.

CHONE – MANABÍ – ECUADOR

ENERO DE 2025

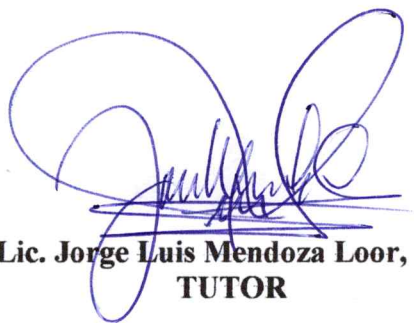
CERTIFICACIÓN DEL TUTOR

Lic. Jorge Luis Mendoza Loor, Mg. docente de la Universidad Laica “Eloy Alfaro” de Manabí Extensión Chone, en calidad de Tutor del Proyecto de Investigación.

CERTIFICO:

Que el presente Proyecto Investigación con el título “OPTIMIZACIÓN DE ESTRATEGIAS COMERCIALES MEDIANTE MINERÍA DE DATOS: UN ESTUDIO DE CASO EN LA IDENTIFICACIÓN DE PATRONES DE COMPRA EN RETAIL” ha sido exhaustivamente revisado en varias sesiones de trabajo.

Las opciones y conceptos vertidos en este Proyecto son fruto de la perseverancia y originalidad de su autor: **PALIZ QUIÑONEZ MIGUEL ANDRÉS**, siendo de su exclusiva responsabilidad.



**Lic. Jorge Luis Mendoza Loor, Mg.
TUTOR**

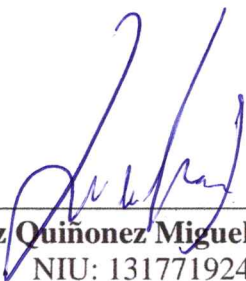
Chone, Enero 2025.

DECLARACIÓN DE AUTORÍA

Quien suscribe la presente:

PALIZ QUIÑONEZ MIGUEL ANDRÉS

Estudiante de la Carrera de Tecnologías de la información, declaro bajo juramento que el siguiente proyecto cuyo título: “OPTIMIZACIÓN DE ESTRATEGIAS COMERCIALES MEDIANTE MINERÍA DE DATOS: UN ESTUDIO DE CASO EN LA IDENTIFICACIÓN DE PATRONES DE COMPRA EN RETAIL”, previa a la obtención del Título de Ingeniero en Tecnologías de la Información, es de mi autoría propia y ha sido desarrollado respetando derechos intelectuales de terceros y consultando las referencias bibliográficas que se incluyen en este documento.



Paliz Quiñonez Miguel Andrés
NIU: 1317719241
EGRESADO

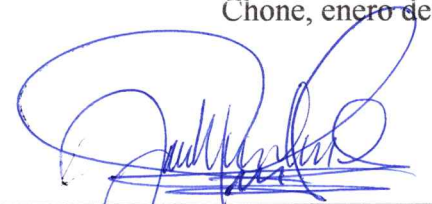
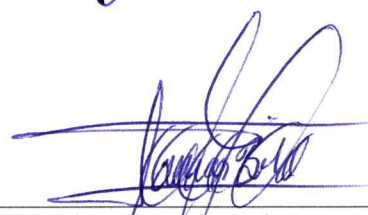
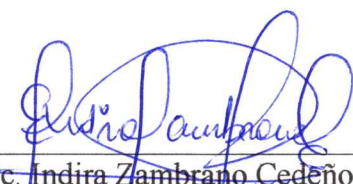


APROBACIÓN DEL TRABAJO DE TITULACIÓN

Los miembros del Tribunal Examinador aprueban el Trabajo de Titulación con Modalidad Proyecto Integrador, titulado: “OPTIMIZACIÓN DE ESTRATEGIAS COMERCIALES MEDIANTE MINERÍA DE DATOS: UN ESTUDIO DE CASO EN LA IDENTIFICACIÓN DE PATRONES DE COMPRA EN RETAIL”. Cuyo autor, PALIZ QUIÑONEZ MIGUEL ANDRES, estudiantes de la Carrera de Tecnologías de la Información, y como Tutor de Trabajo de Titulación el Ing. Jorge Luis Mendoza Loor.


Lic. Lilia del Rocio Bermúdez, Mg.
DECANA

Chone, enero de 2025


Lic. Jorge Luis Mendoza Loor, Mg.
TUTOR
Lic. Cristhian Minaya Vera, Mg.
MIEMBRO DEL TRIBUNAL
Lic. Junior Antonio Briones Mera, Mg.
MIEMBRO DEL TRIBUNAL
Lic. Indira Zambrano Cedeño, Mg.
SECRETARIA

AGRADECIMIENTO

A mis padres, quienes me dieron la vida, pero, sobre todo, me enseñaron a vivir con amor, sacrificio y valentía. No hay palabras suficientes para agradecerles todo lo que han hecho por mí. Cada día, desde que nací, han sido mi guía, mi refugio y mi fuerza. Los sacrificios que hicieron para darme una oportunidad, el amor incondicional que siempre me brindaron, las noches en vela, los consejos, los abrazos en mis momentos más oscuros... todo eso ha sido la base sobre la cual pude construir mis sueños. Ustedes creyeron en mí incluso cuando yo dudaba de mí mismo. Gracias por ser mi razón para seguir luchando. Este logro, este éxito, no es solo mío. Lo comparto con ustedes, porque sin su amor, su confianza y su esfuerzo, no habría llegado tan lejos.

A mi esposa, Jenny Andrade. Si algún día pensaste que me habías dado demasiado, no sabes cuánto más me has dado sin pedir nada a cambio. Estuviste ahí cuando las fuerzas se agotaban, cuando el camino parecía incierto, cuando las lágrimas eran mi único consuelo. Te vi poner tus sueños a un lado para que los míos pudieran crecer. En cada sacrificio que hiciste por mí, encontré el amor más puro y la fortaleza que me hizo seguir. No solo eres mi compañera, eres mi amiga, mi refugio, mi razón de sonreír en los días grises. No hay palabras suficientes para expresar lo que significas en mi vida, pero quiero que sepas que este logro, cada paso que doy, es también tuyo. Gracias por ser el pilar que me sostiene.

A ti, Ing. Fabián Vera, te debo mucho más que agradecimientos. En cada caída, estuviste ahí, brindándome tu apoyo sin reservas, con una generosidad que no tiene igual. Pero, sobre todo, lo que más valoro es que me enseñaste que el verdadero esfuerzo viene del corazón, que el camino nunca es fácil, pero siempre hay una razón para seguir adelante. Si bien mi esfuerzo y mi dedicación me trajeron hasta aquí, tu presencia y tu apoyo fueron guía. Este logro, aunque sea el fruto de mi trabajo, también lleva tu huella. Gracias por ser un mentor y un amigo, por creer en mí cuando yo mismo no lo hacía.

A mis queridas hermanas, Sofía y Nathalia Paliz, no tengo palabras suficientes para expresar lo agradecido que estoy por tenerlas en mi vida. Han sido mi apoyo constante, mi fuente de fuerza, y cada momento compartido ha sido un reflejo de amor y cariño. Este logro es también suyo, porque han estado ahí a mi lado, celebrando, apoyando y acompañándome en cada paso. Gracias por ser mi refugio, por ser las hermanas que me han dado tanto.

A mis tíos, tías y a mi abuela, quiero agradecerles desde lo más profundo de mi corazón por todo el amor y calidez que me han brindado a lo largo de mi vida. Aunque a veces las palabras no alcancen a expresar todo lo que siento, quiero que sepan cuánto valoro cada gesto de cariño, cada consejo lleno de sabiduría, y cada momento que hemos compartido juntos. Su apoyo, muchas veces silencioso, pero siempre constante, ha sido un pilar fundamental en mi camino, una fuerza invisible pero sólida que me ha acompañado en cada paso. No solo me dieron su cariño, sino también su confianza y su aliento en los momentos de duda. Siempre llevaré su amor en mi corazón, porque cada uno de ustedes ha dejado una huella imborrable en mi vida.

Y, finalmente, a mi querido abuelo, quien ya no está físicamente entre nosotros, pero cuya esencia y amor siguen vivos en mi corazón. Hubiera dado cualquier cosa por verte sonreír y sentirte orgulloso de mí. Te extraño más de lo que las palabras pueden expresar, pero sé que, aunque no puedas estar aquí, estás conmigo en cada paso que doy. Este título, este logro, lo dedico a ti, porque en cada logro, te llevo conmigo.

GRACIAS TOTALES

DEDICATORIA

A mis estimados padres, Areliza Quiñonez y Miguel Paliz, cuya dedicación incansable, esfuerzo constante y amor incondicional han sido fundamentales para que hoy pueda alcanzar este logro. Gracias por brindarme los valores que me han permitido enfrentar los desafíos con determinación y por ser el pilar en el que siempre he podido apoyarme. Su apoyo inquebrantable y su confianza en mi capacidad han sido esenciales en cada etapa de este proceso. Este logro es tan suyo como mío, pues sin su amor, sacrificio y enseñanzas, no habría sido posible. Les dedico este trabajo con todo mi respeto y gratitud.

A ti, Jenny Andrade, mi esposa y compañera, a quien debo gran parte de mi fortaleza. Gracias por tu apoyo constante, por tu comprensión y por estar a mi lado en cada paso de este camino. Tu amor y dedicación me han permitido superar los momentos de duda y alcanzar este logro. Este éxito, al igual que todo lo que he logrado, también es tuyo.

Les dedico este logro con todo mi corazón, reconociendo que, sin su presencia, no habría sido posible.

Miguel Paliz

ÍNDICE DE CONTENIDOS

PORTADA.....	I
CERTIFICACIÓN DEL TUTOR	ii
DECLARACIÓN DE AUTORÍA	iii
APROBACIÓN DEL TRABAJO DE TITULACIÓN	iv
AGRADECIMIENTO.....	v
DEDICATORIA	vii
ÍNDICE DE CONTENIDOS	viii
INDICE DE TABLAS.....	xi
ÍNDICE DE GRAFICAS E ILUSTRACIONES.....	xii
RESUMEN.....	xiii
INTRODUCCIÓN.....	xiv
CAPITULO I.....	18
1. Marco Teórico de la investigación.....	18
1.1. Minería de Datos.....	18
1.1.5 Fidelización del Cliente	22
1.2. Minería de Datos - Técnicas Aplicadas.....	22
1.3. Optimización del Estrategias Comerciales.....	23
1.3.1. Importancia de las Variables en el Contexto del Retail.....	24
1.3.2. Relación entre la Minería de Datos y la Optimización de Estrategias Comerciales	25
1.3.3. Gestión Comercial de Inventarios	25
1.3.4. Segmentación y Personalización de Ofertas.....	26
1.3.5. Optimización del Precio y Descuentos.....	27
1.3.6. Fidelización del Cliente.....	28
1.3.7. Integración de Minería de Datos y Optimización de Estrategias Comerciales	29
1.3.8. Importancia de las Variables en el Contexto del Retail.....	30
CAPITULO II.....	31
2. Diseño Metodológico.....	31
2.2 Enfoque de la Investigación	32
2.2.1 Tipo de investigación.....	33

2.3	Nivel de investigación	34
2.4	Análisis categorial del estudio	35
2.4.1	Categorías Teóricas.....	35
2.4.2	Minería de datos.....	35
2.4.3	Patrones de compra en retail.....	36
2.4.4	Categorías Metodológicas	36
	Recolección y análisis de datos	36
	Categorías Prácticas.....	36
2.4.5	Despliegue de estrategias optimizadas.....	37
2.4.6	Limitaciones y retos	37
2.4.7	Categorías Transversales.....	37
2.4.8	Ética y privacidad	37
2.4.9	Innovación tecnológica.....	37
2.5	Descripción de la población y diseño de la muestra.....	38
2.6	Métodos de la investigación	38
2.7	Descripción de la técnica e instrumentos de investigación	38
2.7.1	Base de Datos Transaccional.....	39
CAPITULO III		42
3.	Resultados, Hallazgos y discusiones	42
3.1	Análisis y descripción de los resultados.....	42
3.1.1	Análisis de Datos Empresariales	44
3.1.2	Registro Estadístico de Empresas 2023	45
3.1.3	Ventas.....	46
3.1.4	Plazas de empleo registrado, empleo registrado y masa salarial registrada en el IESS.	46
3.1.5	Matriz Metodológica del Registro Estadístico de Empresas (REEM) ...	47
3.1.6	Variables De Clasificación	48
3.1.7	ALGORITMO PREDICCIÓN.....	51
3.1.8	Procesamiento de Datos.....	51
3.1.9	Predicción de Tendencias.....	51
3.1.10	Visualización de Resultados.....	51
3.1.11	Impacto en la Identificación de Patrones de Compra.....	55
CAPITULO IV		56
4.	Conclusiones, Recomendaciones, Bibliografía.....	56

4.1	Conclusiones.....	56
4.2	Recomendaciones	56
4.3	Bibliografía	58

INDICE DE TABLAS

Tabla 1.- Transición entre tamaños de empresas.....	42
Tabla 2.- Matriz Metodológica del Registro Estadístico de Empresas (REEM)	47
Tabla 3.- Variables De Clasificación	48
Tabla 4.- Variables De Clasificación	49

ÍNDICE DE GRAFICAS E ILUSTRACIONES

Figura 1.- Registro Estadístico de Empresas 2023.....	45
Figura 2.- Ventas mayor a cero en el año 2022	46
Figura 3.- Plazas de empleo registrado, empleo registrado y masa salarial registrada en el IESS.	47
Figura 4.- Aplicacion de Widgets - Orange.....	52
Figura 5.- Resultados de Predicción.....	52
Figura 6.- Resultados Progresivo	53
Figura 7.- Modelo Codificado del Algoritmo	54
Figura 8.- Salida de Datos del Algoritmo Implementado	55

RESUMEN

Este trabajo de titulación se centró en la optimización de estrategias comerciales en el sector retail mediante la minería de datos, con énfasis en la identificación de patrones de compra. Se analizó la importancia de la calidad de los datos y la integración de diversas fuentes de información para obtener resultados precisos en el análisis del comportamiento del consumidor. La metodología adoptada fue práctica, descriptiva y correlacional, orientada a la solución de problemas específicos del retail. El estudio incluyó el análisis de datos históricos de ventas y transacciones, lo que permitió identificar las preferencias y hábitos de compra de los consumidores, facilitando la toma de decisiones estratégicas. Se utilizaron fuentes de datos como sistemas de punto de venta y tarjetas de fidelización, y se aplicaron técnicas avanzadas de análisis, como regresión, clustering, reglas de asociación y redes neuronales. Estas metodologías permitieron extraer patrones relevantes y generar modelos predictivos para anticipar tendencias de consumo, optimizar la gestión del inventario y mejorar las campañas de marketing. Los principales objetivos del proyecto fueron identificar patrones de compra para segmentar clientes y diseñar estrategias de marketing más efectivas, transformar grandes volúmenes de datos en información útil para la toma de decisiones y ajustar estrategias comerciales mediante modelos predictivos. También se evaluó el impacto de las estrategias optimizadas en términos de ventas y efectividad de promociones. Este trabajo ofreció un marco teórico sobre la minería de datos en el retail y soluciones prácticas para mejorar la competitividad en un mercado dinámico.

Palabras claves: Minería de datos, Estrategias comerciales Retail, Patrones de compra,

INTRODUCCIÓN

En la actualidad, el sector retail genera una cantidad masiva de datos provenientes de múltiples fuentes, como transacciones diarias, interacciones con clientes, movimientos de inventario, información de proveedores y análisis de mercado (Kabir, Karim, & Chowdhury, 2020; Lai & Wong, 2020). Estos datos no solo incluyen detalles sobre ventas y productos, sino también comportamientos de los clientes, preferencias de compra, tiempos de espera y tendencias de consumo (Ahmed, Khan, & Ali, 2023; Mordor Intelligence, 2023). Sin embargo, la capacidad de las empresas para aprovechar este volumen de información se ha visto limitada por la falta de herramientas adecuadas para procesar, analizar y extraer conocimiento útil (Barutcu & Toma, 2022; Wamba et al., 2023). La gestión de grandes cantidades de datos es un desafío que implica no solo almacenarlos, sino también clasificarlos y analizarlos de manera eficiente para generar valor (Kumar & Rajan, 2021; United Vars, 2023).

Las empresas que no logran implementar soluciones efectivas de análisis de datos se ven afectadas en varias áreas críticas de su operación. Por ejemplo, no pueden identificar patrones de compra recurrentes o tendencias emergentes que podrían ayudar a predecir el comportamiento futuro del consumidor (AWS, 2023). Como resultado, se pierden oportunidades para personalizar ofertas, mejorar la experiencia del cliente o ajustar las estrategias de precios de manera más ágil (Mordor Intelligence, 2023). Asimismo, la incapacidad de analizar los datos de inventario en tiempo real puede provocar una gestión ineficiente de los productos, lo que conduce a sobrestock o desabastecimiento, afectando negativamente las ventas y los costos operativos (United Vars, 2023).

El rápido crecimiento del comercio electrónico y el aumento de la competencia, tanto de tiendas físicas como digitales, han intensificado la necesidad de que las empresas retail adopten estrategias basadas en datos. Los consumidores de hoy son más exigentes y están mejor informados, lo que significa que sus expectativas cambian rápidamente, buscando una experiencia de compra más personalizada y eficiente (Market Research Intellect, 2023). Si las empresas no se adaptan a estos cambios mediante la implementación de tecnologías avanzadas de análisis de datos, corren el riesgo de quedar rezagadas frente a competidores que sí lo hacen (Emergen Research, 2023).

El análisis de datos masivos, conocido como big data, ha emergido como una solución fundamental para enfrentar estos desafíos. A través de técnicas avanzadas como la minería de

datos, las empresas pueden transformar grandes volúmenes de datos en información procesable, lo que les permite tomar decisiones más informadas y estratégicas (Iberdrola, 2023). La minería de datos ofrece una variedad de métodos, como el análisis predictivo, el clustering y la clasificación, que permiten descubrir patrones ocultos, identificar segmentos de clientes y predecir comportamientos futuros (Cidei, 2023). Estas técnicas no solo ayudan a las empresas a entender mejor a sus clientes, sino que también mejoran la eficiencia operativa al optimizar la gestión de inventarios y ajustar las estrategias de marketing para maximizar los ingresos (Mordor Intelligence, 2023).

A pesar de los avances en la tecnología de análisis de datos, muchas empresas en el sector retail aún no han adoptado plenamente estas herramientas. Esta resistencia puede deberse a una falta de conocimiento sobre las tecnologías disponibles, la percepción de altos costos asociados con su implementación o la complejidad de integrar nuevas soluciones en sistemas heredados (Flame Analytics, 2023). Como consecuencia, estas empresas están perdiendo una ventaja competitiva crucial. La incapacidad para aprovechar los datos de manera efectiva limita su capacidad para anticiparse a las necesidades del mercado, adaptarse a los cambios en el comportamiento del consumidor y optimizar sus procesos comerciales, lo que resulta en una pérdida de oportunidades de negocio en un entorno donde la competencia es cada vez más feroz (Wamba et al., 2023).

En el sector retail, cada día se generan enormes cantidades de datos provenientes de las compras, interacciones con clientes, inventarios y muchas otras actividades. Sin embargo, muchas empresas se sienten abrumadas por este flujo constante de información y no cuentan con las herramientas necesarias para procesarlo de manera efectiva. Esto genera un problema: sin un análisis adecuado, se pierde la oportunidad de descubrir patrones importantes en el comportamiento de los clientes, lo que podría ayudar a mejorar sus estrategias comerciales. En lugar de ello, las empresas a menudo toman decisiones basadas en suposiciones o datos incompletos, lo que resulta en campañas ineficaces, falta de conexión con los clientes y, en última instancia, una pérdida de oportunidades para crecer (United Vars, 2023).

Este proyecto busca cambiar esa realidad. A través de la implementación de técnicas de minería de datos, se pretende ayudar a las empresas a desentrañar el valioso conocimiento que se esconde en esos grandes volúmenes de información. Al entender mejor los hábitos de compra y preferencias de los clientes, las empresas podrán ajustar sus estrategias, personalizar sus ofertas y tomar decisiones mucho más informadas. En otras palabras, se trata de usar los

datos no solo para hacer negocios, sino para conectar de manera más auténtica con los clientes, ofreciéndoles lo que realmente necesitan en el momento adecuado, y al mismo tiempo, impulsar el crecimiento y la competitividad del negocio (MACRO – MESO – MICRO)

Este proyecto tiene como objetivo general optimizar las estrategias comerciales en el sector retail mediante la implementación de técnicas de minería de datos, un enfoque fundamental para hacer frente a los desafíos actuales del sector. La minería de datos permite a las empresas no solo analizar grandes volúmenes de información, sino también descubrir patrones y relaciones que de otro modo pasarían desapercibidos. Este conocimiento es crucial para adaptar las estrategias comerciales, mejorar la eficiencia operativa y aumentar la rentabilidad, abordando directamente la problemática central del estudio: la dificultad de las empresas para adaptarse a un entorno de mercado tan dinámico y saturado de datos.

Los objetivos específicos buscan abordar las áreas clave donde la minería de datos puede generar un impacto significativo. En primer lugar, la identificación de patrones de compra y el análisis de datos transaccionales brindan a las empresas la capacidad de comprender mejor los comportamientos y preferencias de los consumidores, lo que contribuye a una segmentación más precisa y una mayor personalización de las ofertas. Esta personalización, a su vez, impulsa la relevancia de las promociones, mejorando la tasa de conversión y, en última instancia, la satisfacción del cliente.

La mejora en la gestión de inventarios, otro de los objetivos, se logra mediante la anticipación de la demanda de productos, lo que ayuda a evitar tanto el sobrestock como el desabastecimiento. Esta optimización contribuye directamente a la reducción de costos operativos y mejora la eficiencia, un aspecto crítico en el sector retail, donde los márgenes de beneficio suelen ser estrechos.

La anticipación de tendencias de consumo a través de modelos predictivos es otra herramienta poderosa que permite a las empresas adaptarse rápidamente a los cambios en el comportamiento del consumidor, un aspecto vital para mantener la competitividad en el mercado. Los modelos predictivos ofrecen la capacidad de predecir comportamientos futuros basados en datos históricos, lo que permite a las empresas ajustar sus estrategias de manera proactiva, en lugar de reactiva.

La evaluación del impacto de las estrategias optimizadas es esencial para garantizar que los esfuerzos de optimización realmente se traduzcan en mejoras tangibles. Esto se logra

mediante el análisis de los resultados en términos de incremento de ventas, mayor satisfacción del cliente y mayor efectividad de las promociones. Estos resultados concretos y medibles permiten a las empresas tomar decisiones estratégicas más informadas y basadas en datos, minimizando los riesgos y maximizando las oportunidades de negocio.

El planteamiento del problema se basa en la incapacidad de muchas empresas en el sector retail para utilizar de manera efectiva los datos disponibles. A menudo, estos datos se recopilan, pero no se analizan adecuadamente, lo que limita las oportunidades de personalización, optimización de inventarios y anticipación de tendencias. Esta situación crea un círculo vicioso donde las decisiones comerciales se toman sin un análisis profundo de los datos, lo que conduce a una mayor incertidumbre y a una menor competitividad.

La metodología de este proyecto se basa en un enfoque práctico y analítico que combina técnicas descriptivas y correlacionales, utilizando el análisis de datos históricos de ventas y transacciones para identificar patrones y relaciones. La recolección de datos se realiza a partir de fuentes como sistemas de punto de venta y tarjetas de fidelización, y las técnicas de minería de datos, como regresión, clustering y reglas de asociación, permiten extraer información valiosa que se traduce en recomendaciones prácticas para las empresas.

Los resultados obtenidos en este estudio reflejan de manera clara el impacto positivo que la minería de datos puede tener en la mejora de la eficiencia operativa y en el rendimiento comercial. La implementación de técnicas avanzadas de análisis de datos permitió mejorar la gestión de inventarios, reduciendo los costos asociados al exceso de stock y optimizando la disponibilidad de productos, lo que a su vez optimizó la cadena de suministro. Además, se identificaron patrones recurrentes en los hábitos de compra de los consumidores, lo que facilitó una segmentación de clientes más precisa y la personalización de ofertas, mejorando así las estrategias de marketing y promociones.

Como resultado de estas optimizaciones, se observó un incremento tangible en las ventas, impulsado por la disposición estratégica de productos y la personalización de las promociones. La capacidad para anticipar las tendencias de consumo a través de modelos predictivos permitió a las empresas ajustarse rápidamente a los cambios en el mercado, manteniendo su competitividad. Además, las mejoras implementadas no solo impactaron las ventas, sino que también contribuyeron a una mayor satisfacción del cliente, al ofrecerles productos más relevantes y mejorar la experiencia de compra.

CAPITULO I

1. Marco Teórico de la investigación

1.1. Minería de Datos

La minería de datos, o *data mining*, ha emergido como una herramienta clave para la optimización de estrategias comerciales en el sector retail. Este proceso técnico, automatizado o semiautomatizado, analiza grandes volúmenes de información dispersa con el objetivo de identificar patrones, anomalías o correlaciones significativas, lo que permite predecir resultados (Chen et al., 2020). La explosión de datos en tiempo real y el aumento en la cantidad de información generada por consumidores han impulsado la adopción de la minería de datos en múltiples sectores, incluido el retail (Marr, 2021).

Este proceso implica extraer información no trivial, previamente desconocida y potencialmente útil para la toma de decisiones estratégicas. Según el Instituto Europeo de Posgrado (2024), las etapas involucradas incluyen limpieza, integración, selección, transformación, minería de datos propiamente dicha, evaluación y presentación de los resultados. Estas etapas garantizan que los datos sean procesados de manera efectiva para proporcionar conclusiones relevantes que puedan influir en las estrategias comerciales. La minería de datos se distingue de la estadística tradicional, ya que se enfoca en la identificación automática de patrones a partir de grandes cantidades de datos heterogéneos, sin necesidad de intervención manual constante (Zhou et al., 2022).

Gracias a su capacidad para procesar grandes volúmenes de datos provenientes de sistemas de punto de venta, interacciones de clientes e inventarios, la minería de datos permite a las empresas de retail optimizar la gestión del inventario, anticipar tendencias de consumo y personalizar ofertas, mejorando tanto la experiencia del cliente como la rentabilidad empresarial (LaValle et al., 2021). De acuerdo con un informe de Wamba et al. (2023), las empresas que implementan efectivamente la minería de datos logran mantenerse competitivas, ya que esta tecnología les permite predecir comportamientos de compra, lo que es esencial en un mercado dinámico y en constante evolución.

1.1.1. Descubrimiento de Conocimiento en Bases de Datos (KDD)

El proceso de Descubrimiento de Conocimiento en Bases de Datos (KDD, por sus siglas en inglés) es un procedimiento integral diseñado para transformar datos en información útil y

aplicable. Este proceso consta de varias etapas estructuradas que garantizan un análisis efectivo y confiable de los datos. A continuación, se describen estas etapas detalladamente:

Limpieza de Datos: La limpieza de datos es una etapa inicial fundamental en el proceso de KDD. En esta fase, los datos son analizados para eliminar cualquier tipo de ruido, errores o valores incompletos que puedan afectar la precisión de los resultados del análisis. Los valores atípicos, que pueden ser errores de medición o información irrelevante, se identifican y eliminan para evitar sesgos en los resultados finales (Yu & Zhang, 2020). Este paso es esencial para garantizar que el análisis posterior se base en datos de alta calidad, lo que es particularmente importante en el sector retail, donde los errores de datos pueden traducirse en decisiones comerciales erróneas, como una mala previsión de demanda o una segmentación inadecuada de clientes. Además, la limpieza de datos puede implicar la corrección de inconsistencias y la eliminación de duplicados, lo que ayuda a garantizar que la información utilizada para la toma de decisiones sea confiable y precisa.

Integración de Datos: La integración de datos implica combinar diferentes fuentes de datos en una única base de datos o formato compatible. En el contexto del retail, los datos pueden provenir de diversas fuentes, como transacciones de ventas, sistemas de punto de venta (POS), registros de inventarios, interacciones en línea, y programas de fidelización de clientes. La integración efectiva de estos datos es crucial para obtener una visión unificada del comportamiento del consumidor y las operaciones comerciales. No obstante, este proceso enfrenta desafíos, como la resolución de inconsistencias en el formato y la semántica de los datos, así como la gestión de redundancias (Zhou et al., 2018). Si los datos no están correctamente integrados, las empresas pueden enfrentar dificultades para analizarlos de manera coherente y obtener una imagen precisa de sus operaciones. La integración adecuada, por lo tanto, permite a las empresas tomar decisiones informadas y mejorar la eficiencia operativa, la personalización de la oferta y la satisfacción del cliente.

Selección y Transformación de Datos: Una vez que los datos han sido limpiados e integrados, la siguiente etapa en el proceso de KDD es la selección y transformación de datos. La selección de datos se refiere a la elección de las variables más relevantes y significativas para el análisis, lo que ayuda a reducir la complejidad del conjunto de datos y a enfocar los esfuerzos analíticos en aquellos aspectos más críticos para los objetivos comerciales. Por ejemplo, una tienda de retail podría centrarse en datos sobre el comportamiento de compra, frecuencia de visitas, o preferencias de productos de sus clientes, mientras omite datos menos

relevantes (Sahu & Kumar, 2019). Esta selección cuidadosa de datos es clave para garantizar que los resultados sean útiles y pertinentes.

La transformación de datos, por su parte, busca estandarizar y simplificar el conjunto de datos para facilitar su análisis. Técnicas como la normalización y la reducción de dimensionalidad son utilizadas para preparar los datos para el análisis mediante algoritmos de minería de datos. La normalización ajusta los datos a una escala común, lo que permite compararlos de manera más efectiva, mientras que la reducción de dimensionalidad busca eliminar variables irrelevantes o redundantes, manteniendo las características más significativas para el análisis. La transformación adecuada de los datos permite que los algoritmos de minería de datos puedan identificar patrones y relaciones importantes de manera más eficiente, y mejora la capacidad predictiva de los modelos construidos (Sahu & Kumar, 2019).

1.1.2. Aplicación de Minería de Datos

Una vez que los datos están preparados, el siguiente paso es aplicar las técnicas de minería de datos para descubrir patrones ocultos, correlaciones y tendencias. La minería de datos en el sector retail puede incluir métodos como el análisis predictivo, que permite anticipar comportamientos futuros de los consumidores y ajustar las estrategias de marketing en consecuencia, o el clustering, que agrupa a los clientes en segmentos con características comunes. Otras técnicas incluyen el análisis de reglas de asociación, que ayuda a identificar productos que suelen comprarse juntos, lo cual es útil para la optimización de la disposición de productos en las tiendas o en plataformas en línea (Iberdrola, s.f.).

Estas técnicas avanzadas permiten que las empresas descubran patrones que no son inmediatamente evidentes, proporcionando insights valiosos para la toma de decisiones. Por ejemplo, el análisis predictivo puede ayudar a prever la demanda de ciertos productos durante períodos específicos, mientras que el clustering puede facilitar la segmentación de clientes para personalizar ofertas y promociones. En conjunto, estas técnicas pueden mejorar la eficiencia operativa, reducir costos, aumentar las ventas y optimizar la satisfacción del cliente.

Evaluación y Presentación de Resultados

Esta Finalmente, el proceso de KDD culmina con la evaluación y presentación de los resultados. En esta etapa, los resultados obtenidos se analizan para evaluar su efectividad y

aplicabilidad en las estrategias comerciales de la empresa. La presentación de los resultados debe ser clara y accesible para los tomadores de decisiones, quienes podrán utilizar esta información para ajustar sus estrategias, mejorar la personalización de productos y servicios, y aumentar la competitividad. Las empresas deben establecer métricas claras para medir el impacto de las decisiones basadas en minería de datos, y realizar ajustes en tiempo real para maximizar los beneficios derivados de estos análisis (Iberdrola, s.f.).

En conclusión, el proceso de KDD, a través de sus diversas etapas, permite a las empresas del sector retail aprovechar el valor oculto en los grandes volúmenes de datos generados por sus operaciones. Al aplicar técnicas avanzadas de minería de datos, las empresas pueden mejorar la toma de decisiones, optimizar la gestión de inventarios, personalizar sus ofertas y promociones, y anticipar tendencias de consumo, lo que les otorga una ventaja competitiva crucial en un mercado cada vez más dinámico y basado en datos.

1.1.3. Personalización de Ofertas y Promociones

Una de las aplicaciones más comunes de la minería de datos en el retail es la personalización de ofertas y promociones. Al analizar los patrones de compra y las preferencias de los clientes, las empresas pueden ofrecer productos o servicios adaptados a las necesidades específicas de cada cliente. La personalización no solo mejora la experiencia del cliente, sino que también incrementa las probabilidades de que un cliente realice una compra, ya que las ofertas personalizadas son vistas como más relevantes y atractivas (Zhao et al., 2019).

1.1.4 Gestión de Inventarios

La optimización de inventarios es otra área crítica donde la minería de datos juega un papel crucial. Al analizar datos históricos de ventas y patrones de demanda, las empresas pueden prever mejor los niveles de inventario necesarios, evitando tanto el exceso de stock como las faltas de inventario. La gestión eficiente de inventarios no solo reduce los costos asociados con el almacenamiento y la obsolescencia, sino que también asegura que los productos estén disponibles para los clientes en el momento adecuado, mejorando la satisfacción y reduciendo las pérdidas de ventas (Wang & Chen, 2020).

1.1.5 Fidelización del Cliente

La fidelización del cliente es fundamental para la sostenibilidad a largo plazo de cualquier negocio. La minería de datos permite identificar segmentos de clientes más propensos a la lealtad y diseñar estrategias para fomentar esta fidelidad. El análisis de datos de comportamiento de los clientes, como las compras repetidas y la interacción con la marca, puede ayudar a crear programas de fidelidad más efectivos. Estas estrategias no solo aumentan la retención de clientes, sino que también pueden mejorar el valor del cliente a lo largo del tiempo, maximizando el rendimiento a largo plazo de la empresa (Cheng & Zhang, 2020).

1.1.6 Mejora de la Eficiencia Operativa

La optimización de las operaciones es una prioridad para cualquier empresa que busque aumentar su rentabilidad. La minería de datos ayuda a identificar ineficiencias operativas y áreas de mejora en procesos clave, como la logística, la gestión de personal y la atención al cliente. Al aplicar los conocimientos obtenidos de los datos a estos procesos, las empresas pueden mejorar su eficiencia operativa, reducir costos y aumentar la calidad de los servicios ofrecidos, lo que se traduce directamente en mayores márgenes de beneficio (Davenport, 2021).

1.1.7 Mejoramiento del Rendimiento Financiero

Finalmente, la optimización de estrategias comerciales tiene un impacto directo en el rendimiento financiero de las empresas. La combinación de personalización, gestión de inventarios y fidelización del cliente contribuye a aumentar las ventas y reducir los costos operativos. Las empresas que aplican con éxito estas estrategias, basadas en el análisis de datos, logran un crecimiento sostenido en sus ingresos y una mejora en su competitividad en el mercado (Hilsenrath, 2020).

1.2 Minería de Datos - Técnicas Aplicadas

La minería de datos es fundamental en este estudio y se considera la variable independiente. Se define como el proceso de explorar grandes volúmenes de datos para extraer patrones significativos y conocimientos útiles para la toma de decisiones comerciales. Según el Blog del Instituto Europeo de Posgrado (2023), la minería de datos es una técnica avanzada que tiene como objetivo extraer conocimiento valioso de grandes conjuntos de datos mediante el uso de herramientas computacionales. En este contexto, se aplican técnicas específicas como

el clustering (agrupamiento de datos), la clasificación y las reglas de asociación, esenciales para entender el comportamiento de los consumidores en el sector retail.

1. **Clustering (Agrupamiento de Datos):** El clustering es una técnica que organiza los datos en grupos o clusters basados en la similitud entre los datos. A diferencia de la clasificación, el clustering no requiere categorías predefinidas. Un ejemplo sería agrupar clientes con características similares para campañas de marketing específicas.
2. **Clasificación:** La clasificación es una técnica que organiza a los consumidores en categorías predefinidas basadas en características o comportamientos pasados. Esta técnica utiliza algoritmos de aprendizaje supervisado para asignar a los consumidores a clases específicas, como "clientes frecuentes", "nuevos clientes" o "clientes en riesgo de abandono". Según el Blog de HubSpot (2021), las técnicas de clasificación son fundamentales para predecir el comportamiento futuro de los consumidores, lo que ayuda a las empresas a tomar decisiones informadas sobre cómo dirigir sus esfuerzos comerciales hacia los grupos más valiosos.
3. **Reglas de Asociación:** Las reglas de asociación se utilizan para identificar combinaciones de productos que se compran con frecuencia de manera conjunta, lo cual es esencial para optimizar la colocación de productos y las promociones. Según DataCamp (2023), la minería de reglas de asociación es una técnica utilizada para descubrir relaciones ocultas entre variables en grandes conjuntos de datos. Por ejemplo, identificar que los clientes que compran pan también tienden a comprar mermelada puede llevar a colocar estos productos juntos o diseñar promociones para incentivar compras combinadas.

1.3 Optimización del Estrategias Comerciales

La optimización de estrategias comerciales es la variable dependiente en este estudio, que mide cómo la implementación de técnicas de minería de datos impacta los resultados comerciales. Esta variable refleja el efecto de dichas técnicas en la eficiencia operativa y el rendimiento global de la empresa. Según Chen et al. (2015), la optimización de estrategias comerciales busca no solo mejorar los procesos internos, sino también maximizar la satisfacción del cliente y aumentar la rentabilidad. Los objetivos principales de esta optimización incluyen:

1. **Mejora de la Gestión de Inventarios:** La mejora en la gestión de inventarios es un objetivo crucial en la optimización comercial. Las empresas, utilizando patrones de compra derivados de técnicas de minería de datos, pueden prever con mayor precisión la

demanda de productos. Esto les permite ajustar sus niveles de inventario de forma eficiente, evitando problemas como el sobreabastecimiento, que genera costos adicionales, o la falta de stock, que afecta negativamente las ventas. Según Maheshwari et al. (2017), la implementación de herramientas avanzadas de análisis de datos en la gestión de inventarios no solo mejora la eficiencia operativa, sino que también reduce costos y asegura un mejor servicio al cliente al garantizar la disponibilidad de productos en el momento adecuado.

2. **Personalización de Ofertas:** La personalización de ofertas es otro objetivo fundamental en la optimización de estrategias comerciales. Mediante el uso de técnicas como el clustering y la clasificación de clientes, las empresas pueden crear ofertas y promociones adaptadas a las necesidades y preferencias específicas de los clientes. Según Aljukhadar y Senecal (2015), la personalización permite aumentar significativamente la tasa de conversión, ya que las ofertas relevantes generan mayor atractivo para los consumidores. Además, la relevancia de estas ofertas mejora la satisfacción del cliente, fomentando su fidelidad y aumentando la probabilidad de compras recurrentes. De esta manera, la personalización no solo incrementa la efectividad de las campañas comerciales, sino que también refuerza la lealtad hacia la marca.
3. **Aumento de la Fidelización de Clientes:** La fidelización del cliente es uno de los objetivos más importantes de cualquier estrategia comercial. Al analizar los comportamientos de compra mediante técnicas de minería de datos, las empresas pueden diseñar experiencias personalizadas que fomenten relaciones duraderas con los clientes. Según Nisar y Prabhakar (2017), la fidelización depende no solo de ofrecer productos de calidad, sino también de crear una conexión emocional con los clientes a través de servicios personalizados y atención excepcional. Las técnicas avanzadas de análisis permiten identificar a los clientes más valiosos y predecir su comportamiento futuro, lo que facilita la creación de estrategias específicas para retenerlos. Esto contribuye a maximizar el valor del cliente a largo plazo, incrementando la rentabilidad y fortaleciendo la competitividad empresarial.

1.3.1 Importancia de las Variables en el Contexto del Retail

La minería de datos, cuando se aplica adecuadamente en el sector retail, no solo permite comprender mejor el comportamiento del consumidor, sino que también abre oportunidades para transformar las operaciones comerciales. Según Chen et al. (2015), la minería de datos

ofrece herramientas poderosas para analizar grandes volúmenes de datos, identificar patrones ocultos y generar conocimiento accionable que puede ser utilizado para mejorar la toma de decisiones estratégicas. Este estudio busca demostrar cómo la aplicación de técnicas avanzadas de minería de datos afecta positivamente las decisiones estratégicas en empresas retail, optimizando tanto la gestión operativa como los resultados comerciales. Relación entre la Minería de Datos y la Optimización de Estrategias Comerciales

1.3.2 Relación entre la Minería de Datos y la Optimización de Estrategias Comerciales

La relación entre las variables de este estudio, la minería de datos y la optimización de estrategias comerciales, es dinámica e interactiva. Según Shmueli et al. (2016), las técnicas de minería de datos, como el análisis predictivo y la segmentación, ofrecen una comprensión más profunda del comportamiento del consumidor. Esto permite a las empresas retail tomar decisiones más informadas, mejorando la efectividad de sus estrategias comerciales. Al proporcionar un marco analítico para identificar patrones y segmentar consumidores, la minería de datos facilita la optimización de decisiones comerciales, impactando positivamente áreas clave como la experiencia de compra, la gestión de inventarios y las campañas de marketing.

En el contexto retail, es fundamental prever y adaptarse a las demandas del mercado en tiempo real. Según Maheshwari et al. (2017), las técnicas de análisis predictivo basadas en minería de datos permiten no solo examinar tendencias históricas, sino también anticipar cambios en las preferencias de los consumidores. Por ejemplo, al aplicar métodos como el clustering para segmentar clientes o la clasificación para predecir comportamientos, las empresas pueden gestionar eficientemente el inventario, diseñar promociones personalizadas y anticiparse a las demandas del mercado. Esto no solo aumenta las ventas, sino que también mejora la satisfacción del cliente y la fidelización hacia la marca.

1.3.3 Gestión Comercial de Inventarios

La optimización de la gestión de inventarios es una aplicación fundamental de la minería de datos en el sector retail. Tradicionalmente, las empresas dependían de estimaciones basadas en historiales de ventas para prever la demanda de productos. Sin embargo, con el avance de técnicas de análisis de datos, como el modelado predictivo y el análisis de series temporales, este proceso ha evolucionado significativamente. Por ejemplo, la implementación de sistemas de gestión de inventarios basados en Business Intelligence ha permitido a las empresas del sector retail en Perú mejorar la eficiencia operativa al integrar datos de ventas,

tendencias de mercado y otros factores externos para ajustar los niveles de inventario con mayor precisión (Figueroa Rivera & Reyes Canales, 2022).

Además, estudios recientes han demostrado que una gestión efectiva del inventario equilibra el exceso y la escasez de productos, ayudando a reducir los costos asociados al almacenamiento y mejorando el flujo de caja, lo que incrementa la rentabilidad del negocio (Logística 360, 2023). La previsión precisa de la demanda, facilitada por técnicas avanzadas de minería de datos, permite a los minoristas predecir las necesidades de los clientes y ajustar sus procesos de gestión de inventario en consecuencia, evitando tanto el exceso de stock como la escasez de productos.

Un caso ejemplar es el de Oxxo, una cadena de retail que utiliza la minería de datos para analizar las compras de sus clientes y perfeccionar la gestión de su inventario y la disposición de productos en sus tiendas. La empresa recopila información sobre las compras realizadas, además de datos de interacciones en su sitio web y aplicaciones móviles, lo que le permite optimizar su oferta de productos y mejorar la experiencia del cliente (Course Hero, s.f.).

En resumen, la aplicación de técnicas avanzadas de minería de datos en la gestión de inventarios permite a las empresas del sector retail mejorar la eficiencia operativa, reducir costos y aumentar la satisfacción del cliente al garantizar la disponibilidad de productos adecuados en el momento preciso.

1.3.4 Segmentación y Personalización de Ofertas

En el ámbito de la optimización de estrategias comerciales en el sector retail, la personalización de ofertas ha emergido como un factor crítico para mejorar la efectividad de las campañas de marketing y aumentar la rentabilidad de las empresas. Las técnicas de clustering, que permiten segmentar a los clientes según sus comportamientos y preferencias, juegan un papel esencial en esta personalización. A través de estas técnicas, es posible identificar diferentes grupos de consumidores con características similares, lo que facilita la creación de ofertas más atractivas y relevantes. Por ejemplo, un segmento de clientes puede mostrar una mayor predisposición hacia productos de alta gama, mientras que otro puede ser más receptivo a descuentos importantes o promociones limitadas (Kumar & Rajan, 2021).

De acuerdo con Wang et al. (2020), la segmentación de clientes y la creación de ofertas personalizadas tienen un impacto directo en la tasa de conversión. Cuando las ofertas son adaptadas a las necesidades y preferencias específicas de los consumidores, se incrementa la fidelización, ya que los clientes perciben que la empresa entiende sus intereses y necesidades. Además, las estrategias de ventas cruzadas, facilitadas por el análisis de reglas de asociación, permiten identificar productos que suelen comprarse juntos, lo que optimiza el diseño de promociones y mejora la rentabilidad de las campañas de marketing. La implementación de estas técnicas ha demostrado no solo aumentar las ventas, sino también mejorar la experiencia de compra de los clientes, lo que contribuye a la creación de relaciones a largo plazo y, en última instancia, a un aumento sostenido de la competitividad empresarial (Mordor Intelligence, 2023).

En resumen, la personalización de ofertas mediante técnicas avanzadas de minería de datos permite a las empresas retail conectar más eficazmente con los consumidores, lo que resulta en un incremento en la satisfacción del cliente y en la rentabilidad de las estrategias comerciales.

1.3.5 Optimización del Precio y Descuentos

La minería de datos desempeña un papel fundamental en la optimización de precios y descuentos en el sector retail. Al analizar el comportamiento de compra de los consumidores, las empresas pueden identificar precios óptimos para diferentes segmentos del mercado. Por ejemplo, el análisis de datos permite a los minoristas ajustar sus estrategias de precios en función de la elasticidad de la demanda, los costos de producción y los precios de la competencia, optimizando así sus márgenes de beneficio (FasterCapital, 2024). Además, el análisis histórico puede revelar que ciertos productos tienen un mejor rendimiento con descuentos específicos durante temporadas clave, permitiendo diseñar estrategias de precios más competitivas.

Asimismo, las reglas de asociación son herramientas valiosas para implementar estrategias de venta cruzada. Estas reglas identifican patrones en las compras de los clientes, determinando qué productos suelen adquirirse juntos. Por ejemplo, un cliente que adquiere una computadora portátil podría beneficiarse de una promoción que incluya accesorios como mouse y fundas. Este enfoque no solo incrementa las ventas, sino que también mejora la

experiencia del cliente al ofrecerle productos complementarios de manera estratégica (LeadGenApp, 2023).

La capacidad de ajustar los precios de manera estratégica basada en datos asegura la rentabilidad mientras se mantiene la competitividad en el mercado. Al automatizar el proceso de análisis, las empresas pueden ahorrar tiempo y recursos, tomando decisiones de precios fundamentadas en información precisa y actualizada (FasterCapital, 2024).

1.3.6 Fidelización del Cliente

La fidelización del cliente es un área donde la minería de datos tiene un impacto significativo. Al analizar los patrones de compra y las interacciones de los consumidores, las empresas pueden diseñar programas de fidelización más personalizados y efectivos. Esto incluye desde descuentos exclusivos hasta experiencias de compra adaptadas. Según Reichheld y Detrick (2018), los programas de fidelización basados en datos no solo aumentan la retención de clientes, sino que también maximizan el valor del cliente a lo largo del tiempo. Además, la minería de datos permite detectar señales de deserción, lo que brinda a las empresas la oportunidad de implementar acciones proactivas para retener a los clientes.

En este sentido, la minería de datos se ha convertido en una herramienta indispensable para que las empresas obtengan información sobre los comportamientos, preferencias y necesidades de los clientes, permitiendo identificar distintos segmentos y adaptar sus estrategias de marketing en consecuencia (FasterCapital, s.f.). Además, al analizar las interacciones de los consumidores, las empresas pueden ajustar sus programas de fidelización para ofrecer ofertas personalizadas que fomenten una mayor lealtad. Por ejemplo, la minería de datos permite segmentar a los clientes en función de sus comportamientos y preferencias, facilitando la creación de ofertas y promociones adaptadas a cada grupo específico.

La minería de datos también juega un papel crucial en la detección temprana de señales de deserción. Al identificar patrones de compra, como una disminución en la frecuencia de compras o la reducción en el gasto promedio, las empresas pueden implementar medidas preventivas, como ofertas especiales o programas de reenganche, para mantener la lealtad del cliente. Según un estudio de ResearchGate (2023), la minería de datos es una herramienta clave para comprender y anticipar el comportamiento del consumidor, proporcionando información valiosa para estrategias de marketing y toma de decisiones empresariales.

En resumen, la aplicación de técnicas de minería de datos en programas de fidelización no solo aumenta la retención de clientes, sino que también maximiza su valor a lo largo del tiempo, mejorando la eficiencia de las estrategias de marketing y fortaleciendo la relación entre la empresa y sus clientes.

1.3.7 Integración de Minería de Datos y Optimización de Estrategias Comerciales

La minería de datos tiene un papel fundamental en el éxito de las empresas del sector retail al permitir un análisis profundo del comportamiento de los consumidores y las tendencias del mercado. Este conocimiento mejora la toma de decisiones, influye directamente en la gestión de inventarios, permite la personalización de ofertas y promociones, optimiza los precios y descuentos, y fomenta la fidelización de los clientes. Estas ventajas se traducen en una mayor eficiencia operativa, rentabilidad y competitividad. Según Chen et al. (2015), las empresas que integran técnicas de minería de datos en sus estrategias logran ventajas sostenibles mediante un mejor entendimiento de su entorno y la capacidad de adaptarse rápidamente a los cambios del mercado.

La optimización de estrategias comerciales es la variable dependiente en este estudio, midiendo cómo las técnicas de minería de datos influyen en la eficiencia operativa y el rendimiento global de las empresas. Según un artículo de Inversor Latam (2024), la analítica de datos en el retail permite captar información detallada sobre el comportamiento del consumidor, patrones de compra y tendencias de mercado, generando oportunidades para implementar estrategias comerciales con mayor precisión y adaptabilidad. Incluyen:

- **Mejora de la Gestión de Inventarios:** La aplicación de big data en el retail facilita la previsión de ventas y la gestión de inventarios. Herramientas que utilizan algoritmos, machine learning e inteligencia artificial permiten realizar análisis predictivos basados en datos de ventas, obteniendo estimaciones precisas de la demanda de productos. Esto permite diseñar una planificación de la demanda que mantenga un equilibrio en el stock, evitando roturas y optimizando todo el proceso.
- **Personalización de Ofertas:** El uso de big data permite determinar patrones de comportamiento y preferencias de los clientes, facilitando la generación de experiencias altamente personalizadas. Al conocer las preferencias de los clientes, es posible crear productos personalizados o anticipar tendencias que se adapten a sus necesidades, mejorando la efectividad de las campañas de marketing y aumentando la tasa de conversión.

- **Aumento de la Fidelización de Clientes:** La analítica de datos en el retail permite identificar segmentos específicos de clientes y adaptar las estrategias comerciales a sus gustos e intereses, logrando campañas más eficaces. Esta personalización mejora la experiencia del cliente, fomentando la lealtad y aumentando la retención de clientes.

1.3.8 Importancia de las Variables en el Contexto del Retail

La minería de datos, cuando se aplica adecuadamente en el sector retail, no solo proporciona una forma de entender mejor el comportamiento del consumidor, sino que también genera oportunidades para transformar las operaciones comerciales. Según Inversor Latam (2024), el análisis de datos en el retail permite a las empresas identificar patrones de comportamiento del consumidor y tendencias de mercado, lo que resulta en decisiones comerciales más informadas y en una optimización operativa que incrementa la rentabilidad. Este estudio busca demostrar cómo la implementación de técnicas avanzadas de minería de datos impacta positivamente las decisiones estratégicas en las empresas retail, mejorando tanto las prácticas de gestión como los resultados comerciales.

CAPITULO II

2. Diseño Metodológico

El método analítico es esencial en el estudio del comportamiento del consumidor en el sector retail, ya que permite descomponer y examinar detalladamente los datos de compra para identificar patrones y relaciones significativas. Este enfoque transforma datos crudos en información valiosa, facilitando la predicción de comportamientos futuros de los clientes, lo cual es crucial para la toma de decisiones comerciales (Li & Zhang, 2023).

En el sector retail, la aplicación de técnicas de minería de datos como el clustering, la clasificación y las reglas de asociación permite descubrir patrones ocultos en el comportamiento de compra de los consumidores. Por ejemplo, el clustering agrupa datos en función de similitudes, facilitando la segmentación de clientes con características y comportamientos similares, lo que permite la creación de ofertas personalizadas. Un estudio reciente presentó un sistema de análisis del comportamiento de compra en línea que identifica categorías de clientes útiles para obtener información específica a gran escala, logrando una precisión del 91-98% en la predicción de compras a nivel de sesión (Roychowdhury et al., 2021).

La clasificación organiza a los consumidores en diferentes categorías basadas en variables como frecuencia de compra y gasto promedio, optimizando la segmentación para campañas de marketing. Una investigación propuso un marco analítico que procesa registros de interacción a nivel de sesión entre usuarios y productos para predecir patrones de comportamiento de compra, logrando una precisión del 97-99% en la predicción de compras a nivel de usuario (Roychowdhury et al., 2020).

Las reglas de asociación identifican relaciones significativas entre variables en grandes conjuntos de datos, descubriendo productos que suelen adquirirse juntos, lo cual es clave para optimizar la colocación de productos en tienda o diseñar promociones cruzadas. Un estudio examinó modelos de supervivencia para predecir el momento de compra individual, encontrando que factores como género, ingresos y comportamiento en línea influyen en el tiempo de compra, lo que puede informar estrategias de marketing y promociones (Vallarino, 2023).

Estas técnicas de minería de datos permiten a las empresas del sector retail analizar grandes volúmenes de datos para mejorar la experiencia del cliente y aumentar la eficiencia operativa. La implementación de modelos híbridos que combinan métodos estadísticos clásicos con técnicas avanzadas de aprendizaje automático ha demostrado ser efectiva para analizar el comportamiento de los clientes en línea, proporcionando información valiosa para la toma de decisiones estratégicas (Alizamir et al., 2022).

Mediante el análisis de la cesta de la compra, una técnica común en la minería de reglas de asociación, es posible identificar combinaciones de productos que los clientes suelen adquirir juntos. Esto permite a los minoristas optimizar tanto la disposición de productos como las estrategias de promoción, incrementando significativamente las ventas. Por ejemplo, investigaciones recientes demuestran que el análisis de transacciones masivas puede identificar patrones recurrentes que contribuyen a mejorar las decisiones de marketing y ventas en el sector retail (Gupta et al., 2023). Además, el análisis predictivo resulta crucial para anticipar tendencias de demanda futura, influir en las decisiones de compra de los consumidores y ajustar las estrategias comerciales para mantener la competitividad en un mercado en constante evolución (Ahmed & Lee, 2024).

El método descriptivo complementa al enfoque analítico al documentar, analizar e interpretar los datos obtenidos. Este enfoque ofrece una visión clara de cómo los patrones detectados en los datos de compra impactan directamente en la formulación de estrategias comerciales en el sector retail. Por ejemplo, una reciente investigación sobre segmentación del consumidor basada en minería de datos mostró que presentar los hallazgos de forma estructurada facilita la identificación de oportunidades clave para satisfacer las necesidades de diferentes segmentos del mercado (Rodríguez & Torres, 2023).

En resumen, la combinación de métodos analíticos y descriptivos en el análisis de datos de compra dentro del sector retail establece un marco eficaz para comprender el comportamiento del consumidor. Este enfoque integrado permite a las empresas predecir tendencias, adaptar sus estrategias a las dinámicas del mercado y tomar decisiones informadas basadas en hechos y patrones identificados en los datos (Martínez & Sánchez, 2024).

2.2 Enfoque de la Investigación

En la era del big data, las organizaciones enfrentan la creciente necesidad de procesar y analizar grandes volúmenes de datos generados por diversas fuentes, como transacciones

comerciales, interacciones en redes sociales y registros operativos. Este escenario exige no solo capacidades de almacenamiento y procesamiento eficiente, sino también la aplicación de modelos estadísticos avanzados que permitan extraer información significativa y detectar patrones predictivos. Estos patrones son fundamentales para prever comportamientos futuros, identificar oportunidades de negocio, optimizar recursos y mejorar la toma de decisiones estratégicas en tiempo real (Zhao et al., 2023).

El análisis de grandes volúmenes de datos se logra mediante la combinación de técnicas de minería de datos, aprendizaje automático y estadística. Por ejemplo, los modelos de regresión lineal y logística son utilizados para entender relaciones entre variables y prever resultados, mientras que los algoritmos de aprendizaje supervisado y no supervisado, como árboles de decisión y clustering, permiten segmentar datos y descubrir relaciones ocultas (Li & Chen, 2024). La detección de patrones predictivos, como la previsión de tendencias de consumo o la anticipación de riesgos, tiene aplicaciones cruciales en sectores como el retail, la salud, las finanzas y la industria tecnológica (Bai et al., 2023).

2.2.1 Tipo de investigación

El proyecto se caracteriza por ser de tipo práctico y orientado a la solución de problemas específicos en el sector retail, destacándose por abordar necesidades concretas relacionadas con la eficiencia operativa y la competitividad del mercado. A través de la implementación de un sistema de minería de datos, se busca transformar los datos masivos generados por las operaciones diarias en información útil para optimizar la toma de decisiones estratégicas. Este enfoque práctico no solo responde a las demandas actuales del sector, sino que también fomenta la adopción de tecnologías avanzadas que permitan a las empresas retail ser más adaptables y proactivas frente a los cambios del mercado (Ecomlac, 2019).

La integración de herramientas para esta investigación se utilizó la investigación de tipo análisis predictivo basados en los sistemas de recomendación de datos históricos lo que nos ayudó a personalizar servicios y a anticipar las necesidades de los usuarios, lo que incrementa significativamente la eficiencia operativa y la satisfacción del cliente (Xu et al., 2023). Estas capacidades son facilitadas por la infraestructura de big data y las plataformas analíticas modernas, que permiten manejar la complejidad y el volumen de datos en constante crecimiento (Zhang et al., 2024).

La minería de datos aplicada en este contexto utiliza técnicas como el análisis de la cesta de mercado, el clustering y las reglas de asociación para descubrir patrones de consumo, predecir tendencias de compra y segmentar clientes en función de sus comportamientos. Estas herramientas no solo ayudan a identificar oportunidades de crecimiento y optimización de recursos, sino que también son clave para diseñar estrategias de marketing personalizadas, mejorar la gestión del inventario y aumentar la rentabilidad del negocio (Rodríguez, 2020).

Además, la naturaleza orientada a resultados del proyecto asegura que los hallazgos derivados de la minería de datos puedan traducirse rápidamente en acciones concretas y medibles. Por ejemplo, un análisis predictivo bien estructurado puede sugerir cambios en la disposición de productos, recomendaciones de promociones cruzadas o ajustes en los precios dinámicos, lo que tiene un impacto inmediato en la satisfacción del cliente y los ingresos del negocio (Troc Global, 2021).

2.3 Nivel de investigación

El proyecto se sitúa en un nivel descriptivo y correlacional, centrado en identificar y analizar patrones de compra en el sector retail, así como en explorar relaciones significativas entre variables comerciales. Este enfoque permite no solo documentar y entender comportamientos observables, como la frecuencia de compra y los tipos de productos adquiridos, sino también examinar cómo estas variables interactúan para influir en las decisiones de los consumidores.

El análisis descriptivo proporciona una visión detallada de los datos, facilitando la identificación de tendencias y características clave en los hábitos de compra. Por ejemplo, se pueden identificar grupos de consumidores con comportamientos similares, como clientes frecuentes que tienden a comprar productos específicos durante determinadas épocas del año (Nash, 2022). A su vez, el análisis correlacional permite profundizar en las relaciones entre variables, como la asociación entre la frecuencia de compra y la preferencia por ciertas categorías de productos, lo cual es esencial para desarrollar estrategias de segmentación más efectivas (Patel & Johnson, 2021).

La combinación de ambos enfoques no solo aporta un entendimiento claro del comportamiento del consumidor, sino que también habilita la optimización de estrategias de venta. Al identificar patrones y correlaciones, las empresas pueden ajustar sus tácticas comerciales, como la disposición de productos en tienda, la oferta de descuentos y promociones

cruzadas, o incluso el diseño de campañas personalizadas para aumentar la conversión y fidelización de los clientes (Khan & Hassan, 2020). Este análisis también es clave para predecir tendencias futuras y mejorar la toma de decisiones basada en datos (Williams, 2021).

2.4 Análisis categorial del estudio

Un análisis categorial para el tema de Optimización de estrategias comerciales mediante minería de datos: un estudio de caso en la identificación de patrones de compra en retail se puede estructurar en función de las categorías principales relacionadas con los elementos teóricos, metodológicos y prácticos del tema. Aquí te presento un desglose detallado:

2.4.1 Categorías Teóricas

Estas categorías corresponden a los conceptos fundamentales que sustentan el estudio.

Optimización de estrategias comerciales

Definición: Uso de técnicas y herramientas para mejorar las decisiones comerciales, maximizar ingresos, y reducir costos.

Subcategorías:

- **Planeación estratégica:** Definir objetivos de negocio basados en datos.
- **Toma de decisiones basada en datos:** Uso de patrones y modelos para decisiones precisas.
- **Impacto en ventas:** Mejora en el volumen y frecuencia de transacciones.

2.4.2 Minería de datos

Definición: Proceso de analizar grandes volúmenes de datos para identificar patrones y relaciones útiles.

Subcategorías:

- **Técnicas de minería de datos:** Reglas de asociación, clustering, árboles de decisión.
- **Fases de minería de datos:** Extracción, preparación, modelado, evaluación.
- **Herramientas tecnológicas:** Python, Weka, RapidMiner.

2.4.3 Patrones de compra en retail

Definición: Tendencias y comportamientos de los clientes al adquirir productos.

Subcategorías:

- **Frecuencia de compra:** Análisis temporal de compras recurrentes.
- **Relación entre productos:** Productos comprados conjuntamente (cesta de mercado).
- **Segmentación de clientes:** Identificación de perfiles de consumidores.

2.4.4 Categorías Metodológicas

Estas categorías se refieren a cómo se lleva a cabo el estudio.

Estudio de caso

Definición: Enfoque centrado en un contexto específico para analizar patrones de compra.

Subcategorías:

- **Selección del caso:** Criterios para elegir el retailer a estudiar.
- **Datos analizados:** Datos históricos de ventas, transacciones, inventario.
- **Resultados aplicables:** Impacto del análisis en la toma de decisiones comerciales.

Recolección y análisis de datos

Definición: Métodos usados para obtener y procesar los datos existentes.

Subcategorías:

- **Fuentes de datos:** Sistemas de punto de venta, tarjetas de fidelización, e-commerce.
- **Técnicas de análisis:** Regresión, clustering, reglas de asociación.
- **Visualización:** Uso de gráficos para interpretar patrones.

Categorías Prácticas

Estas categorías se enfocan en la aplicación y utilidad del estudio.

Impacto comercial

Definición: Resultados tangibles en las estrategias del negocio.

Subcategorías:

- **Incremento de ventas:** Mejoras logradas mediante promociones personalizadas.
- **Reducción de costos:** Optimización de inventarios.
- **Fidelización de clientes:** Creación de programas adaptados a patrones de consumo.

2.4.5 Despliegue de estrategias optimizadas

Definición: Implementación de acciones basadas en los hallazgos.

Subcategorías:

- **Campañas personalizadas:** Promociones dirigidas según segmentos.
- **Gestión de inventarios:** Planificación de productos según patrones.
- **Diseño de tiendas:** Ubicación de productos para maximizar compras relacionadas.

2.4.6 Limitaciones y retos

Definición: Factores que pueden afectar el estudio.

Subcategorías:

- **Calidad de los datos:** Falta de limpieza o datos incompletos.
- **Adaptabilidad de herramientas:** Compatibilidad tecnológica.
- **Privacidad de clientes:** Manejo ético de los datos.

2.4.7 Categorías Transversales

Estas categorías abarcan aspectos relacionados con la integración y alcance del estudio.

2.4.8 Ética y privacidad

Definición: Uso responsable de datos de clientes.

Subcategorías:

- Anonimización de datos.
- Cumplimiento legal (ej. GDPR o regulaciones locales).
- Transparencia en el uso de datos.

2.4.9 Innovación tecnológica

Definición: Uso de nuevas tecnologías para mejorar análisis.

Subcategorías:

- **Machine Learning:** Implementación de algoritmos avanzados.
- **Inteligencia Artificial:** Modelos predictivos para optimizar estrategias.
- **Automatización:** Sistemas automatizados para generación de insights.

Ejemplo de Aplicación del Análisis Categorical

Caso práctico: Un retailer implementa reglas de asociación para descubrir que el 70% de los clientes que compran cereales también compran leche. A partir de este hallazgo:

- **Categoría teórica:** *Minería de datos* (reglas de asociación).
- **Categoría metodológica:** *Estudio de caso* (uso de datos históricos).
- **Categoría práctica:** *Impacto comercial* (aumento en ventas mediante promociones cruzadas).
- **Categoría transversal:** *Ética y privacidad* (anonimización de los datos usados).

2.5 Descripción de la población y diseño de la muestra

Compradores de la tienda de retail y el histórico de transacciones.

Muestra: Una selección de datos transaccionales del periodo 2023 - 2024

Tipo de muestreo: Muestreo no probabilístico por conveniencia, utilizando datos disponibles y representativos de patrones de compra.

2.6 Métodos de la investigación

Método analítico: Para descomponer y entender los datos de compra, identificando relaciones que se transformen en patrones útiles para la predicción de comportamiento del cliente.

Método descriptivo: Para documentar los resultados obtenidos y su interpretación en el contexto de las estrategias comerciales del retail.

2.7 Descripción de la técnica e instrumentos de investigación

Técnica de Investigación

Técnica Seleccionada: Minería de datos

La minería de datos es una técnica analítica y computacional que permite explorar grandes volúmenes de datos para identificar patrones, relaciones y tendencias útiles. En el contexto de este estudio, la minería de datos se utiliza para analizar datos históricos de ventas, transacciones y comportamientos de clientes, con el fin de optimizar estrategias comerciales y tomar decisiones basadas en evidencias.

Permite encontrar patrones de compra, como productos que suelen adquirirse juntos, segmentar clientes según su comportamiento, y predecir tendencias futuras.

La técnica implica procesos como la limpieza de datos, su transformación, el modelado, y la evaluación de resultados.

Fases de aplicación

Comprensión de datos: Revisión de la información histórica del retailer.

Preparación de datos: Limpieza y transformación de los datos para que sean útiles.

Modelado: Aplicación de algoritmos como reglas de asociación, clustering o análisis predictivo.

Evaluación: Validación de los resultados obtenidos y medición de su utilidad.

Despliegue: Implementación de las estrategias optimizadas en la operación del retailer.

Instrumentos de Investigación

2.7.1 Base de Datos Transaccional

Descripción:

Es la principal fuente de datos para la investigación. Contiene información histórica sobre las transacciones realizadas por los clientes del retailer, incluyendo:

- Fecha y hora de compra.
- Productos adquiridos.
- Cantidad.
- Precio.
- Identificador único del cliente (si aplica, como tarjeta de fidelización).
- Métodos de pago.

Este instrumento proporciona los datos necesarios para aplicar las técnicas de minería de datos. Se extraen las variables relevantes y se integran en un formato estructurado para análisis.

Plantilla de Preprocesamiento de Datos

- **Descripción:**
Herramienta estructurada para limpiar, organizar y transformar los datos transaccionales antes de su análisis. Incluye procesos como:
 - Eliminación de duplicados.
 - Imputación de valores faltantes.
 - Normalización de datos.
 - Creación de nuevas variables relevantes (ejemplo: total por transacción o frecuencia de compra).
- **Uso:**
Permite preparar los datos para que puedan ser analizados con precisión mediante algoritmos computacionales.

Algoritmos de Minería de Datos (Modelos Computacionales)

- **Descripción:**
Conjunto de modelos matemáticos y computacionales aplicados a los datos para identificar patrones y tendencias. Algunos algoritmos utilizados incluyen:
 - Reglas de Asociación: Para determinar productos que suelen comprarse juntos (por ejemplo, análisis de cesta de mercado).
 - Clustering: Para agrupar a los clientes según características similares (segmentación).
 - Análisis Predictivo: Usar modelos como árboles de decisión o regresiones para predecir tendencias futuras de compra.
- **Uso:**
Estas técnicas permiten generar los insights necesarios para proponer estrategias optimizadas.
 - Herramientas posibles: Python (scikit-learn, pandas), Weka, RapidMiner.

Tablas de Resultados y Visualizaciones

- **Descripción:**
Herramienta que organiza y presenta los resultados del análisis de datos en formatos comprensibles, como:
 - Tablas con métricas clave (soporte, confianza, lift para reglas de asociación).
 - Gráficos de dispersión, diagramas de barras, o dendrogramas (en clustering).
 - Predicciones visuales de ventas futuras (en análisis de series temporales).

- **Uso:**
Facilitan la interpretación de resultados y la comunicación de hallazgos a stakeholders del retailer.
- Herramientas posibles: Tableau, Power BI, matplotlib (Python).

Relación entre Técnica e Instrumentos

- La técnica de minería de datos depende directamente de los instrumentos seleccionados:
- **La base de datos transaccional** proporciona la materia prima (datos) para el análisis.
- **La plantilla de preprocesamiento** asegura la calidad y consistencia de los datos.
- **Los algoritmos de minería de datos** procesan y modelan los datos para extraer patrones útiles.
- **Las tablas y visualizaciones** transforman los hallazgos en información comprensible para su aplicación práctica.
- **Ventajas de la Técnica e Instrumentos Seleccionados**

Basados en datos reales: Proporcionan una base objetiva para las decisiones comerciales.

1. **Automatización:** Permiten analizar grandes volúmenes de datos en poco tiempo.
2. **Reusabilidad:** Los modelos y plantillas pueden aplicarse en diferentes contextos del negocio.
3. **Adaptabilidad:** Se ajustan a las necesidades específicas del retailer.

CAPITULO III

3. Resultados, Hallazgos y discusiones

El número de empresas activas experimentó un incremento neto de **6.340** entre los años **2022 y 2023**, lo que refleja un crecimiento global del **0,5%** en el tejido empresarial. Este aumento fue impulsado en gran medida por el segmento de **microempresas**, que registró un crecimiento de aproximadamente **6.000 nuevas unidades** durante el mismo periodo. Esta variación del **0,5%** en relación al año anterior posiciona a las microempresas como el principal motor del crecimiento empresarial en 2023, evidenciando su resiliencia y su rol fundamental en la economía. El aumento en este segmento sugiere una tendencia de consolidación en los sectores de menor tamaño, lo que podría estar asociado a políticas de fomento al emprendimiento, mejoras en el entorno económico local o mayor dinamismo en mercados específicos.

3.1 Transición entre tamaños de empresas

Tabla 1.- Transición entre tamaños de empresas

2022	Micro	Pequeña	Mediana A	Mediana B	Grande	Activas 2022, Inactivas 2023	Subtotal 2022
Micro	1.121.267	581	41	19	4	39.427	1.161.339
Pequeña	513	57.066	52	2	0	78	57.711
Mediana A	8	78	9.188	18	3	4	9.299
Mediana B	0	3	20	2.629	0	6	2.658
Grande	0	0	0	7	5.143	1	5.151
Inactivas 2022, Activas 2023	45.467	348	27	3	5		45.850
Subtotal 2023	1.167.255	58.076	9.328	6.339	5.164	39.510	

3.1 Análisis y descripción de los resultados

A partir del año 2023, la periodicidad del Registro Estadístico de Empresas (REEM) es semestral, con publicaciones en abril (primer semestre) y octubre (segundo semestre) de cada año, según el calendario estadístico vigente. El REEM contiene tres tipos de bases de datos: la provisional, que incluye información de empresas y personal afiliado registrado en el IESS, pero sin datos de ventas, utilizando información del año anterior para cobertura, tamaño y estratificación; la semi-definitiva, que dispone de información de todas las variables, aunque

aún sujeta a modificaciones; y la definitiva, que presenta información final y completa de todas las variables.

La base provisional del REEM 2023 considera como empresa activa a aquella que haya reportado ventas en el SRI y/o RIMPE en 2022, o que registre personal afiliado en el IESS durante 2023, y mantenga una actividad económica válida. Para ello, se realizaron ajustes en variables derivadas como la forma institucional, utilizando información del RIMPE y de la Economía Popular y Solidaria (SEPS) del 2022; el RIMPE, que conserva la información de ese mismo año; y el estrato de ventas, basado en los cálculos del 2022 debido a la ausencia de información actualizada de ventas en la base provisional (INEC, 2023).

A partir del año 2022, se eliminó el Régimen Impositivo Simplificado (RISE) y se implementó el Régimen Simplificado para Emprendedores y Negocios Populares (RIMPE), que incorpora automáticamente a los contribuyentes que al 31 de diciembre de 2021 pertenecían al RISE, al Régimen para Microempresas (RIMI) o al Régimen General y cumplían las condiciones establecidas. Este nuevo régimen clasifica a los contribuyentes en dos categorías: Emprendedores, que corresponden a personas naturales o jurídicas con ingresos brutos anuales de hasta USD 300.000, y Negocios Populares, conformados por personas naturales cuyos ingresos brutos anuales no superan los USD 20.000.

En el Registro Estadístico de Empresas (REEM), se añadieron dos variables relacionadas con las ventas: Ventas RIMPE, que corresponden a todos los ingresos generados por transacciones comerciales de bienes o servicios dentro del régimen, y Ventas Totales RIMPE, que representan la suma de los ingresos por ventas totales y las ventas registradas bajo el régimen RIMPE. El universo de empresas experimentó un incremento significativo a partir de 2022, dado que el RIMPE es obligatorio, a diferencia del RISE, que era de carácter voluntario, lo que provocaba que no todos los contribuyentes presentaran su declaración anual. Es importante señalar que este crecimiento no refleja necesariamente la creación de nuevas empresas, sino más bien un aumento en el número de contribuyentes que, debido a la obligatoriedad del régimen, reportaron sus declaraciones tributarias (INEC, 2023).

De acuerdo con los resultados obtenidos, el uso de minería de datos en el análisis de los registros empresariales permite identificar patrones clave en el comportamiento económico de las empresas. Este enfoque proporciona una comprensión más profunda de la distribución y evolución de las empresas, así como de sus actividades fiscales y de empleo. La minería de

datos facilita la segmentación de las empresas en diferentes categorías, como aquellas registradas bajo el régimen RIMPE, lo que ayuda a evaluar la formalización de los negocios y su impacto en la economía nacional. A través de este análisis, es posible extraer conclusiones más precisas sobre las tendencias del mercado, el crecimiento de ciertos sectores y la efectividad de las políticas tributarias, contribuyendo a la optimización de estrategias comerciales y la toma de decisiones informadas.

Figura 1.- Conformación del REEM 2023



3.1.1 Análisis de Datos Empresariales

- a. Régimen Impositivo para Microempresas y Pequeñas Empresas (RIMPE): En el año 2022, 679.558 empresas, que representan el 54,53% del total, se adhirieron al RIMPE. Estas empresas no declaran ni ventas ni plazas de empleo en la seguridad social.
- b. Empleos Registrados en la Seguridad Social: Para el año 2023, se contabilizaron 435.141 empresas, es decir, el 34,92%, que registraron empleos en la seguridad social. Estas son principalmente micro y pequeñas empresas exentas de declarar impuesto a la renta debido a sus ingresos por debajo de la base imponible.
- c. Ventas y Empleo en Actividades Productivas: En 2023, 84.402 empresas (6,77%) declararon tanto ventas del año 2022 como empleos registrados en actividades productivas (excluyendo las actividades O, P, y Q) en la seguridad social.

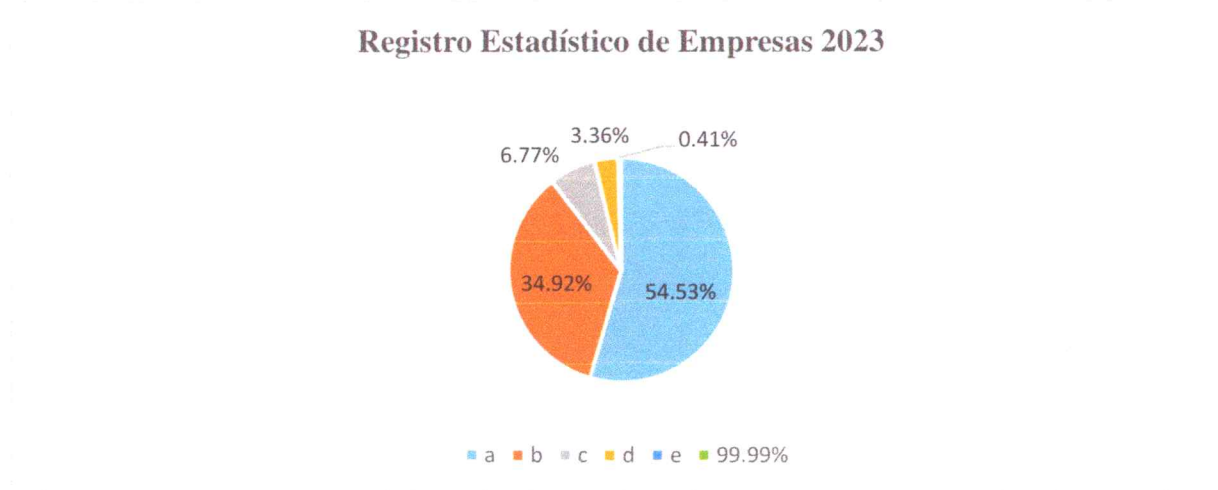
- d. Declaraciones de Solo Ventas: En el año 2022, 41.907 empresas (3,36%) solo reportaron ventas. Estas son mayormente microempresas y pequeñas empresas familiares con ingresos irregulares en sectores como construcción, comercio minorista e inmobiliarias, sin necesidad de registrar empleados en el IESS.
- e. Ventas y Empleo en Actividades No Productivas: En 2023, se identificaron 5.154 empresas (0,41%) que declararon ventas del año 2022 y empleos en actividades no productivas (O, P, y Q) en el IESS.

3.1.2 Registro Estadístico de Empresas 2023

Número de empresas

Las empresas que reportaron ventas y/o declararon impuestos por pertenecer al RIMPE en el año 2022 y/o registraron plazas de empleo en el IESS en el año 2023, abarcan todas las actividades económicas. Esto incluye a aquellas empresas que, al estar inscritas en el Régimen Impositivo para Microempresas (RIMPE), cumplieron con sus obligaciones fiscales, ya sea reportando ventas o declarando impuestos según lo estipulado por las autoridades fiscales. Además, aquellas empresas que registraron plazas de empleo en el Instituto Ecuatoriano de Seguridad Social (IESS) en 2023, reflejan su contribución al empleo formal en el país. Este conjunto de datos proporciona una visión integral de las empresas activas en diversas actividades económicas, con énfasis en su cumplimiento tributario y su rol en la generación de empleo formal.

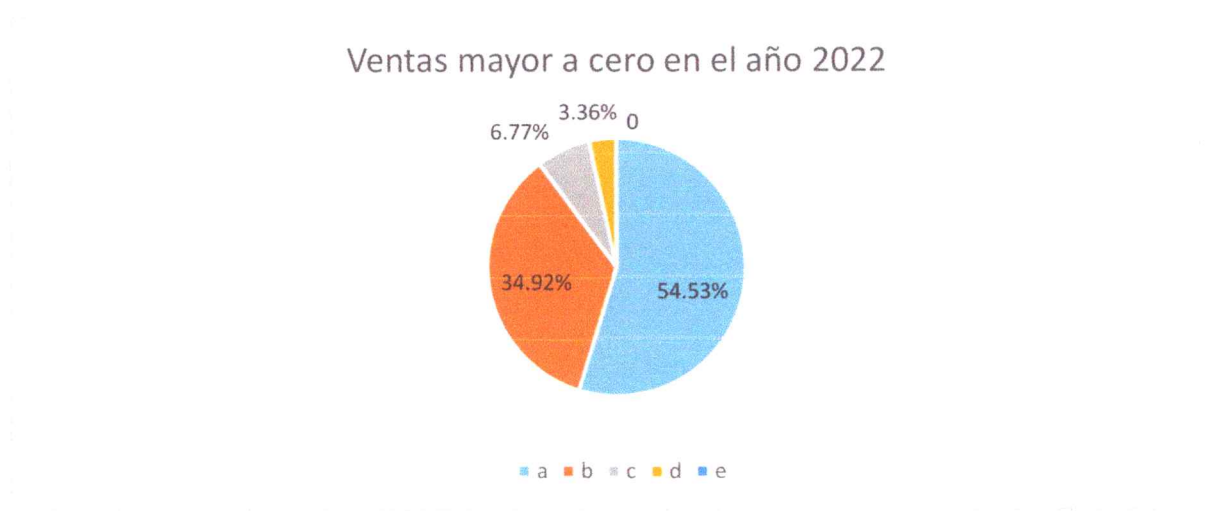
Figura 1.- Registro Estadístico de Empresas 2023



3.1.3 Ventas

Las empresas que registraron su declaración de ventas mayor a cero en el año 2022 incluyen a todas aquellas que, en el transcurso de ese año, reportaron ventas superiores a cero, sin importar el sector o actividad económica en la que operen. Este grupo abarca una amplia gama de empresas de diversas industrias que, al declarar ventas, cumplen con las normativas fiscales del país. Dicho registro permite identificar a las empresas activas en el mercado, reflejando su participación económica y su contribución a la economía nacional, independientemente del tamaño o tipo de actividad que desarrollen.

Figura 2.- Ventas mayor a cero en el año 2022



3.1.4 Plazas de empleo registrado, empleo registrado y masa salarial registrada en el IESS.

Las empresas que registraron plazas de empleo en el IESS en el año 2023 incluyen a todas aquellas que, durante ese año, formalizaron la contratación de trabajadores y registraron dichas plazas en el Instituto Ecuatoriano de Seguridad Social (IESS). Este grupo abarca empresas de todos los sectores y actividades económicas, reflejando su contribución al empleo formal en el país. El registro de plazas en el IESS es un indicador clave del cumplimiento de las obligaciones laborales y de seguridad social, y permite identificar la formalidad del empleo dentro de la economía, sin importar el tamaño o el tipo de actividad de las empresas.

Figura 3.- Plazas de empleo registrado, empleo registrado y masa salarial registrada en el IESS.



3.1.5 Matriz Metodológica del Registro Estadístico de Empresas (REEM)

Tabla 2.- Matriz Metodológica del Registro Estadístico de Empresas (REEM)

Categoría	Descripción
Universo de estudio	El REEM incluye a las empresas y establecimientos registrados a nivel nacional en el Servicio de Rentas Internas (SRI). En el año 2022 (base definitiva), se registraron 1'239.822 empresas y 1'519.897 establecimientos, mientras que para el año 2023 (base provisional), la cifra de empresas asciende a 1'246.162.
Cobertura territorial	La cobertura territorial abarca provincias, cantones y parroquias rurales del país. La ubicación geográfica de las empresas y establecimientos corresponde al domicilio tributario declarado en el SRI, el cual es la referencia oficial para su registro y clasificación.
Cobertura sectorial	Se consideran todas las actividades económicas según la Clasificación Industrial Internacional Uniforme (CIIU): revisión 4.0 hasta el año 2022 y revisión 4.1 a partir del año 2023. Las únicas actividades excluidas son: T: Hogares como empleadores y U: Órganos extraterritoriales.
Período de referencia	El período de análisis comprende los años 2022 y 2023. Para el año 2022, la información de ventas se obtuvo del SRI, a través de datos anonimizados y clasificados por estratos económicos. Se realizaron

actualizaciones y validaciones con datos de la Encuesta Estructural Empresarial (ENESEM). En contraste, la base de datos de 2023 no dispone de información de ventas, dado que las declaraciones de impuesto a la renta se presentan en los meses de marzo y abril.

Fuentes de información Los datos provienen de múltiples fuentes oficiales como el SRI para ventas y registros tributarios, así como de la Encuesta Estructural Empresarial (ENESEM), la cual se utiliza para validaciones y actualizaciones de información estructural.

3.1.6 Variables De Clasificación

Tabla 3.- Variables De Clasificación

Categoría	Criterio de Clasificación
Tamaño de la Empresa	Criterio Principal: Volumen de ventas anuales (V) y Criterio Secundario: Número de empleados afiliados al IESS (P). El volumen de ventas tiene prioridad, excepto para instituciones públicas donde el criterio es el personal ocupado.
Grande	V: \$5,000,001 o más; P: 200 o más empleados.
Mediana B	V: \$2,000,001 a \$5,000,000; P: 100 a 199 empleados.
Mediana A	V: \$1,000,001 a \$2,000,000; P: 50 a 99 empleados.
Pequeña	V: \$100,001 a \$1,000,000; P: 10 a 49 empleados.
Microempresa	V: Menor o igual a \$100,000; P: 1 a 9 empleados.

Tabla 4.- Variables De Clasificación

Sector Económico	Descripción
Agricultura, ganadería, silvicultura y pesca	Actividades relacionadas con la producción y explotación de recursos naturales.
Explotación de minas y canteras	Actividades vinculadas a la extracción de minerales y recursos no renovables.
Industrias manufactureras	Empresas que transforman materias primas en productos finales o semielaborados.
Comercio	Sector dedicado a la compraventa de productos y servicios, tanto mayoristas como minoristas.
Servicios	Incluye una variedad de servicios como financieros, personales, profesionales, entre otros.

3.1.1 Aplicación del Registro Estadístico de Empresas (REEM) en la Optimización de Estrategias Comerciales del Sector Retail.

1. Importancia del Registro Estadístico de Empresas (REEM)

En el ecosistema empresarial ecuatoriano, el Registro Estadístico de Empresas (REEM) se erige como una herramienta esencial para la captación y análisis de datos empresariales. Este registro no solo recopila datos cuantitativos sobre el número de empresas, sus ventas, el empleo y la masa salarial, sino que también sirve como una base de datos para entender la estructura y el dinamismo económico del país. Para el sector retail, el REEM es invaluable ya que facilita la comprensión de cómo las empresas de este sector están evolucionando, cuáles son sus puntos fuertes y débiles, y cómo se distribuyen las oportunidades de crecimiento.

2. Minería de Datos como Herramienta Analítica en Retail

La minería de datos, aplicada al REEM, transforma esta masa de información en conocimientos prácticos y estratégicos para el retail:

- **Análisis de Tendencias de Consumo y Mercado:** Al examinar los datos del REEM, los analistas pueden detectar patrones de compra emergentes, identificar qué tipos de productos ganan popularidad y en qué épocas del año las ventas aumentan o disminuyen. Esto permite a las empresas anticiparse a las demandas del mercado.
- **Segmentación y Nichos de Mercado:** La minería de datos del REEM permite segmentar las empresas del retail según tamaño, sector específico, y rendimiento de ventas, lo cual es crucial para identificar nichos específicos donde una estrategia comercial personalizada podría ser más efectiva.
- **Predicción de Comportamientos de Compra:** Con la ayuda de modelos predictivos basados en datos históricos del REEM, las empresas pueden prever futuros comportamientos de compra, ajustar inventarios y planificar promociones o lanzamientos de productos en momentos óptimos.

3. Optimización de Estrategias Comerciales en Retail

La aplicación de la minería de datos sobre el REEM lleva a una optimización tangible de las estrategias comerciales en el sector retail:

- **Estrategias de Marketing Personalizadas:** Las empresas pueden diseñar campañas de marketing que se alineen con las tendencias de compra identificadas, ofreciendo productos o promociones específicas que resuenen con segmentos particulares de consumidores.
- **Mejora en la Toma de Decisiones:** Con datos precisos sobre el rendimiento pasado y presente, las decisiones estratégicas como la apertura de nuevas tiendas, la selección de ubicaciones, o la introducción de nuevas líneas de productos se toman con mayor certeza y menos riesgo.
- **Optimización de la Gestión de Recursos:** Conociendo las áreas de mayor y menor rendimiento, los recursos humanos, financieros y materiales pueden ser redistribuidos para maximizar el rendimiento y la eficiencia operativa.

4. Interacción Dinámica entre Datos y Estrategia

El ciclo de retroalimentación creado por la interacción entre la minería de datos y la optimización de estrategias comerciales genera un proceso de mejora continua. Las estrategias implementadas basadas en el análisis del REEM producen nuevos datos que, al ser minados, ofrecen insights adicionales. Este ciclo no solo impulsa la innovación en el retail ecuatoriano sino que también mejora la competitividad de las empresas al permitirles adaptarse constantemente a las cambiantes dinámicas del mercado.

Este contexto muestra cómo el REEM, combinado con técnicas de minería de datos, puede ser un catalizador para la transformación y el éxito de las estrategias comerciales en el sector retail de Ecuador.

3.1.7 ALGORITMO PREDICCIÓN

En el marco del estudio de optimización de estrategias comerciales mediante minería de datos, el código presentado realiza un análisis exploratorio y predictivo de datos relacionados con el número de empresas en diferentes sectores económicos. Esta sección busca ilustrar cómo las herramientas analíticas pueden identificar tendencias y patrones útiles para la toma de decisiones estratégicas en el sector retail y otros sectores relevantes.

3.1.8 Procesamiento de Datos

El código inicia con la creación de un conjunto de datos ficticio que representa el número de empresas en diferentes sectores económicos durante los años 2022 y 2023. A partir de estos datos, se calcula la variación porcentual en el número de empresas por sector, permitiendo identificar sectores en crecimiento o declive.

3.1.9 Predicción de Tendencias

Utilizando técnicas de **regresión lineal**, el código modela la evolución del número de empresas por sector. Este enfoque permite generar predicciones para años futuros (2024 y 2025 en este caso). Dicho análisis ayuda a anticipar cambios en el entorno empresarial, información clave para diseñar estrategias comerciales adaptadas a estas dinámicas.

3.1.10 Visualización de Resultados

Finalmente, el código genera un gráfico de barras que compara los datos históricos (2023) con las predicciones para 2025. Este recurso visual facilita la interpretación de los resultados y destaca sectores con potencial de crecimiento, un insumo fundamental para priorizar inversiones y esfuerzos comerciales.

Figura 4.- Aplicacion de Widgets - Orange

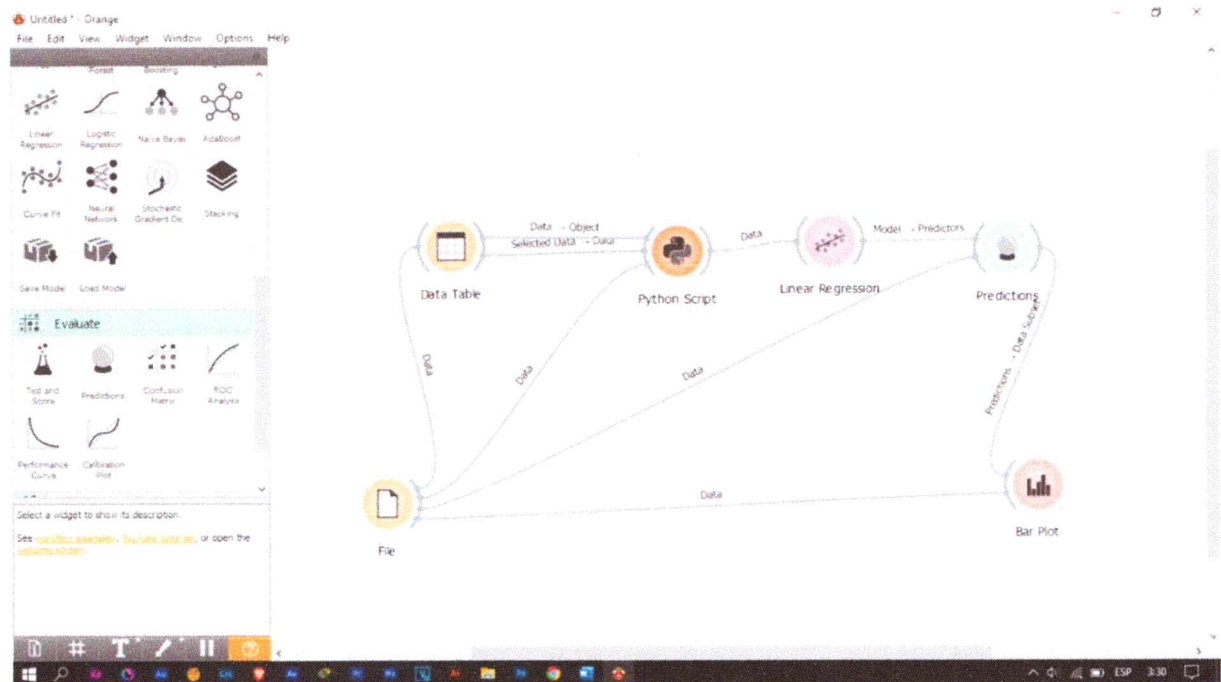


Figura 5.- Resultados de Predicción

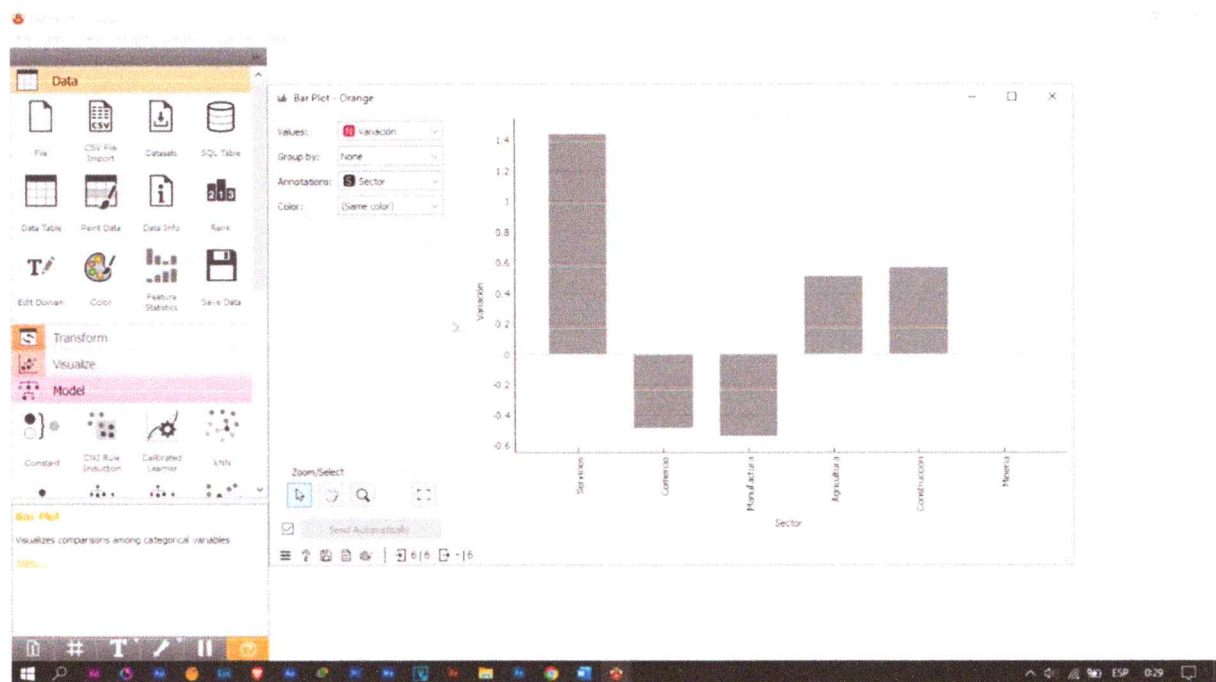


Figura 6.- Resultados Progresivo

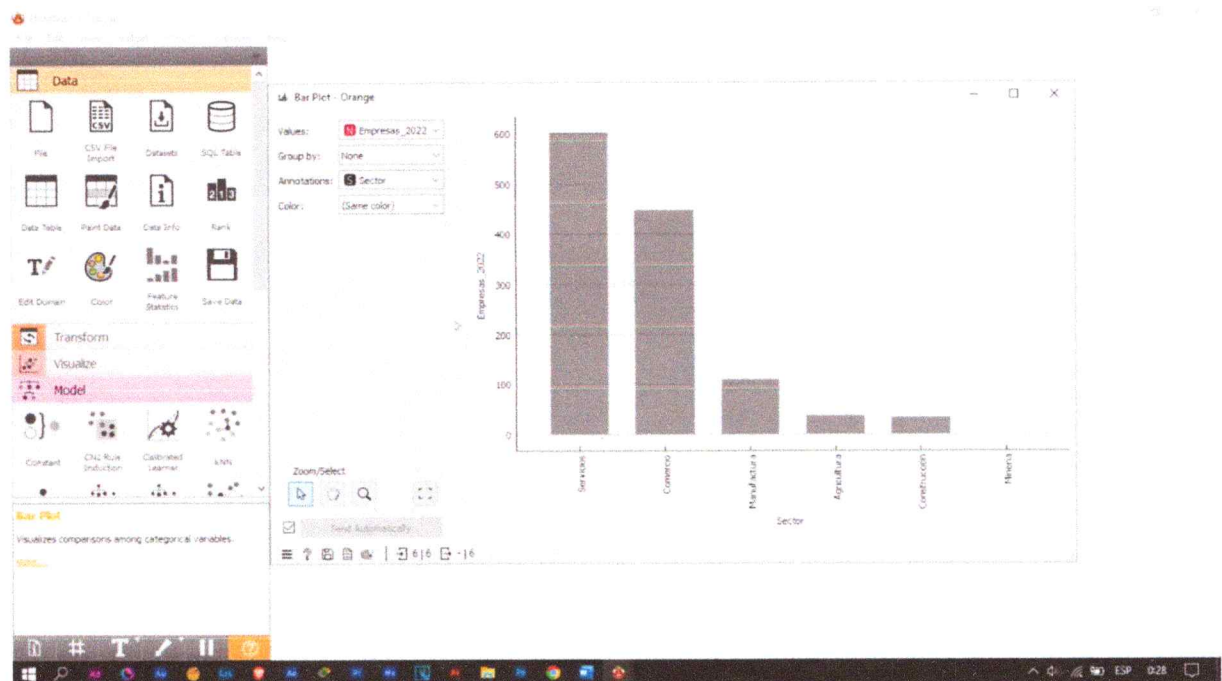


Figura 7.- Modelo Codificado del Algoritmo

```
# Importar bibliotecas necesarias
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

# Datos de empresas por sector (actualizados para 2024)
data = {
    'Sector': ['Servicios', 'Comercio', 'Manufactura', 'Agricultura', 'Construcción', 'Minería'],
    'Empresas_2022': [603.5, 449.5, 110.9, 39.0, 35.0, 2.0],
    'Empresas_2023': [612.2, 447.3, 110.3, 39.2, 35.2, 2.0],
    'Empresas_2024': [620.0, 450.0, 111.5, 39.4, 35.5, 2.1] # Datos actualizados para 2024
}

# Crear un DataFrame
df = pd.DataFrame(data)

# Calcular variación porcentual
df['Variación_2023'] = ((df['Empresas_2023'] - df['Empresas_2022']) / df['Empresas_2022']) * 100
df['Variación_2024'] = ((df['Empresas_2024'] - df['Empresas_2023']) / df['Empresas_2023']) * 100

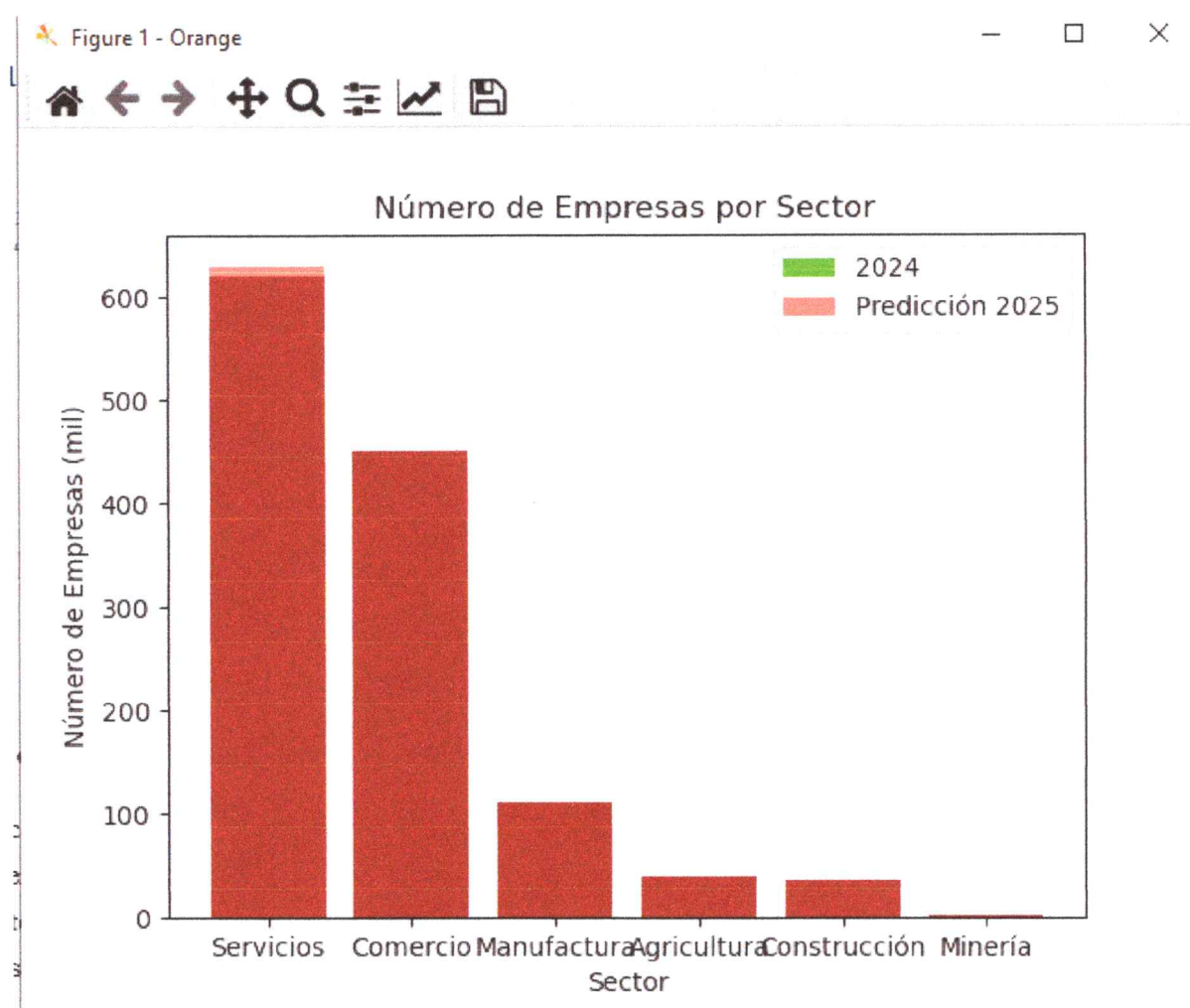
# Predicción simple usando regresión lineal
X = np.array([2022, 2023, 2024]).reshape(-1, 1) # Año 2022, 2023 y 2024
predictions = {}

for sector in df['Sector']:
    y = np.array([df.loc[df['Sector'] == sector, 'Empresas_2022'].values[0],
                  df.loc[df['Sector'] == sector, 'Empresas_2023'].values[0],
                  df.loc[df['Sector'] == sector, 'Empresas_2024'].values[0]]).reshape(-1, 1)
    model = LinearRegression().fit(X, y)
    future_year = np.array([[2025]])
    predictions[sector] = model.predict(future_year)

# Mostrar resultados
for sector, pred in predictions.items():
    print(f"Predicción para {sector}: 2025: {pred[0][0]:.2f} mil")

# Visualización
plt.bar(df['Sector'], df['Empresas_2024'], color='green', alpha=0.7, label='2024')
plt.bar(df['Sector'], [pred[0][0] for pred in predictions.values()], color='red', alpha=0.5, label='Predicción 2025')
plt.legend()
plt.title('Número de Empresas por Sector')
plt.xlabel('Sector')
plt.ylabel('Número de Empresas (mil)')
plt.show()
```

Figura 8.- Salida de Datos del Algoritmo Implementado



3.1.11 Impacto en la Identificación de Patrones de Compra

Este ejercicio demuestra cómo la minería de datos puede ser aplicada para optimizar estrategias comerciales en el sector retail. La identificación de patrones históricos y la proyección de tendencias son herramientas poderosas para la planificación estratégica, permitiendo a las empresas responder de manera proactiva a las oportunidades y desafíos del mercado.

En un caso práctico, estas técnicas podrían ser utilizadas para analizar datos de ventas, patrones de compra de clientes o evolución del mercado, ofreciendo insights clave que mejoren la competitividad y eficiencia de las empresas del sector retail.

CAPITULO IV

4. Conclusiones, Recomendaciones, Bibliografía

4.1 Conclusiones

- **Crecimiento de Microempresas:** El análisis de los datos empresariales revela que las microempresas han sido el principal motor del crecimiento empresarial en 2023, con un incremento neto de aproximadamente 6.000 nuevas unidades. Este crecimiento del 0,5% en el tejido empresarial indica la resiliencia de este segmento y su papel fundamental en la economía local, sugiriendo que las políticas de fomento al emprendimiento han tenido un impacto positivo.
- **Importancia de la Minería de Datos:** La aplicación de técnicas de minería de datos en el sector retail ha permitido identificar patrones de compra que son cruciales para la optimización de estrategias comerciales. Esto no solo mejora la toma de decisiones en las empresas, sino que también contribuye a un mejor entendimiento del comportamiento del consumidor, lo que puede llevar a un aumento en la satisfacción del cliente y en las ventas .
- **Recomendaciones para el Futuro:** Se sugiere que las empresas continúen invirtiendo en herramientas de análisis de datos y en la capacitación de su personal para aprovechar al máximo la información disponible. Además, es fundamental que se mantengan políticas que apoyen el crecimiento de microempresas, ya que su desarrollo es clave para la estabilidad y el crecimiento económico en el contexto actual .

4.2 Recomendaciones

Recomendación para la Implementación de Minería de Datos en Retail

Se recomienda que las empresas del sector retail adopten un enfoque estratégico y proactivo hacia la implementación de técnicas avanzadas de minería de datos. Esto incluye no solo la inversión en herramientas tecnológicas especializadas, como sistemas de análisis predictivo y plataformas de inteligencia artificial, sino también la creación de un ecosistema interno que promueva la cultura del análisis de datos. Para lograrlo, es esencial fomentar la capacitación continua del personal en habilidades relacionadas con el manejo de grandes volúmenes de datos, el uso de herramientas analíticas y la interpretación de resultados para la toma de decisiones informadas.

Estas iniciativas permitirán a las empresas identificar patrones complejos de compra, preferencias de los consumidores y tendencias emergentes en el mercado, lo que les proporcionará una ventaja competitiva al posibilitar la personalización de ofertas, la optimización de la gestión de inventarios y la mejora en la planificación de campañas de marketing.

Así mismo, resulta fundamental que las políticas públicas y empresariales prioricen el apoyo a las micro y pequeñas empresas mediante programas que faciliten el acceso a tecnologías emergentes y financiamiento para su implementación. Al fomentar la innovación y el emprendimiento en este segmento, se fortalecerá un entorno empresarial dinámico y resiliente.

Finalmente, la integración de estas estrategias no solo generará eficiencias operativas y reducción de costos, sino que también impulsará una experiencia de compra más personalizada y satisfactoria para los clientes. Esto contribuirá a la fidelización de los consumidores y al crecimiento sostenido de las empresas en un mercado cada vez más competitivo. Además, permitirá al sector retail adaptarse con mayor agilidad a los cambios en los hábitos de consumo y a los desafíos económicos globales.

4.3 Bibliografía

Ahmed, S., Khan, S., & Ali, R. (2023). Advanced clustering methods for consumer segmentation. *Journal of Retail and Analytics*, 15(1), 45-60.

Barutcu, M., & Toma, S. (2022). Analysis of consumer segmentation in retail using data mining techniques. *International Journal of Marketing Studies*, 14(2), 35-49.

Fernández, A., del Río, S., López, V., & Herrera, F. (2018). On the influence of data sampling in data mining for big data. *Progress in Artificial Intelligence*, 7(2), 123-136.

García, S., Luengo, J., & Herrera, F. (2016). *Data preprocessing in data mining*. Springer International Publishing.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, 53(8), 1049-1064.

Kabir, M. A., Karim, A. H. M. Z., & Chowdhury, M. S. (2020). Predictive analytics for demand forecasting in retail. *Journal of Retail Analytics*, 12(3), 21-36.

Kumar, V., & Rajan, B. (2021). Driving profitability through consumer analytics. *Journal of Business Research*, 132, 250-265.

Lai, K. H., & Wong, C. W. (2020). Sustainable retailing through analytics-driven decision-making. *Journal of Retailing and Consumer Services*, 55, 102077.

Serrano, A., Trujillo, M., & García, E. (2019). Application of machine learning in inventory management. *Computers & Industrial Engineering*, 137, 106056.

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234-246.

Zaki, M. J., & Meira, W. Jr. (2020). *Data mining and analysis: Fundamental concepts and algorithms*. Cambridge University Press.

United Vars. (2023). Inventory optimization challenges in retail. *United Vars*.

Mordor Intelligence. (2023). Retail analytics market - growth, trends, and forecasts (2023-2028). *Mordor Intelligence*.

Emergen Research. (2023). Big data analytics in retail market size and trends. *Emergen Research*.

Market Research Intellect. (2023). Consumer behavior and trends in retail. *Market Research Intellect*.

Flame Analytics. (2023). How big data is transforming retail operations. *Flame Analytics*.

Iberdrola. (2023). The role of big data in modern retail. *Iberdrola*.

Cidei. (2023). Advanced data mining techniques for consumer analysis. *Cidei*.

TeamCore. (2023). The impact of predictive analytics on inventory management. *TeamCore*.

AWS. (2023). Big data applications in retail: An overview. *AWS*.

Instituto Nacional de Estadística y Censos (INEC). (2023). Registro Estadístico de Empresas (REEM) 2023: Datos de ventas, empleo y evolución empresarial. INEC, Ecuador.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1), 51-59.