

APLICACIÓN DE ANALÍTICA DE DATOS PARA OPTIMIZAR VENTAS Y EXPERIENCIA DE CONSUMO EN TIENDAS DE BARRIO

USING DATA ANALYTICS TO OPTIMIZE SALES AND CUSTOMER EXPERIENCE IN NEIGHBORHOOD STORES

Ena Yuritze Barón López¹; Jefferson Olave Rubio²

Universidad Central Colombia

ebaronl@ucentral.edu.co¹ ; jolaver@ucentral.edu.co²

Ena Yuritze Barón López: <https://orcid.org/0000-0002-2659-4500>

Jefferson Olave Rubio: <https://orcid.org/0009-0006-9678-439X>

Recibido: 22-feb-2026

Aceptado: 28-abr-2026

Códigos Clasificación JEL M10, M15, M21

Resumen

Esta investigación tiene como objetivo desarrollar e implementar herramientas de recomendación de productos en analítica de datos para mejorar ventas y experiencias de consumo en tiendas de barrio en Colombia. El estudio se centra en Supermercado La Comarca, ubicado en el barrio Bonanza, considerando las limitaciones tecnológicas y desconfianza hacia los pagos digitales presentes en este tipo de negocios. La propuesta utiliza registros del sistema POS mediante arquitectura orientada al ciclo de vida de los datos, desde su extracción hasta su transformación en información útil para la toma de decisiones. A partir de estos datos se implementó un algoritmo de recomendación basado en similitud de cosenos, integrado a un chatbot capaz de generar ofertas personalizadas según el historial de compras. El análisis se fundamenta en el enfoque Market Basket Analysis y reglas de asociación para identificar patrones de consumo y sugerir productos complementarios, fortaleciendo la fidelización del cliente.

Palabras Clave: Comercio Tradicional Minorista, Chatbot, MBA, Reglas de Asociación, Algoritmo de Recomendación.

Abstract

The objective of this research is to develop and implement product recommendation tools using data analytics to improve sales and customer experiences at neighborhood stores in Colombia. The study focuses on Supermercado La Comarca, located in the Bonanza neighborhood, taking into account the technological limitations and mistrust of digital payments present in this type of business. The proposal utilizes POS system records through a data lifecycle-oriented architecture, from data extraction to its transformation into useful information for decision-making. Based on this data, a recommendation algorithm using cosine similarity was implemented and integrated into a chatbot capable of generating personalized offers based on purchase history. The analysis is based on the Market Basket Analysis approach and association rules to identify consumption patterns and suggest complementary products, thereby strengthening customer loyalty.

Keywords: Traditional Retail Commerce, Chatbot, MBA, Association Rules, Recommendation Algorithm.

Introducción

El sector minorista en Colombia, especialmente las tiendas de barrio o canal tradicional, ha sido esencial en el desarrollo económico de muchas familias y sigue siendo una opción preferida por los consumidores debido a su cercanía y confianza con los tenderos, se calcula que en Colombia hay 450.000 tiendas (Fenalco, 2021). Estos negocios proporcionan empleo a miles de colombianos, incluso después de la pandemia, este es el medio que más personas consideraron como una opción para desarrollarse económicamente. Sin embargo, enfrenta desafíos significativos debido a la competencia de otros formatos minoristas y a la lenta adopción de tecnologías, lo que lo hace vulnerable frente a tiendas de descuento y otros competidores.

El avance tecnológico ha impulsado la integración de canales físicos y digitales en el sector minorista. En los últimos años, numerosos minoristas, independientemente de su tamaño, han incorporado canales en línea, consolidándose como minoristas multicanal. Asimismo, se evidencia un creciente interés de las marcas por fortalecer su presencia digital (Frasquet et al., 2015; Schoenbachler & Gordon, 2002). Actualmente, los negocios multicanal evolucionan hacia la omnicanalidad, entendida como la integración total de los canales de venta y comunicación con el consumidor.

Esta transformación exige que las empresas desarrollen capacidades en áreas estratégicas como: (1) tecnologías y herramientas para facilitar la toma de decisiones; (2) visualización y gestión de la oferta de mercancías; (3) consumo y participación del cliente; (4) recopilación y aprovechamiento de grandes volúmenes de datos; y (5) análisis y rentabilidad. Adicionalmente, resulta relevante considerar el potencial de tecnologías emergentes como el Internet de las cosas, la realidad virtual, la realidad aumentada, la inteligencia artificial, los robots, los drones y los vehículos autónomos (Grewal et al., 2016).

El comercio electrónico ha generado nuevas oportunidades para que los minoristas recopilen y analicen datos de los clientes mediante plataformas digitales. Esto permite comprender los comportamientos de consumo, identificar tendencias y ofrecer recomendaciones personalizadas, convirtiéndose en una herramienta clave para la toma de decisiones y la mejora de la experiencia de compra (Aversa et al., 2021). No obstante, aunque las tiendas de barrio continúan siendo competidores relevantes dentro del consumo masivo, su sostenibilidad futura depende de su capacidad de modernización tecnológica, ya que enfrentan el riesgo de rezago si no se adaptan a las nuevas dinámicas digitales (La República, 2019).

En la actualidad, las tiendas de barrio no aprovechan plenamente los beneficios de la tecnología y de la analítica de datos para mejorar sus ventas, expandir sus negocios y brindar mayores beneficios a sus clientes. Factores como los aspectos culturales de consumidores y tenderos, el desconocimiento en el uso de herramientas tecnológicas y la preferencia arraigada por el dinero en efectivo limitan su avance hacia la adopción de medios digitales (La República, 2021).

En este contexto, el presente proyecto se orienta a la aplicación de herramientas de analítica de datos en beneficio del comercio minorista tradicional de barrio. A partir del modelamiento de datos relacionados con ventas, productos facturados y datos de contacto para domicilios, se busca desarrollar una herramienta de recomendación y segmentación que facilite tanto al tendero como al consumidor los procesos de compra y venta digital. Asimismo, la propuesta permitirá generar recomendaciones de compra basadas en el historial de consumo de los clientes. El proyecto se desarrolla a partir del caso del Supermercado La Comarca, ubicado en el barrio Bonanza.

Revisión de la literatura y contexto

Para las marcas minoristas, los datos se han vuelto cruciales para crear experiencias personalizadas y generar valor para las organizaciones. El valor de los datos es ahora mayor que los activos físicos y porque el uso estratégico de la analítica ha demostrado mayor rentabilidad para las empresas minoristas. La capacidad de transformar datos en acciones concretas ofrece una ventaja competitiva significativa y mejora los resultados financieros en el sector.

Uso y estudio de la analítica de datos en el sector minorista

La literatura también muestra investigaciones sobre el uso de la analítica en el sector minorista. García y Rodríguez (2019) propusieron un diseño integral de una solución digital para innovar el modelo de negocio de la industria del retail integrando analítica, cloud computing e internet. Otra de las investigaciones encontradas, propone un diseño modelo que combina tecnologías con Cloud Computing, con el fin de eliminar dichas restricciones en la planeación y gestión de inventario de las PYME peruanas (López & Guerrero, 2017).

Por otro lado, Iregui Sanclemente (2021) hace un análisis de los beneficios de las PYMES en Bogotá a partir de un estudio de casos. Específicamente en el análisis de modelos analíticos sobre tiendas de barrio se han encontrado los siguientes estudios: Gómez Betancur (2021) desarrolló un modelo para predecir las ventas de la semana siguiente a partir de la información del último año en todas las tiendas para una marca específica. Los datos fueron extraídos directamente de la base de datos de Tienda Registrada, una empresa que se dedica a la captura de información de ventas del canal de las tiendas de barrio, por medio de un sistema POS.

El sector minorista en Colombia y su evolución hacia lo digital

En Colombia, la adopción de procesos digitales ha sido más lenta que en países desarrollados; no obstante, la pandemia del COVID-19 impulsó significativamente la cultura digital, evidenciada en un aumento del 113 % en nuevos compradores online entre 2019 y 2020, lo que reflejó cambios en el comportamiento de consumo (Bancolombia, 2020). Ante este contexto, el sector minorista ha integrado canales digitales junto a los físicos tradicionales. Aunque el consumidor colombiano continúa mostrando preferencia por el canal físico, el uso del canal digital ha crecido de manera importante y forma parte actualmente de sus prácticas de compra.

Panorama del e-commerce en Colombia

El comercio electrónico en Colombia se originó en la década de 1990 y ha experimentado un crecimiento progresivo en las décadas posteriores, impulsado por la expansión de internet, las redes sociales y los sitios de contenido digital, los cuales transformaron las dinámicas de compra y el comportamiento del consumidor. En América Latina, ya se proyectaba un crecimiento significativo del gasto en comercio electrónico desde mediados de la década de 2010, evidenciando el potencial de este canal en la región (Patiño, 2014). No obstante, el punto de inflexión hacia la digitalización ocurrió con la pandemia de COVID-19, periodo en el cual el uso de canales digitales se intensificó de manera notable debido a cambios en los hábitos de consumo, el aumento de la confianza en las transacciones online y la necesidad de mantener el acceso a bienes y servicios durante las restricciones sanitarias.

Este contexto aceleró la transformación digital incluso en organizaciones tradicionalmente resistentes al cambio, incluyendo micro, pequeñas y medianas empresas, que adoptaron tecnologías digitales para garantizar su continuidad operativa (ACIS, 2021; Cámara Colombiana de Comercio Electrónico). Este proceso se vio acompañado por iniciativas del gobierno y del sector privado orientadas a mejorar la infraestructura tecnológica y la penetración de internet, especialmente en ciudades intermedias y pequeñas. A pesar de estos avances, Colombia partía de niveles de digitalización inferiores al promedio mundial —con un índice de 48 sobre 100, aproximadamente un 12,7 % por debajo de la media global—, lo que evidencia que el crecimiento reciente responde tanto a un proceso de aceleración como a la reducción de brechas estructurales en el ecosistema digital (BBVA Research, 2022).

El consumo de productos en canales digitales y los modelos multi y omnicanal

El crecimiento de la venta minorista se está apalancando en el desarrollo de canales digitales y de modelos mixtos como la multicanalidad y la omnicanalidad, provocando que los consumidores cambien sus hábitos y su comportamiento de compra. Es fundamental comprender los factores de la intención de compra en el contexto del consumidor (Kazancoglu & Aydin, 2018). Las empresas de OCR (*Omnichannel Retailing por sus siglas en inglés*) deben comprender la percepción de los consumidores influenciada por las dimensiones de compromiso cognitivo, social, conductual, psicológico y emocional, ya que esto les ayudará a enmarcar estrategias que los involucrarán con mayor éxito (Sharma et al., 2020). Estos nuevos modelos de distribución y comercialización de negocios minoristas implican unas condiciones en cuanto a la logística, el manejo de los datos y la aceptación tecnológica por parte de los consumidores, las condiciones son:

- **Logística del producto y cumplimiento:** Las políticas de devolución representan un elemento estratégico para que las organizaciones aumenten su flexibilidad, competitividad y la confianza del cliente, ya que muchos consumidores las revisan antes de realizar una compra. Sin embargo, en contextos omnicanal su gestión se vuelve más compleja para los minoristas, debido a la posibilidad de devoluciones a través de múltiples canales, lo que genera incertidumbre y mayores desafíos en la logística inversa (Hübner et al., 2016; de Borba et al., 2020).
- **Aceptación tecnológica y seguridad en los datos:** La implementación del modelo omnicanal requiere que organizaciones y consumidores se adapten a nuevas tecnologías, siendo especialmente relevante la adopción y el uso por parte de los clientes, lo que implica superar barreras asociadas a los canales digitales y la aceptación tecnológica. En este contexto, las marcas deben garantizar sistemas seguros que generen confianza y lealtad; no obstante, la solicitud de información personal puede percibirse como una amenaza a la privacidad, provocando desconfianza en los consumidores (Kazancoglu & Aydin, 2018; Piotrowicz & Cuthbertson, 2014).

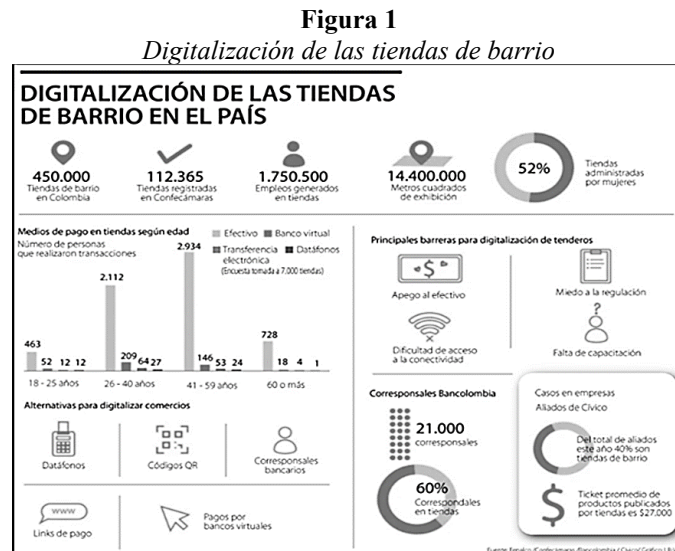
La tecnología en el comercio tradicional

El comercio tradicional corresponde a pequeños negocios independientes ubicados en barrios y urbanizaciones, caracterizados por contar con un espacio físico reducido, atención personal directa por parte del propietario o vendedores, y, en algunos casos, modalidad de autoservicio con una única caja registradora. Generalmente presentan superficies entre 40 y 100 m² y se especializan en la venta de productos de primera necesidad, como panaderías y misceláneas, siendo valorados por los consumidores debido a la cercanía, la familiaridad y la atención personalizada del tendero (Morales, 2012; Lozsán & Nieves, 2022).

Tradicionalmente, los tenderos llevaban registros y controlaban el inventario de forma análoga, pero actualmente han implementado tecnología como el sistema POS para el registro de ventas, especialmente después de la reglamentación de la DIAN (Dirección de Impuestos y Aduanas Nacionales) en Colombia obligara a facturar siempre que la venta del bien y/o prestación del servicio que se registre en el mismo no supere las 5 UVT (Unidad de Valor tributario (DIAN, 2022)). Considerando que el ticket promedio de una tienda es de COP \$ 27000 (La República, 2021) los tenderos se han visto en la obligación de implementar cada vez más el sistema POS.

- Equipo POS que es el sistema de facturación en el punto de venta que garantiza un funcionamiento continuo en su facturación.
- Uso de las Tecnologías de la Información y las Comunicaciones (TIC) para controlar sus inventarios, cuentas, gastos y créditos, e incrementar sus ingresos prestando nuevos servicios.
- Sistema de videovigilancia para poder monitorear los distintos escenarios que se pueden presentar ya sea en el interior o exterior de la tienda.
- Implementar los ‘mini datáfonos’ para permitirle a los usuarios realizar pagos mínimos desde los 2.000 o 5.000 pesos en adelante con sus tarjetas débito o crédito.
- Software para ventas para tiendas, micro mercados, droguerías, almacenes, entre otros, que deseen sistematizar el control de sus procesos, ventas e inventario (ACIS, 2021).

La figura 1 muestra una caracterización de las tiendas de barrio con datos sobre el uso de tecnologías, medios de pago, edades de los consumidores y barreras de digitalización.



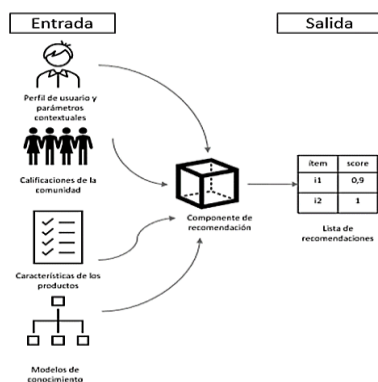
Nota. Fuente: La República 2021.

Modelos analíticos aplicados al comercio minorista.

Mishra et al. (2021) destacan la necesidad de profundizar en el aporte de la analítica de datos para comprender el comportamiento de compra y la toma de decisiones del consumidor. En este contexto, los minoristas utilizan sistemas de recomendación —como motores de búsqueda y algoritmos— que emplean información de la actividad digital para sugerir productos (Cabrera-Sánchez et al., 2020); sin embargo, muchos se basan en datos de popularidad o en marcas patrocinadas (Nguyen et al., 2022), lo que puede generar recomendaciones poco ajustadas

a las preferencias reales. Aunque se ha demostrado que la personalización incrementa las ventas en comercio electrónico, su implementación supone retos para las marcas, especialmente cuando el sistema no permite comprender adecuadamente al consumidor o no se gestiona de forma eficiente el catálogo de productos dentro de la estrategia minorista (Balakrishnan et al., 2018).

Figura 2
Esquema de un sistema de recomendación



Nota. Fuente: elaboración propia a partir de Olguín, De Jesús y Pérez de Celis, 2019.

De acuerdo con Olguín, De Jesús y Pérez de Celis (2019), los sistemas de recomendación pueden clasificarse en relación con el enfoque utilizado para obtener la lista de recomendación. La lista de recomendación es el resultado del orden de los ítems por su utilidad con respecto al usuario. En ese sentido, los sistemas de recomendación consideran los elementos de la lista, basados en las recomendaciones previas de la comunidad o el histórico de elementos considerados por los usuarios. De acuerdo con Olguín, De Jesús y Pérez de Celis (2019) los sistemas híbridos combinan distintos enfoques —como contenido, conocimiento y colaboración— con el fin de compensar las debilidades de cada método y mejorar la precisión de las recomendaciones y se clasifican en 4 categorías:

- Los sistemas de recomendación pueden clasificarse en cuatro tipos principales:
- Los basados en contenido generan sugerencias a partir de las similitudes entre los elementos del catálogo y las preferencias previas del usuario, aunque presentan limitaciones con usuarios nuevos.
- Los basados en conocimiento utilizan información previa y criterios definidos para satisfacer necesidades específicas, pero requieren mecanismos de discriminación cuando varios productos cumplen las mismas condiciones.
- Los basados en colaboración se fundamentan en las valoraciones de otros usuarios, asumiendo que personas con preferencias similares compartirán intereses.

Análisis de la cesta de compra (Market Basket Analysis).

El **Market Basket Analysis** o análisis de la canasta de compra es un método que permite identificar relaciones entre los productos que adquieren los clientes. Al detectar qué artículos se compran conjuntamente con mayor frecuencia, facilita el diseño de promociones dirigidas, recomendaciones y la disposición estratégica de productos en la tienda para aumentar las ventas (Bismi, 2022). Esta técnica se basa en reglas de asociación aplicadas a bases de datos transaccionales, con el objetivo de descubrir patrones de compra y generar conocimiento sobre el comportamiento del consumidor mediante algoritmos de análisis (Raeder & Chawla, 2011).

¿Qué son las reglas de asociación y como las utilizan los minoristas para crear promociones o sugerir las compras? En el sector minorista, las reglas de asociación se emplean para identificar patrones frecuentes en datos provenientes de ventas, lo que permite comprender el comportamiento de compra y diseñar promociones con alta probabilidad de éxito. Este enfoque se aplica comúnmente al análisis de canastas de productos adquiridos por los clientes, facilitando el desarrollo de estrategias de marketing orientadas a mejorar las ventas (Amaris & Rodríguez, 2003). En términos generales, las reglas de asociación constituyen una técnica de minería de datos ampliamente utilizada para detectar relaciones, correlaciones y co-ocurrencias entre conjuntos de datos mediante algoritmos de aprendizaje automático.

El resultado de este análisis se expresa en reglas del tipo “si... entonces”, donde el antecedente corresponde a un elemento presente en los datos y el consecuente al elemento que aparece en combinación con él (Piatetsky-Shapiro, 1996, como se citó en Amaris & Rodríguez, 2003; AprendeIA, 2023). En el contexto de compras, el antecedente es el producto que el consumidor adquiere y el consecuente el que probablemente comprará en conjunto; por ejemplo, si un cliente compra leche, es probable que también compre pan.

Desarrollo de asociaciones a partir de algoritmos

Antes de aplicar cualquier algoritmo de análisis de canasta de compra, es necesario construir una matriz de transacciones que relacione productos y compradores mediante valores binarios (0 y 1), indicando si cada cliente adquirió o no un producto específico. A partir de esta matriz se identifican asociaciones entre productos que suelen comprarse conjuntamente, lo que permite comprender el comportamiento de los consumidores y detectar patrones de compra mediante algoritmos de asociación; en consecuencia, el proceso inicia con la construcción de dicha matriz a partir de las bases de datos de ventas.

Tabla 1
Matriz de transacciones

	Producto 1	Producto 2	Producto 3	Producto 4	Producto 5	Producto 6
Cliente 1	1	1	0	1	0	0
Cliente 2	0	0	1	0	0	0
Cliente 3	1	0	1	1	1	1
Cliente 4	0	1	1	0	0	1

Nota. Fuente: elaboración Propia

Algoritmos usados en la asociación de productos

Los algoritmos Apriori, Eclat y FP-Growth son ampliamente utilizados para identificar asociaciones frecuentes en bases de datos y generar reglas de asociación. El algoritmo Apriori es pionero en este ámbito y opera mediante iteraciones sucesivas que identifican ítems frecuentes a partir de un umbral mínimo de soporte, generando conjuntos cada vez mayores hasta que no se obtienen nuevas combinaciones relevantes (Rocha et al., 2016; Báez Acuña et al., 2018). Por su parte, el algoritmo Eclat realiza búsquedas en profundidad, lo que incrementa su eficiencia al trabajar con conjuntos más pequeños; organiza las transacciones en matrices por ítems y calcula los soportes mediante la intersección de identificadores de transacciones, eliminando aquellos que no cumplen las reglas establecidas (Rocha et al., 2016; Báez Acuña et al., 2018). El algoritmo FP-Growth utiliza una estructura de árbol de patrones frecuentes (FP-tree) que almacena información sobre la ocurrencia de los productos,

permitiendo identificar patrones frecuentes de forma más eficiente sin generar candidatos explícitos (Vergara Orozco et al., 2022).

Adicionalmente, el algoritmo de similitud de coseno es una técnica de análisis de datos empleada para medir la similitud entre vectores, como productos o usuarios, mediante el coseno del ángulo entre ellos; valores cercanos a 1 indican alta similitud, mientras que valores cercanos a -1 reflejan diferencias significativas. En el comercio minorista, esta técnica se utiliza para sistemas de recomendación de productos basados en patrones de compra (Grapheverywhere, 2023).

Metodología

Caso comercio tradicional “Supermercado la Comarca”

Para el desarrollo del proyecto se analizó el caso Supermercado la Comarca, ubicado en el barrio Bonanza que se encuentra en la localidad de Engativá, considerado por sus dimensiones y nivel de servicios como comercio tradicional y opera como una tienda de barrio. El supermercado cuenta con un sistema POS en el que se registran todas las transacciones de la tienda.

Descripción del barrio Bonanza

Bonanza es un barrio está ubicado en la Localidad de Engativá, en la ciudad de Bogotá, ubicado de norte a sur entre las calles 75 y 68, y de oriente a occidente entre la carrera 78 y la Avenida Boyacá. Es un barrio tradicional de clase media con zonas residenciales específicas y vías principales dedicadas al comercio. La tabla 1 muestra datos demográficos de la localidad de Engativá.

Tabla 2
Análisis demográfico y económico de la localidad de Engativá

Datos	Bogotá	Localidad de Engativá
Número de habitantes	7.834.167 habitantes	814.467 habitantes
Tasa de ocupación 2021	Bogotá: 56,6%	Engativá: 56,6%
Tasa de desempleo 2021	Bogotá: 16,2%	Engativá: 11,4%
Tasa de informalidad 2021	Bogotá: 42,4%	Engativá: 37,7%

Nota. Fuente: Alcaldía Mayor de Bogotá, 2022

La localidad de Engativá concentra aproximadamente el 10 % de la población de Bogotá y constituye una de las principales fuentes de fuerza laboral de la ciudad, con una participación cercana al 10,7 %, debido a que la mayoría de sus habitantes se encuentra en edad productiva. Aunque presenta una tasa de ocupación similar al promedio distrital y un desempleo menor, también registra niveles importantes de informalidad (37,7 %), lo que refleja condiciones laborales heterogéneas (ODEB, 2022). En términos socioeconómicos, predomina el estrato 3, seguido del estrato 2, lo que caracteriza a la localidad como mayoritariamente de clase media. Asimismo, se trata principalmente de una zona residencial, con una proporción reducida de uso comercial, industrial o de servicios (Alcaldía Mayor de Bogotá, 2022).

Metodología CRISP-DM

Para abordar el problema del bajo aprovechamiento de la información disponible en las tiendas de barrio, se empleó la metodología CRISP-DM con el propósito de estructurar el análisis de datos y responder al objetivo

del proyecto. Aunque estos negocios registran información financiera y operativa, esta se utiliza de manera limitada para la toma de decisiones y la mejora de la experiencia del cliente. Sin embargo, el creciente uso de canales digitales ha incrementado la disponibilidad de datos que podrían aprovecharse estratégicamente para optimizar inventarios y personalizar la oferta de valor, especialmente en un contexto caracterizado por rezagos en digitalización (Branch, 2020; Fenalco, 2021). A partir de ello, se presentan las etapas de la metodología CRISP-DM.

Tabla 3

Descripción de los pasos de la metodología CRISP DM para el proyecto

Paso	Descripción	Desarrollo
1. Entendimiento del negocio	Exposición de generalidades	Análisis de la situación actual basado en antecedentes sobre el uso de tecnología en tiendas de barrio y en la caracterización del caso “Supermercado La Comarca”, seguido de la definición de los objetivos de minería de datos a partir de la información de ventas y domicilios, y el diseño de la arquitectura de la solución propuesta.
2.Comprensión de los datos	Análisis descriptivo de las variables intervinientes dentro de la base de datos obtenida.	Recopilación y análisis de los datos de domicilios mediante visualización, incluyendo variables de facturación, horarios, volumen de pedidos y georreferenciación de consumidores, así como la verificación de la calidad de la información proveniente del sistema POS.
3. Preparación de datos	Definición de los procesos de extracción, transformación y carga (ETL) para estructurar la base de datos que será utilizada en el modelo propuesto.	Selección de datos, almacenamiento de los datos de los clientes (dirección y teléfono) Limpieza de los datos Construcción de nuevos datos desde el proceso ETL diseñado en Python / Integración de datos de acuerdo con el objetivo del proyecto
4.Modelado	Se establecerán las técnicas a ser utilizadas, así como su implementación y validación a ser desarrollada para obtener el mejor modelo acorde con el objetivo de la investigación.	Selección de técnicas de modelado basadas en sistemas de recomendación, desarrollo de un algoritmo de similitud de cosenos a partir del historial y preferencias de los usuarios, generación de la matriz de predicciones y evaluación del modelo mediante la implementación de un chatbot en WhatsApp como herramienta de interacción con el cliente.
5.Evaluación:	Los resultados obtenidos serán evaluados en su utilidad y pertinencia que le brindarán al campo de análisis en el que estos podrán ser utilizados.	Realización de experimentos controlados. Conjunto de datos de prueba con datos nuevos. Feedback de usuarios Seguimiento de resultados en producción
6.Distribución	Uso de técnicas de visualización descriptiva, como tableros de control, para facilitar la comprensión y el análisis del negocio.	Definición, implementación y validación de las técnicas necesarias para obtener el modelo óptimo que permita alcanzar los objetivos propuestos.

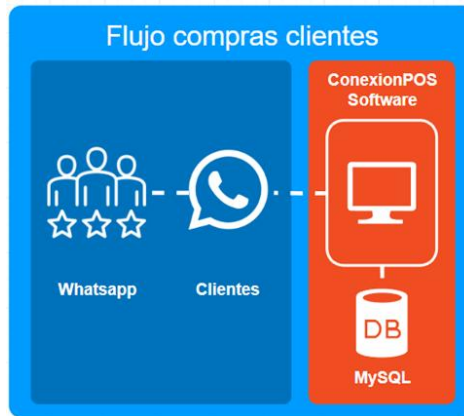
Nota. Fuente: elaboración propia

Entendimiento del negocio

El Supermercado La Comarca es una tienda de barrio que comercializa productos de consumo mediante dos modalidades: atención presencial, donde los clientes seleccionan productos por categorías y realizan el pago en caja, y pedidos a domicilio, gestionados por teléfono, WhatsApp o página web, seguidos de la preparación y entrega. Este último grupo resulta especialmente relevante para el estudio, ya que se dispone de información sobre compras, productos, cantidades y datos de contacto, cuyo aprovechamiento permitiría optimizar la organización del surtido, mejorar la eficiencia del servicio a domicilio, diseñar promociones en períodos de alta demanda, captar

nuevos clientes e implementar desarrollos tecnológicos orientados al incremento de las ventas. En este contexto, se propone una arquitectura de solución alineada con los objetivos del trabajo.

Figura 3
Flujo de compras de los clientes



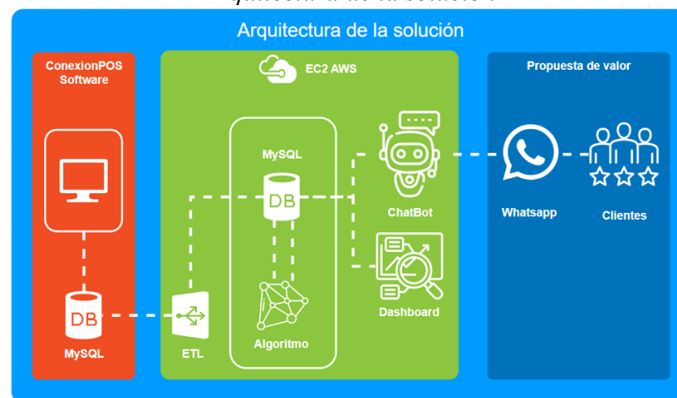
Nota. Fuente: elaboración propia

Arquitectura de datos:

Para el caso puntual del Supermercado La Comarca, la arquitectura descrita presenta varias etapas para el aprovechamiento de los datos almacenados en el sistema POS las cuales consisten en:

1. Almacenamiento de los datos de los clientes (dirección y teléfono)
2. Extracción, transformación y carga de los datos en una nueva base de datos.
3. Implementación de algoritmos de clasificación y recomendación que permita generar valor al cliente
4. Generación de tableros de control para seguimiento de métricas de distribución y ubicación de clientes.
5. Ventas por fecha, clientes frecuentes y de alto consumo, productos de alta y baja rotación entre otros.

Figura 4
Arquitectura de la solución



Nota. Fuente: elaboración propia

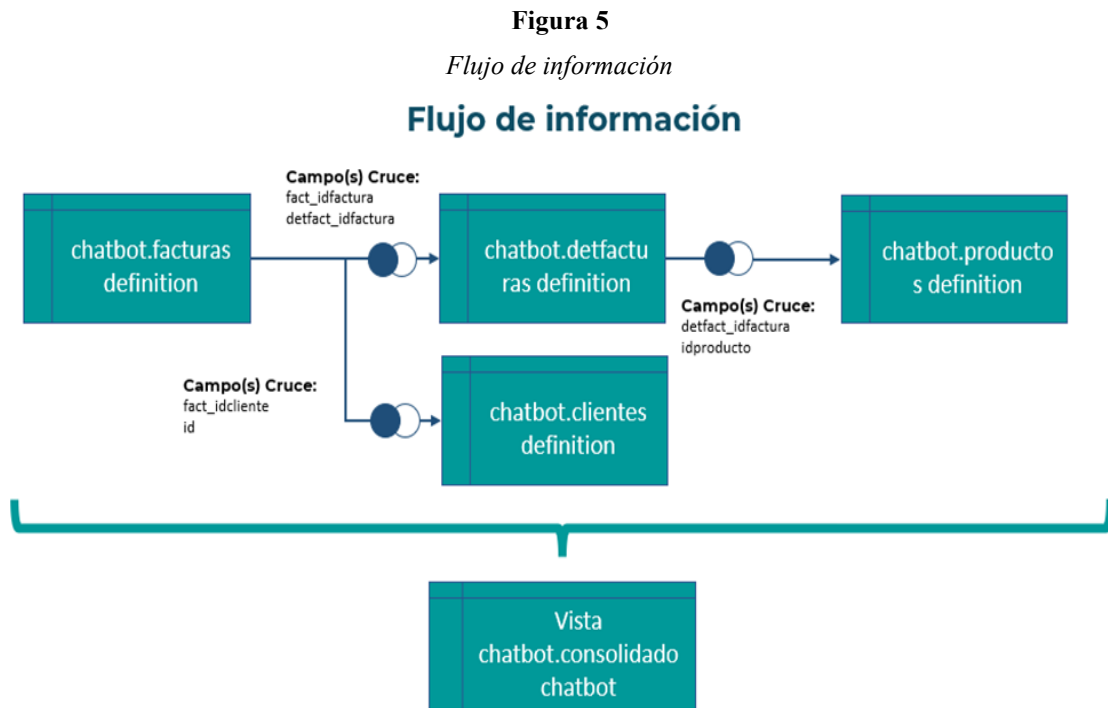
Compresión de los y preparación de los datos

Los datos del Supermercado La Comarca se almacenan en un sistema POS que permite gestionar transacciones, inventarios y compras de clientes, generando información útil para la toma de decisiones. Este

sistema opera sobre una base de datos MySQL con múltiples tablas, por lo que fue necesario implementar un proceso ETL para extraer, transformar y cargar la información relevante requerida para el proyecto.

Estructura, almacenamiento y descripción de los datos:

Dada la extracción de los datos desde la fuente de origen (sistema POS), se relaciona a continuación el flujo (MER) y diseño propuesto:



Nota. Fuente: elaboración propia

Modelado

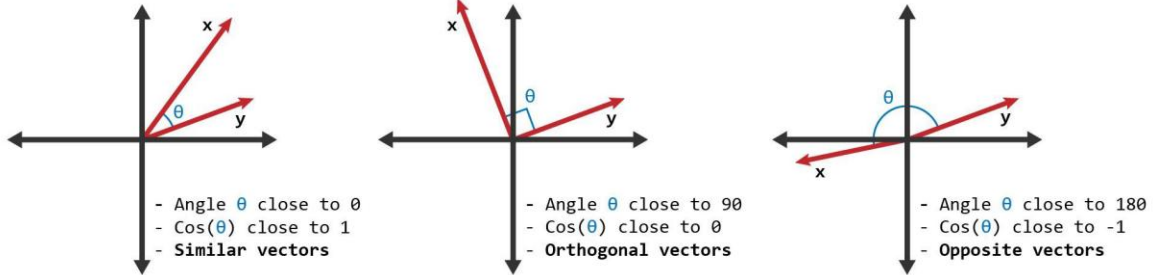
De acuerdo con el objetivo del proyecto, se emplearon sistemas de recomendación basados en reglas de asociación para identificar patrones y hábitos de compra a partir del historial de transacciones, con el propósito de predecir productos de interés para los clientes. Dentro de estos enfoques, se seleccionó el filtrado colaborativo, que permite recomendar productos mediante la identificación de usuarios con preferencias y comportamientos de compra similares.

Descripción del algoritmo de similitud de cosenos

El algoritmo seleccionado fue el de similitud de cosenos, debido a las características y naturaleza de los datos, ya que permite medir la semejanza entre vectores —en este caso, productos representados en un espacio vectorial— mediante el coseno del ángulo entre ellos (Rojas Hernández & Gelvez García, 2016). Este enfoque parte de que valores cercanos a 1 indican alta similitud, valores cercanos a -1 reflejan diferencias, y valores intermedios representan distintos grados de relación entre los elementos (Nguyen & Bai, 2011; Grapheverywhere, 2023). La figura siguiente ilustra este principio a partir de la dirección de los vectores.

Figura 6

Gráfico de vectores con similitudes cercanas a 1, cercanas a 0 y cercanas a -1.



Nota. Fuente: tomado de Karabiber, 2023

La similitud puede tomar valores entre -1 y +1. Los ángulos más pequeños entre vectores producen valores de coseno más grandes, lo que indica una mayor similitud de coseno. Cuando dos vectores tienen la misma orientación, el ángulo entre ellos es 0 y la similitud del coseno es 1. Los vectores perpendiculares tienen un ángulo de 90 grados entre ellos y una similitud de coseno de 0. Los vectores opuestos tienen un ángulo de 180 grados entre ellos y una similitud de coseno de -1.

La formulación de la similitud del coseno en términos del lenguaje de la regla de asociación se puede escribir como:

Figura 7

Fórmula de similitud de cosenos

$$C(A \rightarrow B) = \frac{supp(A, B)}{\sqrt{supp(A) supp(B)}}$$

Nota. La similitud del coseno entre a y b es igual al soporte entre (a,b) / la raíz cuadrada del producto de los soportes individuales de a y de b. Elaboración propia.

El soporte es la probabilidad frecuentista de ocurrencia aparición de un producto. La métrica de soporte mide la proporción de transacciones que contiene un conjunto de elementos, es decir:

Figura 8

Métrica de confianza

$$\frac{\text{número de transacciones con artículos}}{\text{número de transacciones}}$$

Nota 2: La métrica de confianza nos da la probabilidad de que compraremos Y dado que hemos comprado X. Elaboración propia.

En la tabla 3 se interpreta la métrica de confianza

Tabla 4
Interpretación de la métrica de confianza

ID	ID	ID
1	Café, Leche	1
2	Pan, Leche, Naranja	2
3	Pan, Leche	3
4	Pan, Leche, Azúcar	4
5	Pan, Huevo, Leche	5

$$\text{Soporte (leche y café)} = 0,20$$

$$\text{Soporte (leche)} = 1$$

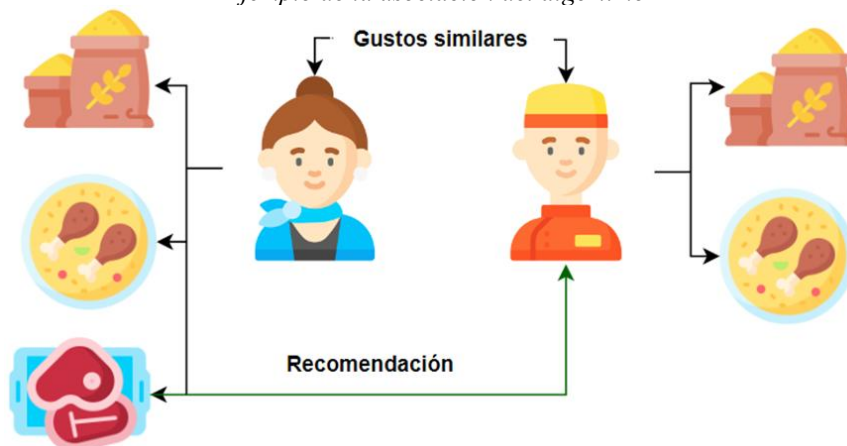
$$\frac{\text{Soporte (leche y café)}}{\text{Soporte (leche)}} = \frac{0,20}{1} = 0.20$$

La probabilidad de comprar leche y café no cambia si condicionamos la compra de leche. Comprar leche no nos dice nada acerca de comprar café.

Operatividad del algoritmo

El funcionamiento del algoritmo puede describirse en tres etapas: primero, el motor de filtrado colaborativo identifica usuarios con patrones de valoración similares al usuario activo; segundo, establece esta similitud a partir del historial, las preferencias y las decisiones de compra; y tercero, utiliza las valoraciones de esos usuarios semejantes para predecir la posible valoración de productos no explorados previamente, permitiendo recomendar artículos que otros consumidores con comportamientos similares ya han adquirido.

Figura 9
Ejemplo de la asociación del algoritmo



Nota. Fuente: elaboración del algoritmo

Resultados

Desarrollo del algoritmo de similitud de cosenos:

Inicialmente y teniendo en cuenta que ya se realizó la importación de librerías, extracción y limpieza de los datos, se dio inicio del análisis sobre la descripción del producto, para tal fin se realizó un conteo por el campo descripción del producto como se muestra a continuación:

Generación de la variable de conteo

```
comercio_tx['prod_descripcion'].value_counts()
```

La siguiente imagen muestra los tipos de datos que son relevantes y no relevantes para el proyecto; existen categorías dentro de los registros como 'SERVICIO A DOMICILIO' que no aportan nada al análisis ni a una futura recomendación.

Figura 10
Variable conteo

SERVICIO A DOMICILIO	12216
SHAMPOO LA JOYA 500 ML	6308
DULCERIA	3461
ATUN VAN CAMPS LOMITOS X 184 GRS	3432
COCA COLA X 1.5 LT	2965
...	
PAPEL HIGIENICO SCOTT RINDEMAX X 4 UNDS	1
SPORADE MANDARINA X 400ML	1
MANGO TOMY X KG	1
PAÑAL HUGGIES ETAPA 3 X 30 UNDS	1
PALOMITAS NATURAL MICRO ONDAS EL FORTIN	1
Name: prod_descripcion, Length: 4989, dtype: int64	

Nota. Fuente: elaboración propia

La variable 'conteo_productosxcliente' permite contar la cantidad de veces que se ha comprado determinado producto, en diferentes facturas. La tabla anterior nos muestra como el cliente con numero de celular 3116886948, compró el producto ATUN VAN CAMPS LOMITOS X 184 GRS en 7 momentos distintos, y ese conteo se ve reflejado en la variable que creamos.

Figura 11
Conteo productos cliente

	fact_fecha_comp	fact_idcliente	prod_idproducto	prod_descripcion	conteo_productosxcliente
223832	2022-09-08	3116886948	737	ATUN VAN CAMPS LOMITOS X 184 GRS	7
229264	2022-09-22	3116886948	737	ATUN VAN CAMPS LOMITOS X 184 GRS	7
238132	2022-10-17	3116886948	737	ATUN VAN CAMPS LOMITOS X 184 GRS	7
245014	2022-11-06	3116886948	737	ATUN VAN CAMPS LOMITOS X 184 GRS	7
248301	2022-11-15	3116886948	737	ATUN VAN CAMPS LOMITOS X 184 GRS	7
250681	2022-11-22	3116886948	737	ATUN VAN CAMPS LOMITOS X 184 GRS	7
269822	2023-01-17	3116886948	737	ATUN VAN CAMPS LOMITOS X 184 GRS	7

Nota. Fuente: elaboración propia

Creación de la matriz usuario producto

Posteriormente se construye la matriz usuario-producto, que registra para cada cliente si ha adquirido o no un producto (1 si lo compró y 0 en caso contrario), independientemente de la cantidad. Esta representación permite cuantificar los valores no vacíos mediante el indicador de sparsity; en este caso, la matriz presenta un 23,57 % de elementos distintos de cero, lo que indica un nivel adecuado de información en el conjunto de datos.

Figura 12
Matriz usuario producto

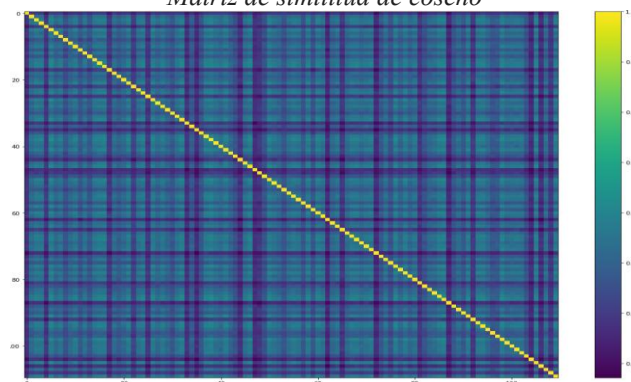
prod_descripcion	3D LIQUIDO 1.8 LTS	3D MULTIUSOS EUCALIPTO 500GR	ABLANDA CARNES X 10 GRS ALÍÑOS SANTA HELENA	ACACIA DE LA INDIA X 15 GRS X 15 GRS	ACEITE 3 EN 1 X 30 ML	ACEITE BON LIFE 1000 CM	ACEITE DE ALMENDRAS OSA X 12 ML	ACEITE DE CANOLA GOURMET 1000 ML	ACEITE DE CITRONELA DELICIAS 125 ML	ACEITE DE OLIVA OLIVETTO 250 ML
fact_idcliente										
3018318595	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
3018401779	1.0	1.0	1.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
3018651677	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
3055843805	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
3056317781	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Nota. Fuente: elaboración propia

Matriz de similitud de cosenos

Matriz de similitud por medio del algoritmo de Similitud del Coseno el cual permite observar y medir la similitud entre los vectores de productos adquiridos por los clientes, es decir, que tan similares son los clientes con respecto a otros con base en los productos comprados, la matriz nos muestra en una escala de 0.3 a 1, unos usuarios que tienen una similitud cercana al 0.8 aproximadamente.

Figura 13
Matriz de similitud de coseno



Nota. Fuente: elaboración propia

Predicción de la compra

La predicción de compra se realizó calculando el peso de cada cliente sobre los productos a partir de la similitud entre usuarios, donde cada rating se multiplica por el factor de similitud correspondiente y la predicción final se obtiene mediante la suma de los pesos normalizados o suma ponderada. Posteriormente, se generan las recomendaciones de productos para cada usuario. Para validar el

funcionamiento del modelo, se definieron dos parámetros —cantidad de productos aleatorios y cantidad de productos seleccionados— que permiten determinar tanto el número de productos con mayor probabilidad de compra como los que serán mostrados al cliente. Este procedimiento incorpora un componente de aleatoriedad que diversifica las recomendaciones, evitando que el sistema sugiera siempre los mismos productos y mejorando la experiencia del usuario.

Figura 14

Resultado de la matriz de predicciones

	Producto	Nombre
0	Producto 1	ATUN VAN CAMPS LOMITOS X 184 GRS
1	Producto 2	BON AIRE CANASTA CANELA X 30 GRS
2	Producto 3	HUEVO EXTRA CAJA X 6

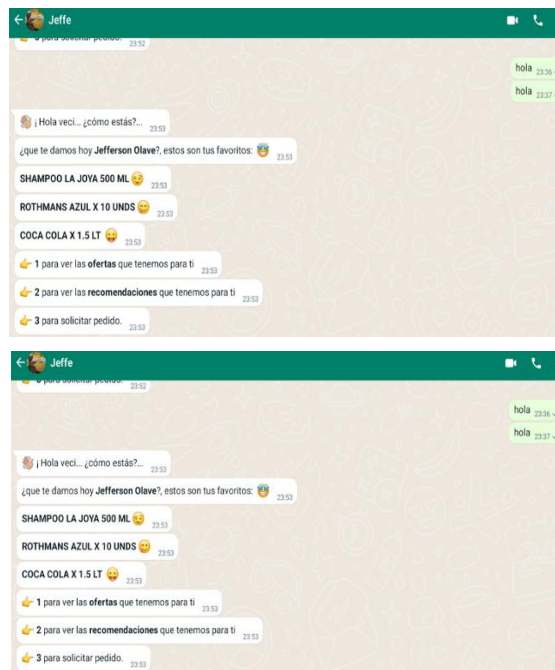
Nota. Fuente: elaboración propia

Herramienta para uso del tendero. Chat bot para la tienda

Para hacer funcionales los resultados del algoritmo se decide realizar un chatbot de atención al cliente vía WhatsApp, que pueda ser una herramienta útil para mejorar la atención e incrementar las ventas. Los chatbot's son programas, algunos con algoritmos de inteligencia artificial, diseñados para interactuar con los clientes a través de una interfaz de chat.

Figura 15

Evidencia del despliegue de WhatsApp



Nota. Fuente: elaboración propia

Este chatbot se alimenta con el histórico de compras del cliente, promociones y recomendaciones. Puntualmente el chatbot puede hacer recomendaciones personalizadas basadas en las compras anteriores del cliente, los productos más vendidos o los productos que están en promoción.

Se espera que, con esta solución, se logre aumentar las ventas al proporcionar recomendaciones relevantes y personalizadas a los clientes toda vez que el algoritmo del que se alimenta puede identificar patrones de compra y sugerir productos complementarios o relacionados que puedan interesar al cliente, aumentando así la probabilidad de aumentar los productos de compra y con ello las ventas.

Discusión

Los sistemas de recomendación se utilizan para apoyar la toma de decisiones de consumidores y empresas y mejorar la experiencia de compra (Cabrera-Sánchez et al., 2020); grandes minoristas como Éxito, Amazon y Mercado Libre los emplean para fortalecer su oferta de valor, mientras que las tiendas de barrio presentan rezagos en su adopción tecnológica (La República, 2021). En este contexto, el proyecto se orienta a aprovechar los datos históricos del tendero —ventas, productos, clientes y domicilios— almacenados en un sistema POS con base MySQL (280 tablas), para lo cual se implementó un proceso ETL en Python que permitió construir una nueva base de datos y desarrollar un algoritmo de recomendación basado en filtrado colaborativo y similitud de coseno, capaz de generar ofertas personalizadas según el perfil del cliente.

El caso de estudio corresponde al Supermercado La Comarca, una tienda tradicional ubicada en el barrio Bonanza, donde se evidenció un bajo aprovechamiento de la información disponible, lo que justificó el uso de analítica de datos como solución al objetivo del proyecto. Asimismo, la revisión de la literatura mostró la ausencia de investigaciones que integren algoritmos de recomendación y chatbots en tiendas de barrio en Bogotá o Latinoamérica con enfoque práctico para incrementar ventas a domicilio, lo que resalta la contribución de este trabajo al proporcionar una herramienta aplicable para el tendero. En este sentido, los algoritmos basados en reglas de asociación se constituyen en herramientas eficaces para identificar relaciones entre productos a partir del historial de compras, en línea con el enfoque estratégico de la canasta de compra.

Conclusiones

El propósito de esta investigación fue desarrollar un chatbot que, a partir de los datos históricos de los clientes de los tenderos, permitiera generar recomendaciones personalizadas para los consumidores. Asimismo, se diseñó una herramienta tecnológica basada en analítica de datos orientada a apoyar a los tenderos en la mejora de la experiencia de consumo y en el fortalecimiento de sus estrategias de venta. En este sentido, el proyecto logró cumplir ambos propósitos mediante la integración de tecnologías de recomendación y herramientas de interacción digital.

El desarrollo de herramientas basadas en analítica de datos representa una oportunidad estratégica para fortalecer la competitividad de las tiendas de barrio dentro del sector minorista. Aunque estos negocios han gestionado tradicionalmente la relación con sus clientes mediante el conocimiento empírico y la experiencia cotidiana, las nuevas dinámicas de consumo exigen una mayor capacidad para interpretar y aprovechar los datos generados en los procesos de compra. En este contexto, el uso de tecnologías de

recomendación permite transformar la información disponible en conocimiento útil para la toma de decisiones y la personalización de la experiencia de consumo.

Los resultados evidencian que la implementación de modelos predictivos y algoritmos de recomendación contribuye a mejorar la interacción entre tenderos y consumidores, facilitando la generación de ofertas más ajustadas a las preferencias y hábitos de compra de los clientes. Asimismo, el desarrollo del chatbot permitió integrar procesos de recomendación automatizada en un entorno de comunicación cotidiano para los consumidores, fortaleciendo la experiencia de compra y promoviendo nuevas posibilidades de comercialización digital para el comercio tradicional.

De igual manera, el estudio demuestra que las estrategias orientadas al análisis de patrones de consumo, como el Market Basket Analysis y los sistemas de recomendación basados en similitud, constituyen herramientas efectivas para identificar asociaciones entre productos y optimizar la oferta de valor del minorista. Esto no solo favorece el incremento potencial de las ventas, sino que también permite al tendero responder de manera más eficiente a los cambios en el comportamiento del consumidor y a las exigencias de un entorno cada vez más omnicanal.

Finalmente, la investigación pone de manifiesto que la transformación digital del comercio tradicional no depende únicamente de la incorporación de tecnología, sino también de la capacidad de los tenderos para adoptar una cultura orientada al uso estratégico de los datos. En consecuencia, el aprovechamiento de la analítica de datos se consolida como un elemento clave para la sostenibilidad y modernización de las tiendas de barrio, permitiéndoles competir en escenarios minoristas caracterizados por consumidores más digitales, informados y dinámicos.

Referencias Bibliograficas

- ACIS. (2021, April). La Cámara Colombiana de Comercio Electrónico presenta el informe de cierre sobre el comercio electrónico en 2020.
- Aversa, J., Hernandez, T., & Doherty, S. (2021). Incorporating big data within retail organizations: A case study approach. *Journal of Retailing and Consumer Services*, 60. <https://doi.org/10.1016/j.jretconser.2021.102447>
- Alcaldía Mayor de Bogotá. (2022). *Diagnóstico Local Engativa*. https://www.integracionsocial.gov.co/images/_docs/entidad/10_Diagnostico_local_Engativa_2021_VF.pdf
- Amaris, M. E., & Rodríguez, R. J. E. (2003). La contribución de las reglas de asociación a la minería de datos. *Tecnura*, 7, 94-109.
- AprendeIA. (2023, May 13). *Reglas de asociación*. <https://aprendeia.com/reglas-de-asociacion/>
- Báez Acuña, J. M., Paredes Cabañas, C. A., Sosa Cabrera, G., & García, M. E. (2018). Descubriendo reglas de asociación en bases de datos del sector retail usando R. In *XXIV Congreso Argentino de Ciencias de la Computación* (Ed.), Descubriendo reglas de asociación en bases de datos del sector retail usando R.

- Bancolombia (2020). *eCommerce en Colombia: cifras, tendencias y retos actuales*. <https://www.bancolombia.com/empresas/capital-inteligente/tendencias/tendencias-globales/ecommerce-colombia-cifras-tendencias-retos>
- BBVA Research. (2022). *DiGiX 2022 Update: A Multidimensional Index of Digitization*. DiGiX 2022 Update: A Multidimensional Index of Digitization.
- Balakrishnan, J., Cheng, C. H., Wong, K. F., & Woo, K. H. (2018). Product recommendation algorithms in the age of omnichannel retailing – An intuitive clustering approach. *Computers and Industrial Engineering*, 115, 459–470. <https://doi.org/10.1016/j.cie.2017.12.005>
- Bismi, B. S., & Azeem, S. (2022). A survey on increasing the capacity of 5g fronthaul systems using ROF. *Optical Fiber Technology*, 74, 103078.
- Branch. (27 de 04 de 2020). *El crecimiento del ecommerce en Colombia*. <https://branch.com.co/en/marketing-digital/el-crecimiento-del-e-commerce-en-colombia-analisis-2020/>
- de Borba, J. L. G., Magalhaes, M. R. de, Filgueiras, R. S., & Bouzon, M. (2020). Barriers in omnichannel retailing returns: a conceptual framework. *International Journal of Retail & Distribution Management*, 49(1), 121–143.
- Cabrera-Sánchez, J. P., Ramos-De-luna, I., Carvajal-Trujillo, E., & Villarejo-Ramos, Á. F. (2020). Online recommendation systems: Factors influencing use in E-commerce. *Sustainability (Switzerland)*, 12(21), 1–15. <https://doi.org/10.3390/su12218888>
- DIAN. (2022). *RESOLUCIÓN 001092*.
- Fenalco. (2021). *Perspectiva económica del tendero*. Fenalco.
- Frasquet, M., Mollá, A., & Ruiz, E. (2015). Identifying patterns in channel usage across the search, purchase and post-sales stages of shopping. *Electronic Commerce Research and Applications*, 14(6), 654–665. <https://doi.org/10.1016/j.elerap.2015.10.002>
- García Rodríguez, N. (2019). *Diseño integral de una solución digital para innovar el modelo de negocio de la industria del retail integrando analítica, cloud computing e internet del todo* [Doctoral dissertation, Universidad Externado de Colombia].
- Gómez Betancur, F. (2021). *Predicción de ventas en las tiendas de barrio Colombianas*. Universidad de Antioquía.
- Grapheverywhere (2023). *Algoritmo de similitud de cosenos*. <https://www.grapheverywhere.com/algoritmo-de-similitud-de-coseno/>
- Grewal, D., Bart, Y., Spann, M., & Zubcsek, P. P. (2016). Mobile advertising: A framework and research agenda. *Journal of interactive marketing*, 34(1), 3-14.
- Hübner, A., Wollenburg, J., & Holzapfel, A. (2016). Retail logistics in the transition from multi-channel to omni-channel. *International Journal of Physical Distribution & Logistics Management*, 46(6/7), 562–583.
- Kazancoglu, I., & Aydın, H. (2018). An investigation of consumers' purchase intentions towards omni-channel shopping: A qualitative exploratory study. *International Journal of Retail & Distribution Management*.

- La República. (2019, August 26). *Tiendas de barrio, el canal tradicional que se sigue reinventando*. <https://www.larepublica.co/consumo/tiendas-de-barrio-el-canal-tradicional-que-se-sigue-reinventando-2899909>
- La Republica. (2021, October 2). *¿Por qué las tiendas de barrio de Colombia se resisten a empezar con la digitalización?*. <https://www.larepublica.co/internet-economy/por-que-las-tiendas-de-barrio-de-colombia-se-resisten-a-entrar-en-la-digitalizacion-3241191>
- Lopez, M., & Guerrero, R. (2017). *Modelo de Inteligencia de Negocios y Analítica en la nube para PYMES del sector retail*. Cloud Business Intelligence & Analytics model for SMEs retail. 4to Congreso Internacional AmITIC. Popayán, Colombia.
- Nguyen, H. V., & Bai, L. (2011). Cosine similarity metric learning for face verification. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6493 LNCS(PART 2), 709–720. https://doi.org/10.1007/978-3-642-19309-5_55
- ODEB. (2022, December 22). El mercado laboral en 19 localidades de Bogotá. *Observatorio de Desarrollo Económico*. <https://observatorio.desarrolloeconomico.gov.co/mercado-laboral-especial/el-mercado-laboral-en-19-localidades-de-bogota-resultados-de-la-encuesta>
- Olguín, G. M., De Jesús, Y. L., & de Celis Herrero, M. C. P. (s. f.). *Métricas de similitud y evaluación para sistemas de recomendación de filtrado colaborativo: Similarity and evaluation metrics for collaborative based recommender systems*.
- Morales, D. P. (2012). Tiendas de barrio en Colombia. *Pensamiento & gestión*, (32), vii-xi.
- Patiño, A. S. (2014). ¿Hacia dónde va el comercio electrónico en Colombia? *Ploutos*, 4(1), 17–24.
- Piotrowicz, W., & Cuthbertson, R. (2014). Retail: Toward Omnichannel Retailing. *International Journal of Electronic Commerce / Summer*, 18(4), 5–15. <https://doi.org/10.2753/JEC1086-4415180400>
- Lopez, M., & Guerrero, R. (2017). *Modelo de Inteligencia de Negocios y Analítica en la nube para PYMES del sector retail*. Cloud Business Intelligence & Analytics model for SMEs retail. 4to Congreso Internacional AmITIC. Popayán, Colombia.
- Lozán, Nieves. (2022, 17 junio). Mercado minorista: significado, clases de tiendas y diferencias entre comercio minorista y mayorista. *Cinco Noticias*. <https://www.cinconoticias.com/mercado-minorista/>
- Mishra, R., Rajesh, Singh, K., & Koles, Bernadett. (2021). Consumer decision-making in omnichannel retailing: Literature review and future research agenda. *Int. J. Consum. Stud*, 45, 147–174. <https://doi.org/10.1111/ijcs.12617>
- Morales, D. P. (2012). Tiendas de barrio en Colombia. *Pensamiento & gestión*, (32), vii-xi
- Nguyen, H. V., & Bai, L. (2010, November). Cosine similarity metric learning for face verification. In *Asian conference on computer vision* (pp. 709-720). Springer Berlin Heidelberg.
- Sharma, M., Gupta, M., & Joshi, S. (2020). Adoption barriers in engaging young consumers in the Omnichannel retailing. *Young Consumers*, 21(2), 193–210.
- Sanclemente, M. I. (2021). *Beneficios de la analítica de datos para PYMES en Bogotá*. https://repository.cesa.edu.co/bitstream/handle/10726/4102/ADM_1020836039_2021_1.pdf
- Schoenbachler, D. D., & Gordon, G. L. (2002). Multi-channel shopping: understanding what drives channel choice. *Journal of Consumer Marketing*.

- Raeder, T., & Chawla, N. V. (2011). Market basket analysis with networks. *Social network analysis and mining, 1*, 97-113.
- Rocha, J. A., Rodríguez, D. M., & Rodríguez, J. E. (2016). A research comparative among association rules algorithms. *Visión Electrónica, 10*(2), 210–217. <https://revistas.udistrital.edu.co/index.php/visele/index>
- Rojas Hernández, A. F., & Gelvez García, N. Y. (2016). *Distributed processing using cosine similarity for mapping Big Data in Hadoop* (UNESCO-UNIR ICT & Education Latam Congress 2016, Ed.). <http://research.unir.net>
- Vergara Orozco, J. C., Vives Calonge, M. C., & Beltrán Cortés, A. M. (2022). *Diseño de un proceso para la reorganización estratégica de productos en Rappi por medio de un análisis de canasta*. Universidad Javeriana.