

**Optimización del marketing digital: revisión sistemática de técnicas de Big Data y  
monitoreo de redes sociales**

Naydaly Mercedes Sanchez Posada

Asesor

Andrés Felipe Solís Pino

Universidad Nacional Abierta y a Distancia UNAD  
Escuela de ciencias Básicas, Tecnología e Ingeniería ECBTI  
Especialización en Ciencia de Datos y Analítica

2025

## Resumen

El crecimiento del comercio electrónico ha impulsado el uso de técnicas de análisis de sentimientos para interpretar las opiniones de los consumidores en plataformas digitales. Sin embargo, persisten desafíos en su adopción, especialmente en contextos emergentes. Este trabajo aborda dicha problemática mediante una revisión sistemática de literatura científica publicadas entre 2020 y 2025, utilizando la metodología PRISMA. Se analizaron 64 estudios que aplican técnicas de ciencia de datos como aprendizaje automático, aprendizaje profundo y análisis basado en aspectos. Los resultados muestran una tendencia hacia la automatización, el uso de modelos híbridos y la integración de redes sociales como fuente de datos. Se concluye que el análisis de sentimientos mejora la personalización, la reputación digital y la toma de decisiones estratégicas, aunque aún enfrenta retos como la ambigüedad del lenguaje y la necesidad de modelos explicativos.

***Palabras claves:*** Big Data, Sentiment Analysis, Ecommerce, Data Mining

### **Abstract**

The exponential growth of e-commerce has driven the adoption of sentiment analysis techniques to interpret consumer opinions on digital platforms. However, challenges remain, especially in emerging contexts. This study addresses that gap through a systematic literature review of scientific publications from 2020 to 2025, using the PRISMA methodology. A total of 64 studies were analyzed, focusing on data science techniques such as machine learning, deep learning, and aspect-based sentiment analysis. The findings reveal a trend toward automation, hybrid model integration, and the use of social media as a data source. The study concludes that sentiment analysis enhances personalization, digital reputation, and strategic decision-making, although issues like language ambiguity and the need for explainable models persist.

***Keywords:*** Big Data, Sentiment Analysis, Ecommerce, Data Mining

## Tabla Contenido

Introducción .....	8
Justificación .....	10
Objetivos.....	12
Objetivo General.....	12
Objetivos Específicos.....	12
Planteamiento del Problema .....	13
Marco Conceptual y Teórico .....	15
Big Data y su Importancia en el Marketing Digital.....	15
Metodologías de Revisión de Literatura.....	16
Planeación de la Revisión Sistemática de la Literatura .....	18
Método Prisma.....	18
Definición de las Preguntas de Investigación.....	19
Estrategia de Búsqueda.....	22
Metodología de Selección de Estudios Primarios.....	23
Criterios de Selección de Estudios Primarios .....	24
Estrategia de Extracción de Datos .....	25
Ejecución de la Revisión Sistemática de Literatura.....	27
Etapa de Ejecución.....	27
Respuestas a las Preguntas de Investigación .....	29
Otros Datos Extraídos del Corpus.....	49
Conclusiones.....	52
Recomendaciones .....	54

Referencias Bibliográficas .....	56
----------------------------------	----

## Lista de Tablas

<b>Tabla 1</b> Preguntas de Investigación Utilizadas para el Mapeo Sistemático .....	20
<b>Tabla 2</b> Cadena de Búsqueda Aplicada a Cada una de las Bases de Datos. ....	23
<b>Tabla 3</b> Criterios de Inclusión y Exclusión de la Investigación. ....	25

## Lista de Figuras

<b>Figura 1</b> <i>Proceso de Realización de Revisión Sistemática de Literatura Basado en PRISMA ...</i>	19
<b>Figura 2</b> <i>Metodología Utilizada en el Mapeo Sistemático de Literatura .....</i>	24
<b>Figura 3</b> <i>Aportes a la Investigación de Cada una de las Bases de Datos .....</i>	28
<b>Figura 4</b> <i>Frecuencia de Publicaciones Sobre las Técnicas de Ciencia de Datos Aplicadas en el Comercio Electrónico Durante los Últimos Cinco años. ....</i>	30
<b>Figura 5</b> <i>Frecuencia de las Técnicas de Análisis de Sentimientos en el Comercio Electrónico.</i>	33
<b>Figura 6</b> <i>Distribución de la Aplicación de las Técnicas de Análisis de Sentimientos en el Comercio Electrónico Según los Artículos Revisados. ....</i>	35
<b>Figura 7</b> <i>Distribución Porcentual de las Tendencias más Importantes de la Ciencia de Datos en el Comercio Electrónico Cuando se Hace uso de Análisis de Sentimientos. ....</i>	38
<b>Figura 8</b> <i>Frecuencia de Publicaciones Sobre los Retos de Ciencia de Datos Aplicadas en el Comercio Electrónico Durante los Últimos Cinco Años. ....</i>	40
<b>Figura 9</b> <i>Impactos más Frecuentes del análisis de Big Data. ....</i>	49
<b>Figura 10</b> <i>Distribución de Otros Lenguajes de Aplicación del Big Data en el Análisis de Sentimientos. ....</i>	51

## Introducción

El crecimiento exponencial del comercio electrónico ha generado una enorme cantidad de datos digitales derivados de las interacciones entre usuarios y plataformas, lo que ha impulsado el uso de técnicas de ciencia de datos para comprender mejor el comportamiento de los consumidores. Entre estas técnicas, destacan el análisis de sentimientos y el monitoreo de redes sociales, que se han convertido en herramientas clave para obtener información valiosa a partir de comentarios, reseñas y publicaciones en línea. Estas herramientas permiten a las empresas identificar patrones de compra, mejorar la experiencia del usuario y tomar decisiones estratégicas basadas en datos.

La literatura reciente revisada entre 2020 y 2025 evidencia un creciente interés en la aplicación de estas técnicas en el ámbito del comercio electrónico. Por ejemplo, el estudio (Ajmeera & Kamakshi, 2024), titulado “Hamiltonian Deep Neural Network Fostered Sentiment Analysis Approach on Product Reviews”, muestra cómo modelos de redes neuronales profundas pueden mejorar la precisión del análisis de sentimientos aplicado a reseñas de productos. De igual manera, (Arra et al., 2023a), en “Opinion Mining of Online Shopping Products Reviews Using Machine Learning”, exploran el uso de algoritmos de *machine learning* para clasificar automáticamente opiniones y extraer información sobre la percepción del consumidor.

Otras investigaciones como la de, (Rahin et al., 2022a) en “Aspect-Based Sentiment Analysis Using SemEval and Amazon Datasets”, han demostrado la efectividad del análisis basado en aspectos (*aspect-based sentiment analysis*), una técnica que permite descomponer los comentarios de los usuarios para identificar emociones asociadas a características específicas del producto. Esto ha abierto la puerta a sistemas de recomendación más personalizados y estrategias de marketing emocional basadas en la voz del cliente.

En paralelo, estudios como el de (Handayani et al., 2022), titulado “Social Network Analysis of Campaign Mobilization of Marketplace on Twitter”, destacan el valor del monitoreo de redes sociales para detectar tendencias emergentes, rastrear menciones de productos en tiempo real y analizar la influencia de actores clave en la difusión de opiniones. Estos enfoques permiten a las empresas actuar rápidamente ante crisis reputacionales, mejorar su posicionamiento y desarrollar campañas más efectivas.

No obstante, la revisión también revela importantes desafíos metodológicos y técnicos. La ambigüedad del lenguaje natural, la escasez de datos etiquetados y la dificultad para interpretar modelos complejos siguen siendo limitaciones recurrentes. Por ejemplo, (Handayani et al., 2022) en “A Comprehensive Survey of Sentiment Analysis Based on User Opinion”, señalan la necesidad de avanzar en la detección de ironía y sarcasmo, así como en el desarrollo de modelos multilingües y explicables.

Este trabajo tiene como propósito realizar una revisión sistemática de literatura científica publicada entre 2020 y 2025, centrada en las técnicas de ciencia de datos aplicadas al análisis de sentimientos y al monitoreo de redes sociales en el comercio electrónico. A partir de un corpus de 64 artículos académicos seleccionados mediante la metodología PRISMA, se busca identificar las técnicas más utilizadas, sus aplicaciones, las tendencias actuales y los principales retos que enfrenta el campo. Esta revisión permitirá establecer un estado del arte actualizado y confiable, útil tanto para la comunidad académica como para profesionales interesados en el aprovechamiento estratégico del análisis de datos en entornos digitales.

## **Justificación**

El análisis de Big Data y el monitoreo de redes sociales se han consolidado como herramientas fundamentales en el marketing digital contemporáneo. Estas tecnologías permiten comprender con mayor precisión los comportamientos, intereses y necesidades de los consumidores, al tiempo que potencian la capacidad de las empresas para anticiparse a tendencias, tomar decisiones estratégicas informadas y mejorar su competitividad en entornos de mercado dinámicos.

La relevancia de este trabajo radica en su capacidad para sistematizar y analizar el conocimiento existente mediante una revisión exhaustiva de la literatura científica relacionada con las técnicas de análisis de Big Data y monitoreo de redes sociales aplicadas al marketing digital. Este enfoque facilita la identificación de herramientas, metodologías y enfoques probados en distintos contextos, brindando una base teórica sólida para futuras investigaciones enfocadas en su viabilidad y adaptación en mercados emergentes (Arora et al., 2022).

Además, esta monografía busca llenar vacíos de conocimiento en la literatura académica sobre transformación digital, estableciendo un marco conceptual y metodológico que sirva de referencia para investigadores, académicos y profesionales del sector. La investigación contribuye no solo a fortalecer la base teórica en torno al Big Data y su aplicación en el marketing digital, sino también a generar insumos relevantes para la toma de decisiones en organizaciones que desean mejorar su posición competitiva a nivel global. En este contexto, el análisis de Big Data ofrece insights clave sobre las preferencias de los usuarios, tendencias emergentes y oportunidades de mercado (Fan et al., 2015).

Por otro lado, los datos generados en redes sociales poseen un enorme potencial para la gestión temprana de crisis corporativas. En un entorno donde las críticas, quejas y valoraciones

negativas pueden escalar rápidamente, la detección temprana de amenazas a la reputación corporativa es crucial. Aunque las técnicas avanzadas de análisis y monitoreo han demostrado ser eficaces en diversos contextos, muchas organizaciones aún enfrentan desafíos en su implementación, ya sea por desconocimiento, falta de personal capacitado o limitaciones tecnológicas (Andrade et al., 2024).

Desde una perspectiva académica, esta investigación pretende contribuir al cuerpo de conocimiento sobre Big Data y su relación con el marketing digital en contextos emergentes, mediante la identificación y análisis de estudios previos. Esto permitirá llenar vacíos existentes en la literatura y ofrecer un marco conceptual útil para investigaciones futuras (Bermúdez-Quimis, 2024).

Finalmente, la pertinencia de este trabajo también se refleja en su alineación con los objetivos estratégicos globales de transformación digital y competitividad empresarial. La adopción de tecnologías avanzadas como el Big Data no solo mejora la eficiencia operativa de las empresas, sino que también fomenta una economía más innovadora y alineada con las dinámicas del mercado global. En este sentido, este estudio no solo representa un aporte académico, sino también una herramienta práctica para impulsar el desarrollo económico y tecnológico de las organizaciones (Cardona & Salinas, 2024).

## **Objetivos**

### **Objetivo General**

Realizar una revisión sistemática de la literatura científica sobre técnicas de análisis de Big Data y monitoreo de redes sociales aplicadas al marketing digital, identificando tendencias, metodologías y su aplicabilidad en diversos contextos.

### **Objetivos Específicos**

Seleccionar y justificar un método adecuado para realizar una revisión bibliográfica sistemática que garantice la exhaustividad y relevancia del análisis.

Diseñar un protocolo detallado para la revisión de la literatura, definiendo criterios de inclusión y exclusión, bases de datos científicas, palabras clave y enfoques temáticos relacionados con Big Data y monitoreo de redes sociales en marketing digital.

Analizar la literatura científica para identificar técnicas específicas de análisis de Big Data y monitoreo de redes sociales, evaluando su aplicabilidad en diferentes contextos del marketing digital.

Determinar las tendencias clave, beneficios y desafíos relacionados con la adopción de estas técnicas, y explorar su viabilidad en contextos emergentes y diversos sectores de la economía digital.

Contextualizar los hallazgos identificando brechas existentes y posibles áreas de aplicación que contribuyan al fortalecimiento del marketing digital en distintos entornos organizacionales.

## Planteamiento del Problema

En el entorno actual, caracterizado por la transformación digital y el incremento exponencial de datos, las redes sociales se han consolidado como una herramienta fundamental en las estrategias de marketing digital. Empresas de todos los sectores utilizan estas plataformas para interactuar con sus audiencias, mejorar su posicionamiento de marca y optimizar sus esfuerzos comerciales. En este contexto, el Big Data ha surgido como una herramienta poderosa que permite analizar grandes volúmenes de datos generados en redes sociales, proporcionando información valiosa para la toma de decisiones estratégicas (Manyika, 2011).

Uno de los problemas principales es la baja adopción de técnicas avanzadas de análisis de datos en entornos digitales emergentes, especialmente en el comercio electrónico. Factores como la falta de acceso a tecnologías, la escasez de profesionales capacitados y un desconocimiento generalizado sobre el potencial del Big Data para optimizar procesos de marketing digital limitan su aplicación práctica. Esto dificulta que muchas organizaciones puedan competir en igualdad de condiciones en un mercado cada vez más orientado a la personalización de experiencias y la predicción del comportamiento de los consumidores (MinTIC, 2020).

Además, la literatura científica y profesional sobre este tema se ha concentrado mayoritariamente en mercados desarrollados, lo que genera una brecha en el conocimiento sobre cómo estas técnicas pueden ser aplicadas efectivamente en escenarios de comercio digital en regiones con menor nivel de adopción tecnológica. Por ejemplo, mientras que en dichos contextos se implementan técnicas avanzadas como análisis de sentimientos en tiempo real y modelado predictivo, en otros sectores esta integración sigue siendo incipiente debido a barreras estructurales, formativas y presupuestales (Avella Chaparro & Rodríguez, 2024). Esta falta de evidencia localizada crea la necesidad de una revisión sistemática que identifique las técnicas

más documentadas de análisis de Big Data y monitoreo de redes sociales aplicadas al marketing digital, especialmente en contextos de comercio electrónico y transformación digital (Avella Chaparro & Rodríguez, 2024)

Otra dificultad relevante es la gestión de crisis en redes sociales. En un entorno donde los consumidores expresan sus opiniones en tiempo real, las marcas deben contar con capacidades analíticas que les permitan responder de forma oportuna. Según Forrester Research (2020), el 76% de los usuarios en plataformas sociales espera respuestas rápidas a sus inquietudes, lo que subraya la importancia de técnicas de análisis como el monitoreo de redes, análisis de sentimientos y minería de opiniones para anticipar riesgos reputacionales (Reyes Hidalgo, s. f.).

En este sentido, la problemática central radica en la falta de conocimiento, formación y adopción efectiva de técnicas de análisis de Big Data y monitoreo de redes sociales en las estrategias digitales de sectores que participan del comercio electrónico. Esto no solo limita su competitividad, sino que también representa una subutilización del enorme potencial que los datos ofrecen para la innovación, la personalización y la optimización del marketing digital. Esta monografía busca contribuir a cerrar esa brecha mediante una revisión sistemática de literatura, que permita comprender las dinámicas actuales de estas tecnologías y brindar una base para su aprovechamiento en entornos digitales altamente competitivos y en constante evolución (González, 2019).

En este contexto, surge la necesidad de identificar qué enfoques metodológicos han demostrado ser más efectivos en la literatura científica reciente para abordar los desafíos del comercio electrónico mediante el análisis de datos. Por ello, esta investigación se orienta a responder la siguiente pregunta: ¿Cuáles técnicas de análisis de sentimientos, basadas en Big Data y monitoreo de redes sociales, son más efectivas para mejorar la competitividad?

## Marco Conceptual y Teórico

### Big Data y su Importancia en el Marketing Digital

El término Big Data hace referencia al procesamiento y análisis de grandes volúmenes de datos estructurados y no estructurados. Estas técnicas permiten extraer patrones, tendencias y conocimiento útil para la toma de decisiones estratégicas. En el contexto del marketing digital, Big Data proporciona la capacidad de analizar el comportamiento del consumidor, personalizar experiencias, optimizar campañas publicitarias y mejorar la retención de clientes. También permite una segmentación avanzada de audiencias mediante la integración de datos demográficos, psicológicos y conductuales (Ogbeide et al., 2024; Y. Yang et al., 2024).

1. Big Data se caracteriza por las llamadas 5V (Kitchin, 2014a):
2. Volumen: grandes cantidades de datos generados de manera continua.
3. Velocidad: rapidez en la creación y procesamiento de los datos.
4. Variedad: diversidad de formatos (texto, imagen, vídeo, etc.).
5. Veracidad: calidad y precisión de los datos.
6. Valor: capacidad de traducir los datos en conocimiento útil.

Estas características hacen del Big Data una herramienta estratégica para mejorar la experiencia del cliente y lograr una mayor fidelización mediante propuestas personalizada (Chen et al., 2014).

#### Monitoreo de redes sociales y análisis de sentimientos

El monitoreo de redes sociales consiste en el uso de herramientas y tecnologías para observar, recopilar y analizar datos en tiempo real provenientes de plataformas como Facebook, Twitter, Instagram, TikTok y LinkedIn. Estas redes constituyen fuentes valiosas de datos que

incluyen comentarios, valoraciones, emociones y percepciones de los usuarios sobre productos, marcas o servicios (Kitchin, 2014b; Li & Bernoff, 2008).

Entre los tipos de análisis más destacados se encuentran:

1. Análisis de sentimientos: permite detectar emociones (positivas, negativas o neutras) asociadas a mensajes o marcas (Zeng, 2010).

2. Análisis de redes: identifica conexiones entre usuarios o comunidades de influencia (Kaplan & Haenlein, 2010).

3. Análisis de influencia: detecta líderes de opinión capaces de amplificar mensajes clave (Li & Bernoff, 2008).

Estas técnicas se aplican con el apoyo de herramientas especializadas como Hootsuite, Brandwatch, Sprout Social y Talkwalker, que son ampliamente utilizadas para la inteligencia digital (Davenport, 2014). Relación entre Big Data, redes sociales y comercio electrónico.

La integración de Big Data con el monitoreo de redes sociales potencia la capacidad de análisis predictivo en entornos de comercio electrónico. El flujo constante de datos que se genera en redes sociales permite a las empresas identificar patrones de comportamiento del consumidor, medir el impacto de sus campañas digitales y anticipar crisis de reputación. Así, se promueve la creación de campañas personalizadas en tiempo real y se optimiza el retorno de la inversión publicitaria (Davenport, 2014; Tsai et al., 2014).

### **Metodologías de Revisión de Literatura**

Las revisiones de literatura científica pueden realizarse mediante metodologías como la revisión sistemática, la revisión bibliométrica y el metaanálisis.

La revisión sistemática establece un protocolo riguroso con criterios de inclusión y exclusión, garantizando la objetividad y replicabilidad del análisis (Kitchenham, 2007).

La revisión bibliométrica analiza de forma cuantitativa las publicaciones, palabras clave, autores, tendencias y redes de colaboración científica (Garza Garza, 2010).

Este trabajo adopta un enfoque combinado, que integra elementos de revisión sistemática y bibliométrica para explorar el estado del arte sobre análisis de Big Data, monitoreo de redes sociales y análisis de sentimientos en el comercio electrónico. Aportaciones teóricas a la investigación: Esta monografía se fundamenta en los siguientes marcos teóricos:

1. Teoría de la personalización (Pine et al., 1999): los consumidores valoran experiencias personalizadas, y el análisis de Big Data permite adaptar los mensajes y contenidos a cada perfil.

2. Teoría de la adopción tecnológica (Davis & others, 1989): analiza los factores que influyen en la adopción de tecnologías en contextos organizacionales. Esta teoría resulta especialmente pertinente para este estudio, ya que permite interpretar las barreras y facilitadores en la implementación de técnicas de Big Data y monitoreo de redes sociales en el marketing digital, particularmente en contextos emergentes como el colombiano. En coherencia con esta perspectiva, la presente monografía adopta la metodología PRISMA, tal como fue planteado en el anteproyecto, para realizar una revisión sistemática de literatura que permita identificar las técnicas más documentadas y su aplicabilidad en dichos entornos.

## **Planeación de la Revisión Sistemática de la Literatura**

Este capítulo describe la planificación metodológica para la revisión sistemática de literatura. Se detallan las estrategias de búsqueda, criterios de selección, preguntas de investigación y técnicas de extracción de datos, siguiendo el enfoque PRISMA.

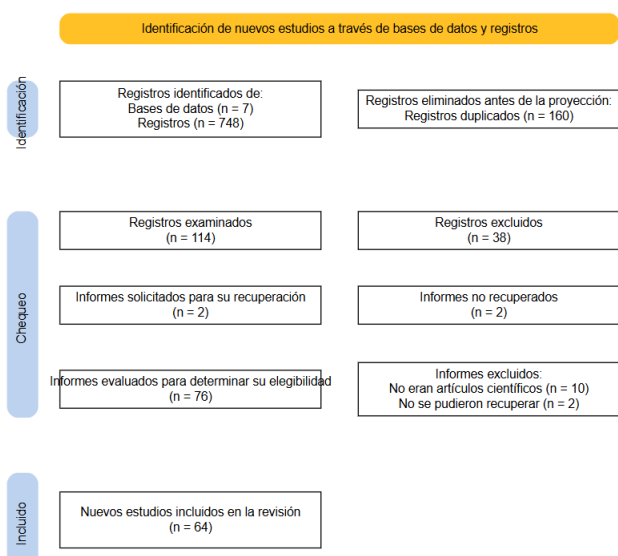
### **Método Prisma**

La revisión sistemática de literatura es una metodología reconocida que permite recopilar y analizar de forma rigurosa y estructurada el conocimiento existente sobre un tema específico. En este trabajo, se adoptó el enfoque PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) como guía metodológica para garantizar la transparencia, calidad y replicabilidad del proceso de búsqueda, selección y síntesis de los artículos científicos revisados (Moher et al., 2009). Este estudio se estructuró en cinco etapas, como se observa en la Figura 1: (1) Definición de las preguntas de investigación, alineadas con los objetivos de la monografía y el enfoque temático (Big Data, análisis de sentimientos, monitoreo de redes sociales y comercio electrónico). Estas preguntas orientaron la selección de fuentes y el análisis posterior. (2) Diseño y ejecución de la estrategia de búsqueda, construida bajo el enfoque PICOC (Población, Intervención, Comparación, Outcome, Contexto). Se aplicaron cadenas de búsqueda específicas en bases de datos científicas reconocidas como IEEE Xplore, Scopus, Web of Science, Dimensions, The Lens, ScienceDirect, ACM Digital Library y Google Scholar. La búsqueda se limitó a artículos publicados entre 2020 y 2025, en inglés o español, con revisión por pares. (3) Selección y revisión inicial de los artículos, aplicando filtros por título, resumen y palabras clave. Se eliminaron duplicados y se excluyeron documentos que no abordaban directamente el análisis de sentimientos en el comercio electrónico. (4) Aplicación de criterios de inclusión y exclusión, definidos previamente para garantizar la pertinencia, calidad y relevancia de los estudios. Se

incluyeron artículos que presentaran aplicaciones concretas de técnicas de ciencia de datos en el análisis de sentimientos, y se excluyeron revisiones secundarias, artículos sin acceso completo o sin validación empírica. (5) Extracción y análisis de datos, mediante matrices de codificación que permitieron organizar la información, identificar patrones, clasificar técnicas y aplicaciones, y responder a las preguntas de investigación. Este proceso también permitió generar visualizaciones como gráficos de evolución temporal y combinaciones metodológicas.

## Figura 1

### *Proceso de Realización de Revisión Sistemática de Literatura Basado en PRISMA*



A continuación, se presenta el proceso de elaboración del mapeo sistemático.

### **Definición de las Preguntas de Investigación**

Las preguntas de investigación planteadas en esta monografía buscan establecer en qué contexto se están aplicando las técnicas de ciencia de datos, particularmente el análisis de sentimientos, en el ámbito del comercio electrónico. El objetivo es identificar vacíos en el conocimiento, tendencias emergentes y enfoques metodológicos predominantes, con el fin de

comprender cómo estas técnicas están siendo utilizadas para mejorar la experiencia del cliente, optimizar estrategias comerciales y fortalecer la reputación digital de las plataformas. En la Tabla 1 se presentan cada una de las preguntas junto con su motivación. A partir de estas, se seleccionó, analizó y categorizó la información extraída de los artículos científicos revisados, lo que permitió estructurar los hallazgos de manera coherente y alineada con los objetivos del estudio.

**Tabla 1**

*Preguntas de Investigación Utilizadas para el Mapeo Sistemático*

Pregunta de investigación	Motivante	Respuestas esperadas
¿Cuál es la frecuencia de publicaciones sobre las técnicas de ciencia de datos aplicadas en el comercio electrónico?	Identificar la cantidad de estudios recientes y observar si hay un incremento en la producción científica sobre el tema, lo cual indicaría su relevancia actual.	Número de publicaciones por año. Base de datos científica más frecuente. Revistas o conferencias más citadas.
¿Cuáles son las técnicas de ciencia de datos aplicadas en el comercio electrónico empleando análisis de sentimientos?	Reconocer qué enfoques están siendo usados actualmente en el análisis de sentimientos y cómo se aplican al e-commerce.	Machine Learning Deep Learning NLP Modelos híbridos
¿Cómo se están aplicando las técnicas de ciencia de datos con el análisis de sentimientos para el comercio electrónico?	Comprender de qué manera estas técnicas están ayudando en áreas clave del comercio electrónico como la personalización, la atención al cliente o el análisis de reputación.	Casos de aplicación: Recomendadores Atención al cliente automatizado Evaluación de reseñas Seguimiento reputacional

Pregunta de investigación	Motivante	Respuestas esperadas
¿Cuáles son las tendencias más importantes de la ciencia de datos en el comercio electrónico cuando se hace uso de análisis de sentimientos?	Detectar hacia dónde se están dirigiendo las investigaciones y las tecnologías más emergentes en este cruce temático.	Tendencias: Uso de IA generativa Análisis multimodal (texto + imagen) Integración en CRM y ERP Automatización del marketing
¿Cuáles son los retos más importantes de la ciencia de datos con el análisis de sentimientos para el comercio electrónico?	Identificar limitaciones actuales que enfrentan las empresas al implementar estas técnicas.	Retos: Calidad de los datos Ambigüedad del lenguaje Costos tecnológicos Falta de talento analítico
¿Cómo se está haciendo el monitoreo de redes sociales mediante la ciencia de datos aplicada al comercio electrónico?	Explorar cómo las empresas están extrayendo información útil de redes sociales y qué herramientas o procesos aplican.	Herramientas y prácticas: Social Listening Dashboards en tiempo real APIs (Twitter, Meta) Plataformas como Brandwatch, Talkwalker
¿Cómo mejora el análisis de sentimientos el comercio electrónico?	Establecer el valor agregado que ofrece esta técnica al rendimiento, reputación o satisfacción del cliente.	Beneficios: Mejora de experiencia de usuario Personalización de contenidos Reducción de churn Optimización de estrategias de marketing
¿Cómo las empresas de comercio electrónico pueden aprovechar la minería de opiniones para mejorar su reputación digital?	Evidenciar cómo el análisis de reseñas y comentarios en línea puede ser usado para gestionar la imagen corporativa.	Prácticas: Identificación de aspectos críticos Respuestas automatizadas Campañas correctivas Insights para desarrollo de productos

Pregunta de investigación	Motivante	Respuestas esperadas
¿Qué impacto tiene el análisis de Big Data en la identificación de tendencias y patrones en la compra en el comercio electrónico?	Analizar cómo el Big Data permite descubrir patrones de consumo que ayudan en la toma de decisiones empresariales.	Impactos: Segmentación avanzada Predicción de demanda Diseño de promociones Forecasting de ventas

### **Estrategia de Búsqueda**

Para llevar a cabo la revisión sistemática de literatura, se diseñó una estrategia de búsqueda (Tabla 2) estructurada con base en el enfoque PICOC (Población, Intervención, Comparación, Outcome y Contexto), el cual permite organizar de forma lógica y coherente los conceptos clave del estudio (Souza et al., 2018). Esta metodología facilitó la identificación de investigaciones relevantes sobre el uso de técnicas de ciencia de datos, análisis de sentimientos y monitoreo de redes sociales en el contexto del comercio electrónico.

La búsqueda se ejecutó en ocho bases de datos científicas reconocidas por su cobertura temática y rigurosidad académica: IEEE Xplore, Scopus, Web of Science, ScienceDirect, Dimensions, The Lens, ACM Digital Library y Google Scholar. Estas fuentes fueron seleccionadas por su relevancia en los campos de la ingeniería, la informática, el análisis de datos y el marketing digital.

La cadena de búsqueda fue aplicada a los campos de título, resumen y palabras clave en todas las plataformas, con excepción de ScienceDirect, donde se utilizó una versión reducida debido a la limitación de operadores lógicos impuesta por el sistema. Esta adaptación mantuvo los términos esenciales para asegurar la coherencia de los resultados.

Adicionalmente, se realizó una búsqueda manual en Google Scholar con el fin de complementar los resultados obtenidos en las bases académicas tradicionales. Esta búsqueda se

centró en artículos altamente citados, recientes y pertinentes al enfoque temático de la monografía.

Finalmente, se estableció una ventana temporal de cinco años (2020–2025), con el objetivo de garantizar la actualidad de los hallazgos y su pertinencia en un entorno digital caracterizado por su rápida evolución.

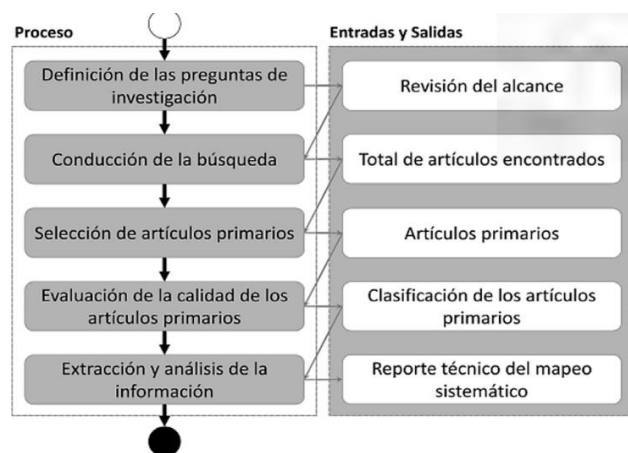
## **Tabla 2**

### *Cadena de Búsqueda Aplicada a Cada una de las Bases de Datos*

Cadena de búsqueda
<p>("Big Data" OR "Data Science" OR "Data Mining" OR "Data Analytics") AND ("Social Media Monitoring" OR "Social media analytics" OR "Social media measurement" OR "Social media mining" OR "Social media intelligence" OR "Social network analysis" OR "Sentiment Analysis" OR "Opinion Mining" OR "emotion AI" OR "Consumer sentiment") AND ("Ecommerce" OR "Electronic Commerce" OR "Online Retail" OR "Digital Commerce" OR "Online shopping" OR "e-business" OR "internet commerce")</p>

## **Metodología de Selección de Estudios Primarios**

Para garantizar un proceso riguroso y sistemático en la revisión de literatura, se adoptaron las directrices metodológicas propuestas por Petersen et al. (2008), adaptadas al enfoque PRISMA. El objetivo fue asegurar la trazabilidad, coherencia y validez de los estudios seleccionados, permitiendo una síntesis confiable del conocimiento científico sobre el análisis de sentimientos aplicado al comercio electrónico. El proceso se desarrolló de forma iterativa y estructurada, como se muestra en la Figura 2.

**Figura 2***Metodología Utilizada en el Mapeo Sistemático de Literatura*

*Nota.* Tomado de ResearchGate, s.f., recuperado de <https://www.researchgate.net/>

**Criterios de Selección de Estudios Primarios**

Siguiendo las directrices metodológicas propuestas por Petersen et al. (2008), se establecieron criterios de inclusión y exclusión (ver Tabla 3) con el fin de garantizar la pertinencia, calidad y relevancia de los estudios seleccionados para esta revisión sistemática. Estos criterios permitieron filtrar de manera rigurosa los artículos científicos, asegurando que los documentos analizados respondieran adecuadamente a las preguntas de investigación planteadas (Yassin et al., 2018).

Un estudio fue considerado relevante si cumplía con todos los criterios de inclusión. Por el contrario, si un artículo presentaba al menos una condición de exclusión, fue descartado del análisis.

**Tabla 3***Criterios de Inclusión y Exclusión de la Investigación.*

Criterios de inclusión
Artículos publicados entre 2019 y 2024.
Estudios escritos en inglés o español.
Investigaciones que apliquen técnicas de análisis de sentimientos en el contexto del comercio electrónico.
Artículos con revisión por pares, publicados en revistas científicas o conferencias indexadas.
Estudios que presenten aplicaciones, modelos o resultados empíricos.
Criterios de exclusión
Estudios secundarios, revisiones narrativas o artículos sin validación empírica.
Documentos sin acceso completo al texto.
Publicaciones duplicadas o versiones preliminares del mismo estudio.
Artículos que no aborden directamente el análisis de sentimientos o su aplicación en e-commerce.

**Estrategia de Extracción de Datos**

La estrategia de extracción de datos se diseñó con el propósito de recopilar y organizar la información relevante para responder de manera sistemática a cada una de las preguntas de investigación planteadas. Para ello, se construyó una matriz de análisis que relaciona cada pregunta con su motivación y las respuestas esperadas, lo que permitió establecer categorías temáticas y criterios de codificación homogéneos.

Este enfoque facilitó la identificación de patrones, técnicas, aplicaciones, beneficios y retos documentados en los artículos revisados. La extracción se realizó de forma manual, asegurando la trazabilidad de cada dato con su fuente original. Además, se registraron elementos complementarios como el año de publicación, la base de datos de origen, el tipo de técnica utilizada (por ejemplo, aprendizaje automático, procesamiento de lenguaje natural o modelos híbridos), y el tipo de aplicación (como sistemas de recomendación, análisis de reputación o personalización de contenidos).

La información extraída fue organizada en función de su correspondencia con las preguntas de investigación, lo que permitió una síntesis estructurada de los hallazgos. Esta estrategia también permitió identificar tendencias emergentes, desafíos metodológicos y oportunidades de aplicación del análisis de sentimientos en el comercio electrónico.

## **Ejecución de la Revisión Sistemática de Literatura**

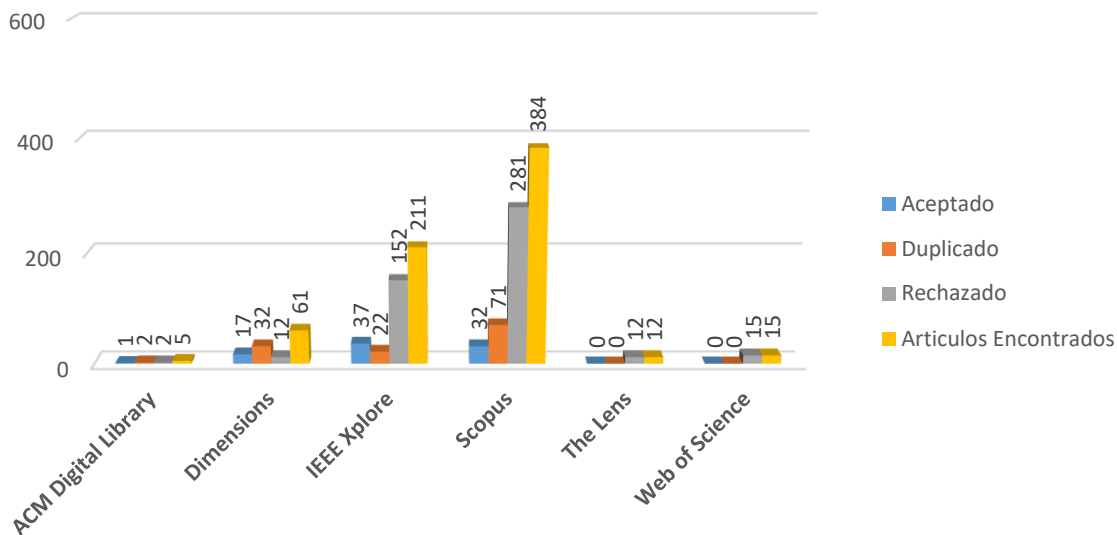
### **Etapas de Ejecución**

En esta etapa se llevó a cabo la implementación del protocolo metodológico definido previamente, estructurado en tres fases consecutivas. Para facilitar la organización y trazabilidad del proceso, se utilizó la herramienta en línea Parsifal (<https://parsif.al/>), especializada en revisiones sistemáticas. En la fase 1, se aplicó la cadena de búsqueda en las bases de datos seleccionadas, lo que permitió recuperar un total de 748 registros iniciales. Posteriormente, se aplicaron los criterios de inclusión y exclusión sobre los títulos, resúmenes y palabras clave de cada artículo, además de eliminar duplicados. Como resultado, se obtuvieron 114 documentos considerados como “candidatos a estudios primarios”. Luego en la fase 2, se realizó una revisión completa del texto de cada uno de los artículos preseleccionados. Esta evaluación permitió verificar su pertinencia temática, calidad metodológica y alineación con los objetivos del estudio. Tras este proceso, se seleccionaron 64 artículos que cumplieron con todos los criterios establecidos, los cuales fueron denominados “estudios primarios”.

En la fase 3, se procedió a la extracción de datos, caracterización de los estudios y codificación de la información relevante en función de las preguntas de investigación. Este proceso incluyó la identificación de técnicas utilizadas, aplicaciones prácticas, beneficios reportados, retos metodológicos y tendencias emergentes. En los casos en que se identificaron estudios duplicados, se priorizó la versión con mayor índice de citación y acceso completo. Los estudios secundarios fueron excluidos del análisis principal, aunque se utilizaron como referencia comparativa para validar algunos hallazgos.

**Figura 3**

*Aportes a la Investigación de Cada una de las Bases de Datos*



En la Figura 3 se detalla la contribución de cada una de las bases de datos consultadas, así como la cantidad de artículos aceptados, duplicados y rechazados. Se observa que la base de datos Scopus fue la que mayor número de artículos recuperó (384 en total), aunque también presentó el mayor volumen de documentos rechazados (281) y duplicados (71). Por su parte, IEEE Xplore Digital Library fue la fuente que más artículos primarios aportó al estudio, con un total de 37 aceptados. En contraste, bases como The Lens y Web of Science no contribuyeron con estudios aceptados, ya que todos los artículos encontrados en estas plataformas fueron descartados tras la aplicación de los criterios de selección. Esta distribución evidencia la importancia de aplicar un proceso riguroso de depuración y validación, así como la relevancia de ciertas bases de datos en la construcción del corpus documental.

A continuación, se da paso a la solución de las preguntas de investigación planteadas anteriormente de acuerdo a los resultados obtenidos a partir de los estudios selección.

## Respuestas a las Preguntas de Investigación

Una vez aplicada la estrategia de búsqueda en las bases de datos seleccionadas, se obtuvieron inicialmente 748 registros. En la primera fase del proceso de selección, se realizó un análisis preliminar de los títulos, resúmenes y palabras clave, lo que permitió reducir el conjunto a 114 estudios potencialmente relevantes. Posteriormente, tras la lectura completa de los textos, se seleccionaron 64 artículos que cumplieran con todos los criterios establecidos, conformando así el conjunto final de estudios primarios incluidos en la revisión.

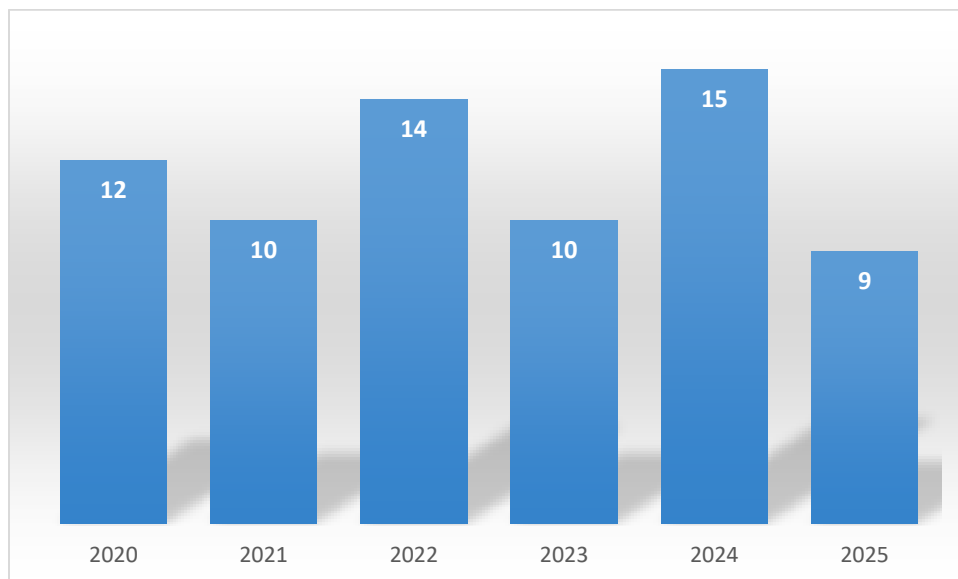
*1. ¿Cuál es la frecuencia de publicaciones sobre las técnicas de ciencia de datos aplicadas en el comercio electrónico?*

Con base en los datos recopilados entre los años 2020 y 2025, se observa que la frecuencia de publicaciones sobre técnicas de ciencia de datos aplicadas en el comercio electrónico ha sido variable. El pico en 2024 podría correlacionarse con la popularización de modelos de lenguaje grandes (LLMs) y su accesibilidad a través de APIs, lo que facilitó nuevas investigaciones (esta información no proviene de las fuentes proporcionadas y puede requerir verificación independiente). El año con mayor número de publicaciones fue 2024, con un total de 15, mientras que el año con menor producción fue 2025, con solo 9 publicaciones. El promedio anual de publicaciones durante este periodo fue de aproximadamente 11,7.

Esta oscilación en la frecuencia sugiere que, si bien existe un interés sostenido en el tema, la producción académica puede estar influenciada por factores externos como avances tecnológicos, prioridades investigativas o coyunturas del mercado. La representación gráfica de estos datos permite visualizar de manera clara los picos y descensos en la actividad investigativa, facilitando así una comprensión más precisa del comportamiento temporal de las publicaciones en este campo.

**Figura 4**

*Frecuencia de Publicaciones Sobre las Técnicas de Ciencia de Datos Aplicadas en el Comercio Electrónico Durante los Últimos Cinco años.*



*2. ¿Cuáles son las técnicas de ciencia de datos aplicadas en el comercio electrónico empleando análisis de sentimientos?*

A partir de la revisión sistemática de 64 artículos científicos publicados entre 2020 y 2025, se identificaron diversas técnicas de ciencia de datos aplicadas al análisis de sentimientos en el comercio electrónico. Estas técnicas tienen como objetivo principal interpretar las emociones, valoraciones y percepciones que los consumidores expresan mediante reseñas y comentarios, permitiendo así que las plataformas electrónicas puedan optimizar sus procesos comerciales y de atención al cliente.

Las técnicas identificadas se agrupan en dos grandes enfoques. El primero corresponde a los métodos basados en léxicos, que utilizan diccionarios de polaridad para asignar valores sentimentales a palabras o frases. Este enfoque fue mencionado en 16 artículos (25%) y se

caracteriza por su escalabilidad y bajo costo computacional. Por ejemplo, (Liu et al., 2024) combinaron léxicos sentimentales con redes neuronales para analizar opiniones de usuarios en plataformas chinas, logrando una mejora del 8.5% en la precisión respecto al uso exclusivo de léxicos. El uso de recursos como SentiWordNet y VADER permite realizar análisis de polaridad sin necesidad de entrenamiento supervisado, lo que resulta útil en contextos con escasez de datos etiquetados.

El segundo enfoque corresponde a los métodos basados en aprendizaje automático, que incluyen tanto técnicas tradicionales como modelos avanzados de aprendizaje profundo. El machine learning clásico fue utilizado en 19 artículos (30%), destacándose algoritmos como Support Vector Machines (SVM), Naïve Bayes y árboles de decisión. Estas técnicas siguen siendo ampliamente empleadas debido a su facilidad de implementación y eficiencia computacional.(Arra et al., 2023a), por ejemplo, aplicaron SVM y Naïve Bayes para clasificar reseñas de productos en Amazon, alcanzando una precisión promedio del 87%.

Por otro lado, el uso de deep learning se encontró en 21 estudios (33%), con modelos como redes neuronales convolucionales (CNN), redes recurrentes (LSTM, BiLSTM) y transformers. Estas técnicas permiten modelar el contexto profundo de los textos y detectar sentimientos implícitos. Un caso destacado es el de (Ajmeera & Kamakshi, 2024), quienes propusieron una red neuronal hamiltoniana para mejorar la precisión del análisis en reseñas complejas, logrando una tasa de acierto del 92.3%, superior a la obtenida con modelos tradicionales.

Además, el procesamiento del lenguaje natural (PLN) fue identificado en 30 artículos (47%). Esta técnica facilita la extracción de significado a partir de texto no estructurado, permitiendo identificar entidades, relaciones semánticas y estructuras gramaticales relevantes.(H.

K et al., 2023). integraron PLN con modelos clasificadores para analizar reseñas de productos, obteniendo una mejora del 15% en la precisión del etiquetado emocional frente a métodos sin PLN.

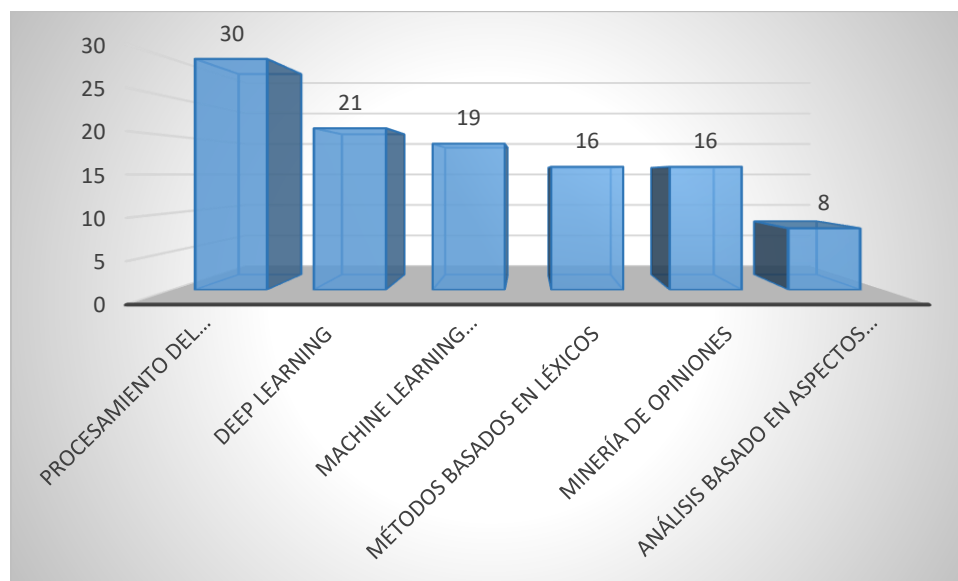
También se identificó el uso de minería de opiniones, presente en 16 artículos (25%), enfocada en extraer términos relevantes y emociones específicas desde grandes volúmenes de texto sin necesidad de etiquetas manuales. Esta técnica resulta especialmente útil en entornos donde la generación de datos es constante y masiva.

Una técnica más especializada es el análisis basado en aspectos (Aspect-Based Sentiment Analysis), utilizada en 8 artículos (13%). Esta técnica descompone las reseñas en componentes específicos del producto (como batería, precio o servicio), asignando polaridades a cada uno. (Rahin et al., 2022a) aplicaron este enfoque utilizando los datasets de Amazon y SemEval, logrando una precisión del 89% en la asignación de polaridad a los aspectos evaluados por los usuarios. Este tipo de análisis es particularmente valioso para identificar fortalezas y debilidades específicas de los productos, y para orientar mejoras en el diseño y la funcionalidad.

En conjunto, estas técnicas permiten transformar información subjetiva en datos estructurados que pueden alimentar modelos predictivos, sistemas de recomendación y estrategias de marketing personalizadas. La diversidad de enfoques y su aplicación en distintos contextos evidencian el dinamismo del campo y la importancia del análisis de sentimientos como herramienta clave en el comercio electrónico.

**Figura 5**

*Frecuencia de Técnicas de Análisis de Sentimientos en el Comercio Electrónico*



3. *¿Cómo se están aplicando las técnicas de ciencia de datos con el análisis de sentimientos para el comercio electrónico?*

Las técnicas de ciencia de datos aplicadas al análisis de sentimientos en el comercio electrónico se utilizan principalmente para la clasificación automática de reseñas, siendo esta la aplicación más común identificada en 36 de los 64 artículos revisados. Esta tarea consiste en categorizar las opiniones de los usuarios según su polaridad (positiva, negativa o neutral), utilizando modelos de aprendizaje supervisado y métodos semánticos (L. Yang et al., 2020); (Sharm et al., 2022).

Otra aplicación relevante es la predicción del comportamiento del consumidor, presente en 19 estudios, en la cual los algoritmos de ciencia de datos permiten anticipar patrones de compra basándose en las emociones expresadas en reseñas previas. Por ejemplo, algunos

artículos utilizan modelos de redes neuronales o análisis semántico profundo para anticipar la intención de compra o abandono de un producto (Gao, 2024; Gubbala & Kaur, 2025).

La detección de emociones específicas como alegría, enojo o decepción también ocupa un papel destacado, mencionada en 17 artículos. Esta técnica va más allá de la polaridad básica y permite refinar la segmentación emocional de los consumidores, lo que es clave en la personalización de campañas publicitarias (Qu, 2022; Zuo, 2024a).

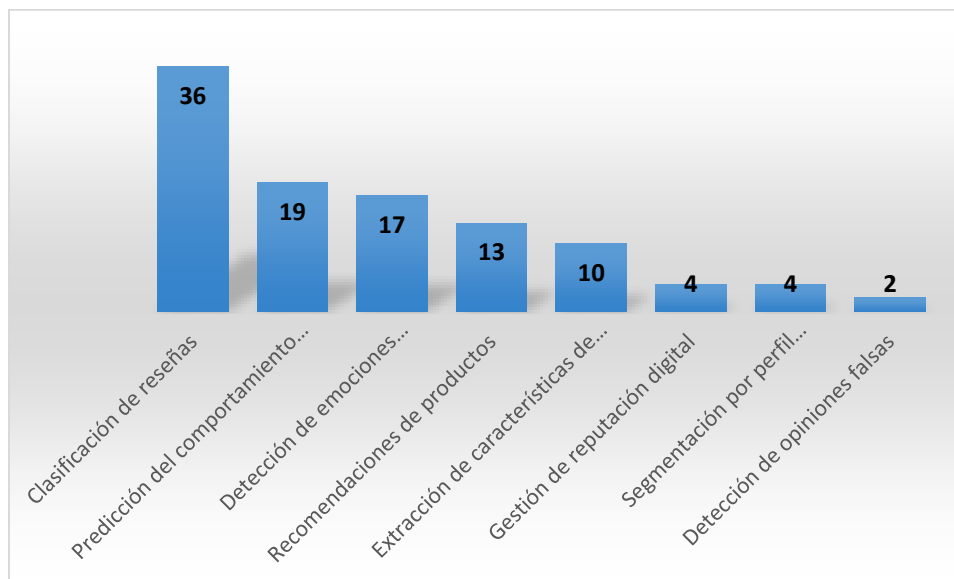
El uso del análisis de sentimientos para recomendaciones de productos se observa en 13 estudios, donde las valoraciones emocionales influyen directamente en los sistemas de recomendación. Esto mejora la experiencia del usuario al sugerir productos con base en opiniones similares o aspectos valorados positivamente (Ma et al., 2025a; Patidar & Patel, 2022).

También se encontraron aplicaciones enfocadas en la extracción de características o aspectos de productos (10 artículos), como desempeño, diseño, funcionalidad o servicio, utilizando técnicas como el análisis basado en aspectos (Ahmad et al., 2025a; Rahin et al., 2022a).

Otras aplicaciones menos frecuentes incluyen la gestión de la reputación digital (4 artículos), la segmentación de usuarios por perfil emocional (4 artículos), y la detección de opiniones falsas o engañosas (solo 2 estudios), lo cual es relevante para mejorar la confiabilidad de las plataformas de comercio electrónico (Hassan & Islam, s. f.; M et al., 2025a).

**Figura 6**

*Distribución de la Aplicación de las Técnicas de Análisis de Sentimientos en el Comercio Electrónico Según los Artículos Revisados*



4. *¿Cuáles son las tendencias más importantes de la ciencia de datos en el comercio electrónico cuando se hace uso de análisis de sentimientos?*

En la revisión de los 64 artículos del corpus, se identificaron varias tendencias predominantes en el uso de técnicas de ciencia de datos aplicadas al análisis de sentimientos en el contexto del comercio electrónico. Estas tendencias reflejan una evolución desde enfoques descriptivos hacia técnicas predictivas y prescriptivas, impulsadas por la necesidad de comprender mejor al consumidor digital, automatizar procesos y personalizar la experiencia de compra.

Una de las tendencias más destacadas es el uso de algoritmos de Deep Learning, especialmente redes neuronales profundas como BiLSTM, CNN y transformers. Estas técnicas han ganado protagonismo debido a su capacidad para modelar relaciones complejas y

contextuales en el lenguaje natural, superando la precisión de los métodos tradicionales. Su adopción responde a la creciente disponibilidad de datos no estructurados y a la necesidad de interpretar sentimientos implícitos en textos informales. Por ejemplo, (Sharm et al., 2022) demostraron que los modelos de deep learning aplicados a reseñas de Amazon mejoran significativamente la precisión frente a enfoques clásicos.

Otra tendencia relevante es el análisis basado en aspectos (aspect-based sentiment analysis), que permite descomponer las opiniones en componentes específicos del producto (como precio, calidad o servicio) y evaluar los sentimientos asociados a cada uno. Esta técnica ha ganado popularidad porque ofrece insights más detallados y accionables para la mejora continua de productos y servicios. (Ahmad et al., 2025a) destacan su utilidad en contextos multilingües, mientras que (Rahin et al., 2022a) lograron una precisión del 89% al aplicarla sobre los datasets de Amazon y SemEval. Esta tendencia se explica por la necesidad de ir más allá de la polaridad general y entender qué atributos específicos generan satisfacción o insatisfacción.

Una tercera tendencia es la automatización en tiempo real del análisis de opiniones, impulsada por la demanda de respuestas inmediatas ante el comportamiento del consumidor digital. Esta automatización se ha visto favorecida por el desarrollo de herramientas de análisis en la nube y por la integración de modelos de análisis de sentimientos directamente en las plataformas de e-commerce. (Gao, 2024) propuso un modelo de análisis de comportamiento basado en big data que permite predecir patrones de compra y abandono en tiempo real. Actualmente, esta tendencia está transformando la forma en que las empresas monitorean la experiencia del cliente, permitiéndoles actuar de manera proactiva ante señales de insatisfacción o interés.

También se observa una tendencia hacia el uso de modelos híbridos, que combinan métodos estadísticos, reglas lingüísticas y redes neuronales. Estos modelos buscan equilibrar precisión y explicabilidad, lo cual es esencial en aplicaciones comerciales sensibles como las recomendaciones de productos. (Ma et al., 2025a) desarrollaron un modelo conjunto de representación de aspectos y calificaciones que mejora la calidad de las recomendaciones al integrar aprendizaje profundo con técnicas de correlación semántica. Esta tendencia responde a la necesidad de construir sistemas robustos que no solo sean precisos, sino también comprensibles para los usuarios y responsables de negocio.

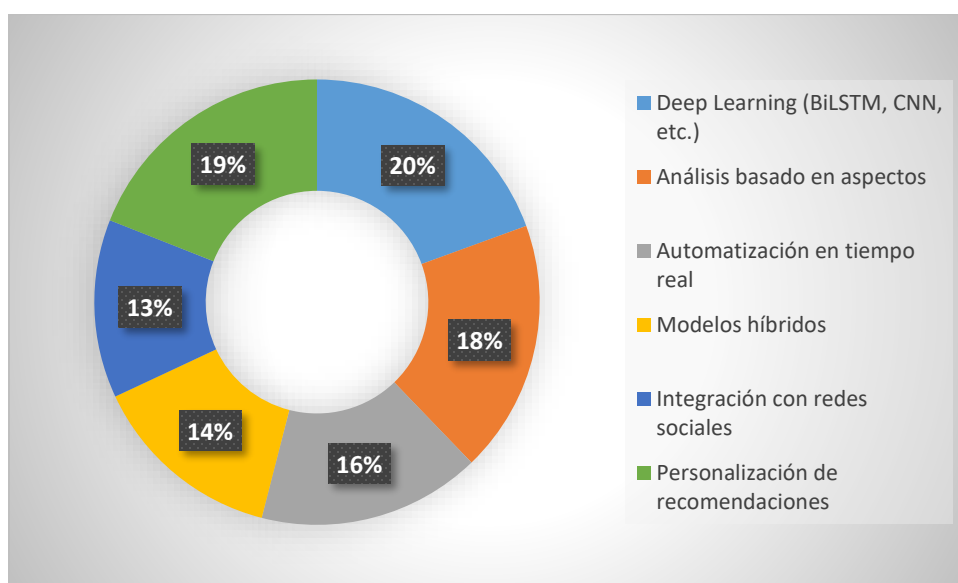
Otra tendencia emergente es la integración de redes sociales como fuente de datos. Los consumidores ya no se limitan a opinar en las plataformas de venta, sino que comparten sus experiencias en entornos como Twitter, Instagram o Facebook. Esta integración ha permitido enriquecer los análisis con datos emocionales más espontáneos y en tiempo real. (Huang & Ye, 2024) propusieron un método de optimización de marketing basado en redes sociales y análisis de sentimientos, mientras que (Kundra et al., 2024a) desarrollaron un sistema de análisis emocional digital que mejora la predicción de recomendaciones. Esta tendencia se explica por la necesidad de capturar la voz del cliente en múltiples canales y contextos.

Finalmente, se destaca la tendencia hacia la personalización de recomendaciones basada en los sentimientos detectados en las opiniones. Las empresas están utilizando estos análisis para adaptar sus catálogos y ofertas a los intereses emocionales de sus clientes, lo que ha demostrado mejorar la experiencia del usuario y aumentar las tasas de conversión. (Panwar et al., 2021) diseñaron un sistema de recomendación que integra análisis de sentimientos con preferencias emocionales del usuario, logrando una mejora significativa en la precisión de las sugerencias.

Estas tendencias reflejan un cambio de paradigma en el análisis de sentimientos aplicado al comercio electrónico, donde la inteligencia artificial no solo interpreta lo que los consumidores dicen, sino que también anticipa lo que podrían necesitar, permitiendo a las empresas ofrecer experiencias más inteligentes, contextuales y emocionalmente relevantes.

### Figura 7

*Distribución Porcentual de las Tendencias más Importantes de la Ciencia de Datos en el Comercio Electrónico Cuando se Hace uso de Análisis de Sentimientos.*



5. *¿Cuáles son los retos más importantes de la ciencia de datos con el análisis de sentimientos para el comercio electrónico?*

El análisis de sentimientos aplicado al comercio electrónico enfrenta múltiples retos que limitan su precisión, escalabilidad y aplicabilidad en contextos reales. Uno de los más señalados por los estudios es la ambigüedad del lenguaje natural utilizado por los consumidores, lo que dificulta la correcta identificación de emociones. Interpretar los matices del lenguaje en

opiniones sobre productos es una tarea compleja que exige modelos sofisticados de comprensión semántica (Rahin et al., 2022a).

Otro de los retos destacados es la variabilidad en los comportamientos de los consumidores, lo que obliga a que los modelos se adapten dinámicamente a cambios en tendencias, vocabulario y percepciones. (Huang & Ye, 2024), el cual enfatiza que los modelos deben actualizarse constantemente para seguir siendo relevantes ante cambios rápidos en la opinión pública.

Además, se identifican problemas de calidad y ruido en los datos, (Kundra et al., 2024a) donde se destaca que el exceso de información irrelevante o contradictoria en las opiniones puede afectar el rendimiento de los algoritmos de predicción.

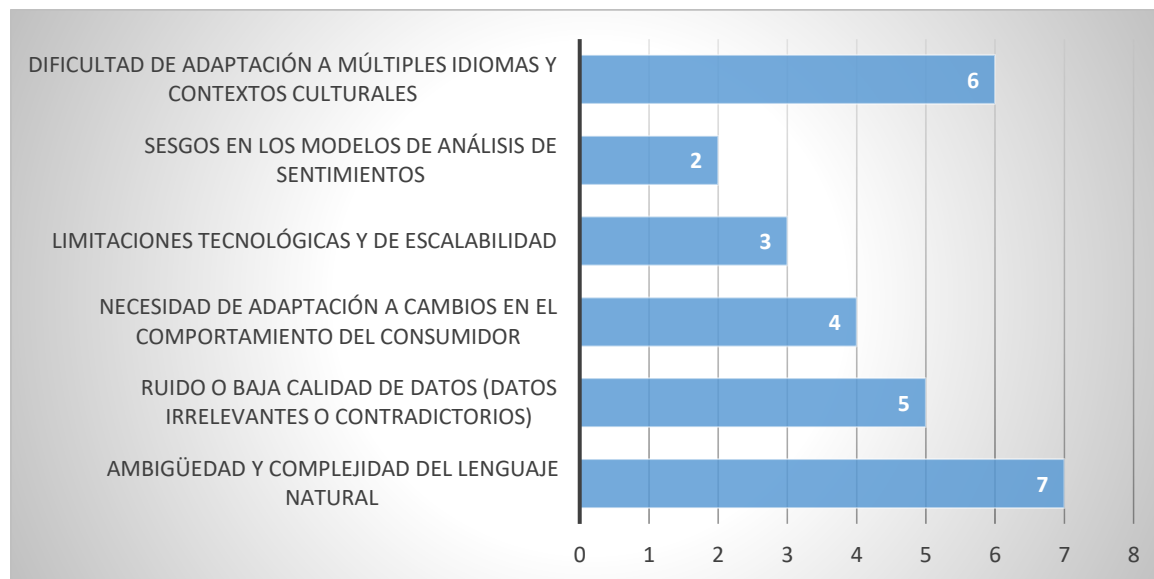
También se discuten las limitaciones técnicas en cuanto al procesamiento eficiente de grandes volúmenes de datos, especialmente en entornos de comercio electrónico a gran escala. Esto requiere infraestructuras avanzadas y modelos computacionalmente costosos, lo cual no siempre es viable para todas las organizaciones.

Finalmente, otro reto frecuente es la necesidad de reducir el sesgo algorítmico, sobre todo cuando los modelos han sido entrenados con datos desbalanceados. Si no se controla este factor, pueden generarse resultados poco representativos o incluso discriminatorios, lo que afecta la experiencia del usuario y la toma de decisiones empresariales.

En resumen, los retos más importantes identificados son: la ambigüedad del lenguaje, la calidad del dato, la necesidad de adaptación constante, las limitaciones técnicas y el control de sesgos. Estos factores deben abordarse cuidadosamente al implementar soluciones de análisis de sentimientos en comercio electrónico.

## Figura 8

*Frecuencia de Publicaciones Sobre los Retos de Ciencia de Datos Aplicadas en el Comercio Electrónico Durante los Últimos Cinco Años.*



6. *¿Cómo se está haciendo el monitoreo de redes sociales mediante la ciencia de datos aplicada al comercio electrónico?*

El monitoreo de redes sociales en el comercio electrónico se ha convertido en una práctica esencial para comprender las opiniones, emociones y comportamientos de los consumidores. Este proceso se realiza principalmente mediante el análisis de sentimientos aplicado a datos textuales generados en plataformas como Twitter, Facebook, Instagram y foros especializados. Según (Columbres, 2025), estas plataformas representan una fuente vital de contenido generado por los usuarios, el cual es procesado mediante técnicas de ciencia de datos para extraer información útil que oriente la toma de decisiones empresariales.

El análisis de sentimientos se apoya en técnicas de procesamiento del lenguaje natural (PLN), minería de texto y aprendizaje automático. (L. Yang et al., 2020) destacan que el proceso

comienza con la recolección de datos mediante APIs o técnicas de web scraping, seguido por el preprocesamiento del texto para eliminar ruido y normalizar el lenguaje. Posteriormente, se extraen características relevantes y se aplican modelos de clasificación como máquinas de soporte vectorial (SVM), redes neuronales recurrentes (LSTM, BiLSTM) y transformers, que permiten identificar la polaridad emocional de los comentarios. Una tendencia destacada es la implementación de sistemas de análisis en tiempo real, que permiten a las empresas reaccionar de forma inmediata ante cambios en la percepción del cliente. (R. K & S, 2021a) explican que estos sistemas se apoyan en arquitecturas de Big Data como Spark MLlib, que facilitan el procesamiento eficiente de grandes volúmenes de datos. Además, (Handayani et al., 2022) muestran cómo el análisis de redes sociales (SNA) permite mapear las interacciones entre usuarios, identificar influenciadores clave y comprender la dinámica de difusión de opiniones en campañas de marketing digital.

El monitoreo también se ha sofisticado con la incorporación de modelos híbridos que combinan redes neuronales con mecanismos de atención y análisis de grafos. (Huang & Ye, 2024) proponen un enfoque basado en redes neuronales gráficas para optimizar estrategias de marketing de precisión, integrando vectores de sentimiento derivados de redes sociales. Asimismo se destacan el uso de visualización de sentimientos geolocalizados, lo cual permite analizar la intensidad emocional por región y adaptar las estrategias comerciales a contextos específicos.

En cuanto a sus aplicaciones, el monitoreo de redes sociales se utiliza para mejorar la experiencia del cliente, personalizar ofertas, gestionar la reputación de marca, detectar crisis potenciales, optimizar campañas de marketing y detectar reseñas falsas o spam. (M et al., 2025a) abordan la detección de reseñas fraudulentas como una extensión del análisis de sentimientos,

mientras que (Mugil et al., 2023) destacan su utilidad para consolidar datos de clientes y monitorear la percepción de marca en tiempo real.

En síntesis, el monitoreo de redes sociales mediante ciencia de datos en el comercio electrónico constituye un proceso complejo y multidimensional que transforma grandes volúmenes de datos no estructurados en conocimiento estratégico. Este enfoque permite a las empresas comprender mejor a sus clientes, anticiparse a sus necesidades y fortalecer su posicionamiento en un entorno digital altamente competitivo.

#### 7. ¿Cómo mejora el análisis de sentimientos el comercio electrónico?

El análisis de sentimientos ha emergido como una herramienta estratégica fundamental en el comercio electrónico, al permitir transformar grandes volúmenes de opiniones de usuarios en información procesable. Esta técnica, basada en el procesamiento del lenguaje natural (PLN) y el aprendizaje automático, facilita tanto a consumidores como a empresas la toma de decisiones más informadas y efectivas.

Desde la perspectiva del consumidor, el análisis de sentimientos permite filtrar productos óptimos, identificar fortalezas y debilidades, y tomar decisiones de compra más informadas, lo que reduce las malas experiencias y mejora la satisfacción general (Lou, 2024). Además, al clasificar automáticamente las reseñas como positivas, negativas o neutrales, se facilita la comprensión rápida del consenso general sobre un producto (Graovac et al., 2020).

Para las empresas, esta técnica ofrece múltiples beneficios. En primer lugar, permite comprender con mayor profundidad las preferencias del cliente, lo que facilita el diseño de productos más alineados con sus expectativas (Bawack et al., 2022). Asimismo, el análisis de sentimientos contribuye a mejorar la calidad del servicio, optimizar estrategias de marketing y fortalecer la reputación digital de las marcas (Casas-Valadez et al., 2020).

Otro aspecto clave es la capacidad del análisis de sentimientos para detectar patrones y tendencias en tiempo real. Esto permite a las empresas anticiparse a cambios en la demanda, ajustar sus estrategias y responder de manera proactiva a las preocupaciones de los clientes (Columbres, 2025). Además, herramientas como Spark ML Pipeline han demostrado ser eficaces para manejar grandes volúmenes de datos y extraer valor estratégico de ellos (R. K & S, 2021a).

El análisis de sentimientos también se ha integrado en sistemas de recomendación, mejorando su precisión al incorporar la polaridad emocional de las reseñas. Por ejemplo, el análisis basado en aspectos (ABSA) puede descomponer las opiniones para identificar emociones asociadas a características específicas del producto. Estudios como el de (Rahin et al., 2022b) han aplicado este enfoque, logrando una precisión del 89% en la asignación de polaridad a los aspectos evaluados por los usuarios. Esta mejora en la comprensión granular del sentimiento, sumada a la información de las reseñas, contribuye a que los sitios de comercio electrónico que permiten valoraciones experimenten un aumento significativo en los ingresos por ventas, alcanzando hasta el 108.4%. Esto, a su vez, permite a las empresas ofrecer sugerencias más personalizadas y relevantes, lo que incrementa la satisfacción del cliente y las tasas de conversión (Ma et al., 2025b). Además, (Kundra et al., 2024b) han desarrollado un sistema de análisis emocional digital que mejora la predicción de recomendaciones, impulsando aún más estas capacidades.

En el ámbito de la gestión de la reputación y la confianza del consumidor, esta técnica permite detectar reseñas falsas o manipuladas, lo que protege tanto a los usuarios como a las plataformas de comercio electrónico (M et al., 2025a). Asimismo, facilita la identificación de emociones negativas recurrentes, lo que permite a las empresas priorizar la resolución de problemas críticos (Qu, 2022).

Finalmente, el análisis de sentimientos contribuye a la innovación y al desarrollo sostenible del comercio electrónico. Al proporcionar una comprensión más precisa de la voz del cliente, permite a las empresas adaptar sus productos, servicios y estrategias de manera continua, fortaleciendo su competitividad en un entorno digital dinámico (Kundra et al., 2024a).

8. ¿Cómo las empresas de comercio electrónico pueden aprovechar la minería de opiniones para mejorar su reputación digital?

Las empresas de comercio electrónico pueden aprovechar la minería de opiniones, también conocida como análisis de sentimientos, como una herramienta estratégica para mejorar su reputación digital. Esta técnica transforma grandes volúmenes de comentarios de clientes en información valiosa y accionable, permitiendo una gestión proactiva de la percepción pública y la imagen de marca.

Una de las formas fundamentales en que se mejora la reputación es a través del monitoreo y la comprensión continua de la percepción de marca. La minería de opiniones permite a las empresas monitorear constantemente las percepciones y sentimientos expresados por los clientes en diversas plataformas en línea. Esto facilita diagnosticar la posición actual en cuanto a la satisfacción del cliente y observar cómo cambian las percepciones a lo largo del tiempo, evaluando el éxito general de la marca al monitorear la retroalimentación positiva y negativa. Las empresas pueden comprender detalladamente la percepción pública y el sentimiento de los clientes hacia sus productos y servicios. Además, pueden observar los pensamientos, opiniones y reseñas de los clientes en las redes sociales para evaluar la imagen de la marca (Sultan et al., 2020), (Zuo, 2024b), (Ko et al., 2024). El monitoreo sistemático de la retroalimentación permite detectar tendencias y problemas emergentes que podrían afectar la reputación (Zhao et al., 2021a).

La minería de opiniones también es crucial para la identificación rápida y la resolución de problemas o crisis. Posibilita detectar rápidamente tendencias de sentimiento negativo o comentarios críticos, permitiendo una respuesta proactiva a posibles crisis de reputación antes de que escalen y causen un daño mayor. Las empresas pueden abordar rápidamente la retroalimentación negativa para mitigar daños y demostrar responsabilidad. Al identificar los puntos débiles o las sugerencias expresadas en las opiniones, las empresas pueden mejorar sus productos, lo que conduce a una mayor satisfacción del cliente y, por lo tanto, a una mejora de la reputación. Identificar y resolver problemas directamente demuestra una gestión activa que afecta la percepción del usuario (Ahmad et al., 2025b). Esta capacidad de manejar eficazmente grandes conjuntos de datos de sentimientos significa que las empresas pueden reaccionar rápidamente a comentarios negativos, gestionar crisis de reputación o amplificar el boca a boca positivo (R. K & S, 2021b).

Un pilar fundamental para la reputación es la mejora dirigida de productos y servicios. Al analizar las reseñas, las empresas pueden identificar fortalezas y debilidades específicas de sus ofertas, lo que les permite realizar ajustes que se alineen con las expectativas del cliente. Esta capacidad de respuesta directa a la retroalimentación contribuye a una percepción más positiva de la marca. Priorizar el control de calidad es fundamental para evitar la proliferación de emociones negativas y proteger la reputación. La minería de opiniones permite percibir la experiencia real de los compradores con sus productos y servicios, captando sus ventajas y desventajas para realizar mejoras específicas en la calidad (Catalas et al., 2020). Saber exactamente qué le gusta o no le gusta al cliente de sus productos es crucial para identificar los puntos fuertes que se pueden promocionar y las áreas que necesitan mejora (Arra et al., 2023b).

La construcción de confianza y el fomento de la lealtad del cliente son resultados directos del aprovechamiento de la minería de opiniones. Demostrar que las opiniones de los clientes son valiosas y que se actúa en consecuencia ayuda a generar confianza (Zhao et al., 2021a). Esto es crucial, ya que los consumidores confían en las reseñas en línea para tomar decisiones de compra. Al asegurar que la información en sus plataformas sea genuina y al protegerse de reseñas maliciosas (detectando contenido engañoso), las empresas mantienen la integridad de su reputación y fomentan la confianza del consumidor (M et al., 2025b), (Kumar et al., 2024). La detección y eliminación de reseñas falsas es crucial para combatir opiniones manipuladas que pueden dañar la reputación (Chinnasamy et al., 2024). Una alta proporción de recomendaciones positivas indica una reputación favorable (Rahin et al., 2022b).

La minería de opiniones también proporciona información para la toma de decisiones estratégicas y de marketing. Ofrece información valiosa para decisiones estratégicas, desde el desarrollo de productos hasta la formulación de campañas de marketing más alineadas con las emociones del consumidor. Permite adaptar ofertas según los intereses emocionales de los clientes y evaluar las estrategias de campaña. Asimismo, contribuye a formular activamente estrategias de negocio para mantener el comercio y desempeña un papel importante en la evaluación de ventas y la optimización de los enfoques de marketing (R. K & S, 2021b). La capacidad de analizar grandes volúmenes de datos permite a las empresas obtener una visión completa del sentimiento del cliente a gran escala (Wang et al., 2023).

Finalmente, la mejora de los sistemas de recomendación es otro beneficio directo. Las valoraciones emocionales influyen directamente en los sistemas de recomendación, mejorando la experiencia del usuario al sugerir productos con base en opiniones similares o aspectos valorados positivamente. Al basarse en un análisis exhaustivo de las opiniones, las recomendaciones se

adaptan mejor a los intereses del cliente, mejorando la experiencia del usuario y generando mayor confianza.

*9. ¿Qué impacto tiene el análisis de Big Data en la identificación de tendencias y patrones en la compra en el comercio electrónico?*

La minería de opiniones se ha consolidado como una herramienta clave para las empresas de comercio electrónico que buscan fortalecer su reputación digital mediante el análisis sistemático de las percepciones de los consumidores. A través del procesamiento de grandes volúmenes de datos generados por los usuarios, esta técnica permite identificar patrones de comportamiento, emociones predominantes y aspectos críticos que influyen en la experiencia de compra.

Estudios recientes han demostrado que el análisis de sentimientos aplicado a reseñas de productos no solo permite evaluar la satisfacción del cliente, sino también anticipar tendencias emergentes en el mercado. Por ejemplo, el trabajo (Zhao et al., 2021b) propone un enfoque novedoso de ponderación de términos y selección de características que mejora la precisión en la clasificación de sentimientos, lo cual es esencial para detectar cambios sutiles en la percepción del consumidor. Esta capacidad predictiva permite a las empresas ajustar sus estrategias de marketing y desarrollo de productos de manera proactiva.

Asimismo, el uso de modelos híbridos como los propuestos por (Ajmeera & Kamakshi, 2024), basados en redes neuronales profundas, ha demostrado ser eficaz para extraer patrones complejos de grandes conjuntos de datos de reseñas. Estos modelos permiten identificar no solo la polaridad del sentimiento, sino también las características específicas del producto que generan mayor impacto emocional, lo que resulta fundamental para la mejora continua de la oferta comercial.

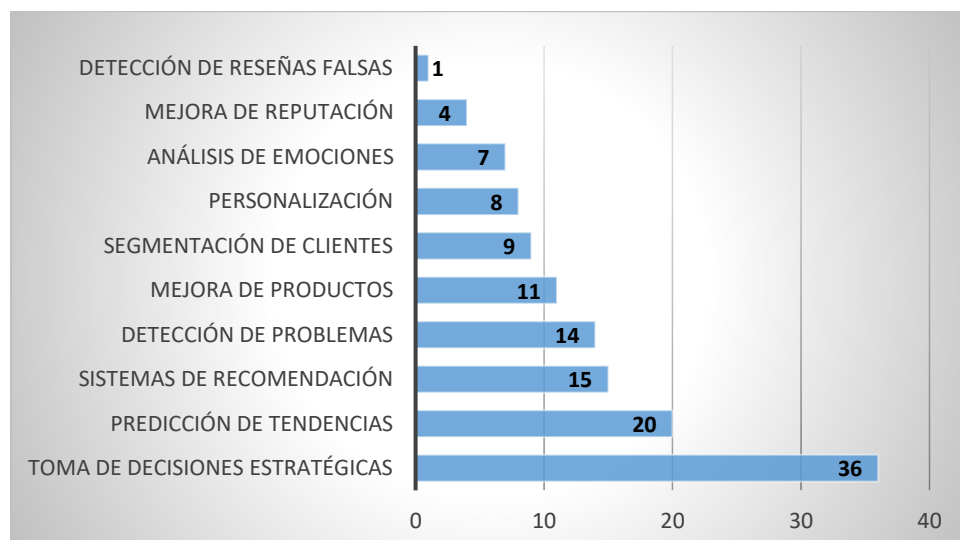
Otro aspecto relevante es la integración de la minería de opiniones con sistemas de recomendación personalizados. Investigaciones como la de (Zhou & Fang, 2021) destacan cómo el análisis de sentimientos puede alimentar motores de recomendación más precisos, al considerar no solo el historial de compras, sino también las emociones expresadas en las reseñas. Esta personalización contribuye significativamente a la fidelización del cliente y a la construcción de una reputación digital positiva.

Además, la minería de opiniones se ha utilizado para detectar patrones de desconfianza o insatisfacción que podrían afectar la credibilidad de la plataforma. El estudio de (Indrawati & Putri, 2021), centrado en el análisis de confianza en marketplaces, demuestra cómo el procesamiento de grandes volúmenes de texto permite identificar preocupaciones recurrentes de los usuarios, como problemas de entrega o falta de transparencia, que deben ser abordados para preservar la reputación de la marca.

Estos hallazgos evidencian que la minería de opiniones no solo es una herramienta de monitoreo pasivo, sino una fuente activa de inteligencia estratégica. Su implementación permite a las empresas de comercio electrónico comprender mejor a sus clientes, anticiparse a sus necesidades y responder de manera ágil a los desafíos reputacionales, consolidando así su posicionamiento en un entorno digital altamente competitivo.

## Figura 9

### *Impactos más Frecuentes del análisis de Big Data*



*Nota.* Identificación de tendencias y patrones en la compra en el comercio electrónico.

## Otros Datos Extraídos del Corpus

Además de las respuestas a las preguntas de investigación, se identificaron datos complementarios en los artículos revisados que enriquecen el mapeo sistemático. Estos datos no responden directamente a las preguntas planteadas, pero ofrecen una visión más amplia sobre el estado actual del análisis de sentimientos en el comercio electrónico.

### Lenguajes de programación y herramientas utilizadas

Aunque no todos los artículos especifican los lenguajes de programación empleados, se identificó el uso de herramientas y entornos como Spark ML Pipeline (R. K & S, 2021a), lo que sugiere el uso de lenguajes como Scala o Python. También se hace referencia al uso de TANAGRA para visualización y análisis de datos (Khan & Karthikeyan, 2023), una herramienta que opera sobre entornos Windows y se basa en Java. En estudios que aplican redes

neuronales profundas, es razonable inferir el uso de frameworks como TensorFlow o PyTorch, aunque no siempre se menciona explícitamente.

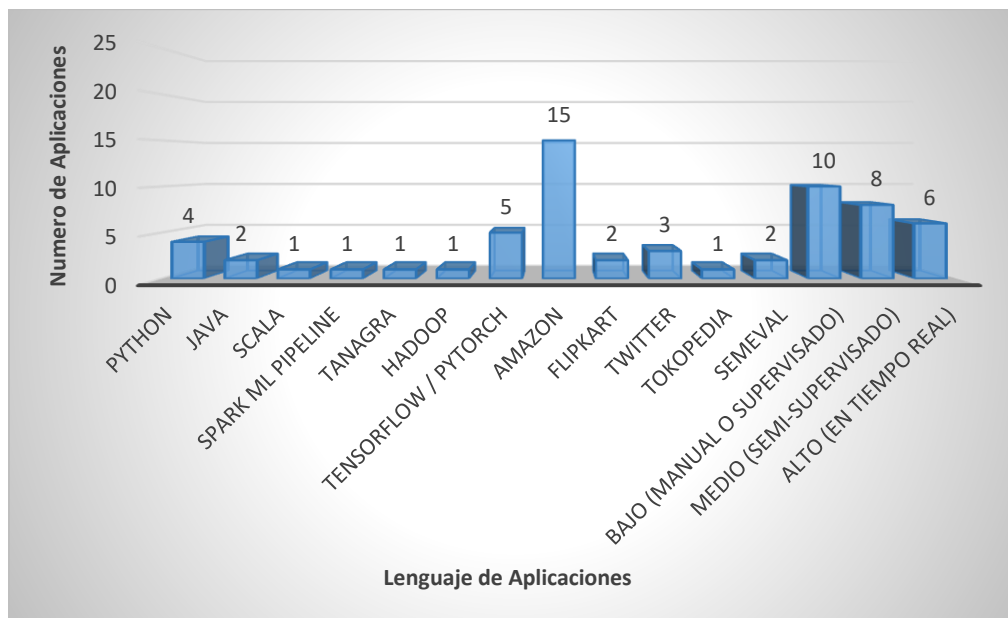
Fuentes de datos: Las fuentes de datos más comunes fueron las reseñas de productos en plataformas como Amazon y Flipkart, utilizadas en múltiples estudios (Nancy Deborah et al., 2023; Sharm et al., 2022). También se identificó el uso de redes sociales como Twitter para el análisis de sentimientos, especialmente en estudios que combinan análisis de comportamiento con estrategias de marketing digital (Handayani et al., 2022; Huang & Ye, 2024). Esta diversidad de fuentes refleja una tendencia hacia la integración de datos estructurados y no estructurados para obtener una visión más completa del consumidor.

Nivel de automatización: Algunos artículos proponen sistemas con un alto grado de automatización. Por ejemplo, el estudio de (Gao, 2024) presenta un algoritmo de predicción de comportamiento del usuario basado en big data, diseñado para operar en tiempo real. Otros trabajos, como el de (Kundra et al., 2024a), desarrollan sistemas de análisis emocional digital que se integran directamente en plataformas de recomendación, lo que sugiere un nivel medio a alto de automatización. Sin embargo, también se identificaron enfoques más manuales o semi-supervisados, como los utilizados en la detección de reseñas falsas (M et al., 2025a), lo que indica que el campo aún presenta una diversidad de niveles de madurez tecnológica.

Plataformas de comercio electrónico analizadas: La mayoría de los estudios se centraron en plataformas ampliamente utilizadas como Amazon, Flipkart y Tokopedia. Estas plataformas ofrecen grandes volúmenes de datos públicos, lo que facilita su uso en investigaciones académicas. En algunos casos, se utilizaron datasets específicos como SemEval, especialmente en estudios de análisis basado en aspectos (Rahin et al., 2022a).

**Figura 10**

*Distribución de Otros Lenguajes de Aplicación del Big Data en el Análisis de Sentimientos.*



## Conclusiones

Esta revisión sistemática permitió identificar, clasificar y analizar 64 estudios relevantes sobre el uso de técnicas de Big Data y monitoreo de redes sociales aplicadas al marketing digital entre 2020 y 2024. Se evidenció un crecimiento sostenido en la producción científica, con un enfoque predominante en el análisis de sentimientos, machine learning y visualización de datos. Esta investigación contribuye al campo al mapear de manera estructurada las técnicas, herramientas y enfoques más empleados en este período, revelando tendencias emergentes, vacíos temáticos y desafíos metodológicos. En particular, ofrece una hoja de ruta clara para investigadores que buscan explorar la intersección entre la analítica de datos y las estrategias de mercadeo digital, así como para profesionales interesados en implementar soluciones basadas en inteligencia de datos.

La revisión sistemática realizada permitió identificar que el análisis de sentimientos y el monitoreo de redes sociales han adquirido un rol estratégico en el marketing digital, especialmente en el contexto del comercio electrónico. Se evidenció que técnicas como *Machine Learning*, *Deep Learning*, el procesamiento del lenguaje natural (PLN) y los modelos híbridos son ampliamente utilizados para interpretar las emociones y opiniones de los consumidores, facilitando la personalización de contenidos y la toma de decisiones basadas en datos (Ajmeera & Kamakshi, 2024; Arra et al., 2023).

Asimismo, se constató una tendencia creciente hacia la automatización de procesos mediante sistemas inteligentes capaces de integrar múltiples fuentes de información, como redes sociales, reseñas de productos y datos de CRM, lo que permite una visión más integral del comportamiento del cliente (Handayani et al., 2022). No obstante, persisten retos técnicos

relevantes, como la ambigüedad del lenguaje, la detección de ironía y sarcasmo, la escasez de datos etiquetados y la necesidad de modelos multilingües y explicables (Rahin et al., 2022).

Desde una perspectiva teórico-práctica, este trabajo contribuye a consolidar el conocimiento existente sobre el uso de Big Data en el marketing digital, proporcionando un panorama actualizado de las técnicas más efectivas y los desafíos emergentes. Además, se identificaron vacíos en la literatura, particularmente en lo relacionado con la evaluación empírica de herramientas automatizadas y la medición del impacto en indicadores reales de negocio.

## Recomendaciones

En función de los hallazgos obtenidos, se recomienda impulsar el desarrollo de modelos explicables que permitan interpretar las decisiones tomadas por los algoritmos, especialmente en contextos donde la transparencia es crítica. La explicabilidad debe convertirse en un criterio fundamental para la adopción de soluciones de inteligencia artificial en marketing digital, tal como lo sugieren Arora et al. (2022) en su propuesta sobre inteligencia artificial confiable. También es esencial fortalecer las capacidades técnicas del talento humano, mediante programas de formación especializados en analítica de datos, procesamiento de lenguaje natural y uso ético de tecnologías basadas en datos. La inversión en capacitación contribuirá a reducir la dependencia tecnológica y mejorar la apropiación de estas herramientas en las organizaciones.

Asimismo, se recomienda avanzar en la construcción de modelos multilingües y culturalmente adaptados que permitan ampliar el alcance del análisis de sentimientos a mercados más diversos. El enfoque actual, mayoritariamente anglocéntrico, limita la generalización de los resultados y excluye importantes sectores de consumidores, como los de América Latina (Bermúdez-Quimis, 2024). De igual manera, se sugiere priorizar investigaciones orientadas a la detección de opiniones falsas o manipuladas, dado que este tipo de contenidos impacta directamente la reputación digital de las marcas (CRITIQUESHIELD, 2022). Otra recomendación clave consiste en fomentar la integración de estas tecnologías con sistemas corporativos existentes como CRM, ERP o plataformas de atención al cliente, para aprovechar al máximo la información generada por los consumidores en tiempo real.

Finalmente, se insta a realizar pruebas piloto en contextos empresariales reales, con el propósito de evaluar de manera más precisa la utilidad, aplicabilidad y retorno de inversión de las técnicas identificadas en esta revisión. Esta validación práctica permitirá cerrar la brecha

entre el conocimiento académico y las necesidades del sector productivo, impulsando una transformación digital más efectiva, inclusiva y basada en evidencia.

### Referencias Bibliográficas

- Ahmad, W., Khan, H. U., Alarfaj, F. K., & Alreshoodi, M. (2025a). Aspect-Based Sentiment Analysis: A Comprehensive Review and Open Research Challenges. *IEEE Access*, *13*, 65138-65182. <https://doi.org/10.1109/access.2025.3555744>
- Ahmad, W., Khan, H. U., Alarfaj, F. K., & Alreshoodi, M. (2025b). Aspect-Based Sentiment Analysis: A Comprehensive Review and Open Research Challenges. *IEEE Access*, *13*, 65138-65182. <https://doi.org/10.1109/access.2025.3555744>
- Ajmeera, N., & Kamakshi, P. (2024). Hamiltonian deep neural network fostered sentiment analysis approach on product reviews. *Signal, Image and Video Processing*, *18*(4), 3483-3494. <https://doi.org/10.1007/s11760-024-03014-6>
- Andrade, G. J. S., León, S. D. C., & Gusqui, E. D. M. (2024). La preservación de la privacidad y la salvaguardia de datos personales en el contexto de las redes sociales: Un análisis de su impacto en el derecho a la intimidad. *Dilemas contemporáneos: Educación, Política y Valores*.
- Arora, K., Faisal, M., & others. (2022). The Use of Data Science in Digital Marketing Techniques: Work Programs, Performance Sequences and Methods. *Startupreneur Business Digital (SABDA Journal)*, *1*(2), 143-155. <https://doi.org/10.34306/sabda.v1i2.110>
- Arra, A., Yeboah, J., & Nti, I. K. (2023a). Opinion Mining of Online Shopping Products Reviews Using Machine Learning. *2023 International Conference on Computational Science and Computational Intelligence (CSCI)*, 270-276. <https://doi.org/10.1109/csci62032.2023.00048>

- Arra, A., Yeboah, J., & Nti, I. K. (2023b). Opinion Mining of Online Shopping Products Reviews Using Machine Learning. *2023 International Conference on Computational Science and Computational Intelligence (CSCI)*, 270-276.  
<https://doi.org/10.1109/csci62032.2023.00048>
- Avella Chaparro, J., & Rodríguez, L. (2024). Big Data y marketing digital: Brechas y estrategias en el contexto colombiano. *Colombian Journal of Data Analytics*, 9(2), 45-62.
- Bawack, R. E., Wamba, S. F., Carillo, K. D. A., & Akter, S. (2022). Artificial intelligence in E-Commerce: A bibliometric study and literature review. *Electronic Markets*, 32(1), 297-338. <https://doi.org/10.1007/s12525-022-00537-z>
- Bermúdez-Quimis, N. C. (2024). Análisis de Big Data en las tecnologías de la información. *Revista Científica Arbitrada de Investigación en Comunicación, Marketing y Empresa REICOMUNICAR. ISSN 2737-6354.*, 7(14), 676-682.
- Cardona, G. H., & Salinas, J. G. (2024). Análisis de Big Data como estrategia de marketing personalizado en LH Computer. *593 Digital Publisher CEIT*, 9(6), 1261-1276.
- Casas-Valadez, M. A., Faz-Mendoza, A., Medina-Rodríguez, C. E., Castaneda-Miranda, R., Gamboa-Rosales, N. K., & Lopez-Robles, J. R. (2020). The synergic relationship between e-commerce and Sentiment Analysis: A content analysis of published articles between 2007 and 2020. *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, 1-5.  
<https://doi.org/10.1109/icdabi51230.2020.9325689>
- Cataltas, M., Dogramaci, S., Yumusak, S., & Oztoprak, K. (2020). Extraction of Product Defects and Opinions from Customer Reviews by Using Text Clustering and Sentiment Analysis.

2020 *IEEE International Conference on Big Data (Big Data)*, 4529-4534.

<https://doi.org/10.1109/bigdata50022.2020.9377851>

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, 19(2), 171-209.

Chinnasamy, P., Ayyasamy, R. K., Madhukar, G., Karthik, G., Bhargav, T., & Nayak, D. Y. K. (2024). Comment Analyzer by Sentimental Analysis through Natural Language Processing. *2024 10th International Conference on Communication and Signal Processing (ICCSP)*, 1123-1128. <https://doi.org/10.1109/iccsp60870.2024.10544106>

Columbres, M. R. (2025). Exploring Trends and Techniques in Sentiment Analysis for Online Product Ratings: A Comprehensive Review. *International Journal of Research and Scientific Innovation*, XII(II), 312-324. <https://doi.org/10.51244/ijrsi.2025.12020027>

Davenport, T. H. (2014). *Big data at work: Dispelling the myths, uncovering the opportunities*. Harvard Business Review Press.

Davis, F. D. & others. (1989). Technology acceptance model: TAM. *Al-Suqri, MN, Al-Aufi, AS: Information Seeking Behavior and Technology Adoption*, 205(219), 5.

Fan, S., Lau, R. Y. K., & Zhao, J. L. (2015). Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix. *Big Data Research*, 2(1), 28-32. <https://doi.org/10.1016/j.bdr.2015.02.006>

Gao, X. (2024). User Behavior Analysis and Prediction Algorithm Design of E-commerce Platform Based on Big Data Analysis. *2024 Asia-Pacific Conference on Software Engineering, Social Network Analysis and Intelligent Computing (SSAIC)*, 141-146. <https://doi.org/10.1109/ssaic61213.2024.00033>

- Garza Garza, R. (2010). Método sistemático para la revisión de la literatura. *Revista Internacional de Investigación e Innovación Educativa*, 7, 33-45.
- González, F. (2019). Big data, algoritmos y política: Las ciencias sociales en la era de las redes digitales. *Cinta de moebio*, 65, 267-280.
- Graovac, J., Radovic, M., & Girgin, B. A. (2020). ML-SPD: Machine Learning based Sentiment Polarity Detection. *2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, 1-7. <https://doi.org/10.1109/inista49547.2020.9194633>
- Gubbala, S., & Kaur, M. (2025). A Comprehensive Review: Enhanced Customer Centric Predictive Analysis in E-Commerce. *2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, 95-101. <https://doi.org/10.1109/idciot64235.2025.10915119>
- Handayani, F., Evelina, L. W., & Tinihada, A. F. (2022). Social Network Analysis of Campaign Mobilization of Marketplace on Twitter. *2022 International Conference on Information Management and Technology (ICIMTech)*, 241-245. <https://doi.org/10.1109/icimtech55957.2022.9915260>
- Hassan, R., & Islam, R. (s. f.). *Detection of fake online reviews using semi-supervised and supervised learning*.
- Huang, Y., & Ye, S. (2024). An E-Commerce Precision Marketing Optimization Method Based on Graph Neural Network and Social Media Sentiment Analysis. *2024 4th International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI)*, 192-197. <https://doi.org/10.1109/cei63587.2024.10871730>
- Indrawati, & Putri, N. P. O. I. Y. (2021). Indonesian Marketplace Trust Analysis Using Text Mining: A Case of Tokopedia. *2021 International Conference Advancement in Data*

*Science, E-Learning and Information Systems (ICADEIS)*, 1-6.

<https://doi.org/10.1109/icadeis52521.2021.9702080>

K, H., A, A. A., L, B., P, H., & A, A. (2023). Analyzation of Sentiments of Product Reviews using Natural Language Processing. *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 1851-1854.

<https://doi.org/10.1109/icaccs57279.2023.10112777>

K, R., & S, M. V. (2021a). An Implementation of Hybrid Enhanced Sentiment Analysis System using Spark ML Pipeline: A Big Data Analytics Framework. *International Journal of Advanced Computer Science and Applications*, 12(5).

<https://doi.org/10.14569/ijacsa.2021.0120540>

K, R., & S, M. V. (2021b). An Implementation of Hybrid Enhanced Sentiment Analysis System using Spark ML Pipeline: A Big Data Analytics Framework. *International Journal of Advanced Computer Science and Applications*, 12(5).

<https://doi.org/10.14569/ijacsa.2021.0120540>

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), 59-68.

Khan, G., & Karthikeyan, M. P. (2023). Implementing TANAGRA for the Visualization and Analytical Interpretation of E-commerce Datasets. *2023 International Conference on Communication, Security and Artificial Intelligence (ICCSAI)*, 187-191.

<https://doi.org/10.1109/iccsai59793.2023.10421548>

Kitchenham, B. (2007). *Guidelines for performing systematic literature reviews in software engineering*.

- Kitchin, R. (2014a). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage Publications.
- Kitchin, R. (2014b). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage Publications.
- Ko, C. S., Lee, H., & Kim, T. (2024). *VoC Analytics Based on Web Crawling and Robotic Process Automation* (No. 03). ICIC International 学会.  
<https://doi.org/10.24507/icicelb.15.03.287>
- Kumar, L. K., Thatha, V. N., Udayaraju, P., Siri, D., Kiran, G. U., Jagadesh, B. N., & Vatambeti, R. (2024). Analyzing Public Sentiment on the Amazon Website: A GSK-Based Double Path Transformer Network Approach for Sentiment Analysis. *IEEE Access*, 12, 28972-28987. <https://doi.org/10.1109/access.2024.3368441>
- Kundra, H., Vishnu, C., Vaishnavi, B. V., Kasireddy, C. R., Akash, G., & Jindam, S. (2024a). Digital Emotions using Sentiment Analysis for Predictive Insights on Customer Recommendations. *2024 IEEE 5th India Council International Subsections Conference (INDISCON)*, 1-6. <https://doi.org/10.1109/indiscon62179.2024.10744376>
- Kundra, H., Vishnu, C., Vaishnavi, B. V., Kasireddy, C. R., Akash, G., & Jindam, S. (2024b). Digital Emotions using Sentiment Analysis for Predictive Insights on Customer Recommendations. *2024 IEEE 5th India Council International Subsections Conference (INDISCON)*, 1-6. <https://doi.org/10.1109/indiscon62179.2024.10744376>
- Li, C., & Bernoff, J. (2008). *Groundswell: Winning in a world transformed by social technologies*. Harvard Business Review Press.
- Liu, P., Zhang, K., Dong, X., & Wang, P. (2024). A Big Data-Kano and SNA-CRP Based QFD Model: Application to Product Design Under Chinese New E-commerce Model. *IEEE*

*Transactions on Engineering Management*, 71, 4246-4260.

<https://doi.org/10.1109/tem.2022.3227094>

Lou, P. (2024). Assessing and Recommending Brand Features of Electronic Products through Online Reviews and BiLSTM Models. *2024 IEEE 2nd International Conference on Sensors, Electronics and Computer Engineering (ICSECE)*, 1484-1488.

<https://doi.org/10.1109/icsece61636.2024.10729421>

M, A., M.E., S. S., P, C. S., M, K., & D, N. (2025a). CRITIQUESHIELD: Fake Review Detection. *2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI)*, 919-924. <https://doi.org/10.1109/icmsci62561.2025.10894384>

M, A., M.E., S. S., P, C. S., M, K., & D, N. (2025b). CRITIQUESHIELD: Fake Review Detection. *2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI)*, 919-924. <https://doi.org/10.1109/icmsci62561.2025.10894384>

Ma, J., Wang, R., Wei, T., Liu, K., Zhang, H., & Lu, X. (2025a). Correlation-Driven Explainable Recommendation With Aspect and Rating Boosted Representation Learning: A Unified Joint-Ranking Framework. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 55(4), 2489-2502. <https://doi.org/10.1109/tsmc.2024.3522980>

Ma, J., Wang, R., Wei, T., Liu, K., Zhang, H., & Lu, X. (2025b). Correlation-Driven Explainable Recommendation With Aspect and Rating Boosted Representation Learning: A Unified Joint-Ranking Framework. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 55(4), 2489-2502. <https://doi.org/10.1109/tsmc.2024.3522980>

Manyika, A. H., J. ., Chui, M. ., Brown, B. ., Bughin, J. ., Dobbs, R. ., Roxburgh, C. ., & Byers. (2011). *Big data: The next frontier for innovation, competition, and productivity*.

MinTIC. (2020). *Informe de transformación digital en Colombia*.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *Open Medicine*.

Mugil, P., Jeevanantham, C., Yadav, Y. N. N., Harikaran., M., & Dhanapal, R. (2023). A Secure and Efficient Product Analysis for Trust based E-Commerce Application. *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, 1370-1376. <https://doi.org/10.1109/icscds56580.2023.10104925>

Nancy Deborah, R., Alwyn Rajiv, S., Vinora, A., Gobinath, A., Vigneswaran, S., & Shiivadhanush, K. V. (2023). Sentiment Analysis and Machine Learning Algorithm Implementation on Flipkart Product Customer Reviews. *2023 International Conference on Advanced Computing Technologies and Applications (ICACTA)*, 1-6. <https://doi.org/10.1109/icacta58201.2023.10393161>

Ogbeide, S. O., Ogbeifun, R. O., & others. (2024). BIG DATA AND FIRMS' FINANCIAL PERFORMANCE IN A DEVELOPING COUNTRY: A LITERATURE REVIEW APPROACH. *SAPIENTIA FOUNDATION JOURNAL OF EDUCATION, SCIENCES AND GENDER STUDIES*, 6(3).

Panwar, M., Wadhwa, A., & Pippal, S. (2021). An Overview: Exploratory Data Analytics on E-Commerce Dataset. *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, 91-93. <https://doi.org/10.1109/icac3n53548.2021.9725765>

Patidar, R., & Patel, S. (2022). Design & Implementation of Product Recommendation Solution using Sentiment Analysis. *2022 International Conference on Edge Computing and Applications (ICECAA)*, 233-239. <https://doi.org/10.1109/icecaa55415.2022.9936480>

- Pine, B. J., Gilmore, J. H., & others. (1999). *The experience economy*. Boston: Harvard Business School Press.
- Qu, L. (2022). Influence of Emotional Factors (Positive, Negative) on the Usefulness of Product Reviews Based on Big Data. *Wireless Communications and Mobile Computing*, 2022, 1-13. <https://doi.org/10.1155/2022/2958718>
- Rahin, S. A., Hasib, T., & Hassan, M. (2022a). Aspect-Based Sentiment Analysis Using SemEval and Amazon Datasets. *2022 Fifth International Conference of Women in Data Science at Prince Sultan University (WiDS PSU)*, 85-90. <https://doi.org/10.1109/wids-psu54548.2022.00029>
- Rahin, S. A., Hasib, T., & Hassan, M. (2022b). Aspect-Based Sentiment Analysis Using SemEval and Amazon Datasets. *2022 Fifth International Conference of Women in Data Science at Prince Sultan University (WiDS PSU)*, 85-90. <https://doi.org/10.1109/wids-psu54548.2022.00029>
- Reyes Hidalgo, C. (s. f.). *REPUTACIÓN ORGANIZACIONAL Y REDES SOCIALES: UN ANÁLISIS A LOS DESAFÍOS Y ESTRATEGIAS*.
- Sharm, N., Jain, T., Narayan, S. S., & Kandakar, A. C. (2022, julio 29). Sentiment Analysis of Amazon Smartphone Reviews Using Machine Learning & Deep Learning. *2022 IEEE International Conference on Data Science and Information System (ICDSIS)*. 2022 IEEE International Conference on Data Science and Information System (ICDSIS), Hassan, India. <https://doi.org/10.1109/icdsis55133.2022.9915917>
- Sultan, M. F., Jabeen, M., & Mannan, M. A. (2020). Sentiment Analysis through Big Data in online Retail Industry: A Conceptual Quantitative Study on linkage of Big-Data and

- Assortment Proactive of Online Retailers. *KIET Journal of Computing and Information Sciences*, 3(2), 16. <https://doi.org/10.51153/kjcis.v3i2.47>
- Tsai, C.-W., Lai, C.-F., & Vasilakos, A. V. (2014). Big data analytics: A survey. *Journal of Big Data*, 2(1), 1-32.
- Wang, Z., Ji, Y., Zhang, T., Li, Y., Wang, L., & Qu, S. (2023). Product competitiveness analysis from the perspective of customer perceived helpfulness: A novel method of information fusion research. *Data Technologies and Applications*, 57(4), 437-464. <https://doi.org/10.1108/dta-03-2022-0124>
- Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning. *IEEE Access*, 8, 23522-23530. <https://doi.org/10.1109/access.2020.2969854>
- Yang, Y., Chen, M., Blankenship, Y., Lee, J., Ghassemlooy, Z., Cheng, J., & Mao, S. (2024). Positioning using wireless networks: Applications, recent progress and future challenges. *arXiv preprint arXiv:2403.11417*.
- Zeng, et al., Daniel. (2010). Social media analytics and intelligence. *IEEE Intelligent Systems*, 25(6), 13-16.
- Zhao, H., Liu, Z., Yao, X., & Yang, Q. (2021a). A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach. *Information Processing & Management*, 58(5), 102656. <https://doi.org/10.1016/j.ipm.2021.102656>
- Zhao, H., Liu, Z., Yao, X., & Yang, Q. (2021b). A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach.

*Information Processing & Management*, 58(5), 102656.

<https://doi.org/10.1016/j.ipm.2021.102656>

Zhou, Y., & Fang, F. (2021). Analysis of User Behavior through Network Big Data and Intelligent Information System. *2021 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, 807-810.

<https://doi.org/10.1109/aeeca52519.2021.9574265>

Zuo, S. (2024a). E-Commerce Platform User Evaluation Sentiment Analysis and Emotion-Oriented Marketing Strategy Research Based on Big Data Technology. *International Journal of High Speed Electronics and Systems*.

<https://doi.org/10.1142/s0129156425400907>

Zuo, S. (2024b). E-Commerce Platform User Evaluation Sentiment Analysis and Emotion-Oriented Marketing Strategy Research Based on Big Data Technology. *International Journal of High Speed Electronics and Systems*.

<https://doi.org/10.1142/s0129156425400907>