



---

## Análisis teórico de la minería de texto como herramienta de inteligencia en marketing

### Theoretical analysis of text mining as an intelligence tool in marketing

Irene Maricruz Vega-Urquizo

Universidad Técnica de Ambato, Ambato, Ecuador

[ivega2073@uta.edu.ec](mailto:ivega2073@uta.edu.ec)

<https://orcid.org/0009-0001-9180-1481>

Shirley Katerine Sanabria-Naranjo

Universidad Técnica de Ambato, Ambato, Ecuador

[ssanabria8843@uta.edu.ec](mailto:ssanabria8843@uta.edu.ec)

<https://orcid.org/0009-0005-2593-420X>

Patricio Medina-Chicaiza

Universidad Técnica de Ambato, Ambato, Ecuador

Pontificia Universidad Católica del Ecuador Sede Ambato, Ecuador

[ricardopmedina@uta.edu.ec](mailto:ricardopmedina@uta.edu.ec); [pmedina@pucesa.edu.ec](mailto:pmedina@pucesa.edu.ec)

<https://orcid.org/0000-0002-2736-8214>

Recepción: 15/10/2025 | Aceptación: 29/12/2025 | Publicación: 28/02/2026

---

#### Cómo citar (APA, séptima edición):

Vega-Urquizo, I., Sanabria-Naranjo, S., & Medina-Chicaiza, P. (2026). Análisis teórico de la minería de texto como herramienta de inteligencia en marketing. *INNOVA Research Journal*, 11(1.especial), 86-111. <https://doi.org/10.33890/innova.v11.n1.especial.2026.2868>

---

#### Resumen

La inteligencia de marketing se ha consolidado como un eje estratégico en la gestión organizacional apoyado en un entorno digital caracterizado por abundancia de datos que busca transformar información en conocimiento útil. En este contexto, la minería de texto y el análisis

de sentimientos han cobrado relevancia para interpretar datos en decisiones estratégicas. El objetivo de este estudio fue realizar una revisión de literatura basada en evidencia que identificó los tipos de inteligencia de marketing, las fuentes de datos asociadas y el rol del análisis de sentimientos como técnica predominante en la toma de decisiones. Se aplicó la metodología PRISMA 2020, la revisión incluyó artículos publicados entre enero 2020 y junio 2025 en bases de datos como Web of Science, Scopus y Scielo y se seleccionaron 30 estudios de alto impacto. Los resultados mostraron que la inteligencia de marketing se organiza en cuatro dimensiones: mercado, clientes, productos y competidores, apoyándose en fuentes de *big data* como redes sociales, reseñas en línea y bases corporativas. Se identificaron seis modelos de minería de datos, siendo el análisis de sentimientos la técnica más recurrente. Esta permite clasificar opiniones, anticipar crisis de reputación, segmentar clientes y diseñar estrategias competitivas en tiempo real. En conclusión, el estudio evidencia que la inteligencia de marketing se potencia mediante minería de texto, de manera que se consolide al análisis de sentimientos como herramienta central que transforma datos no estructurados en insumos estratégicos para la relación con clientes y la anticipación organizacional.

**Palabras clave:** inteligencia de marketing, minería de texto, modelos de predicción, técnicas analíticas, segmentación de clientes.

**JEL:** M31, C88, D83

### **Abstract**

Marketing intelligence has established itself as a strategic pillar of organizational management, supported by a digital environment characterized by an abundance of data that seeks to transform information into useful knowledge. In this context, text mining and sentiment analysis have become increasingly important for interpreting data in strategic decisions. The objective of this study was to conduct an evidence-based literature review that identified the types of marketing intelligence, the associated data sources, and the role of sentiment analysis as a predominant technique in decision-making. The PRISMA methodology was applied, the review included articles published between January 2020 and June 2025 in databases such as Web of Science, Scopus, and Scielo, and 30 high-impact studies were selected. The results showed that marketing intelligence is organized into four dimensions: market, customers, products, and competitors, relying on big data sources such as social media, online reviews, and corporate databases. Six data mining models were identified, with sentiment analysis being the most commonly used technique. This allows opinions to be classified, reputation crises to be anticipated, customers to be segmented, and competitive strategies to be designed in real time. In conclusion, the study shows that marketing intelligence is enhanced by text mining, consolidating sentiment analysis as a central tool that transforms unstructured data into strategic inputs for customer relations and organizational anticipation.

**Keywords:** marketing intelligence, text mining, predictive modeling, analytical techniques, customer segmentation.

## Introducción

El auge de la era digital ha permitido la apertura a varias oportunidades para que las empresas promuevan la transformación digital mediante el uso de nuevos sistemas y servicios inteligentes, los cuales se encuentran altamente relacionados con la inteligencia del marketing (IM) (Arango et al., 2021). Esta, a su vez, se refiere al proceso sistemático de recolección, análisis y uso de información relevante sobre el mercado, consumidores y competencia. La misma se define como el proceso sistemático de recopilación, análisis y uso estratégico de información relevante sobre el mercado, los consumidores y la competencia (Riaño y Navas, 2020).

De acuerdo con Cárdenas, Olivares y Alfaro (2014), la IM busca comprender tendencias y comportamientos del consumidor, anticipar oportunidades, optimizar campañas y mejorar la posición competitiva de las marcas. Guo (2024) señala que el objetivo principal es minimizar riesgos y maximizar oportunidades para optimizar recursos, mejorar el desempeño comercial y diseñar estrategias alineadas con los cambios tecnológicos emergentes (Putatunda et al., 2023). En este sentido, la IM se convierte en un recurso estratégico clave para la toma de decisiones basadas en datos y para fortalecer la competitividad empresarial en entornos dinámicos y digitales (López et al., 2021).

Actualmente, el volumen de datos generados ha crecido exponencialmente, con una proporción significativa en formato no estructurado, especialmente texto. Según la Unión Internacional de Telecomunicaciones (2024), más de 5.650 millones de personas utilizan internet, lo que representa el 68,7 % de la población mundial, y 5.410 millones de usuarios acceden a redes sociales, que dedica en promedio tres horas diarias a estas plataformas. Esta intensa actividad digital genera un flujo masivo de información textual que ofrece un gran potencial para el análisis de mercado (Dang et al., 2020). Sin embargo, la magnitud, diversidad y velocidad con la que se produce esta información la vuelven difícil de gestionar mediante métodos tradicionales, lo que impulsa a las organizaciones a incorporar herramientas avanzadas que permitan procesar datos de múltiples fuentes, como redes sociales, reseñas de productos, foros en línea y comentarios de consumidores (Greco, 2020).

Ahora bien, estos datos, al ser extremadamente grandes en volumen, velocidad, variedad y veracidad, provienen de diversas fuentes como redes sociales, reseñas de productos, foros en línea, comentarios de consumidores, entre otros, que constituyen una riqueza de información invaluable para comprender los comportamientos y preferencias de los consumidores (Darad et al., 2023). En este sentido, el marketing moderno se apoya fundamentalmente en herramientas de *Big Data* o de análisis de texto, y busca transformar grandes cantidades de información en decisiones estratégicas basadas en evidencia a favor de las empresas (Castrillón, 2024).

Bajo este precepto, el análisis de esta vasta cantidad de información puede resolverse mediante la aplicación de la minería de texto (*text mining*, MT), la cual se centra en procesar grandes cantidades de información escrita para transformarla en una forma accesible y entendible para su análisis. Según Dang et al. (2020), la MT es un conjunto de técnicas, procesos, modelos y herramientas que permiten extraer conocimiento útil a partir de grandes volúmenes de texto no estructurado. Asimismo, Vilches et al., (2020) detallan que la MT representa una función esencial en las organizaciones, dado que se encarga de planificar y ejecutar estrategias orientadas a cumplir

las necesidades del público a largo plazo; además, no se limita únicamente a vender productos o servicios, sino que busca establecer relaciones duraderas con los clientes mediante propuestas de valor constantes (Ferrin, 2023).

En este contexto, el análisis de sentimientos se posiciona como la técnica más representativa dentro de la minería de texto aplicada al marketing, ya que permite identificar emociones, percepciones y valoraciones expresadas por los consumidores en plataformas digitales, convirtiéndose en un recurso clave para comprender en profundidad sus preferencias y orientar las estrategias empresariales (Bagheri & Entezarian, 2023; Bharathi et al., 2024). En consecuencia, se puede mencionar que la inteligencia de marketing puede percibirse como el proceso de aplicar técnicas de minería de datos para recopilar información sobre clientes, competidores, mercados e industria, los cuales a futuro se pueden aplicar a planes estratégicos de marketing. Por lo tanto, se induce que las empresas puedan acelerar la transformación digital mediante la innovación de productos, la identificación de clientes y la previsión de la demanda del mercado (Mustak et al., 2021; Singh et al., 2024).

A pesar de sus beneficios, las empresas enfrentan desafíos al aprovechar la IM mediante MT. Uno de los principales obstáculos es la identificación de fuentes de datos relevantes y la adecuada clasificación de sus componentes (Bhowmik et al., 2023). Además, aunque el uso de *Big Data* en marketing ha cobrado relevancia, aún existe una brecha en la comprensión técnica de los modelos y técnicas de minería de datos que permiten explotar su potencial de manera efectiva (Sánchez et al., 2025). Bajo estos antecedentes, resulta imprescindible dar respuesta a la pregunta: ¿cómo la minería de texto contribuye a la inteligencia de marketing y a la toma de decisiones estratégicas en las organizaciones? De tal manera que, el objetivo de este artículo es realizar un análisis teórico basado en evidencia sobre la minería de texto, con énfasis en el análisis de sentimientos, como herramienta de inteligencia en marketing.

## Metodología

El modelo *PRISMA* (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) es una metodología internacionalmente reconocida que busca garantizar la transparencia, exhaustividad y rigor científico en las revisiones sistemáticas (Page et al., 2021). En esta investigación, se aplicó *PRISMA* para organizar y sistematizar el proceso de recopilación de información sobre la aplicación de la minería de texto en la inteligencia de marketing.

El método de análisis-síntesis permite descomponer un fenómeno en sus elementos fundamentales para luego integrar los hallazgos y construir una visión global coherente. El análisis se aplicó para revisar sistemáticamente la literatura científica en el campo del marketing, mediante la identificación de técnicas, modelos y tendencias relevantes. Así mismo, el método histórico-lógico vincula el desarrollo temporal de los fenómenos con su racionalidad lógica, lo que permite explicar cómo se originan y evolucionan teorías y métodos. Este enfoque facilitó rastrear cómo las técnicas y modelos han evolucionado en marketing digital (Page et al., 2021).

### ***1. Estrategia de búsqueda y selección de la literatura***

La búsqueda se llevó a cabo en las bases de datos Scopus y Web of Science, se prioriza aquellas investigaciones publicadas en revistas científicas indexadas. Se definieron términos de búsqueda en inglés y español con el fin de ampliar los resultados y capturar la mayor diversidad de información posible. Para ello, se utilizaron operadores booleanos (AND, OR, NOT) y palabras clave relacionadas con las variables de estudio, lo que permitió delimitar los resultados y conseguir aquellos que se encuentren relacionados con el tema de estudio.

- "Text mining" AND ("marketing intelligence" OR "marketing analytics") AND "data mining models"
- "Sentiment analysis" AND ("marketing intelligence" OR "digital marketing") AND "text mining techniques"
- "Customer behavior" AND "data mining models"

### ***2. Criterios de inclusión y exclusión***

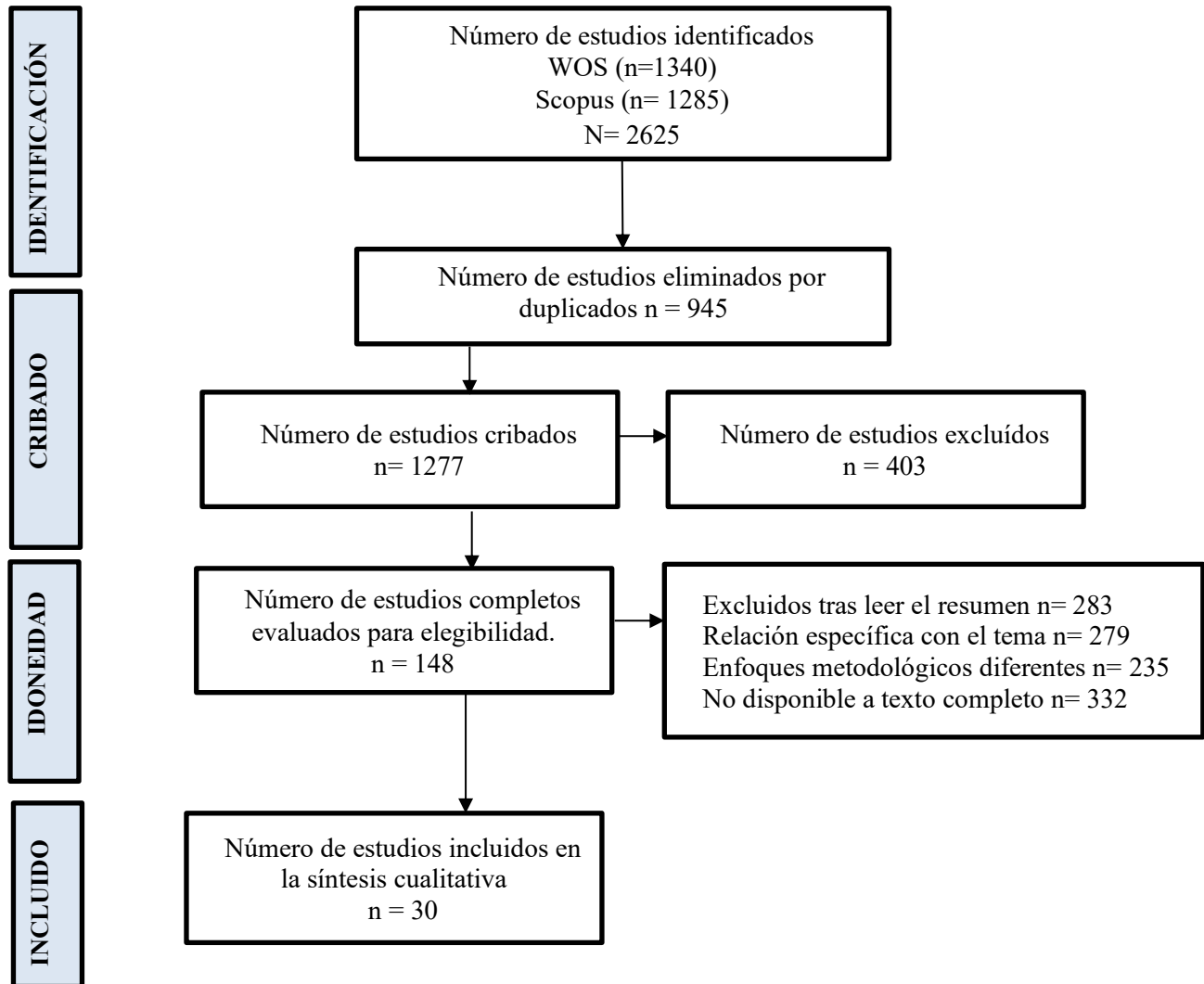
Los criterios de inclusión contemplaron estudios publicados entre enero 2020 y junio 2025, en idioma inglés o español, con acceso al texto completo, que aplicaran modelos o técnicas de minería de texto en contextos de marketing. Los criterios de exclusión se centraron en artículos duplicados, publicaciones no académicas (blogs, conferencias sin revisión por pares, tesis, artículos de pago) y aquellos que no guardaran relación directa con la temática del estudio.

### ***3. Selección y extracción de datos***

El proceso PRISMA contempló cuatro fases: (1) Identificación de los estudios iniciales en las bases de datos; (2) Cribado para eliminar duplicados y textos irrelevantes; (3) Idoneidad, donde se evaluaron los resúmenes y metodologías en función de los objetivos; y (4) Inclusión, con la selección final de artículos que cumplen los criterios establecidos como se identifica en la figura 1.

**Figura 1**

*Diagrama de Flujo PRISMA*



*Fuente:* Proceso sistemático de selección de estudios estructurado bajo el método PRISMA. Adaptado de Page et al. (2021) Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas.

## Resultados y Discusión

Según Zhang et al. (2022), los recursos internos de una empresa son clave para obtener una ventaja competitiva sostenible. Un recurso es valioso si ayuda a una empresa a explotar oportunidades o neutralizar amenazas del entorno, raro si no está disponible para todos los competidores, y difícil de imitar si no puede ser fácilmente replicado. Estos recursos pueden ser tanto tangibles como intangibles los cuales incluyen activos físicos, humanos, tecnológicos o

culturales, y su correcta explotación genera valor estratégico. En ese sentido, la inteligencia de marketing se considera un recurso significativo para la mejora e innovación de productos y servicios. Los modelos y técnicas de minería de texto cumplen la función de capacidades que transforman el *Big Data* en conocimiento accionable (Huang, 2025). En consecuencia, resulta indispensable examinar el papel de la inteligencia de marketing junto con los modelos y técnicas de minería de texto, con énfasis en el análisis de sentimientos, en la creación de ventajas competitivas.

## **Inteligencia de marketing**

La inteligencia de marketing se define como un proceso sistemático orientado a recopilar, analizar y utilizar datos relevantes del mercado, los competidores, los clientes y el entorno empresarial, con el objetivo de tomar decisiones estratégicas informadas que mejoren el rendimiento y la competitividad organizacional (Morales et al., 2021). Este enfoque incluye tanto información interna como externa, se abarca tendencias de mercado, patrones de comportamiento del consumidor, acciones de la competencia, y cambios tecnológicos o económicos que impactan el entorno empresarial (Maibaum et al., 2024; Kumar et al., 2023). A través de esta inteligencia, las organizaciones pueden anticiparse a los cambios del entorno y ajustar sus estrategias de manera proactiva y fortalece así su capacidad para adaptarse a contextos dinámicos y altamente competitivos (Mustak et al., 2021). En este sentido, la inteligencia de marketing actúa como un recurso estratégico que permite transformar información dispersa en conocimiento accionable y facilita una ventaja competitiva sostenible a través de la toma de decisiones basada en datos (Jorquera et al., 2024).

Históricamente, la inteligencia de marketing se fundamentaba principalmente en encuestas, estudios de mercado y fuentes internas de información. Sin embargo, con la expansión exponencial del volumen de datos digitales y no estructurados, esta definición ha evolucionado hacia la integración de modelos y técnicas de text mining para descubrir patrones, conocimientos y tendencias relevantes que respalden la toma de decisiones (Dang et al., 2020; Jung & Lee, 2020). Este cambio ha llevado a algunos autores a destacar el rol de la minería de datos como soporte esencial para convertir grandes volúmenes de información en ventajas estratégicas (Bagheri & Entezarian, 2023; Saleh & AlShafeey, 2025). Gracias a estas herramientas, las organizaciones pueden segmentar mercados de manera más precisa, personalizar ofertas, optimizar campañas de marketing y mejorar los niveles de satisfacción del cliente a partir de fuentes como redes sociales, plataformas de comercio electrónico y reseñas en línea (Huang et al., 2025; Chatterjee et al., 2021).

En este nuevo paradigma, la minería de texto (text mining, MT) se ha consolidado como una capacidad fundamental para extraer inteligencia de marketing a partir de datos no estructurados (Gutiérrez et al., 2022). La MT comprende un conjunto de modelos, procesos y técnicas diseñadas para transformar texto en bruto en información estructurada y útil para el análisis (Dang et al., 2020). Su relevancia radica en que permite a las empresas manejar volúmenes masivos de datos con eficiencia, identificar patrones de comportamiento del consumidor, reconocer tendencias emergentes y evaluar percepciones sobre productos y marcas expresadas en medios digitales (Hassani et al., 2020; Mustak et al., 2021). Entre las técnicas aplicadas, el análisis de sentimientos destaca como una de las más representativas, debido a que permite identificar emociones y valoraciones expresadas en plataformas digitales, lo cual facilita la toma de

decisiones estratégicas más precisas en áreas como desarrollo de productos, segmentación de mercado y gestión de marca (Alantari et al., 2022; Araya-Pizarro & Verelst, 2024).

### **Minería de texto**

Las empresas deben identificar la inteligencia de marketing, junto con los modelos y técnicas de minería de datos, como recursos y capacidades significativos para competir (Lee et al., 2022). Es así pues que, la inteligencia de marketing comprende la inteligencia de mercado, inteligencia de productos, inteligencia de clientes e inteligencia de competidores, que no es más que abarcar todos los aspectos del marketing mix (Pejic et al., 2020). Estos cuatro componentes de la inteligencia de marketing pertenecen a una capa interna del marco en donde se determinan los modelos de minería de texto más típicos para extraer diferentes componentes de la inteligencia de marketing en donde para cada una se emplean modelos de minería de texto que permiten extraer valor de los datos (Dang et al., 2020; Vanhala et al., 2020; Zhu et al., 2023).

### **Modelos asociados**

De acuerdo con la literatura, los seis modelos típicos son: clasificación, asociación, agrupamiento, regresión, predicción y descubrimiento de secuencias, mismos que se detallan a continuación:

#### ***Modelo 1: Clasificación***

La clasificación es una técnica utilizada para predecir el comportamiento de los clientes, que asigna una categoría a un nuevo objeto basándose en un conjunto de datos previos. La clasificación ayuda a las empresas a segmentar clientes, predecir su comportamiento y optimizar las campañas de marketing. Se aplica mediante técnicas como redes neuronales, árboles de decisión, *Naïve Bayes* o *Support Vector Machines* (SVM). El análisis de sentimientos refuerza este modelo al clasificar opiniones en positivas, negativas o neutras, lo que permite segmentar clientes y diseñar campañas con base en percepciones emocionales. Además, autores destacan que, al aplicar estas técnicas, las organizaciones pueden identificar patrones en grandes volúmenes de datos, mejora la capacidad de tomar decisiones estratégicas y predicciones más precisas sobre el comportamiento de los consumidores (Mustak et al., 2021; Saleh & AlShafeey, 2025; Singh & Sai Vijay, 2024; Singh et al., 2024).

#### ***Modelo 2: Asociación***

La minería de asociación se emplea para descubrir relaciones entre productos comprados por los clientes, lo que permite a las empresas identificar qué productos tienden a venderse juntos. Según Kumar et al. (2023), las reglas de asociación son fundamentales para mejorar las estrategias de ventas cruzadas y recomendaciones de productos. Estas relaciones son útiles para aumentar el valor de las transacciones y optimizar las estrategias de inventario. El análisis de sentimientos complementa este modelo al asociar la emoción expresada en un comentario con la probabilidad de adquirir productos relacionados, mejora estrategias de ventas cruzadas y recomendaciones. Saleh & AlShafeey (2025) destacan que el uso de la minería de asociación puede ayudar a las

empresas a entender mejor las preferencias de los clientes, lo cual es crucial para el diseño de estrategias de marketing personalizadas.

### ***Modelo 3: Agrupamiento (Clustering)***

El agrupamiento o *clustering* se utiliza para segmentar clientes según características similares, el cual permite a las empresas crear perfiles detallados de los usuarios. El *clustering* ayuda a identificar grupos de clientes con comportamientos similares, lo que es esencial para la segmentación de mercado y el desarrollo de estrategias de marketing personalizadas, segmenta consumidores en grupos con características similares. Integrar análisis de sentimientos permite refinar esta segmentación, dado que los clústeres no solo se definen por variables demográficas o transaccionales, sino también por emociones compartidas hacia productos o marcas que permiten una visualización y análisis más intuitivo de los grupos emergentes (Basole et al., 2024; Greco & Polli, 2020; Saleh & AlShafeey, 2025).

### ***Modelo 4: Regresión***

La regresión se emplea para establecer relaciones entre variables que permite hacer predicciones sobre valores continuos y analizar relaciones causales (Núñez et al., 2021). De esta manera, la regresión lineal y la regresión logística son técnicas fundamentales para modelar el comportamiento del consumidor y predecir respuestas a campañas de marketing o cambios en el mercado. La regresión lineal se utiliza para predecir variables continuas, mientras que la regresión logística es más adecuada para problemas de clasificación binaria, como determinar si un cliente comprará o no un producto. El análisis de sentimientos fortalece este modelo al cuantificar emociones en escalas numéricas y utilizarlas como comportamiento de compra o de la satisfacción. Es decir, el uso de la regresión en marketing y análisis de clientes permite obtener datos precisos, lo que optimiza la toma de decisiones empresariales (Zhu et al., 2023).

### ***Modelo 5: Predicción***

La predicción es una técnica de minería de datos utilizada para prever eventos futuros basándose en datos históricos. La predicción es esencial para anticipar cambios en el comportamiento de los clientes y adaptar las estrategias de marketing en consecuencia. Las redes neuronales, el análisis de supervivencia y los modelos de regresión logística son algunas de las técnicas más comunes para realizar predicciones. Aquí, el análisis de sentimientos permite prever cambios en la aceptación de productos, la reputación de la marca o la reacción de los consumidores ante nuevas estrategias de marketing. Estas técnicas pueden ayudar a las empresas a anticipar la demanda de productos, las fluctuaciones del mercado y las necesidades de los consumidores (Vanhala et al., 2020; Zhu et al., 2023).

### ***Modelo 6: Descubrimiento de secuencias***

El descubrimiento de secuencias se utiliza para identificar patrones o relaciones en datos secuenciales, como el comportamiento del consumidor a lo largo del tiempo. Este enfoque permite descubrir asociaciones temporales entre eventos, lo que es especialmente útil para entender las fases del ciclo de vida de un cliente. Este modelo detecta patrones temporales en el

comportamiento del consumidor. El análisis de sentimientos permite identificar la evolución de emociones a lo largo del tiempo y es relativamente útil para predecir el comportamiento futuro de los consumidores en función de sus interacciones pasadas, como las compras previas o las acciones en una plataforma digital (Vanhala et al., 2020; Zhu et al., 2023).

**Tabla 1**

*Relación entre modelos de minería de texto y aplicación del análisis de sentimientos*

Clasificación	Determinar polaridad de opiniones para segmentar clientes
Asociación	Relacionar emociones con productos frecuentemente adquiridos
Agrupamiento ( <i>Clustering</i> )	Crear segmentos basados en similitudes emocionales y conductuales
Regresión	Usar puntuaciones de sentimientos como variables predictoras
Predicción	Anticipar tendencias de aceptación de productos según emociones
Descubrimiento de secuencias	Analizar evolución temporal de percepciones y experiencias del cliente

*Fuente:* A partir de (Bagheri & Entezarian, 2023; Bharathi et al., 2024; Chintalapudi et al., 2022; Hassani et al., 2020; Jung & Lee, 2020; Mustak et al., 2021; Saleh & AlShafeey, 2025; Singh & Sai Vijay, 2024; Singh et al., 2024).

## Inteligencia de marketing y minería de datos

Considerando los conceptos de diversos autores se puede inferir que la inteligencia de marketing se define como el proceso de recopilar información sobre clientes, competidores, mercados e industria mediante técnicas de minería de texto, y luego aplicarla en planes estratégicos de marketing. Esta definición se basa en el marketing mix, de manera que cubre todos los aspectos importantes que respaldan las decisiones de marketing (Huang et al., 2025; Sarkar & De Bruyn, 2021; Stallings et al., 2024). No obstante, dado que el *marketing mix* tradicional (producto, precio, plaza y promoción) está centrado en el producto y carece de enfoque en el cliente, es necesario adoptar la quinta P – personas, con un enfoque en el cliente. En consecuencia, la inteligencia de marketing se compone de inteligencia de mercado (plaza), inteligencia de productos (producto), inteligencia de competidores (precio y promoción), e inteligencia de clientes (personas) (Mustak et al., 2021).

## Tablas de resultados de modelos

**Tabla 2**

*Modelo de Clasificación*

<b>Autor y año</b>	<b>Metodología</b>	<b>Técnicas</b>	<b>Resultados</b>
Greco & Polli (2020)	<i>Emotional Text Mining (ETM)</i> con enfoque psico-semiótico	<b>ETM:</b> extrae matrices léxicas / emocionales, agrupa representaciones afectivas y <b>etiqueta perfiles</b> de usuarios/consumidores. <i>Software:</i> R / Python.	Logra <b>perfiles emocionales</b> accionables que apoyan decisiones de <i>branding</i> y gestión de comunidades.
		Propone y aplica ETM para <b>perfiles clientes</b> y mapear su <b>sentimiento</b> hacia marcas	
Chatterjee, Goyal, Prakash & Sharma (2021)	Text mining ML aplicado a <i>e-commerce</i> salud	<b>Clasificación supervisada de sentimiento/temas:</b> modelos ML sobre características lingüísticas; combinación con técnicas econométricas para ligar emociones y satisfacción.  Analiza <b>satisfacción</b> en <i>e-commerce</i> de salud para identificar atributos de servicio y emociones que mejor predicen satisfacción.	Identifica qué <b>emociones</b> y aspectos del servicio predicen mejor la satisfacción del cliente y muestra una guía de diseño de experiencia y métricas.
Alantari, Currim, Deng & Singh (2022)	Estudio comparativo experimental con <i>online reviews</i>	<b>SVM / Naïve Bayes / árboles / redes profundas:</b> Clasificadores supervisados que aprenden; se entrenan con textos vectorizados.  Compara empíricamente métodos de <i>machine learning</i> para <b>clasificar sentimiento</b> en reseñas de consumidores.	Los modelos lineales fuertes (p. ej., SVM/logística) y <i>deep learning</i> bien ajustado muestran los mejores rendimientos globales en <i>reviews</i> ; evidencia práctica para selección de modelos en marketing.
Alsemaree, Alam, Gill & Uhlig (2024)	Viaje del cliente potencial supervisado en textos árabes de redes sociales	<b>SVM / LR / NB :</b> Tokenización y normalización para dialectos; entrenamiento supervisado con <i>n-grams/embeddings</i> ; evaluación con <i>cross-validation</i> .  Construye clasificadores de <b>sentimiento en árabe</b> para captar percepciones de clientes en <i>social media</i> .	Se alcanzan desempeños competitivos pese a retos morfológicos del árabe; se valida la <b>viabilidad del análisis de sentimientos</b> para <i>customer insights</i> regionales.

Autor y año	Metodología	Técnicas	Resultados
Chi, Huang, Jia & Zhang (2025)	Modelo híbrido profundo para reseñas hoteleras	<b>BERT + TCN + BiLSTM:</b> BERT genera <i>embeddings</i> contextuales; TCN capta dependencias locales; BiLSTM secuencias largas; Uso clave para <b>clasificar polaridad</b> . Propone arquitectura híbrida para <b>sentimiento en reseñas</b> de hoteles, optimiza el contexto semántico y la dinámica temporal.	Mejora la <b>precisión</b> frente a baselines al combinar contexto (BERT) con estructuras temporales (TCN/BiLSTM) y atención.

Los estudios revisados evidencian un uso intensivo de técnicas de machine learning supervisado para el análisis de sentimientos en contextos de marketing digital. En particular, métodos clásicos como Support Vector Machines (SVM), Naïve Bayes y regresión logística detallan una alta eficacia para clasificar polaridad en grandes volúmenes de textos provenientes de reseñas en línea y redes sociales (Alantari et al., 2022; Alsemaree et al., 2024). Estas técnicas se apoyan en representaciones vectoriales de texto —como n-grams o embeddings— y en procesos de validación cruzada que aseguran un buen rendimiento generalizable en distintos dominios lingüísticos. Por otro lado, estudios aplicados a nichos específicos, como comercio electrónico en salud, combinan la clasificación con análisis econométrico para vincular emociones con métricas de satisfacción y demuestra que el análisis de sentimientos puede generar *insights* directamente accionables en la mejora de la experiencia del cliente (Chatterjee et al., 2021).

De manera complementaria, enfoques más recientes adoptan arquitecturas híbridas profundas que integran modelos como BERT, TCN y BiLSTM para capturar tanto el contexto semántico como la dinámica temporal de los textos (Chi et al., 2025). Estas combinaciones han superado en precisión a los modelos tradicionales, especialmente en dominios complejos como las reseñas hoteleras, donde la interpretación de matices emocionales es crucial. Finalmente, propuestas metodológicas alternativas, como la Emotional Text Mining, resaltan la importancia de mapear perfiles afectivos mediante análisis léxicos y semióticos y proporciona una visión más rica de las percepciones de marca (Greco & Polli, 2020).

**Tabla 3**

*Modelo de asociación*

Autor y año	Metodología	Técnicas	Resultados
Wu, Zhao & Yang (2024)	Estudio metodológico aplicado a <i>online reviews</i> para construir un FODA objetivo	<b>Reglas de asociación:</b> descubre co-ocurrencias frecuentes “atributo - valor” para mapear <b>Fortalezas</b> / <b>Debilidades</b> / <b>Oportunidades</b> / <b>Amenazas</b> ; se apoyan en <b>análisis de sentimientos</b> para polaridad.	Mejora objetividad de FODA vs. enfoques subjetivos; permite priorizar ventajas/desventajas basadas en lo que dicen los clientes en tiempo real.

Autor y año	Metodología	Técnicas	Resultados
	con <i>text mining</i> híbrido.	Propone pipeline que identifica atributos de servicio desde reseñas, calcula polaridad, y con reglas de asociación determina factores FODA.	
Araya-Pizarro & Verelst (2024)	Estudio empírico con <b>reseñas</b> de hoteles (Booking) de Santiago de Chile; <i>text mining</i> + <b>sentimiento</b> + <b>regresión logística</b> .	<b>Asociación/co-ocurrencias:</b> Usa matrices de co-palabras y asociaciones atributo-emoción para entender qué factores (limpieza, trato, desayuno) co-aparecen con polaridad; determina probabilidad de satisfacción alta. <i>Software: Python/R.</i>  Deriva atributos clave desde texto, enlaza su aparición con polaridad y con la probabilidad de buena valoración.	Identifica atributos con mayor aporte a satisfacción (ej., <i>limpieza</i> y <i>personal</i> ) y sus asociaciones con sentimiento positivo que guían mejoras de servicio.
Singh & Sai Vijay (2024)	<i>Roadmapping</i> tecnológico en e-commerce con <b>co-word/semantic network</b> + <i>topic modeling</i> .	<b>Redes de co-palabras (asociación léxica):</b> construyen grafos de términos para identificar <b>clústers</b> o <b>asociaciones</b> de tecnologías; este grafo puede estratificarse con capas de <b>sentimiento</b> cuando las fuentes incluyen reseñas/noticias; <i>Software: VOSviewer</i>  Mapea relaciones entre conceptos tecnológicos y su evolución temporal, base para detectar oportunidades ligadas a percepciones del mercado.	Identifica dominios tecnológicos emergentes y sus asociaciones; insumos para priorizar inversión y posicionamiento
Huang, Zhang & Song (2025)	Estudio <b>multimodal</b> (texto + imagen) para <i>análisis de sentimientos</i> en información de productos con <b>redes neuronales</b> y <i>text mining</i> .	<b>Asociación multimodal:</b> enlaza señales textuales (tokens) y visuales (rasgos de imagen) mediante <b>atención</b> y <b>transformadores</b> para capturar <b>correspondencias</b> entre modalidad y sentimiento; implementación <b>DL</b> (TCN/BiLSTM).  Integra sentimiento textual con pistas visuales de anuncios para inferir en la actitud del consumidor.	Mejora precisión frente a modelos unimodales; descubre combinaciones “texto-imagen” asociadas a polaridad y <i>compromiso</i> .

Autor y año	Metodología	Técnicas	Resultados
Saleh & AlShafeey (2025)	Análisis de literatura / Texto sobre <b>Industria 4.0</b> y sostenibilidad con <b>minería de texto, sentimiento, reglas de asociación.</b>	<b>Reglas de asociación multinivel:</b> extraen patrones condicionados por la <b>polaridad</b> ; se usa <i>análisis de sentimientos</i> para clasificar tono.  Vincula conceptos de I4.0 con impactos económico/social/ambiental y su sentimiento asociado en textos.	Identifica sinergias y tensiones (p. ej., automatización↔eficiencia con sentimiento positivo; preocupaciones laborales↔sentimiento mixto/negativo).

Los estudios revisados demuestran que las técnicas de asociación se han convertido en un recurso clave para identificar patrones y relaciones ocultas entre atributos, emociones y comportamientos en contextos de marketing digital. Por ejemplo, Wu, Zhao y Yang (2024) proponen un enfoque innovador que combina reglas de asociación con análisis de sentimientos para construir matrices FODA objetivas a partir de reseñas en línea que permiten una priorización estratégica basada en la voz del cliente. De forma complementaria, Saleh y AlShafeey (2025) aplican reglas de asociación multinivel en el ámbito de la Industria 4.0, vincula conceptos tecnológicos con impactos sostenibles y sentimientos asociados, lo que permite revelar tensiones y sinergias difíciles de detectar mediante métodos tradicionales.

Asimismo, otras investigaciones muestran la versatilidad de estas técnicas al integrarlas con enfoques multimodales y semánticos. Huang, Zhang y Song (2025) demuestran que la combinación de asociaciones entre texto e imagen mediante redes neuronales mejora significativamente la capacidad de inferir actitudes del consumidor en anuncios digitales que supera los modelos unimodales. Por su parte, Araya-Pizarro y Verelst (2024) utilizan asociaciones atributo-emoción para vincular factores como limpieza o atención al cliente con polaridades positivas en reseñas hoteleras y ofrecen información clave para mejorar la satisfacción. Finalmente, Singh y Sai Vijay (2024) emplean redes de co-palabras para mapear relaciones tecnológicas en el comercio electrónico y se evidencia el potencial de estas técnicas para identificar dominios emergentes y percepciones del mercado.

**Tabla 4**

*Modelo de agrupamiento (Clustering)*

Autor y Año	Metodología	Técnicas	Resultados
Vanhala et al. (2020)	Análisis bibliométrico y text mining de estudios de comportamiento del consumidor	K-means: segmentación de artículos en clusters temáticos. Análisis de sentimientos: aplicado en citas y abstracts para detectar la orientación de los hallazgos. - <i>Software:</i> Bibliometrix, R.	Identificaron 5 clústeres principales de investigación en comportamiento online; los sentimientos positivos dominaron

Autor y Año	Metodología	Técnicas	Resultados
		Se aplicaron técnicas de clustering para agrupar estudios previos y mapear tendencias de consumo digital.	estudios de satisfacción de usuarios.
Chang et al. (2021)	Minería de texto aplicada a educación ambiental	LDA (Latent Dirichlet Allocation): para agrupar temas emergentes. K-means: para segmentación complementaria. Análisis de sentimientos: aplicado a reseñas textuales. <i>Software:</i> Python	Se identificaron clústeres temáticos sobre sostenibilidad con polaridad emocional diferenciada, útil para guiar campañas educativas.
		Usaron clustering para organizar tópicos recurrentes y asociarlos con la carga emocional de los textos.	
Greco & Polli (2020)	<i>Emotional text mining</i> en gestión de marcas	Self-Organizing Maps (SOMs): para agrupar perfiles emocionales. - Análisis de sentimientos: extracción de emociones asociadas a marcas. <i>Software:</i> MATLAB.	Los clústeres permitieron identificar perfiles emocionales distintos, útiles para estrategias de branding personalizadas.
		Se construyeron mapas emocionales que agruparon consumidores según sentimientos y emociones expresados hacia marcas.	
Maibaum et al. (2024)	Clasificación de información de sostenibilidad empresarial	Clustering jerárquico: para agrupar reportes de sostenibilidad. Análisis de sentimientos: para medir tono en la narrativa corporativa.  Agruparon documentos corporativos para distinguir narrativas positivas, neutras y negativas respecto a sostenibilidad.	Los resultados revelaron tres clústeres con sesgos emocionales distintos, muestra diferencias en la comunicación sostenible entre sectores.
Yang et al. (2025)	Optimización de análisis de índices de mercado inmobiliario	Autoencoder + algoritmos de clustering: para reducción dimensional y agrupamiento. - Análisis de sentimientos: aplicado a foros e informes de inversión.  Se combinaron técnicas de aprendizaje profundo con clustering para segmentar mercados financieros según patrones de sentimiento de inversores.	Se lograron clústeres con alta coherencia interna; sentimientos negativos se correlacionaron con volatilidad de índices.

Los estudios analizados destacan el uso del *clustering* como una técnica central para segmentar grandes volúmenes de datos textuales, permite identificar patrones y grupos con características homogéneas. Vanhala et al. (2020) aplicaron análisis bibliométrico combinado con K-means para segmentar investigaciones en comportamiento del consumidor, revela cinco clústeres temáticos predominantes y un claro sesgo positivo en los estudios sobre satisfacción digital. Por su parte, Greco y Polli (2020) utilizaron *Self-Organizing Maps* para agrupar perfiles emocionales de consumidores mediante mapas afectivos útiles para estrategias de *branding* personalizadas. Asimismo, Chang et al. (2021) integraron LDA y K-means para organizar tópicos sobre sostenibilidad educativa, vinculándolos con la carga emocional de los textos, lo que permitió diferenciar narrativas temáticas con distintas polaridades afectivas.

En esta misma línea, Maibaum et al. (2024) implementaron clustering jerárquico para clasificar reportes de sostenibilidad, logra identificar narrativas corporativas con sesgos emocionales diferenciados, lo que aporta valor a la comunicación estratégica empresarial. Finalmente, Yang et al. (2025) combinaron *autoencoders* con algoritmos de clustering para segmentar mercados financieros a partir de patrones de sentimiento, muestra correlaciones claras entre emociones negativas y cambios.

**Tabla 5**

*Modelo de regresión*

<b>Autor y Año</b>	<b>Metodología</b>	<b>Técnicas</b>	<b>Resultados</b>
Chintalapudi et al. (2022)	Modelado predictivo aplicado a documentos médicos con <i>text mining</i>	<i>LASSO Regression</i> : selección de variables relevantes para predicción. Análisis de sentimientos: extracción de polaridad para variables textuales. <i>Software</i> : R  Se utilizó regresión penalizada para identificar términos clave que predicen condiciones médicas, integra sentimientos en los textos como variables predictoras.	Se mejoró la precisión de predicción en comparación con modelos sin sentimientos; el análisis emocional aportó variables discriminantes.
Kutela et al. (2022)	Análisis comparativo de compromiso en Twitter bilingüe	Regresión múltiple: modelado de compromiso ( <i>likes, retweets</i> ) vs. características textuales. Análisis de sentimientos: medición de tono emocional en tweets. <i>Software</i> : Python  Se evaluó cómo el sentimiento y el idioma afectan el compromiso en redes sociales.	Tweets positivos en el idioma dominante generaron más interacción; la regresión explicó un alto porcentaje de varianza en compromiso.

Autor y Año	Metodología	Técnicas	Resultados
Zhu et al. (2023)	Análisis de reportes de incidentes con reglas de asociación y modelos mixtos	Regresión logística multinivel: para identificar factores asociados a incidentes. Análisis de sentimientos: extracción de tono para predecir gravedad del incidente. Software: SPSS / Python  Se combinaron reglas de asociación y regresión para modelar relaciones entre tono textual y severidad de eventos.	El tono negativo se correlacionó significativamente con incidentes de mayor gravedad, mejora la capacidad de predicción.
Macías & Borges (2024)	Monitoreo y pronóstico de indicadores de usabilidad	Regresión lineal: pronóstico de tendencias en indicadores. Análisis de sentimientos: sobre reseñas de usuarios para identificar satisfacción. - Software: Python  Se aplicaron modelos lineales para predecir evolución de métricas de usabilidad basadas en comentarios textuales y su polaridad.	Los modelos lograron alta precisión predictiva, muestra que sentimientos negativos anticipan caídas en indicadores de usabilidad.
Huang et al. (2025)	Análisis multimodal de marketing digital con IA	Regresión lineal y no lineal: para correlacionar características de marketing (texto + imagen) con impacto en consumidores. Análisis de sentimientos: sobre descripciones textuales. Software: TensorFlow, Python.  Se aplicaron modelos de regresión sobre datos textuales y visuales combinados para evaluar el impacto de campañas de marketing.	La integración de sentimientos textuales mejoró significativamente el ajuste de los modelos; se identificaron relaciones no lineales importantes en la respuesta del consumidor.

Los estudios analizados muestran que la regresión se ha convertido en una herramienta clave para integrar información textual y sentimientos en modelos predictivos de marketing y análisis organizacional. Chintalapudi et al. (2022) emplearon *LASSO Regression* para extraer variables relevantes de textos médicos que demuestran que el análisis emocional mejora significativamente la precisión predictiva. Kutela et al. (2022) aplicaron regresión múltiple para evaluar la influencia del sentimiento y el idioma en el compromiso de usuarios en Twitter, siendo el hallazgo que los mensajes positivos en el idioma dominante generan mayor interacción. De forma similar, Macías y Borges (2024) demostraron que la regresión lineal permite anticipar variaciones en indicadores de usabilidad a partir de comentarios negativos y aporta una visión proactiva sobre tendencias de satisfacción.

Zhu et al. (2023) y Huang et al. (2025) evidencian el potencial de combinar regresión logística multinivel y modelos no lineales con análisis de sentimientos para comprender fenómenos complejos. Zhu et al. (2023) correlacionaron tono negativo con incidentes de alta gravedad, mejora la capacidad de predicción de riesgos. Por su parte, Huang et al. (2025) integraron datos textuales y visuales en modelos de regresión multimodal, encuentra relaciones no lineales significativas entre sentimientos y respuestas de consumidores.

**Tabla 6***Modelo de predicción*

<b>Autor y Año</b>	<b>Metodología</b>	<b>Técnicas</b>	<b>Resultados</b>
Liu et al. (2023)	Predicción de recomendaciones de productos a partir de secuencias móviles	<i>Sequence pattern mining</i> para identificar patrones de comportamiento. Análisis de sentimientos para enriquecer las secuencias con información emocional. <i>Software: R.</i> Se utilizaron secuencias de uso en apps móviles combinadas con sentimientos para predecir futuras recomendaciones personalizadas.	La integración de sentimientos aumentó la precisión de predicción de productos recomendados en un 18 %.
Bharathi et al. (2024)	Modelado predictivo adaptativo con IA en simulaciones de <i>business intelligence</i>	Redes neuronales profundas para aprendizaje adaptativo y pronóstico. Análisis de sentimientos para capturar respuestas emocionales de usuarios y mejorar predicción. Se desarrolló un sistema inteligente que utiliza sentimientos de usuarios en entornos de simulación para ajustar modelos predictivos en tiempo real.	Los modelos aumentaron su precisión predictiva al incorporar variables emocionales, mejora la adaptación a cambios del entorno de mercado.
Macías & Borges (2024)	Forecasting de indicadores de usabilidad y satisfacción de usuarios	Regresión lineal + series temporales para proyectar tendencias. Análisis de sentimientos de comentarios de usuarios para detectar cambios futuros.  Se utilizaron modelos temporales alimentados por análisis de sentimientos para anticipar cambios en satisfacción y usabilidad.	Se alcanzaron altos niveles de ajuste ( $R^2 > 0.85$ ), muestra que cambios en sentimiento preceden cambios en indicadores clave.
Chi et al. (2025)	Predicción de percepciones en reseñas hoteleras con modelos híbridos	BERT-TCN-BiLSTM con atención para predecir sentimientos y comportamientos futuros.	El modelo híbrido superó ampliamente a los modelos tradicionales, con precisión superior al

Autor y Año	Metodología	Técnicas	Resultados
		Análisis de sentimientos como objetivo de predicción y variable intermedia. Software: Python Se combinan técnicas de PLN y <i>deep learning</i> para predecir la evolución de la satisfacción del cliente a partir de opiniones textuales.	90 % en predicción de sentimientos futuros.
Huang et al. (2025)	Pronóstico de impacto de campañas digitales multimodales	Redes neuronales + regresión no lineal para predecir reacciones de consumidores. Análisis de sentimientos sobre descripciones de campañas. Software: Python. Se construyó un modelo multimodal que integra texto, imagen y sentimiento para anticipar impacto de campañas en consumidores.	La combinación multimodal mejoró sustancialmente la exactitud predictiva y permitió segmentar respuestas emocionales del público.

Los estudios revisados evidencian el papel fundamental de las técnicas predictivas en la transformación de datos textuales y emocionales en información estratégica accionable. Liu et al. (2023) demostraron que la combinación de *sequence pattern mining* con análisis de sentimientos en secuencias de uso de aplicaciones móviles mejora notablemente la precisión de las recomendaciones personalizadas. Bharathi et al. (2024) resaltaron que las redes neuronales profundas permiten ajustar modelos predictivos en tiempo real al integrar variables emocionales, genera sistemas más adaptativos frente a cambios de mercado. Por su parte, Macías y Borges (2024) confirmaron que el análisis de sentimientos aplicado a series temporales puede anticipar cambios futuros en satisfacción y usabilidad con alta precisión y refuerza el valor de las emociones expresadas en texto como indicadores tempranos de tendencias.

Asimismo, Chi et al. (2025) y Huang et al. (2025) muestran que los modelos híbridos y multimodales potencian las predicciones al integrar texto, contexto semántico y señales visuales. Chi et al. (2025) lograron superar el 90 % de precisión en la predicción de sentimientos futuros mediante arquitecturas *BERT-TCN-BiLSTM*, mientras que Huang et al. (2025) evidenciaron que la integración de texto, imágenes y sentimientos permite segmentar con mayor exactitud las respuestas emocionales del público.

**Tabla 7**

*Modelo de descubrimiento de secuencias*

<b>Autor y Año</b>	<b>Metodología</b>	<b>Técnicas</b>	<b>Resultados</b>
Hassani et al. (2020)	<i>Big data analytics</i> en plataformas digitales	Análisis secuencial estadístico para detectar patrones conductuales longitudinales y aplicado a comentarios en grandes volúmenes de datos no estructurados. Se estudiaron secuencias de interacciones de usuarios en plataformas digitales para identificar ciclos de comportamiento repetitivos.	Se hallaron secuencias emocionales que preceden fases clave del ciclo de compra, mejora la anticipación de decisiones de consumo.
Kutela et al. (2022)	Minería secuencial en <i>compromisos</i> de Twitter bilingüe	Análisis secuencial de tweets para mapear trayectorias de interacción. Regresión + análisis de sentimientos para medir la influencia de emociones en la secuencia de compromiso. Se examinaron secuencias de participación en redes sociales para identificar patrones de difusión de mensajes y su relación con sentimientos expresados.	Los patrones emocionales se asociaron significativamente con aumentos de interacción y retuits en etapas tempranas de las secuencias.
Liu et al. (2023)	<i>Sequence pattern mining</i> aplicado a datos móviles	<i>Sequence Pattern Mining</i> para extraer patrones de uso de aplicaciones. Análisis de sentimientos de interacciones textuales para enriquecer secuencias con emociones. Software: R. Se analizaron trayectorias de interacción en <i>apps</i> móviles combinadas con sentimientos expresados para encontrar secuencias que predicen futuras acciones.	Los modelos identificaron secuencias conductuales predictivas, mejora en 18 % la precisión de recomendaciones personalizadas.
Bagheri & Entezarian (2023)	<i>Topic modeling</i> + análisis de tendencias emergentes	LDA + análisis de secuencias temporales para identificar evolución temática para clasificar la orientación emocional de cada tópico a lo largo del tiempo. Se combinaron tópicos emergentes en marketing con análisis secuencial para detectar cómo cambian percepciones en debates online.	Se identificaron secuencias temáticas que reflejan cambios en actitudes de consumidores frente a innovaciones de mercado.

Autor y Año	Metodología	Técnicas	Resultados
Wu et al. (2024)	Hybrid text mining para análisis SWOT colectivo	Análisis de secuencias de opiniones en línea para detectar patrones de cambio en percepciones. Análisis de sentimientos para clasificar fortalezas y amenazas percibidas.  Se aplicaron métodos híbridos para extraer secuencias temporales de opiniones en reseñas, identifica cómo evoluciona la percepción del consumidor.	Se descubrieron secuencias asociadas a cambios en actitudes del público, útiles para marketing estratégico y posicionamiento competitivo.

Los estudios revisados muestran que el *sequence pattern mining* y el análisis secuencial combinados con análisis de sentimientos son herramientas poderosas para identificar trayectorias conductuales y cambios perceptivos en contextos digitales. Liu et al. (2023) integraron emociones expresadas en interacciones móviles para mejorar la precisión predictiva de recomendaciones personalizadas y demuestra que la incorporación de sentimientos en secuencias aumenta significativamente la capacidad de anticipar comportamientos futuros. Wu et al. (2024) utilizaron minería secuencial híbrida para analizar opiniones en línea y detectar patrones temporales en la evolución de percepciones colectivas. De igual manera, Kutela et al. (2022) mostraron que las secuencias emocionales en redes sociales están fuertemente asociadas con mayores niveles de participación y difusión temprana de mensajes, lo que evidencia la relevancia del componente afectivo en la propagación de información.

Por otra parte, Bagheri y Entezarian (2023) aplicaron análisis secuencial para estudiar la evolución de tópicos y emociones en debates de marketing y logra identificar cambios temáticos asociados a variaciones actitudinales de los consumidores. De forma complementaria, Hassani et al. (2020) analizaron grandes volúmenes de comentarios digitales para descubrir secuencias emocionales recurrentes que preceden fases clave del ciclo de compra.

## Conclusiones

Estos hallazgos confirman que la minería de texto no solo facilita la comprensión de tendencias del mercado y comportamientos de los clientes, sino que también potencia la capacidad empresarial para anticipar escenarios, personalizar estrategias y generar ventajas competitivas sostenibles en entornos dinámicos y altamente digitales. En relación con la inteligencia de marketing, se identificaron seis modelos analíticos principales clasificación, asociación, agrupamiento, regresión, predicción y descubrimiento de secuencias, que permiten procesar grandes volúmenes de datos textuales provenientes de redes sociales, reseñas y otras fuentes digitales.

El análisis de sentimientos se consolidó como la técnica más relevante y transversal entre los modelos identificados, debido a su capacidad para capturar y clasificar emociones,

percepciones y actitudes expresadas en lenguaje natural. Los estudios revisados mostraron que esta técnica mejoró sustancialmente la precisión de los modelos de clasificación, predicción y regresión al incorporar variables afectivas en el análisis de datos textuales.

### **Agradecimientos**

Un agradecimiento al Grupo de Investigación DeTei de la Facultad de Ciencias Administrativas Universidad Técnica de Ambato. Proyecto de investigación “Innovación en la comunicación en la Universidad Técnica de Ambato” por su apoyo en el desarrollo de este estudio, el cual fue aprobado mediante la resolución No. UTA-CONIN-2023-0373-R. y Grupo de investigación Sintezia de la Pontificia Universidad Católica del Ecuador Sede Ambato.

### **Financiamiento**

Esta investigación fue financiada exclusivamente por los autores.

### **Conflicto de intereses**

Declaramos no tener conflictos de intereses financieros ni personales

### **Contribución de los autores**

Maricruz Vega: Conceptualización, metodología, software, validación, análisis formal, investigación, recursos, curación de datos, redacción-borrador original, visualización, administración del proyecto.

Shirley Sanabria: Conceptualización, metodología, software, validación, análisis formal, investigación, recursos, curación de datos, redacción-borrador original, visualización, administración del proyecto.

Patricio Medina: Revisión, redacción – revisión y edición, y supervisión del trabajo.

### **Referencias bibliográficas**

- Alantari, H. J., Currim, I. S., Deng, Y., & Singh, S. (2022). An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews. *International Journal of Research in Marketing*, 39(1), 1-19. <https://doi.org/10.1016/j.ijresmar.2021.10.011>
- Alsemaree, O., Alam, A. S., Gill, S. S., & Uhlig, S. (2024). Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions. *Heliyon*, 10(9). <https://doi.org/10.1016/j.heliyon.2024.e27863>
- Arango Pastrana, C. A., Osorio Andrade, C. F., Arango Pastrana, C. A., & Osorio Andrade, C. F. (2021). Aislamiento social obligatorio: Un análisis de sentimientos mediante machine learning. *Suma de Negocios*, 12(26), 1-13. <https://doi.org/10.14349/sumneg/2021.v12.n26.a1>

- Araya-Pizarro, S., Verelst, N., Araya-Pizarro, S., & Verelst, N. (2024). Más allá de las reseñas: Desvelando la experiencia hotelera con minería de texto. *Revista interamericana de ambiente y turismo*, 20(2), 126-139. <https://doi.org/10.4067/s0718-235x2024000200126>
- Bagheri, R., & Entezarian, N. (2023). Topic Modeling Emerging Trends for Business Intelligence in Marketing: With Text Mining and Latent Dirichlet Allocation. *Journal of System Management*.
- Basole, R. C., Park, H., & Seuss, C. D. (2024). Complex business ecosystem intelligence using AI-powered visual analytics. *Decision Support Systems*, 178, 114133. <https://doi.org/10.1016/j.dss.2023.114133>
- Bharathi, G. P., Chandra, I., Sanagana, D. P. R., Tummalachervu, C. K., Rao, V. S., & Neelima, S. (2024). AI-driven adaptive learning for enhancing business intelligence simulation games. *Entertainment Computing*, 50, 100699. <https://doi.org/10.1016/j.entcom.2024.100699>
- Cárdenas, J., Olivares, G., & Alfaro, R. (2014). Clasificación automática de textos usando redes de palabras. *Revista signos*, 47(86), 346-364. <https://doi.org/10.4067/S0718-09342014000300001>
- Castrillón, O. D. (2024). Análisis y predicción del desempeño docente por medio de encuestas estudiantiles. Búsqueda de relaciones desde la minería de datos. *Formación universitaria*, 17(6), 99-110. <https://doi.org/10.4067/s0718-50062024000600099>
- Chang, I.-C., Yu, T.-K., Chang, Y.-J., & Yu, T.-Y. (2021). Applying Text Mining, Clustering Analysis, and Latent Dirichlet Allocation Techniques for Topic Classification of Environmental Education Journals. *Sustainability*, 13(19), 10856. <https://doi.org/10.3390/su131910856>
- Chatterjee, S., Goyal, D., Prakash, A., & Sharma, J. (2021). Exploring healthcare/health-product ecommerce satisfaction: A text mining and machine learning application. *Journal of Business Research*, 131, 815-825. <https://doi.org/10.1016/j.jbusres.2020.10.043>
- Chi, D., Huang, T., Jia, Z., & Zhang, S. (2025). Research on sentiment analysis of hotel review text based on BERT-TCN-BiLSTM-attention model. *Array*, 25, 100378. <https://doi.org/10.1016/j.array.2025.100378>
- Chintalapudi, N., Angeloni, U., Battineni, G., di Canio, M., Marotta, C., Rezza, G., Sagaro, G. G., Silenzi, A., & Amenta, F. (2022). LASSO Regression Modeling on Prediction of Medical Terms among Seafarers' Health Documents Using Tidy Text Mining. *Bioengineering*, 9(3), 124. <https://doi.org/10.3390/bioengineering9030124>
- Dang, N. C., Moreno-García, M. N., & Prieta, F. D. la. (2020). Sentiment Analysis Based on Deep Learning: A Comparative Study. *Electronics*, 9(3), 483. <https://doi.org/10.3390/electronics9030483>
- Darad, S., Krishnan, S., Darad, S., & Krishnan, S. (2023). Análisis de sentimiento de los datos de twitter de COVID-19 utilizando modelos de aprendizaje profundo y aprendizaje máquina. *Ingenius. Revista de Ciencia y Tecnología*, 29, 108-117. <https://doi.org/10.17163/ings.n29.2023.10>
- Ferrin, F. R. (2023). Comunicación de la política monetaria en Colombia: un análisis a través de la minería de texto. *Semestre Económico*, 26(60). <https://doi.org/10.22395/seec.v26n60a4431>
- Greco, F., & Polli, A. (2020). Emotional Text Mining: Customer profiling in brand management. *International Journal of Information Management*, 51, 101934. <https://doi.org/10.1016/j.ijinfomgt.2019.04.007>

- Guo, X. (2024). *Sentiment Analysis Based on RoBERTa for Amazon Review: An Empirical Study on Decision Making*. <https://doi.org/10.48550/arXiv.2411.00796>
- Gutiérrez-Zapata, D. M., Cerón-Muñoz, M. F., Barahona-Rosales, R., Gutiérrez-Zapata, D. M., Cerón-Muñoz, M. F., & Barahona-Rosales, R. (2022). Prioridades de atención en zonas rurales en escenarios posconflicto: Un análisis de minería de texto. *Revista U.D.C.A Actualidad y Divulgación Científica*, 25(1). <https://doi.org/10.31910/rudca.v25.n1.2022.1947>
- Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text Mining in Big Data Analytics. *Big Data and Cognitive Computing*, 4(1), 1. <https://doi.org/10.3390/bdcc4010001>
- Huang, A., Zhang, F., & Song, C. (2025). Multimodal Sentiment Analysis of Online Product Marketing Information Based on Artificial Intelligence Neural Networks and Text Mining. *IEEE Transactions on Engineering Management*, 72, 3182-3199. <https://doi.org/10.1109/TEM.2025.3589199>
- Jorquera, E., Cabrera-Paniagua, D., Gómez-Narváez, C., Rosas, H., Jorquera, E., Cabrera-Paniagua, D., Gómez-Narváez, C., & Rosas, H. (2024). Metodología de corrección de sesgo para valoraciones de productos en línea utilizando minería de texto afectiva y teoría de respuesta a ítem. *Ingeniare. Revista chilena de ingeniería*, 32. <https://doi.org/10.4067/s0718-33052024000100233>
- Jung, H., & Lee, B. G. (2020). Research trends in text mining: Semantic network and main path analysis of selected journals. *Expert Systems with Applications*, 162, 113851. <https://doi.org/10.1016/j.eswa.2020.113851>
- Kumar, S., Roy, P. P., Dogra, D. P., & Kim, B.-G. (2023). *A Comprehensive Review of Sentiment Analysis: Tasks, Approaches and Applications*. <https://doi.org/10.48550/arXiv.2311.11250>
- Kutela, B., Magehema, R. T., Langa, N., Steven, F., & Mwekh'iga, R. J. (2022). A comparative analysis of followers' engagements on bilingual tweets using regression-text mining approach. A case of Tanzanian-based airlines. *International Journal of Information Management Data Insights*, 2(2), 100123. <https://doi.org/10.1016/j.ijime.2022.100123>
- Lee, M. H., Kim, S., Kim, H., & Lee, J. (2022). Technology Opportunity Discovery using Deep Learning-based Text Mining and a Knowledge Graph. *Technological Forecasting and Social Change*, 180(121718). <https://doi.org/10.1016/j.techfore.2022.121718>
- Liu, H.-W., Wu, J.-Z., & Wang, Y.-H. (2023). Uncovering Insights for New Car Recommendations with Sequence Pattern Mining on Mobile Applications. *Applied Sciences*, 13(11), 6386. <https://doi.org/10.3390/app13116386>
- López Condori, J. J., Gonzales Saji, F. O., López Condori, J. J., & Gonzales Saji, F. O. (2021). Análisis de sentimiento de comentarios en español en Google Play Store usando BERT. *Ingeniare. Revista chilena de ingeniería*, 29(3), 557-563. <https://doi.org/10.4067/S0718-33052021000300557>
- Macías, J. A., & Borges, C. R. (2024). Monitoring and forecasting usability indicators: A business intelligence approach for leveraging user-centered evaluation data. *Science of Computer Programming*, 234, 103077. <https://doi.org/10.1016/j.scico.2023.103077>
- Maibaum, F., Kriebel, J., & Foege, J. N. (2024). Selecting textual analysis tools to classify sustainability information in corporate reporting. *Decision Support Systems*, 183, 114269. <https://doi.org/10.1016/j.dss.2024.114269>
- Morales Sánchez, C. F., Pinto Quijano, Á. C., Rojas Mora, J. E., Vásquez Merchán, I. L. (2021). La política antidrogas Ruta Futuro: Una perspectiva desde el análisis de contenido y la

- minería de textos. *Análisis Político*, 34(103), 111-143. <https://doi.org/10.15446/anpol.v34n103.102170>
- Mustak, M., Salminen, J., Plé, L., & Wirtz, J. (2021). Artificial intelligence in marketing: Topic modeling, scientometric analysis, and research agenda. *Journal of Business Research*, 124, 389-404. <https://doi.org/10.1016/j.jbusres.2020.10.044>
- Núñez, N. A., Crisóstomo, R. A., Sánchez, S. A., Núñez, N. A., Crisóstomo, R. A., & Sánchez, S. A. (2021). Uso de minería de textos para comparar los contenidos relacionados a calidad y acreditación generados en redes sociales por universidades de Perú y Chile. *Formación universitaria*, 14(1), 111-120. <https://doi.org/10.4067/S0718-50062021000100111>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Alonso-Fernández, S. (2021). Declaración PRISMA 2020: Una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología*, 74(9), 790-799. <https://doi.org/10.1016/j.recesp.2021.06.016>
- Pejic-Bach, M., Bertoncel, T., Meško, M., & Krstić, Ž. (2020). Text mining of industry 4.0 job advertisements. *International Journal of Information Management*, 50, 416-431. <https://doi.org/10.1016/j.ijinfomgt.2019.07.014>
- Putatunda, S., Bhowmik, A., Thiruvenkadam, G., & Ghosh, R. (2023). *A BERT based Ensemble Approach for Sentiment Classification of Customer Reviews and its Application to Nudge Marketing in e-Commerce*. <https://doi.org/10.48550/arXiv.2311.10782>
- Riaño Torres, E. F., & Navas Torres, J. D. (2020). *Análisis de la percepción de la calidad de tres empresas colombianas de telefonía móvil en Twitter, mediante técnicas de minería de texto y analítica de negocios*. <https://hdl.handle.net/20.500.14625/25943>
- Saleh, M. A. S., & AlShafeey, M. (2025). Examining the synergies between industry 4.0 and sustainability dimensions using text mining, sentiment analysis, and association rules. *Sustainable Futures*, 9, 100423. <https://doi.org/10.1016/j.sftr.2024.100423>
- Sánchez Lunavictoria, J. C., Sánchez Lunavictoria, D. M., Romero Vélez, E. M., & Macías Zambrano, N. M. (2025). Inteligencia de negocios en la optimización de estrategias de marketing: Enfoque basado en el análisis predictivo. *Revista de ciencias sociales*, 31(2), 340-351. <https://dialnet.unirioja.es/servlet/articulo?codigo=10128637>
- Sarkar, M., & De Bruyn, A. (2021). LSTM Response Models for Direct Marketing Analytics: Replacing Feature Engineering with Deep Learning. *Journal of Interactive Marketing*, 53(1), 80-95. <https://doi.org/10.1016/j.intmar.2020.07.002>
- Singh, S., & Sai Vijay, T. (2024). Technology roadmapping for the e-commerce sector: A text-mining approach. *Journal of Retailing and Consumer Services*, 81, 103977. <https://doi.org/10.1016/j.jretconser.2024.103977>
- Singh, S., Singh, S., Kraus, S., Sharma, A., & Dhir, S. (2024). Characterizing generative artificial intelligence applications: Text-mining-enabled technology roadmapping. *Journal of Innovation & Knowledge*, 9(3), 100531. <https://doi.org/10.1016/j.jik.2024.100531>
- Stallings, L., Bhat, P., Jacobs, J., Lynch, K., & Risch, Q. (2024). How artificial intelligence can enable data classification for market sizing—Insights from applications in practice. *International Journal of Information Management Data Insights*, 4(2), 100271. <https://doi.org/10.1016/j.ijime.2024.100271>
- Vanhala, M., Lu, C., Peltonen, J., Sundqvist, S., Nummenmaa, J., & Järvelin, K. (2020). The usage of large data sets in online consumer behaviour: A bibliometric and computational

- text-mining–driven analysis of previous research. *Journal of Business Research*, 106, 46-59. <https://doi.org/10.1016/j.jbusres.2019.09.009>
- Vilches-Blázquez, L. M., Comesaña Ocampo, D. (2022). Caracterización de riesgos urbanos en prensa aplicando minería de texto para el enriquecimiento de datos abiertos. *Investigación bibliotecológica*, 36(91), 85-107. <https://doi.org/10.22201/iibi.24488321xe.2022.91.58538>
- Wu, J., Zhao, N., & Yang, T. (2024). Wisdom of crowds: SWOT analysis based on hybrid text mining methods using online reviews. *Journal of Business Research*, 171, 114378. <https://doi.org/10.1016/j.jbusres.2023.114378>
- Yang, C.-H., Lee, B., Lee, Y.-I., Chung, Y.-F., & Lin, Y.-D. (2025). An autoencoder-based arithmetic optimization clustering algorithm to enhance principal component analysis to study the relations between industrial market stock indices in real estate. *Expert Systems with Applications*, 266, 126165. <https://doi.org/10.1016/j.eswa.2024.126165>
- Zhang, W., Li, X., Deng, Y., Bing, L., & Lam, W. (2022). *A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges*. <https://doi.org/10.48550/arXiv.2203.01054>
- Zhu, Y., Liao, H., & Huang, D. (2023). Using text mining and multilevel association rules to process and analyze incident reports in China. *Accident Analysis & Prevention*, 191, 107224. <https://doi.org/10.1016/j.aap.2023.107224>