
Análisis de sentimientos en hashtags de consumo sostenible en la red social X

Sentiment analysis of sustainable consumption hashtags on social media platform X

Darío Alejandro Segovia Hernández¹ y Elsa Ortega Rodríguez²

Sumario: 1. Introducción, 2. Fundamentos y fases de la minería de textos, 3. Metodologías y aplicaciones en el análisis de opinión, 4. La influencia de las redes sociales en el consumo sostenible, 5. Metodología, 6. Resultados, 7. Conclusiones, Fuentes de información

Resumen

Este artículo analiza cómo la minería de textos y el análisis de sentimientos pueden mejorar las estrategias de marketing enfocadas en el consumo sostenible. Basado en datos de redes sociales del segundo trimestre de 2024, se identificaron palabras y combinaciones de términos frecuentemente asociados con los hashtags #consumoresponsable, #productosverdes y #sostenibilidad, revelando temas de interés como el cambio climático y la responsabilidad social. Estos términos reflejan las preocupaciones y valores que comparten los usuarios, lo que ayuda a las empresas a entender mejor el contexto en que se discute el consumo sostenible. El análisis de sentimientos muestra una predominancia de "confianza" junto a emociones de "tristeza" y "miedo", lo que sugiere tanto una visión positiva del consumo sostenible como una preocupación sobre sus desafíos. Esto ofrece oportunidades para que las empresas ajusten sus mensajes y aborden las expectativas de sus consumidores. Asimismo, el análisis de clúster identifica tres o cuatro grupos temáticos, lo que permite segmentar audiencias y crear campañas dirigidas, aumentando así la relevancia de las marcas en el mercado sostenible.

Palabras clave: consumo sostenible, minería de textos, análisis de sentimientos, responsabilidad social, segmentación de audiencias

¹ Estudiante de la maestría en Administración por el Instituto de Investigaciones y Estudios Superiores de las Ciencias Administrativas de la Universidad Veracruzana, México. zS22024651@estudiantes.uv.mx ORCID: <https://orcid.org/0009-0009-6977-5318>

² Docente e investigadora en el Instituto de Investigaciones y Estudios Superiores de las Ciencias Administrativas de la Universidad Veracruzana, México. eortega@uv.mx. ORCID: <https://orcid.org/0000-0002-1088-276X>

Abstract

This article examines how text mining and sentiment analysis can enhance marketing strategies focused on sustainable consumption. Based on social media data from the second quarter of 2024, it identifies frequently associated words and term combinations with the hashtags #responsibleconsumption, #greenproducts, and #sustainability, revealing areas of interest such as climate change and social responsibility. These terms reflect the concerns and values shared by users, helping companies to better understand the context in which sustainable consumption is discussed. The sentiment analysis shows a predominance of "trust" alongside emotions like "sadness" and "fear," suggesting both a positive view of sustainable consumption and a concern about its challenges. This provides opportunities for companies to tailor their messages to meet consumer expectations. Additionally, the cluster analysis identifies three or four thematic groups, allowing for audience segmentation and targeted campaigns, thus increasing the relevance of brands within the sustainable market.

Keywords: sustainable consumption, text mining, sentiment analysis, social responsibility, audience segmentation

1. Introducción

Las redes sociales han ganado una gran importancia como canal de comunicación, destacándose por su rapidez en la difusión de información; lo que ha relegado medios tradicionales como la televisión y la radio. Shum (2017) las describe como una nueva estructura social que permite interactuar y compartir contenidos digitales, desde comentarios hasta fotos y videos.

Lo atractivo de estas plataformas radica en la posibilidad de conectar con personas de todo el mundo, facilitando interacciones antes difíciles de lograr. Además, internet ofrece un espacio para expresar ideas libremente y encontrar comunidades con ideologías similares. Este fenómeno ha generado grandes volúmenes de información, abriendo oportunidades para el análisis de datos a través de *big data*.

La minería de opinión o *data mining* surge como una técnica para analizar estos datos textuales no estructurados, asignando valor cuantitativo a opiniones y sentimientos. Esta herramienta permite identificar el grado de polaridad (positivo, neutro o negativo) de las entidades analizadas, y se aplica en redes sociales para resumir las opiniones de los usuarios sobre temas clave, como el consumo sostenible.

El consumo sostenible ha cobrado relevancia en la agenda global, impulsado por la creciente conciencia sobre el impacto ambiental y social de las decisiones de compra. Las redes sociales, como X (anteriormente Twitter), juegan un papel crucial en la difusión de información y la formación de opiniones sobre prácticas sostenibles (Hernández et al., 2024). En este sentido, el análisis de sentimientos en redes sociales es una herramienta útil para comprender las actitudes de los consumidores hacia la sostenibilidad (Baviera, 2017).

El objetivo de este trabajo es demostrar el uso de técnicas de análisis textual en temas sociales, como el consumo sostenible, utilizando la plataforma X y tres hashtags: #consumoresponsable, #productosverdes y #sostenibilidad. Este estudio proporciona información valiosa a organizaciones y marcas interesadas en promover prácticas responsables, ofreciendo una visión de cómo los consumidores valoran la sostenibilidad y qué aspectos generan apoyo o críticas.

2. Fundamentos y fases de la minería de textos

La minería de textos es una técnica de análisis que facilita la extracción de conocimiento de grandes colecciones de documentos. Aunque complementa la minería de datos, que trabaja con bases de datos estructuradas, se ha consolidado como una herramienta versátil aplicada en muchas áreas. Este proceso consta de cinco fases clave.

La primera fase es definir el propósito del estudio, lo que permite establecer cómo recopilar los datos. La segunda fase consiste en identificar y seleccionar los textos relevantes para el análisis. Si los datos provienen de plataformas digitales, es necesario un procesamiento adicional debido a su naturaleza no estructurada.

La tercera fase es el preprocesamiento del texto, que incluye convertir el texto a minúsculas, eliminar signos de puntuación y palabras vacías, como artículos y pronombres. También se realiza la tokenización, que divide el texto en palabras o tokens, y la lematización, que reduce las palabras a su raíz. Esto facilita la interpretación y análisis de los datos.

En la cuarta fase se selecciona el modelo de análisis más adecuado, utilizando técnicas estadísticas y herramientas de minería de textos. La última fase es el análisis de resultados, donde se interpreta la información obtenida para confirmar los objetivos del estudio.

La minería de textos está estrechamente relacionada con la lingüística computacional, que se enfoca en los aspectos gramaticales y sintácticos de los textos. Aunque ambas disciplinas se complementan, cada una tiene objetivos distintos.

Entre las técnicas más utilizadas se encuentran el análisis de sentimientos, que examina la carga emocional de los textos para identificar emociones positivas, negativas o neutras. Las redes sociales son un ámbito ideal para este tipo de análisis. Un ejemplo es el trabajo de Valero (2017), que utiliza el servicio de streaming de X para analizar el tono de los mensajes mediante hashtags y un “diccionario de sentimientos”.

Otra técnica útil es la creación de nubes de palabras: representaciones gráficas que muestran las palabras más frecuentes en blogs o páginas web. Estas nubes organizan las palabras por tamaño e intensidad de color, destacando las más recurrentes (Miente, 2019).

3. Metodologías y aplicaciones en el análisis de opiniones

La minería de textos, como subdisciplina de la minería de datos, extrae información valiosa de grandes volúmenes de documentos textuales al identificar patrones y relaciones entre términos (Martínez et al., 2019). A diferencia de la minería de datos, que trabaja con datos estructurados, la minería de textos se enfoca en datos no estructurados y enfrenta desafíos como la clasificación y comprensión del lenguaje natural en diferentes contextos.

Según Sancho (2023), este proceso permite identificar conceptos clave y tendencias en colecciones textuales, sin necesidad de conocer las palabras exactas. Se apoya en técnicas como la recuperación de información, la extracción de datos y el aprendizaje automático. Entre las técnicas más comunes en el análisis de opiniones se encuentran el análisis de sentimientos, el clustering y los bigramas. El análisis de sentimientos, según Anturi-Martínez y otros (2019), detecta el tono emocional de los textos, mientras que el clustering, según Valero (2017), agrupa documentos por similitud sin una clasificación previa. Los bigramas, por su parte, combinan palabras adyacentes para capturar expresiones que aportan un sentido completo.

Para implementar un modelo de análisis, es crucial seleccionar la técnica adecuada según los objetivos de la investigación. Una de las primeras etapas es la recopilación de un *corpus* de documentos relevantes (John, 2021), como tuits filtrados por *hashtags* en redes sociales. Posteriormente, los resultados se analizan y se presentan de manera visual mediante gráficos o nubes de palabras, lo que facilita su interpretación. Dado que la minería de textos utiliza técnicas de aprendizaje automático, es útil en el reconocimiento de patrones y el análisis de información novedosa, lo que amplía su aplicación en diversas áreas.

4. La influencia de las redes sociales en el consumo sostenible

El consumo sostenible, que implica adoptar hábitos para reducir el impacto ambiental y promover el bienestar social, se ha convertido en una prioridad para muchos consumidores y empresas. En la era digital, la información en línea desempeña un papel crucial en la difusión de prácticas sostenibles y en la toma de decisiones de compra más conscientes. Las plataformas digitales, como redes sociales, blogs y sitios web de noticias, han facilitado el acceso a información sobre sostenibilidad, permitiendo a los consumidores conocer productos y servicios alineados con sus valores ecológicos y sociales (Sanahuja & Tezanos-Vázquez, 2017).

Las redes sociales se han consolidado como un espacio clave para discutir y promover el consumo sostenible, permitiendo a los usuarios compartir experiencias y opiniones (Hernández, 2024). Plataformas como X, Facebook e Instagram dan voz a los consumidores para expresar sus ideas sobre una variedad de temas, incluyendo el consumo sostenible. Hashtags como #consumoresponsable y #productosverdes se utilizan para etiquetar publicaciones relacionadas con prácticas y productos ecológicos (Salas, 2018).

El análisis de estas publicaciones puede proporcionar información valiosa sobre cómo los consumidores perciben y valoran la sostenibilidad en sus decisiones de compra (Navarro et al., 2017). Investigaciones anteriores han demostrado que las redes sociales pueden influir en las actitudes y comportamientos de los consumidores al proporcionar un espacio para la retroalimentación y la difusión de información (Lagla et al., 2017). Sin embargo, también existen desafíos, como la proliferación de desinformación y la necesidad de interpretar el contexto en el que se expresan las opiniones (Baviera, 2017).

El análisis de sentimientos en estos contextos puede revelar patrones y tendencias en las percepciones del consumo sostenible, así como identificar áreas de preocupación y oportunidades para mejorar las estrategias de comunicación y marketing.

5. Metodología

Para llevar a cabo el análisis de sentimientos sobre el consumo sostenible en la red social X, se debe seguir un procedimiento estructurado para garantizar la correcta recolección y procesamiento de los datos. En primer lugar, es necesario registrarse en la plataforma (<https://www.x.com>) y crear una X App. Este paso permite a los desarrolladores acceder a los contenidos a través de programas. El registro puede demorar unas semanas, ya que debe justificarse y evaluarse el perfil para su aceptación. Una vez aprobado, la plataforma otorga claves y tokens de identificación, necesarios para acceder a la aplicación y extraer la información.

El análisis textual de los tuits que mencionan los hashtags #consumoresponsable, #productosverdes y #sostenibilidad comienza con un conteo de frecuencias de las palabras, lo que permite seleccionar el vocabulario más relevante. Luego, se generará una nube de palabras para visualizar las reacciones de los usuarios sobre el consumo sostenible en México. A continuación, se realizará la minería de opinión, clasificando las emociones de los tuits en categorías: positivo, negativo y neutral.

El vocabulario se seleccionará a través de filtros, eliminando palabras poco frecuentes y preposiciones. Según la teoría de minería de opinión, el investigador podrá decidir el número de palabras a utilizar en el análisis, asegurando que el vocabulario seleccionado represente un porcentaje significativo del total.

Para llevar a cabo este análisis, se empleará el *software* estadístico RStudio, que dispone de una librería especializada para la extracción de datos a través de la API de la plataforma. La librería “rtweet” facilita la conexión con X y la descarga de los tuits. Durante este proceso, se eliminarán enlaces HTTP, emoticonos y números para centrar el análisis únicamente en el contenido textual. Los tuits se almacenarán en un archivo CSV, separados por diagonales y “limpios” de caracteres innecesarios.

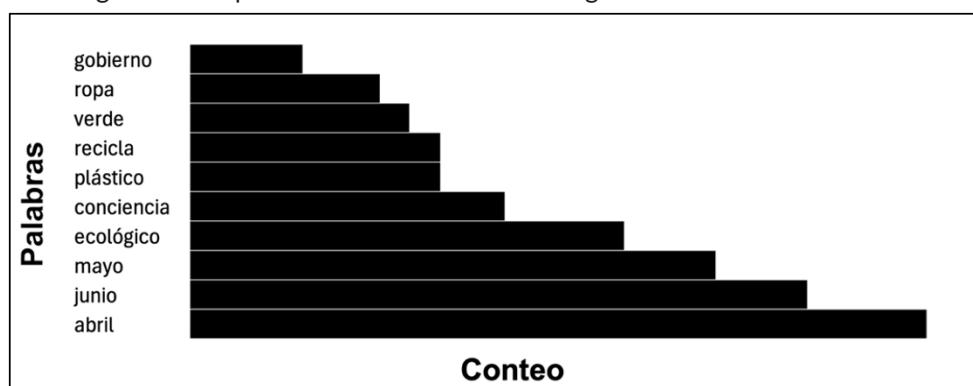
Una vez clasificadas las opiniones, se realizará un análisis estadístico comparando los tuits del segundo trimestre de 2024, con el fin de determinar si existen diferencias significativas en las percepciones sobre el consumo sostenible. El análisis de correspondencia se utilizará para representar los datos en una tabla de contingencia con dos variables cualitativas u ordinales. Se creará un mapa para mostrar la relación y el grado de asociación entre estas variables.

Finalmente, tras la limpieza de los datos recolectados, se aplicará el análisis de sentimientos, utilizando la técnica de los bigramas. Esto permitirá un análisis más detallado de las relaciones entre las palabras y ayudará a entender mejor las actitudes hacia el consumo sostenible.

6. Resultados

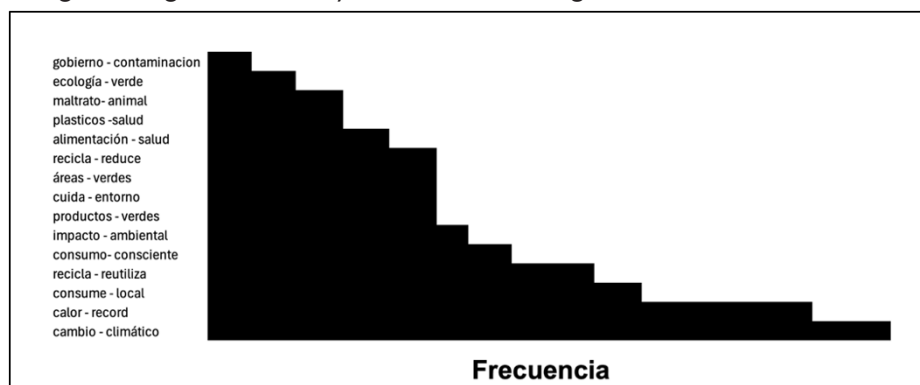
En la figura 1 se muestran las diez palabras más frecuentes utilizadas por los usuarios en el segundo trimestre de 2024 en relación con los hashtags: #consumoresponsable, #productosverdes y #sostenibilidad. Estas palabras reflejan las temáticas y términos que más resuenan entre los usuarios en sus publicaciones sobre consumo sostenible, proporcionando una visión preliminar de las áreas de interés y las percepciones compartidas en redes sociales respecto a estos temas.

Figura 1. Diez palabras más frecuentes del segundo trimestre del año 2024



En la figura 2 se presentan los pares de palabras más frecuentemente utilizados en los tuits realizados durante el segundo trimestre de 2024 en relación con los hashtags: #consumoresponsable, #productosverdes y #sostenibilidad. Estos pares de palabras, o bigramas, proporcionan un contexto más profundo sobre las asociaciones comunes entre términos clave y permiten identificar temas y conceptos que los usuarios tienden a relacionar en sus conversaciones sobre prácticas de consumo sostenible.

Figura 2. Bigramas con mayor frecuencia del segundo trimestre del año 2024



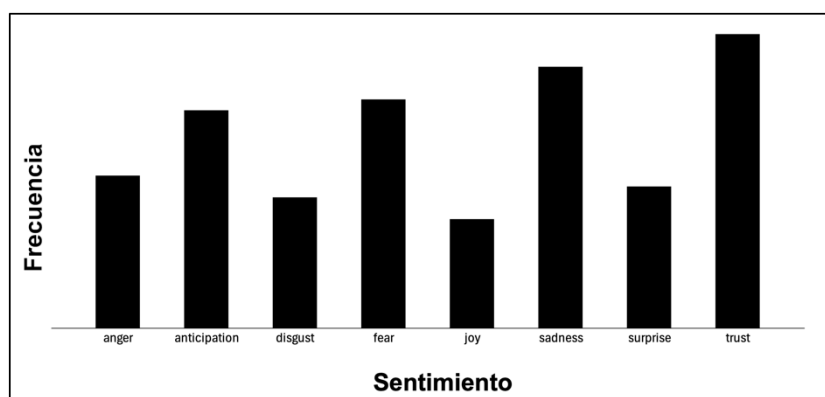
Utilizando gráficos de palabras para visualizar los términos más frecuentes en las publicaciones, se pudo observar que los hashtags más utilizados incluían #consumoresponsable, #productosverdes, y #sostenibilidad, los cuales eran predecibles dado el enfoque de nuestra recolección de datos. Sin embargo, otros términos emergieron con notable frecuencia, tales como #cambioclimático, #ecología, y #responsabilidadsocial, indicando que los usuarios no solo discuten el consumo sostenible de manera aislada, sino que lo relacionan con problemáticas ambientales más amplias y la responsabilidad social corporativa. Estos hallazgos se hacen evidentes en la figura 3.

Figura 3. Nube de palabras para el segundo trimestre del año 2024



En la figura 4 se observa que el código ha clasificado los tuits del mes de abril predominantemente con el sentimiento de “confianza”. Este es el sentimiento que prevalece, sugiriendo una percepción relativamente positiva o de seguridad hacia temas relacionados con el consumo sostenible. El segundo sentimiento con mayor carga emocional es “tristeza”, lo cual podría estar asociado a preocupaciones o frustraciones de los usuarios sobre el impacto ambiental o los desafíos en prácticas sostenibles. En tercer lugar se identifica el sentimiento de “miedo”, reflejando posibles inquietudes respecto a riesgos o consecuencias futuras vinculadas al consumo y sostenibilidad.

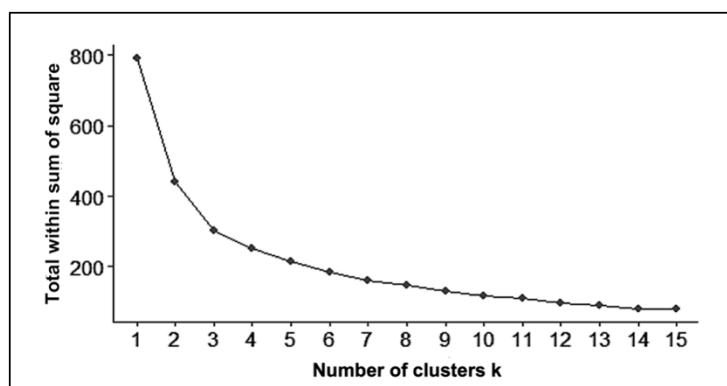
Figura 4. Carga de sentimientos de los tuits durante el segundo trimestre del año 2024



En la figura 5 se muestra una red semántica construida a partir de los tuits publicados en el segundo trimestre del año 2024. En esta red se destaca una alta frecuencia de conexiones entre las palabras “productos” y “sostenibles”, lo que indica una asociación frecuente de estos términos en las conversaciones sobre consumo responsable. Asimismo, se

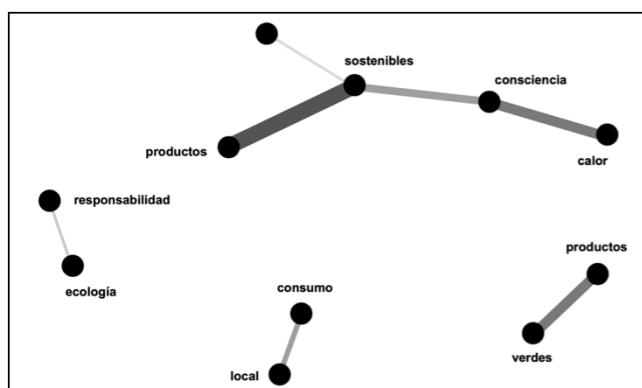
observa una relación significativa entre las palabras “consciencia” y “calor”, posiblemente reflejando un interés en temas de impacto ambiental y climático, donde los usuarios expresan la necesidad de una mayor consciencia frente a las olas de calor registradas en ese mismo año.

Figura 5. Red semántica del segundo trimestre del año



En la figura 6 se identifica que el número óptimo de clústeres para segmentar los datos obtenidos es de 3 o 4. Esto sugiere una estructura en la que los tuits sobre consumo sostenible se agrupan en tres o cuatro temas principales, cada uno representando distintas perspectivas o enfoques en las discusiones. Esta segmentación permite analizar en mayor detalle las tendencias y variaciones en las opiniones de los usuarios sobre prácticas y productos sostenibles.

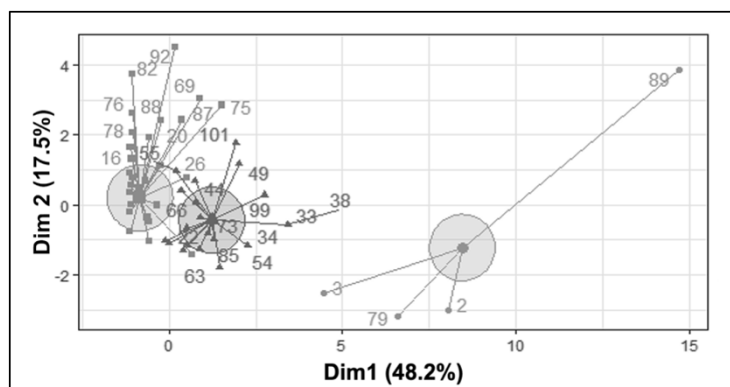
Figura 6. Número óptimo de clústeres para el segundo trimestre del año 2024



En la figura 7, tras aplicar la técnica de agrupamiento k-means, se observa la formación de tres clústeres que se alinean con la recomendación del modelo en la Figura 6. Esta agrupación confirma que los datos se estructuran naturalmente en tres grupos; lo cual

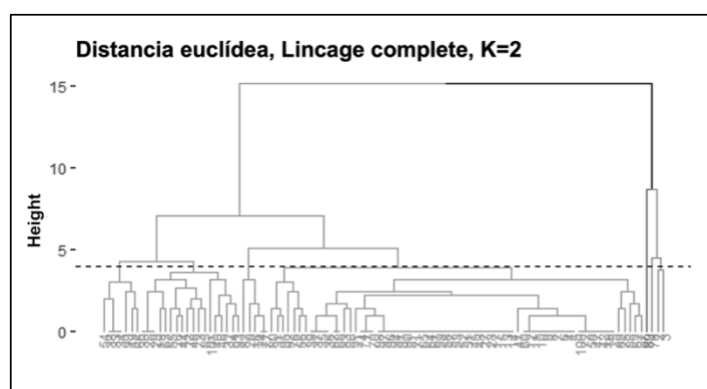
facilita un análisis más detallado de los temas predominantes y las distintas perspectivas de los usuarios en torno al consumo sostenible.

Figura 7. Resultado de clústeres a través de k-means



En la figura 8 se presenta un clúster jerárquico que muestra la formación de diversos grupos. Esta estructura permite visualizar las relaciones entre los temas de manera más detallada, revelando subgrupos que comparten similitudes en los datos. La organización jerárquica ayuda a identificar niveles de afinidad entre las palabras y temas, ofreciendo una perspectiva más profunda sobre cómo se conectan y agrupan los conceptos en torno al consumo sostenible. Figura 8. Clúster Jerárquico para el segundo trimestre del año 2024

Figura 8. Clúster Jerárquico para el segundo trimestre del año 2024



7. Conclusiones

A partir del análisis de los datos del segundo trimestre de 2024, se identificaron palabras y combinaciones de términos frecuentes en las publicaciones sobre consumo sostenible. Estos términos permiten a las marcas identificar los temas que resuenan con

los consumidores, revelando no sólo un interés en el consumo responsable, sino también en problemáticas ambientales más amplias, como el cambio climático y la responsabilidad social. Al entender estas asociaciones y preferencias, las empresas pueden ajustar su comunicación y estrategias de *marketing* para alinearse mejor con los intereses y valores de su audiencia.

El análisis de sentimientos aporta una perspectiva valiosa sobre la percepción de los consumidores respecto al consumo sostenible. La predominancia de sentimientos como “confianza” refleja una receptividad positiva, que puede ser aprovechada por marcas que buscan proyectarse como líderes en sostenibilidad. Por otro lado, los sentimientos “tristeza” y “miedo” indican que existe también una preocupación significativa entre los consumidores sobre el impacto ambiental y los desafíos que enfrenta el consumo sostenible, lo cual sugiere una oportunidad para que las empresas aborden estos temas en sus mensajes, reforzando el valor de sus productos en términos de sostenibilidad y mitigación de riesgos ambientales.

El análisis de clúster mostró una estructura temática organizada en tres o cuatro grupos, lo que permite segmentar y adaptar las estrategias de marketing de acuerdo con las preferencias específicas de cada subgrupo de consumidores. Estos clústeres ayudan a las empresas a identificar patrones y conectar mensajes clave con los diferentes intereses y preocupaciones de los usuarios. Además, la segmentación jerárquica permite detectar subgrupos y relaciones entre temas, facilitando el desarrollo de campañas dirigidas que aborden los aspectos de sostenibilidad de una manera que resuenen de forma auténtica y efectiva con distintos segmentos del mercado.

Las implicaciones de los resultados obtenidos en el análisis de consumo sostenible en la red social X son cruciales para diversas industrias, especialmente para aquellas que buscan destacar en sostenibilidad y adaptar sus estrategias de marketing. La predominancia de sentimientos positivos entre los usuarios indica un mercado receptivo y dispuesto a apoyar productos y servicios que promuevan prácticas responsables con el medio ambiente.

Desde una perspectiva de marketing, esto sugiere que las empresas que invierten en sostenibilidad no solo pueden mejorar su reputación, sino también captar una base de clientes leales que valoran el impacto ambiental de sus decisiones de compra. Esta tendencia resalta la importancia de desarrollar campañas de *marketing* que destaquen los compromisos genuinos con la sostenibilidad, utilizando mensajes que conecten emocionalmente con los consumidores.

En conjunto, la minería de textos y el análisis de sentimientos ofrecen a las empresas de consumo sostenible herramientas para comprender y anticipar las percepciones y actitudes de su audiencia, optimizando así sus campañas de marketing y reforzando el valor de sus productos desde una perspectiva de responsabilidad ambiental y social.

Fuentes de información

- Anturi-Martínez, J., Paz-Realpe, J., Timaná-Peña, J., & Cobos, C. (2019). Clasificadores para el análisis de sentimientos en Twitter: Una revisión. *Revista Ibérica de Sistemas y Tecnologías de Información*, E23, 353–366.
- Baviera, T. (2017). Técnicas para el análisis del sentimiento en Twitter: Aprendizaje automático supervisado y SentiStrength. *Dígitos*, 1(3), 33–50.
- Hara, K., Saito, D., & Shouno, H. (2015). Analysis of function of rectified linear unit used in deep learning. En 2015 International Joint Conference on Neural Networks (IJCNN) (pp. 1–8). IEEE.
- Hernández, M., Hernández-Cordero, S., Unar-Munguía, M., Gómez-Arias, W., Lozano-Hidalgo, E., & Teruel-Belismelis, G. (2024). La red social X como escenario para el debate político y propagandístico en torno a los productos comestibles ultraprocesados en México: Un análisis de redes sociales. *Revista Economía y Política*, 40, 58–77.
- John, T. (2021). Análisis de opinión sobre tuits del COVID-19 generados por usuarios ecuatorianos. *CEDAMAZ Revista del Centro de Estudio y Desarrollo de la Amazonia*, 70–77.
- Lagla, G., Chisag, J., Moreano, J., Pico, O., & Pulloquina, R. (2017). La influencia de las redes sociales en los estudiantes universitarios. *Boletín Redipe*, 6(4), 56–65.
- Martínez, P., Aguilar, J., Zezzatti, C., Franco, E., & Gallegos, J. (2019). Determinación de criminales potenciales en análisis de textos: Caso de estudio. *Programación Matemática y Software*, 11(1), 9–14.
- Muente, G. (2019, septiembre 14). Nube de palabras: Qué es y para qué sirve. *Rock Content*. <https://rockcontent.com/es/blog/nube-de-palabras/>
- Navarro, M., Fleites, G., Santos, J., Zepeda, H., & Moreno, R. (2017). Minería de datos aplicada al pensamiento ambiental universitario, en la oferta educativa de posgrados 2017 de la BUAP. *Innovaciones Tecnológicas en las Ciencias Computacionales*, 9.
- Rodríguez, V. (2014). Las redes sociales y su incidencia en la sociedad actual. *AEQUITAS-Virtual Facultad de Ciencias Jurídicas*, 8(21).
- Salas, H. (2018). Marketing ecológico: La creciente preocupación empresarial por la protección del medio ambiente. *Fides et Ratio - Revista de Difusión Cultural y Científica de la Universidad La Salle en Bolivia*, 15(15), 151–170.
- Sanahuja, J., & Tezanos-Vázquez, S. (2017). Del milenio a la sostenibilidad: Retos y perspectivas de la Agenda 2030 para el desarrollo sostenible. *Política y Sociedad*, 54(2), 533–555. <https://doi.org/10.5209/POSO.519>
- Sancho, A. (2023). Clasificación de textos con redes neuronales para detección de hablantes en una conversación (Proyecto Fin de Carrera/Trabajo Fin de Grado). E.T.S.I. de Sistemas Informáticos, Universidad Politécnica de Madrid, Madrid.

- Shum, Y. (2021, febrero 7). Internet, digital, redes sociales y media mundial. Yimin Shum.
<https://yiminshum.com/internet-social-media-mundo-2021/>
- Valero, A. (2017). Técnicas estadísticas en minería de textos (Trabajo Fin de Grado inédito).
Universidad de Sevilla, Sevilla, España.