




## Análisis de sentimientos utilizando ChatGPT: Una Revisión Sistemática de la Literatura

Reducindo J. <sup>1,\*</sup> 

[jhoan.reducindo@unmsm.edu.pe](mailto:jhoan.reducindo@unmsm.edu.pe)

Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional Mayor de San Marcos

Calero H. <sup>2</sup> 

[harrinzon.calero@unmsm.edu.pe](mailto:harrinzon.calero@unmsm.edu.pe)

Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional Mayor de San Marcos

Fernández C. <sup>3</sup> 

[cristian.fernandez@unmsm.edu.pe](mailto:cristian.fernandez@unmsm.edu.pe)

Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional Mayor de San Marcos

Ramos E. <sup>4</sup> 

[nelly.ramos@unmsm.edu.pe](mailto:nelly.ramos@unmsm.edu.pe)

Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional Mayor de San Marcos

### Resumen:

Este artículo realiza una revisión sistemática de la literatura sobre el análisis de sentimientos utilizando ChatGPT, con el objetivo de explorar exhaustivamente las técnicas, avances y desafíos en este campo. La metodología de investigación empleada se estructura en tres etapas: planificación, ejecución y análisis de resultados. En el proceso de revisión de literatura, se formularon preguntas clave y se llevó a cabo una búsqueda exhaustiva en motores académicos, seleccionando artículos relevantes entre 2021 y 2024. Además, se desarrolló una aplicación práctica para demostrar cómo ChatGPT clasifica opiniones de usuarios por tópicos y proporciona un análisis con gráficos interactivos de porcentajes, mostrando insights valiosos. Por último se concluye que ChatGPT no solo es un modelo de lenguaje grande (LLM), sino también una herramienta valiosa y complementaria para el análisis de sentimientos, capaz de comprender y generar texto contextual en múltiples idiomas, lo que lo convierte en una opción prometedora en este campo.

Palabras Claves: ChatGPT, Análisis de sentimientos, Grandes modelos de Lenguaje (LLM), NLP

**Abstract:**

This article conducts a systematic review of the literature on sentiment analysis using ChatGPT, with the aim of exhaustively exploring the techniques, advances and challenges in this field. The research methodology used is structured in three stages: planning, execution and analysis of results. In the literature review process, key questions were asked and an exhaustive search was carried out in academic engines, selecting relevant articles between 2021 and 2024. In addition, a practical application was developed to demonstrate how ChatGPT classifies user opinions by topics and provides analysis with interactive percentage graphs, showing valuable insights. Finally, it is concluded that ChatGPT is not only a large language model (LLM), but also a valuable and complementary tool for sentiment analysis, capable of understanding and generating contextual text in multiple languages, making it a promising option. in this field.

**Keywords:** ChatGPT, Sentiment Analysis, Large language models (LLM), NLP



## 1. Introducción

Los modelos de IA generativa han sido parte de la inteligencia artificial durante mucho tiempo, con innovaciones tempranas como los Modelos Ocultos de Markov (HMMs) y los Modelos de Mezcla Gaussiana (GMMs) que se originaron en la década de 1950. La introducción de Chat GPT por OpenAI en noviembre de 2022 ha incrementado notablemente la atención y la adopción global de la IA generativa [8]. Esto resalta la rápida aceptación de ChatGPT, además subraya el creciente interés y la importancia de la IA generativa en la actualidad en los distintos campos de estudio.

Asimismo, se puede decir que el análisis de sentimientos se centra en identificar los sentimientos, emociones u opiniones expresadas en textos, como publicaciones en redes sociales, conversaciones o reseñas de usuarios [5, 4]. A su vez, el análisis de sentimientos se ha convertido en una herramienta fundamental para entender cómo funciona el mercado y anticipar las tendencias que pueden surgir. Evaluar los sentimientos permite obtener insights valiosos sobre el ánimo general del mercado, lo cual facilita tomar decisiones estratégicas y bien fundamentadas. Además, las capacidades de conversación de Chat GPT pueden emplearse para transmitir y clarificar descubrimientos detallados en análisis de sentimientos tanto a usuarios con experiencia como a principiantes, facilitando el acceso y la aplicabilidad de los estudios en diverso campos [4]. Por lo que utilizar esta herramienta para una investigación orientada al análisis de sentimientos puede representar una oportunidad para muchas empresas que buscan conocer las reacciones a sus productos y/o servicios.

Además, los modelos generativos de IA, como ChatGPT y su capacidad para comprender y generar texto contextual, han cambiado fundamentalmente la forma en que interactuamos con la tecnología. Estos modelos no solo interpretan el lenguaje humano, sino que también son capaces de generar respuestas contextuales coherentes, lo que los hace ideales para aplicaciones que requieren generación y comprensión de texto natural. En los últimos años, los modelos de lenguajes grandes (LLM, por sus siglas en inglés) como ChatGPT han demostrado ser efectivos en tareas multilingües y análisis de sentimientos [5]. El objetivo de este documento es explorar exhaustivamente cómo se aborda la tarea de análisis de sentimientos en la actualidad, además de brindar una idea clara de cómo realizar el análisis de sentimientos utilizando ChatGPT para proporcionar información precisa y relevante para el análisis de sentimientos.

En este trabajo de investigación, se realiza un análisis exhaustivo de la literatura existente sobre métodos y modelos existentes para resolver problemas de análisis de sentimientos. Además, se presenta una aplicación práctica de análisis de sentimientos, la cual clasifica las opiniones de los usuarios por tópicos, como positivas, negativas o neutras y proporciona un porcentaje adecuado para cada tópico. Finalmente, se enfatiza que estos avances brindan oportunidades futuras para expandir aún más las capacidades de análisis de sentimiento.

## 2. Metodología

Por ser un trabajo académico que obedece a cumplir el objetivo de cierre del curso, la propuesta de artículo puede ajustarse a un promedio de 12 a 20 páginas. La metodología para este trabajo es descriptiva basada en la revisión sistemática de artículos científicos.

La metodología para llevar a cabo la revisión sistemática se basa en el trabajo de Kitchenham & Brereton [3], los cuales



ECT

Revista Científica Emprendimiento  
Científico Tecnológico

ISSN: 2810 – 8493

<https://revista.ectperu.org.pe/index.php/ect/index>

proponen las siguientes tres etapas:



- **Planificación:** Se plantean las interrogantes relacionadas con la investigación, se determinan las palabras claves que se usarán durante la búsqueda de artículos. Se establecen los criterios que se utilizarán para incluir o excluir los artículos de investigación.
- **Desarrollo:** Se determina qué artículos cumplen con los criterios establecidos. Se recopilan las técnicas y métodos para determinar cuáles son las más adecuadas para abordar la problemática.
- **Resultados:** Se extraen y analizan los datos relacionados con la información recopilada durante la fase de desarrollo.

### 3. Desarrollo

#### 3.1. Formulación de las preguntas de investigación

A continuación se presenta la formulación de las preguntas de investigación:

**Pregunta de investigación 01:** ¿Qué técnicas y métodos se utilizan actualmente para realizar análisis de sentimiento?

**Pregunta de investigación 02:** ¿Cuáles son los principales avances en el análisis de sentimiento utilizando modelos de lenguaje como ChatGPT?

**Pregunta de investigación 03:** ¿Cómo se compara el desempeño de ChatGPT con otros modelos de lenguaje en la tarea de análisis de sentimiento?

#### 3.2. Bases de datos y cadenas de búsqueda utilizadas

Fuente	Cadena de búsqueda
Scopus	(TITLE-ABS-KEY(ChatGPT) AND TITLE-ABS-KEY (Analysis) AND TITLE-ABS-KEY (Sentiment)) AND ( LIMIT-TO ( DOCTYPE , "ar" ). Range: 2021-2024.
ScienceDirect	ChatGPT With Analysis Sentiments
Web of science	TS=("ChatGPT" AND "Sentiment" AND "Analysis"). Range: 2021-2024. Document type: Article
Proquest	GPT AND analysis AND sentiment. Últimos 5 años
Wiley	GPT AND analysis AND sentiment. Últimos 5 años



ECT

Revista Científica Emprendimiento  
Científico Tecnológico

ISSN: 2810 – 8493

<https://revista.ectperu.org.pe/index.php/ect/index>

### 3.3. Criterios de inclusión y exclusión



Criterios de inclusión	Criterios de exclusión
<ul style="list-style-type: none"> <li>• Artículos que sean de acceso abierto</li> <li>• Abordan una o más interrogantes de la investigación propuesta.</li> <li>• Artículos publicados en revistas científicas revisadas por pares, de revistas o conferencias.</li> <li>• Artículos comprendidos entre los años 2021-2024.</li> </ul>	<ul style="list-style-type: none"> <li>• Artículos que tengan antigüedad mayor a 5 años.</li> <li>• Artículos escritos en idiomas diferentes al inglés o español</li> <li>• Artículos que no mencionan específicamente a Chat GPT o modelos de lenguaje similares en el contexto del análisis de sentimientos.</li> </ul>

**3.4. Resultados de la investigación**

ID	Título del Artículo	Año Publicación	Metodología y/o Técnica	Resultados obtenidos
01	[16] Generative approach to Aspect Based Sentiment Analysis with GPT Language Models	2023	Se utilizaron datos en ruso etiquetados automáticamente con edición manual y datos en inglés etiquetados automáticamente. Luego se aplicaron modelos GPT en modos few-shot y fine-tuning para realizar la extracción de tripletes de sentimiento basados en aspectos (ASTE) en dominios abiertos. Y finalmente se compararon las estrategias de few-shot y fine-tuning.	En textos en ruso, el modelo OpenAI GPT-3 text-davinci-003 ha mostrado los mejores resultados. Al ser el modelo de generación de instrucciones más grande, puede realizar extracción de tripletes en modo de pocos disparos y lograr un F1-score de 0,52 en datos de prueba de dominio mixto. El ajuste fino de GPT-2 en un conjunto de entrenamiento más grande ha mostrado excelentes resultados con un punto de control grande GPT-2 ajustado que alcanza un F1-score de 0,74 y una base GPT-2 más pequeña que alcanza una puntuación F1 de 0,71.



<b>02</b>	[3] Sentiment analysis of COP9-related	<b>2024</b>	Se realizó una evaluación en dos etapas. En la primera etapa, se compararon varias técnicas de	En la primera etapa, BERT logró el F1-score más alto (IMDB: 0.9380, Sentiment140: 0.8114), seguido de
-----------	--	-------------	--	---





	tweets: a comparative study of pre-trained models and traditional techniques		análisis de sentimiento (basadas en léxico, aprendizaje automático, Bi-LSTM, BERT y GPT-3) en conjuntos de datos estándar IMDB y Sentiment140 utilizando métricas de evaluación estándar como exactitud, puntuación F1 y precisión. En la segunda etapa, se aplicaron las técnicas de mejor rendimiento de la primera etapa a tweets relacionados con la conferencia COP9 parcialmente anotados.	GPT-3 (IMDB: 0.9119, Sentiment140: 0.7913) y Bi-LSTM (IMDB: 0.8971, Sentiment140: 0.7778). En la segunda etapa, GPT-3 tuvo el mejor desempeño en el análisis de sentimiento en tweets relacionados con la conferencia COP9 parcialmente anotados, con un F1-score de 0,8812.
03	[10] LLMs and NLP Models in Cryptocurrency Sentiment Analysis: A Comparative Classification Study	2024	Este artículo utiliza un conjunto de datos de 31.037 filas y seis columnas, sobre Crypto News obtenido de Kaggle. Posteriormente se le realiza un preprocesamiento. Para el entrenamiento de BERT y FinBERT se seleccionó aleatoriamente 5 mil filas. Luego en el caso de GPT-4, primero se generó el ajuste del modelo, segundo la predicción, tercero la comparación con las etiquetas de los datos y finalmente se integró los resultados, mostrando sus métricas (exactitud, precisión, recuperación y F1-score). En el caso de BERT y FinBert fue similar excepto que en vez de ajustar el modelo se le entrenó con una muestra de las filas seleccionadas.	GPT-4 obtiene el mejor desempeño con una precisión de 86.7%, seguido por FinBert con 84.3% y BERT con 83.3%. En conclusión este estudio muestra que GPT-4 es el más apto para interpretar y categorizar con precisión los sentimientos extraídos de artículos de noticias de criptomonedas.



<b>04</b>	[7] Assessing Look-Ahead Bias in Stock Return Predictions Generated By GPT Sentiment	<b>2023</b>	Este estudio cuenta con dos conjuntos de titulares de noticias sobre acciones de empresas obtenidos de la web tomando como guía RavenPack y de los datos de Thomson Reuters respectivamente.	La anonimización mejora el rendimiento de una estrategia comercial impulsada por los resultados de análisis de sentimiento de GPT 3.5 a titulares de noticias de acciones. El efecto de distracción interfiere negativamente en
-----------	--	-------------	--	---



	Analysis		Luego se realiza la anonimización de dichos titulares cambiando el nombre de la empresa por “Blahblahblah”, además de los productos relevantes que también son reemplazados. Y finalmente se realiza el análisis de sentimiento con GPT 3.5 Turbo con el mensaje de que olvide todas sus instrucciones anteriores y evalúe las noticias como un experto financiero.	GPT 3.5 en su trabajo de evaluar el sentimiento de las noticias. En conclusión la anonimización mejora el análisis de sentimiento de GPT tanto fuera como dentro de la muestra.
05	[11] A Comparison of ChatGPT and Fine-Tuned Open Pre-Trained Transformers (OPT) Against Widely Used Sentiment Analysis Tools: Sentiment Analysis of COVID-19 Survey Data	2023	Este estudio usa dos conjuntos de datos etiquetados como positivo, negativo o neutral por expertos. Luego se realizó el análisis de sentimiento con ocho herramientas ampliamente utilizadas en el ámbito de la salud. Adicionalmente se usó Few-shot learning, OPT y Zero-shot learning, GPT 3.5. Finalmente se realizó la comparación entre todas las herramientas.	Los que más destacaron fueron ChatGPT y OPT. Pero ChatGPT obtuvo mejores resultados a comparación de los otros modelos. En el conjunto de datos de NIH, ChatGPT superó a OPT en un 6% en precisión y un 7% en F-score. De igual manera en el conjunto de Stanford superó con una precisión de 6% y un F-score de 4%.



<b>06</b>	[15] ChatGPT and finetuned BERT: A comparative study for developing intelligent design support systems	<b>2024</b>	Los autores nos muestran una comparación entre el modelo ChatGPT y un modelo BERT en la clasificación y generación. La metodología del estudio incluyó la selección de los modelos BERT y ChatGPT, seguido de la creación y preprocesamiento de un conjunto de datos etiquetados específicos del dominio. Luego, se ajustó el modelo BERT y se evaluó el rendimiento de ChatGPT con y sin indicaciones específicas. Finalmente, se realizó	Los resultados obtenidos por los autores muestran que ChatGPT es comparable al modelo BERT en la tarea de clasificación a nivel de oraciones, pero tiene dificultades en secuencias cortas, además el modelo Bert muestra mejores resultados a nivel de mapeo de definiciones y de clasificación, requiere acceso a conjunto de datos etiquetados de mayor volumen mientras GPT requiere menor volumen de datos para la personalización. Por otro lado, ChatGPT, con aprendizaje en contexto
-----------	--	-------------	--	--



			una comparación del rendimiento de ambos modelos en tareas de clasificación y generación de texto mediante pruebas de secuencias y análisis de resultados.	y prompts personalizados, es efectivo en clasificación con pocos ejemplos y generación de contenido informativo.
07	[4] Enhancing Restaurant Management through Aspect-Based Sentiment Analysis and NLP Techniques.	2024	Este estudio presenta una metodología para extraer y analizar el sentimiento de los clientes a través de reseñas online. Además se utiliza el Análisis de Sentimientos Basado en Aspectos (ABSA) mediante técnicas avanzadas de Procesamiento de Lenguaje Natural (PLN), destacando el uso de modelos como BART y DeBERTa para la extracción y clasificación de once atributos específicos. Por otra parte, se emplea el modelo ChatGPT en sus versiones 3.5 y 4.0 para discernir y clasificar estos atributos dentro de las reseñas completas, aprovechando su capacidad para generar respuestas humanas simuladas.	Los resultados destacan que ChatGPT, especialmente la versión 4.0, logró el F1 Score más alto comparado con otros modelos en la clasificación automatizada de atributos y sentimientos en reseñas de restaurantes. Este rendimiento superior sugiere que ChatGPT puede ser una herramienta efectiva para analizar y categorizar la retroalimentación de los clientes, proporcionando insights valiosos para la gestión de restaurantes y la optimización de la satisfacción del cliente.



<p><b>08</b></p>	<p>[17] Mitigating Class Imbalance in Sentiment Analysis through GPT-3-Generated Synthetic Sentences</p>	<p><b>2024</b></p>	<p>En este estudio de clasificación de sentimientos, se aplicó una metodología exhaustiva que comenzó con el preprocesamiento de datos, incluyendo la eliminación de palabras no inglesas y la generación de revisiones sintéticas mediante GPT-3 para balancear el conjunto de datos. Se emplearon nueve modelos estándar que abarcan tanto el aprendizaje automático como el aprendizaje profundo, para realizar la tarea de análisis de sentimientos. La evaluación se centró en métricas como la Exactitud y el F1-score,</p>	<p>Los resultados del estudio mostraron que al generar revisiones sintéticas de alta calidad utilizando el modelo Davinci de GPT-3 con ajuste fino, se mejoró significativamente la capacidad de clasificación de sentimientos. Por otro lado el modelo Naïve Bayes Multinomial mostró la mayor precisión de 75.12% en este conjunto de datos. Estos resultados muestran que el modelo GPT-3 tiene potencial para mitigar el desequilibrio de datos en el análisis de sentimientos.</p>
------------------	--	--------------------	---	---



			evidenciando mejoras significativas en la precisión al utilizar datos balanceados.	
<b>09</b>	[6] Efficacy of ChatGPT in Cantonese Sentiment Analysis: Comparative Study	<b>2023</b>	Los autores exploran la eficacia de Chat GPT en el análisis de sentimientos en cantonés, comparando su rendimiento con métodos tradicionales basados en léxicos y enfoques de aprendizaje automático en el contexto de asesoramiento en línea. El estudio analizó transcripciones de un servicio de asesoramiento en Hong Kong.	El estudio propuesto por los autores encontró que las variantes GPT-3.5 y GPT-4 de ChatGPT lograron una precisión superior al 90% en la detección de sentimientos en textos en cantonés, superando significativamente a los métodos basados en lexicones y a los modelos de aprendizaje automático ajustados.
<b>10</b>	[5] Transforming sentiment analysis in the financial domain with chatGPT	<b>2023</b>	Los autores realizan una evaluación del sentimiento en el dominio financiero mediante el uso de ChatGPT. El estudio se centra en evaluar las capacidades de ChatGPT 3.5 para el análisis de sentimientos en noticias. Este enfoque ayuda a las instituciones financieras a mejorar la toma de decisiones en estrategias de inversión, gestión de riesgos y optimización de carteras, buscando potencialmente mayores rendimientos.	Los resultados obtenidos, muestran que el modelo GPT-P4N obtuvo una precisión del 0.784 en la clasificación de sentimientos en el dominio financiero. La tasa de precisión, recall y F1-score fueron evaluadas en 0.804, 0.784 y 0.789 respectivamente. Además, se observó que el modelo GPT-P6N logró una precisión del 0.652 en la predicción de la dirección del mercado.



11	[1] A Wide Evaluation of ChatGPT on Affective Computing Tasks	2024	El artículo ofrece un estudio exhaustivo de los modelos ChatGPT (GPT-3.5 y GPT-4) en 13 tareas de computación afectiva. Estas tareas incluyen la extracción de aspectos, la clasificación de la polaridad de los aspectos, la extracción de opiniones, el análisis de sentimientos, la clasificación de la intensidad de los sentimientos, la clasificación de la	Los resultados demuestran que, aunque los modelos ChatGPT sobresalen en tareas relacionadas con los sentimientos, especialmente en las que implican emociones negativas y detección de toxicidad, tienen dificultades en tareas que requieren la detección de señales implícitas, como la medición del compromiso, la evaluación de la personalidad y la
----	---	------	---	--





			intensidad de las emociones, la detección de la tendencia al suicidio, la detección de la toxicidad, la evaluación del bienestar, la medición del compromiso, la evaluación de la personalidad, la detección del sarcasmo y la detección de la subjetividad.	detección del sarcasmo. En general, GPT-4 supera a GPT-3.5, y ambos modelos obtienen resultados comparables o mejores que los modelos tradicionales de PNL en escenarios específicos. Sin embargo, el modelo RoBERTa suele superar a ambos modelos ChatGPT en muchas tareas debido a su ajuste especializado.
12	[20] Can ChatGPT Be Served as the Sentiment Expert? An Evaluation of ChatGPT on Sentiment and Metaphor Analysis	2024	El estudio evalúa el potencial de ChatGPT para el análisis de sentimientos y la comprensión de metáforas. La metodología se centra en cuatro tareas representativas de análisis de sentimientos: análisis de sentimientos de subjetividad, análisis de sentimientos a nivel de aspecto, detección de humor, reconocimiento de metáforas. Para llevar a cabo estas tareas, el estudio empleó cinco conjuntos de datos ampliamente evaluados en el análisis de sentimientos en chino: SMP2020 (Usual y Virus), SMP2020 (Humor), ASAP y Metaphor. Los modelos de vanguardia utilizados para la comparación incluye CMCNN, Bi-LSTM+Attention, CapsNet-BERT, entre otros. Los resultados se midieron usando precisión, recall, micro-F1 y exactitud.	Los resultados del estudio muestran que ChatGPT tiene un desempeño comparable e incluso superior en algunas tareas de análisis de sentimientos cuando se le compara con otros modelos de última generación. En los conjuntos de datos Usual y Virus para el análisis de subjetividad, ChatGPT logró un Macro-F1 de 82.16% y 80.20%, respectivamente, superando a varios modelos tradicionales como Bi-LSTM+Attention y DMM-CNN. En tareas de análisis de aspectos, aunque ChatGPT mostró una buena capacidad de identificación y captura de información contextual, no superó a modelos especializados como ACSA-generation y AC-MIMLLN. En el reconocimiento de metáforas, ChatGPT superó a varios modelos basados en RoBERTa y XLNet con un Macro-F1 de 85.71%. Sin embargo, en la tarea de detección de humor, ChatGPT se quedó rezagado respecto a modelos como BSI e IASPS, indicando que aún tiene margen de mejora en la comprensión y detección del humor.



13	[8] Exploring the	2023	El estudio "Exploring the Flexibility	Los resultados del estudio demostraron
----	-------------------	------	---------------------------------------	--



<p>Flexibility and Accuracy of Sentiment Scoring Models through a Hybrid KNN-RNN-CNN Algorithm and ChatGPT</p>	<p>and Accuracy of Sentiment Scoring Models through a Hybrid KNN-RNN-CNN Algorithm and ChatGPT" desarrolló un modelo de puntuación de sentimientos combinando algoritmos KNN, RNN y CNN junto con ChatGPT. Se seleccionó el tema del cambio climático, y se recopilaron datos de diversas fuentes, incluyendo artículos y ensayos existentes, para crear un conjunto de datos de entrenamiento. Se evaluaron y compararon los algoritmos KNN, CNN y RNN, tanto de manera independiente como combinada, ajustándolos para garantizar una evaluación precisa de los sentimientos en los textos. Los ensayos de estudiantes se utilizaron como entrada, y las oraciones generadas por ChatGPT se compararon y evaluaron utilizando estos algoritmos de análisis de sentimientos.</p>	<p>que la combinación de los algoritmos KNN, RNN y CNN con ChatGPT logró una precisión impresionante del 88.17% en el análisis de sentimientos. Individualmente, el algoritmo KNN obtuvo una precisión del 83.24% con una medida F de 0.82, el RNN logró una precisión del 85.68% con una medida F de 0.84, y el CNN alcanzó una precisión del 87.92% con una medida F de 0.87. La combinación de estos algoritmos permitió superar las limitaciones individuales de cada uno, proporcionando un análisis de sentimientos preciso para textos complejos. El KNN clasifica datos con características similares, el RNN procesa datos secuenciales reteniendo información a nivel de palabra, y el CNN reconoce patrones complejos en textos.</p>
--	---	---



<b>14</b>	[19] From Big to Small Without Losing It All: Text Augmentation with ChatGPT for Efficient Sentiment Analysis	<b>2023</b>	El estudio empleó un enfoque innovador para mejorar la eficiencia y rendimiento de modelos de análisis de sentimientos a través de la aumentación de datos utilizando ChatGPT. La metodología se centró en generar datos sintéticos para entrenar modelos más pequeños y menos intensivos en recursos, haciéndolos competitivos con sus contrapartes más grandes. El proceso de aumentación de datos se realizó utilizando el modelo GPT-3.5 de OpenAI a través de su API,	Los resultados mostraron que el uso de datos aumentados generados por ChatGPT mejoró significativamente el rendimiento de los modelos de análisis de sentimientos. En el conjunto de datos PerSenT, los modelos aumentados lograron mejores métricas de precisión y F1 macro en comparación con los modelos base. Por ejemplo, RoBERTa-small mostró una mejora del F1 macro del 36% al 40% y en precisión del 38% al 41% cuando se utilizaron datos aumentados. Para RoBERTa-base, la precisión aumentó
-----------	---	-------------	--	---



		<p>aplicando dos estrategias principales: parafraseo y generación inspiracional. En el parafraseo, se generaron representaciones variadas de textos originales para mantener la relevancia contextual. En la generación inspiracional, se creó contenido completamente nuevo basado en el tema original pero manteniendo el sentimiento, lo cual amplió el alcance de los datos. Se utilizaron cuatro prompts específicos para cubrir estas estrategias, aplicados a muestras de dos conjuntos de datos de análisis de sentimientos: PerSenT y MultiEmo. Los modelos entrenados incluyeron RoBERTa-small, RoBERTa-base y XtremeDistil, ajustados con combinaciones de datos originales y aumentados, evaluando su rendimiento en términos de precisión y F1 macro.</p>	<p>del 39% al 46% y el F1 macro del 38% al 43%. El modelo XtremeDistil, siendo el más eficiente en términos de recursos, también mostró mejoras notables con la precisión aumentando del 43% al 46%. En el conjunto de datos MultiEmo, los resultados fueron igualmente positivos, con mejoras en precisión y F1 macro para todos los modelos. RoBERTa-small, por ejemplo, mejoró su precisión del 78% al 85% y su F1 macro del 78% al 84%.</p>
--	--	--	---



<b>15</b>	[14] Using Large Language Models to Improve Sentiment Analysis in Latvian Language	<b>2024</b>	<p>El estudio empleó un enfoque innovador utilizando modelos de lenguaje de gran tamaño (LLMs) y la ingeniería de prompts para mejorar el análisis de sentimientos en letón. La metodología consistió en crear un nuevo conjunto de datos, LVReddit, utilizando datos de Reddit en letón. Para la recopilación de datos, se utilizó la API de Pushshift.io debido a las limitaciones de la API oficial de Reddit. Posteriormente, se aplicó el paquete langdetect para filtrar los posts en letón. La ingeniería de prompts se realizó iterativamente, desarrollando</p>	<p>El mejor prompt alcanzó una precisión del 82.0% en el conjunto de datos de validación, superando significativamente los resultados anteriores en análisis de sentimientos para el idioma letón. En comparación, los métodos basados en lexicones lograron una precisión del 43.0%. La precisión del mejor prompt en el conjunto de datos etiquetados manualmente fue del 70.4%, mientras que la precisión entre los anotadores humanos fue del 74.0%. Los resultados en otros conjuntos de datos existentes también mostraron mejoras notables. Por ejemplo, en el conjunto de datos</p>
-----------	--	-------------	--	---



			<p>y evaluando prompts en inglés y letón para el modelo GPT-3.5-turbo. Se probaron 24 prompts diferentes, utilizando una parte del conjunto de datos letón de tweets para la validación. Los datos se etiquetaron automáticamente mediante la API de OpenAI, y las respuestas que no podían ser interpretadas correctamente se clasificaron como "Neutral". Para la validación, se creó un subconjunto de datos etiquetados manualmente por dos anotadores humanos. Finalmente, se realizaron experimentos con diferentes estrategias de ingeniería de prompts para maximizar la precisión del modelo en la clasificación de sentimientos.</p>	<p>latvian-tweet-sentiment-corpus, se logró una precisión del 82.0%, comparado con el 60.4% obtenido con métodos anteriores.</p>
<b>16</b>	[18] Applying BERT and ChatGPT for Sentiment	<b>2023</b>	<p>La metodología usada para el análisis de sentimientos consta de 6 pasos: Recolección de datos de artículos científicos, limpieza de datos, selección de modelo, clasificación del texto, visualización e inspección y validación del texto clasificado.</p>	<p>En este artículo no se obtiene un resultado puntual como tal, sin embargo, de la data obtenida con el análisis se puede determinar la distribución de puntuación de análisis de sentimiento, las tendencias y validar si es que ChatGTP y BERT muestran o no un resultado similar.</p>



17	[2] Fine-grained Affective Processing Capabilities	2023	<p>Se utilizó ChatGPT para realizar análisis de sentimiento a través de la asignación de valores VAD a situaciones y palabras emocionales.</p> <p>Se realizaron experimentos conversacionales para evaluar el rendimiento del modelo en el análisis de sentimientos a partir de texto situacional y palabras emocionales.</p>	<p>ChatGPT demostró una buena capacidad para etiquetar el sentimiento en dimensiones de valencia, excitación y dominancia.</p> <p>El modelo pudo representar emociones en términos de categorías de emoción y dimensiones afectivas, además de realizar una obtención de emociones básica</p>
----	--	------	---	---





<b>18</b>	[9] Sentiment analysis A survey on design framework	<b>2023</b>	<p>Se abordan varias técnicas y métodos utilizados en el análisis de sentimiento, incluyendo enfoques basados en procesamiento de lenguaje natural (PNL), extracción de características, clasificación de sentimientos, entre otros.</p> <p>Se menciona también la importancia de técnicas avanzadas de preprocesamiento de datos, como la normalización estándar, y la extracción de palabras clave utilizando métricas de centralidad del grado para obtener una colección representativa y sentimental de palabras.</p>	<p>Como resultado se obtiene el desarrollo de un marco genérico para el diseño de modelos de análisis de sentimiento eficaces, identificación de lagunas de investigación en el análisis de sentimiento, como la falta de técnicas efectivas de preprocesamiento de datos y la extracción inexacta de palabras clave..</p>
<b>19</b>	[13] Sentiment analysis methods, applications, and challenges: A systematic literature review	<b>2024</b>	<p>El documento utiliza la revisión sistemática de la literatura (SLR) como metodología, siguiendo el marco PRISMA. Este enfoque permite identificar, recopilar, seleccionar y analizar estudios sobre análisis de sentimientos (SA). Se mencionan diversas técnicas de SA: enfoques basados en léxico, aprendizaje automático tradicional, aprendizaje profundo y enfoques híbridos, utilizadas para clasificar y analizar opiniones y sentimientos en textos.</p>	<p>El análisis de la literatura revela que se prefieren los léxicos generales sobre los específicos, aunque deben actualizarse continuamente. Los métodos basados en léxico enfrentan desafíos como frases negativas, errores ortográficos, sinónimos y jerga. Diversos enfoques de evaluación (precisión, puntuación F1, ROC, AUC, RMSE, MAE) son utilizados para evaluar el SA, proporcionando una visión general de los enfoques, desafíos y futuras direcciones en el campo del análisis de sentimientos.</p>



<b>20</b>	[12] Sentiment Analysis Comprehensive Reviews Recent Advances and Open Challenges	<b>2023</b>	En el artículo se aborda el análisis del sentimiento desde el punto de vista de redes neuronales y sistemas de aprendizaje. Se mencionan diferentes enfoques, como modelos de pre entrenamiento, aprendizaje contrastivo, métodos basados en	Los resultados muestran una revisión detallada de los avances recientes en el análisis de sentimiento, destacando la importancia de la integración de información multimodal para mejorar la comprensión del lenguaje y la detección de sentimientos. Se mencionan modelos
-----------	---	-------------	--	--



			alineación multimodal, entre otros. Se destaca el uso de modelos como GPT-3.5 y Chat-GPT para tareas de comprensión del lenguaje natural y análisis de sentimiento. Además, se mencionan redes neuronales como VistaNet y MulT, diseñadas para abordar la alineación multimodal y la clasificación de sentimientos.	como Chat-GPT y MulT, que han demostrado eficacia en la tarea de análisis de sentimiento
--	--	--	---	--

- a. Resultados de la pregunta de investigación 1: ¿Qué técnicas y métodos se utilizan actualmente para realizar análisis de sentimiento?

ID	Respuesta a la pregunta de investigación
01	El artículo utiliza GPT-3 text-davinci-003, GPT-3.5-turbo y GPT-2
02	Para el análisis de sentimiento de tweets relacionados con la conferencia COP9 utilizaron métodos tradicionales de aprendizaje automático y técnicas modernas basadas en modelos pre-entrenados como BERT y GPT-3.
03	Para el análisis de sentimiento de noticias sobre criptomonedas, en este artículo se usó GPT-4, FinBERT y BERT por ser modelos de última generación.
05	Las ocho herramientas mayormente utilizadas en el ámbito de la salud para el análisis de sentimiento son LIWC2015, SentiStrength, TextBlob, VADER, Stanza, TweetEval, Pysentimiento y NLPTown. Adicionalmente se usó OPT y ChatGPT (GPT 3.5).
07	Las técnicas como el Análisis de Sentimientos Basado en Aspectos (ABSA) y modelos de Procesamiento de Lenguaje Natural (PLN) como BART y DeBERTa. Son herramientas que permiten la extracción y clasificación precisa de atributos específicos como la calidad de la comida, el servicio, el ambiente, el costo y la ubicación en el contexto del análisis de sentimiento en restaurantes. Además, las versiones 3.5 y 4.0 de ChatGPT se puede utilizar para discernir y clasificar estos atributos dentro de las reseñas completas, aprovechando su capacidad para generar respuestas que simulan respuestas humanas.



**08**

En el análisis de sentimientos realizado en el estudio, se utilizaron tanto enfoques de aprendizaje automático tradicional como técnicas de aprendizaje profundo. Entre los modelos empleados se incluyeron Support Vector Machines (SVM), Naïve Bayes multinomial, árboles de decisión, Adaboost, y modelos de redes neuronales recurrentes (RNN), como LSTM, GRU y BiLSTM. Estos modelos fueron entrenados y evaluados utilizando métricas como la Exactitud Balanceada y el F1-score Macro, que son adecuadas para



	manejar conjuntos de datos desbalanceados, además se utilizó GPT para las revisiones sintéticas y balanceo de datos con lo cual se mejoró el rendimiento del modelo de clasificación de análisis de sentimiento, lo que demuestra la versatilidad de ChatGPT tanto en el análisis de sentimiento como en el preprocesamiento de datos para mejorar el rendimiento de modelos de clasificación.
16	Según el artículo, actualmente se usan modelos de lenguaje preentrenados como BERT y ChatGPT, técnicas de aprendizaje automático y enfoques basados en lexicones. Se utilizan herramientas de programación como Python y bibliotecas específicas como transformers y SHAP para realizar análisis de sentimiento de manera efectiva
17	Se emplean enfoques basados en léxico, aprendizaje automático tradicional (Naive Bayes, SVM, DT, LR), aprendizaje profundo (CNN, RNN, Transformer) y enfoques híbridos (combinaciones como SVM y Relief) para analizar sentimientos en texto, cruciales para AI y aplicaciones en negocios, gobierno, salud y redes sociales.
18	Actualmente, se utilizan diversas técnicas y métodos en el análisis de sentimiento, incluyendo técnicas de preprocesamiento avanzadas, extracción de palabras clave basada en métricas de centralidad del grado, y modelos de aprendizaje automático para la clasificación de sentimientos.
19	Se emplearon modelos de lenguaje grandes como ChatGPT para asignar valores VAD a situaciones y palabras emocionales. Se realizaron análisis de sentimientos multidimensionales (Valencia, Excitación, Dominancia) a partir de descripciones situacionales y palabras emocionales.
20	Se utilizan modelos de preentrenamiento, aprendizaje contrastivo, redes neuronales multimodales y enfoques basados en alineación multimodal para realizar análisis de sentimiento.

**Respuesta a la pregunta de investigación 1:** Las técnicas y métodos actuales para el análisis de sentimiento incluyen enfoques de aprendizaje automático y aprendizaje profundo, utilizando modelos como SVM, Naïve Bayes, RNN, LSTM, GRU, BiLSTM, Árboles de decisión; Herramientas clásicas de análisis de sentimiento, como LIWC2015, SentiStrength, TextBlob y Vader y por último modelos de lenguaje avanzados como GPT-3, GPT-3.5, GPT-4, BERT y sus modelos relacionados como FinBERT..

- b. Resultados de la pregunta de investigación 2: ¿Cómo se compara el desempeño de ChatGPT con otros modelos de lenguaje en la tarea de análisis de sentimiento?

ID	Respuesta a la pregunta de investigación
----	--



02

GPT-3 solo requería una configuración mínima mas no dependía de ningún conjunto de datos específico a



	<p>diferencia de los otros modelos como Bi-LSTM y BERT. Sin embargo, logró un rendimiento F1-Score relativamente bueno (IMDB: 0.91, Sentiment140: 0.79) superando al modelo Bi-LSTM (IMDB: 0.90, Sentiment140: 0.78). Mientras que BERT logró los valores más altos de F1-Score de 0,94 y 0,81. Por lo tanto, los resultados lograron un F1-score relativamente mejor que todos los demás modelos de sentimiento, excepto BERT, para los conjuntos de datos estándar utilizados.</p>
<b>03</b>	<p>En el análisis de sentimiento de noticias de criptomonedas el modelo GPT-4 fue superior con 86.7% en precisión a FinBert que obtuvo 84.3% y a BERT con 83.3%.</p>
<b>05</b>	<p>En los dos conjuntos de datos sobre opiniones sobre el Covid-19, ChatGPT superó a los otros modelos en análisis de sentimiento, con una superioridad sobre el modelo OPT de 6% en precisión y un 7% en F-score en el primer conjunto y una precisión de 6% y un F-score de 4% en el segundo.</p>
<b>06</b>	<p>El Modelo BERT sobresale en clasificación de frases y mapeo de definiciones, pero requiere grandes conjuntos de datos etiquetados. Por otro lado, ChatGPT, con aprendizaje en contexto y prompts personalizados, es efectivo en clasificación con pocos ejemplos y generación de contenido informativo, su efectividad disminuye con preguntas específicas siendo inferior al modelo Bert en este aspecto.</p>
<b>09</b>	<p>Los estudios realizados por los autores, revelaron que las versiones GPT-3.5 y GPT-4 de ChatGPT demostraron una precisión superior al 90% en la detección de sentimientos en textos en cantonés, destacándose por encima de métodos basados en lexicones y modelos de aprendizaje automático ajustados. En el ámbito financiero, el modelo GPT-P4N alcanzó una precisión de 0.784 en la clasificación de sentimientos, con evaluaciones de precisión, recall y F1-score de 0.804, 0.784 y 0.789 respectivamente. Además, el modelo GPT-P6N logró una precisión del 0.652 en la predicción de la dirección del mercado.</p>
<b>10</b>	<p>Los autores nos muestran que el modelo GPT-P4N obtuvo una precisión del 0.784 en la clasificación de sentimientos en el dominio financiero. Las métricas de precisión, recall y F1-score dieron como resultado 0.804, 0.784 y 0.789 respectivamente, lo que demuestra una efectividad notable de este modelo en la clasificación de sentimientos en el dominio de las finanzas. Además, se observó que el modelo GPT-P6N logró una precisión del 0.652 en la predicción de la dirección del mercado. lo que refleja mayor rendimiento frente a los otros modelos en ese aspecto. Esto resalta la importancia de considerar los modelos GPT como una buena alternativa de solución en la tarea de análisis de sentimientos.</p>
<b>16</b>	<p>La validación de resultados con ChatGPT mostró coherencia con los modelos preentrenados como BERT, indicando un desempeño comparable y robusto.</p>
<b>17</b>	<p>ChatGPT, desarrollado por OpenAI, sobresale en análisis de sentimientos debido a su fluidez y precisión en comprender el lenguaje humano. Supera a muchos modelos en esta tarea y su combinación con técnicas híbridas mejora aún más el rendimiento</p>



**Respuesta a la pregunta de investigación 2:** ChatGPT, especialmente en sus versiones más recientes como GPT-3.5 y GPT-4, ha demostrado ser altamente efectivo en la tarea de análisis de sentimiento. Aunque BERT sigue siendo muy fuerte en clasificación de frases y mapeo de definiciones con grandes conjuntos de datos etiquetados, ChatGPT destaca por su capacidad de entender y generar texto de manera fluida y precisa con una configuración mínima y sin la necesidad de grandes conjuntos de datos específicos. En aplicaciones específicas como el análisis de noticias de criptomonedas y opiniones sobre el Covid-19, ChatGPT ha superado a otros modelos, incluyendo BERT y FinBERT. Además, su desempeño robusto en diversos dominios y su capacidad de validación consistente lo posicionan como una excelente opción para tareas de análisis de sentimiento.

Esta comparativa muestra que, aunque no es perfecto y puede tener desventajas en ciertos aspectos específicos, ChatGPT ofrece un rendimiento sobresaliente y competitivo en la mayoría de las aplicaciones de análisis de sentimiento.

- c. Resultados de la pregunta de investigación 3: ¿Cuáles son los principales avances en el análisis de sentimiento utilizando modelos de lenguaje como ChatGPT?

ID	Respuesta a la pregunta de investigación
04	El avance que se visualiza en este artículo es que GPT puede ser utilizado para el análisis de sentimiento en el contexto de predicciones de retornos de acciones, destacando avances significativos como la mitigación de sesgos a través de la anonimización.
11	Los modelos ChatGPT sobresalen en tareas relacionadas con el análisis de sentimientos, clasificación de polaridad de aspectos, extracción de opiniones y clasificación de la intensidad de las emociones, mostrando un rendimiento comparable o superior a métodos de NLP tradicionales y otros modelos pre entrenados como RoBERTa. En particular, GPT-4 muestra un rendimiento notable en la detección de emociones extremadamente negativas, la evaluación del bienestar y la detección de toxicidad. No obstante, los modelos presentan limitaciones en tareas con señales implícitas como la evaluación de la personalidad, la detección de sarcasmo y la detección de subjetividad.



**12**

Los avances en el análisis de sentimientos utilizando modelos de lenguaje como ChatGPT incluyen su capacidad para lograr un rendimiento competitivo en comparación con los modelos más avanzados, acercándose al juicio humano en tareas de clasificación de sentimientos, humor e identificación de metáforas. Además, se encontró que la efectividad de ChatGPT en el análisis de sentimiento puede mejorarse significativamente mediante la optimización de los patrones de instrucciones (prompting). Sin embargo, a pesar de estos avances, todavía existe una brecha en comparación con la precisión del juicio humano, especialmente en la interpretación de humor y metáforas.



<b>13</b>	Uno de los avances en el análisis de sentimientos utilizando modelos de lenguaje como ChatGPT incluye un rendimiento competitivo y una alta precisión, demostrando un 88.17% de precisión cuando se combina con algoritmos como KNN, RNN y CNN. Gracias a su entrenamiento en grandes volúmenes de datos, ChatGPT puede captar características contextuales y matices en el lenguaje, lo que mejora significativamente su análisis de sentimientos.
<b>14</b>	Actuales avances en el análisis de sentimientos utilizando modelos de lenguaje como ChatGPT incluyen su capacidad para generar datos sintéticos que mejoran significativamente el rendimiento de modelos más pequeños, haciéndolos competitivos con modelos más grandes y costosos. ChatGPT puede captar contextos y matices lingüísticos complejos, lo que aumenta la precisión del análisis de sentimientos. La integración de datos sintéticos generados por ChatGPT también ayuda a abordar el desequilibrio de clases en los conjuntos de datos de entrenamiento, y el uso de plantillas de sugerencias específicas ha mostrado mejorar aún más la precisión.
<b>15</b>	El artículo presenta avances significativos en el uso de modelos de lenguaje grande (LLM) como ChatGPT para el análisis de sentimiento en letón. Este estudio demuestra que mediante la ingeniería de prompts, se puede mejorar considerablemente la precisión del análisis de sentimiento en esta lengua. Los resultados obtenidos indican que los LLMs no solo pueden reemplazar parcialmente a los etiquetadores humanos, sino también mejorar la eficiencia y la calidad de los análisis de sentimiento en idiomas con recursos limitados, consolidando el potencial de estos modelos para tareas de procesamiento de lenguaje natural en diversos idiomas.
<b>16</b>	Los avances incluyen la capacidad de los modelos para proporcionar interpretaciones más precisas y coherentes de sentimientos en texto, aprovechando el aprendizaje profundo y el entrenamiento en grandes corpus de datos. Estos modelos pueden manejar mejor la variabilidad del lenguaje humano en diferentes contextos.
<b>17</b>	Los modelos como ChatGPT mejoran la comprensión del contexto y las relaciones internas en el texto, permitiendo un análisis de sentimientos más preciso. Ofrecen análisis en tiempo real, enfoque en el cliente y ventaja competitiva, destacándose por su eficiencia y precisión en diversas aplicaciones.
<b>20</b>	Los avances incluyen la capacidad de comprender el lenguaje natural y realizar tareas de análisis de sentimiento, aunque con limitaciones en la detección de sentimientos sutiles e implícitos como el sarcasmo.



**Respuesta a la pregunta de investigación 3:** Los avances en el análisis de sentimientos utilizando modelos de lenguaje como ChatGPT incluyen mejoras significativas en precisión, manejo de sesgos, generación de datos sintéticos y optimización de instrucciones. Estos avances permiten a ChatGPT ofrecer un análisis de sentimientos más preciso y eficiente, acercándose al juicio humano en muchas tareas, aunque todavía existen áreas donde se requiere mejorar, especialmente en la detección de señales implícitas y matices sutiles del lenguaje.

#### 4. Aplicación práctica

Con el fin de complementar esta investigación y demostrar la factibilidad del uso de ChatGPT en el análisis de sentimientos se desarrolló un proyecto enfocado en implementar una herramienta innovadora de análisis de sentimientos de reseñas de servicios de alojamiento web en línea utilizando GPT-3.5. La herramienta extrae reseñas de la página web TrustPilot, donde los usuarios pueden dejar reseñas y calificaciones sobre empresas. La empresa seleccionada en este caso de estudio fue SiteGround España, que se dedica al servicio de alojamiento de hosting. Posteriormente, se analiza el sentimiento en las reseñas y se presentan los resultados de manera clara y visualmente atractiva. La solución incluye tres etapas importantes: extracción de reseñas, análisis de sentimiento y visualización de resultados.

##### 4.1. Extracción de Reseñas (Web Scraping)

Para la extracción de reseñas, se utilizó un backend desarrollado con Django y bibliotecas de Python como BeautifulSoup y Requests. Los datos obtenidos se almacenaron en MongoDB. Los pasos específicos para esta etapa incluyen la preparación del entorno y el desarrollo del script de extracción de reseñas

En la implementación de scripts en Python, para realizar el web scraping de plataformas de reseñas, se utilizó BeautifulSoup para el parsing de HTML y Requests para obtener el contenido de las páginas web.

```
class ScrapeReviewsView(APIView):  
  
    def scrape_reviews(self, url):  
        response = requests.get(url)  
        soup = BeautifulSoup(response.content, 'html.parser')  
        # Ajustar la clase según la estructura HTML de la página de reseñas  
        #reviews = soup.find_all(class_='review')  
        reviews = soup.find_all(class_='styles_reviewContent__0Q2Tg')  
        extracted_reviews = []  
        for review in reviews:  
            #encontrar la etiqueta con con la clase donde está el texto  
  
            #Analiza el comentario en negrita, etiqueta h2 de la página  
            review_text = review.find('h2').text  
  
            extracted_reviews.append(review_text)  
  
        return extracted_reviews
```



ECT

Revista Científica Emprendimiento  
Científico Tecnológico

ISSN: 2810 – 8493

<https://revista.ectperu.org.pe/index.php/ect/index>

Imagen 1: Función para extraer reseñas.

## 4.2. Análisis de Sentimiento con GPT-3.5 turbo instruct

Para el análisis de sentimiento de las reseñas, se desarrolló un backend utilizando Node.js y Express.js. Se integró el modelo GPT-3.5 turbo instruct a través de la API de OpenAI, definiendo prompts específicos para el análisis de sentimiento y clasificación de los comentarios.

```
import { openai } from "../config.js"

export const consultarSentimiento = async(req, res) =>{
  try{
    //Obteniendo el comentario
    let {comentario} = req.body

    if (!comentario) {
      return res.status(400).json({ error: 'El comentario es requerido' });
    }

    //consulta de con modelos medios en openai
    const response = await openai.completions.create({
      model: "gpt-3.5-turbo-instruct",

      prompt: `SiteGround España es una empresa que brinda el servicio de alojamiento web
      Tiene como funciones la atención al cliente y el soporte tecnico
      Realiza un análisis de sentimiento del siguiente comentario : "${comentario}", Proporcione el
      sentimiento y el tema del comentario
      ejemplo: {"sentimiento": "positivo", "topico": "atención al cliente"} ,
      En caso el análisis no se pueda determinar Output: {"sentimiento": "neutral", "topico": "indeterminado"}`,
      max_tokens: 50,
      temperature: 0,
    })

    console.log("##### Nuevo comentario #####")
    console.log("Comentario :", comentario)
    console.log("Resultado de chatGPT: ", response)
    let sentiment = response.choices[0].text.trim().toLowerCase();

    console.log("Resultado de la consulta: ",sentiment)

    let data= JSON.parse(sentiment);

    //envia el campo sentiment al api scrapping
    return res.status(200).json({sentiment: data});
  }catch(error){
    console.log(error)
    return res.status(500).json({mensaje:'Error al listar sentimientos',error: error});
  }
}
```

Imagen 2: Función para realizar el análisis de sentimientos.

```
//consulta de con modelos medios en openai
const response = await openai.completions.create({
  model: "gpt-3.5-turbo-instruct",
  prompt: `SiteGround España es una empresa que brinda el servicio de alojamiento web
  Tiene como funciones la atención al cliente y el soporte tecnico
  Realiza un análisis de sentimiento del siguiente comentario : "${comentario}", Proporcione el
  sentimiento y el tema del comentario
  ejemplo: {"sentimiento": "positivo", "topico": "atención al cliente"} ,
  En caso el análisis no se pueda determinar Output: {"sentimiento": "neutral", "topico": "indeterminado"}`,
  max_tokens: 50,
  temperature: 0,
})
```



ECT

Revista Científica Emprendimiento  
Científico Tecnológico

ISSN: 2810 – 8493

<https://revista.ectperu.org.pe/index.php/ect/index>

Imagen 3: Prompt usado para pedir a ChatGPT el análisis de sentimiento.

#### 4.3. Visualización de Resultados

Para la visualización de los resultados del análisis de sentimiento, se desarrolló una interfaz de usuario utilizando Angular, Bootstrap y Chart.js, proporcionando una visualización interactiva y amigable de los datos de sentimiento.

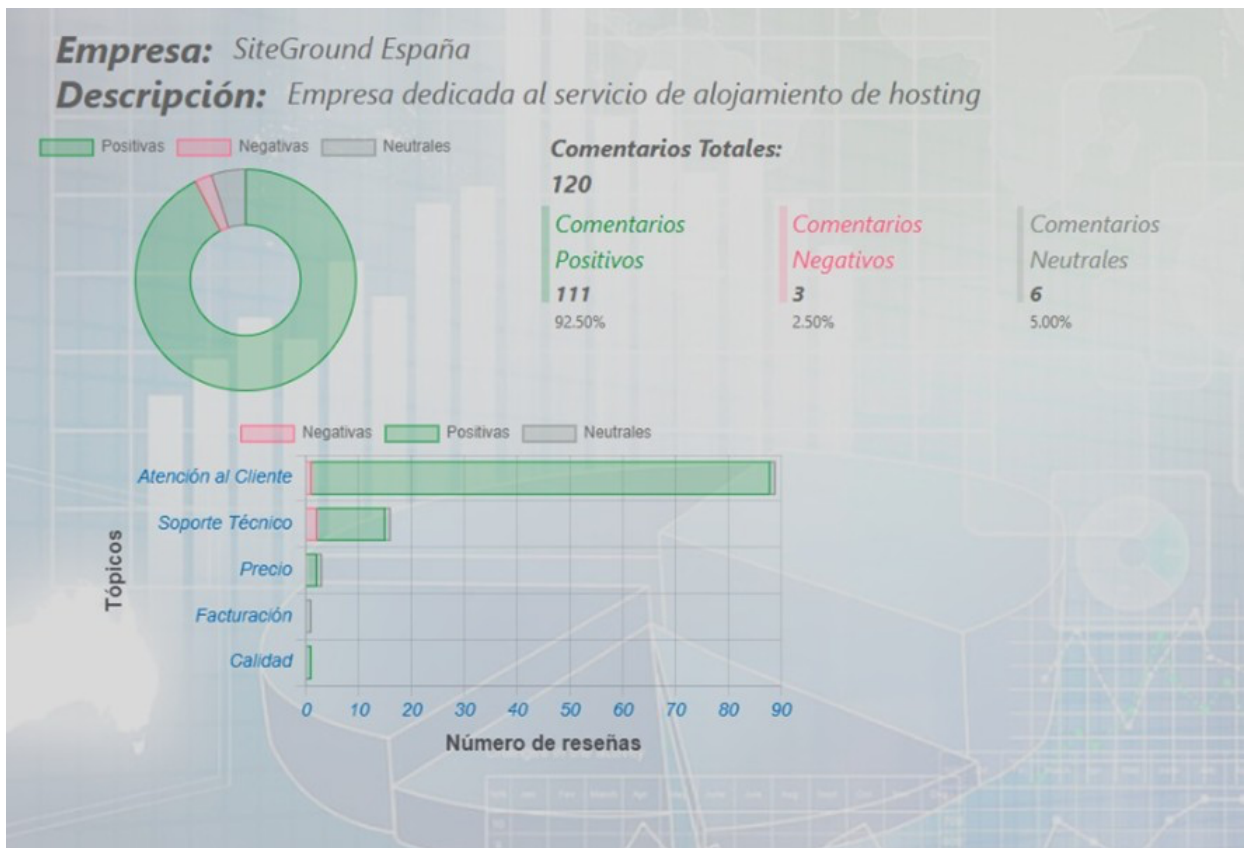


Imagen 4: Interfaz para mostrar los resultados del análisis de sentimientos

El proyecto ha demostrado ser una herramienta eficaz para el análisis de sentimientos en reseñas de servicios de alojamiento web utilizando tecnologías avanzadas como GPT-3.5. A través de la implementación de un sistema integrado que abarca desde la extracción de datos hasta la visualización de resultados, se ha logrado proporcionar insights valiosos para empresas como SiteGround España, permitiéndoles entender mejor las percepciones y opiniones de sus usuarios.

#### 5. Conclusiones

- **Con respecto a la pregunta de investigación 1:** “¿Qué técnicas y métodos se utilizan actualmente para realizar análisis de sentimiento?”, la revisión de literatura revela que el análisis de sentimiento se beneficia ampliamente de la innovación en modelos de lenguaje pre entrenados como BERT y GPT, así como de técnicas avanzadas de aprendizaje automático y profundo. Estos métodos no solo permiten una clasificación precisa de sentimientos en una variedad de contextos y dominios, desde noticias financieras hasta opiniones sobre eventos globales, sino que también facilitan el desarrollo de enfoques híbridos que mejoran la comprensión y precisión del análisis emocional en textos complejos.



- **Con respecto a la pregunta de investigación 2:** “¿Cómo se compara el desempeño de ChatGPT con otros modelos de lenguaje en la tarea de análisis de sentimiento? ”, en base a la revisión de artículos, se concluye que ChatGPT





muestra un rendimiento competitivo en la tarea de análisis de sentimiento, superando a varios modelos de lenguaje en diferentes contextos y dominios específicos. Si bien modelos como BERT pueden alcanzar los puntajes más altos en métricas como F1-Score, ChatGPT destaca por su capacidad de manejar datos con menos ejemplos etiquetados y su efectividad en la generación de contenido informativo a través de prompts personalizados. Esto lo hace especialmente útil en aplicaciones donde la adaptabilidad y la comprensión contextual son cruciales, aunque enfrenta desafíos en la interpretación de sutilezas emocionales como el sarcasmo. En conjunto, estos hallazgos subrayan a ChatGPT como una alternativa robusta y efectiva en el análisis de sentimiento, ofreciendo precisión y versatilidad en diversas aplicaciones prácticas.

- **Con respecto a la pregunta de investigación 3:** “¿Cuáles son los principales avances en el análisis de sentimiento utilizando modelos de lenguaje como ChatGPT?”, la revisión de literatura revela que los avances en el análisis de sentimiento utilizando modelos de lenguaje como ChatGPT muestran un progreso significativo en la capacidad de estos modelos para realizar tareas complejas de clasificación emocional y evaluación de sentimientos. ChatGPT, especialmente en sus variantes más avanzadas como GPT-4, ha demostrado ser competitivo en la detección de emociones extremas y la clasificación de sentimientos en una variedad de contextos lingüísticos. Sin embargo, persisten desafíos en la interpretación de humor, metáforas y otros aspectos sutiles del lenguaje, donde aún no alcanza la precisión del juicio humano. La optimización de patrones de instrucciones y la generación de datos sintéticos son estrategias prometedoras para mejorar aún más la efectividad de ChatGPT en esta área, señalando un camino hacia aplicaciones más robustas y precisas en el análisis de sentimientos.

## 6. Referencias bibliográficas

- [1] Amin, M. M., Mao, R., Cambria, E., & Schuller, B. W. (2024). A wide evaluation of ChatGPT on affective computing tasks. *IEEE Transactions on Affective Computing*. <https://arxiv.org/abs/2308.13911>.
- [2] Broekens, J., Hilpert, B., Verberne, S., Baraka, K., Gebhard, P., & Plaat, A. (2023). Fine-grained affective processing capabilities emerging from large language models. In *2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII)*. <https://arxiv.org/abs/2309.01664v1>.
- [3] Elmitwalli, Sherif y Mehegan, John (2024). “Sentiment analysis of COP9-related tweets: a comparative study of pre-trained models and traditional techniques”. *Front. Big Data, Sec. Data Science Volume 7* <https://doi.org/10.3389/fdata.2024.1357926>.
- [4] Carrasco, P., & Dias, S. (2024). Enhancing Restaurant Management through Aspect-Based Sentiment Analysis and NLP Techniques. *Procedia Computer Science*, 237, 129-137. <https://doi.org/10.1016/j.procs.2024.05.088>
- [5] Fatouros, G., Soldatos, J., Kouroumalis, K., Makridakis, G., & Kyriazis, D. (2023). Transforming sentiment analysis in the financial domain with ChatGPT. *Machine Learning With Applications*, 14, 100508.



ECT

Revista Científica Emprendimiento  
Científico Tecnológico

ISSN: 2810 – 8493

<https://revista.ectperu.org.pe/index.php/ect/index>

<https://doi.org/10.1016/j.mlwa.2023.100508>



- [6] Fu, Z., Hsu, Y. C., Chan, C. S., Lau, C. M., Liu, J., & Yip, P. S. F. (2023). Efficacy of ChatGPT in Cantonese Sentiment Analysis: A Comparative Study. *JMIR. Journal Of Medical Internet Research/Journal Of Medical Internet Research*. <https://doi.org/10.2196/51069>
- [7] Glasserman, Paul y Lin, Caden (2024). "Assessing Look-Ahead Bias in Stock Return Predictions Generated by GPT Sentiment Analysis". *Journal of Financial Data Science* Volume 6, Issue 1, Pages 25. 10.3905/jfds.2023.1.14
- [8] Hariguna, T., & Ruangkanjanases, A. (2023). Exploring the Flexibility and Accuracy of Sentiment Scoring Models through a Hybrid KNN-RNN-CNN Algorithm and ChatGPT. *HighTech and Innovation Journal*, 4(2), 315-326. <https://doi.org/10.28991/HIJ-2023-04-02-06>.
- [9] Kaur, G., Kumar, A., & Sharma, A. (2023). Sentiment analysis: A survey on design framework, challenges and applications. *Journal of King Saud University - Computer and Information Sciences*. <https://doi.org/10.1016/j.jksuci.2023.03.017>.
- [10] Konstantinos I. Roumeliotis, Nikolaos D. Tselikas y Dimitrios K. Nasiopoulos (2024). "LLMs and NLP Models in Cryptocurrency Sentiment Analysis: A Comparative Classification Study", *MDPI Big Data and Cognitive Computing; Basel Tomo 8, N.º 6*: 63. <https://doi.org/10.3390/bdcc8060063>.
- [11] Lossio-Ventura JA et al., (2024). "A Comparison of ChatGPT and Fine-Tuned Open Pre-Trained Transformers (OPT) Against Widely Used Sentiment Analysis Tools: Sentiment Analysis of COVID-19 Survey Data". *JMIR Ment Health; Volume 11, Issue 1: Article number e50150* doi: 10.2196/50150
- [12] Lu, Q., Sun, X., Long, Y., Gao, Z., Feng, J., & Sun, T. (2023). Sentiment analysis: Comprehensive reviews, recent advances, and open challenges. *IEEE Transactions on Neural Networks and Learning Systems*. <https://doi.org/10.1109/TNNLS.2023.3294810>.
- [13] Mao, Y., Liu, Q., & Zhang, Y. (2024). Sentiment analysis methods, applications, and challenges: A systematic literature review. *Journal Of King Saud University. Computer And Information Sciences/Mağalaġ Ğam'aġ Al-malġk Saud : Ûlm Al-ħasib Wa Al-ma'lumat*, 102048. <https://doi.org/10.1016/j.jksuci.2024.102048>.
- [14] Purvins, P., Urtans, E., & Caune, V. (2024). Using large language models to improve sentiment analysis in Latvian language. *Baltic Journal of Modern Computing*, 12(2), 165-175. <https://doi.org/10.22364/bjmc.2024.12.2.03>.
- [15] Qiu, Y., & Jin, Y. (2024). ChatGPT and Finetuned BERT: A Comparative Study for Developing Intelligent Design Support Systems. *Intelligent Systems With Applications*, 21, 200308. <https://doi.org/10.1016/j.iswa.2023.200308>.
- [16] Stanislav Chumakova, Anton Kovantseva, Anatoliy Surikov (2023). Generative approach to Aspect Based Sentiment



ECT

Revista Científica Emprendimiento  
Científico Tecnológico

ISSN: 2810 – 8493

<https://revista.ectperu.org.pe/index.php/ect/index>

Analysis with GPT Language Models. 12th International Conference Young Scientist Conference on Computational Science (YSC 2023). Procedia Computer Science 229 (2023) 284–293. <https://doi.org/10.1016/j.procs.2023.12.030>.



- [17] Suhaeni, C., & Yong, H. (2023). Mitigating Class Imbalance in Sentiment Analysis through GPT-3-Generated Synthetic Sentences. *Applied Sciences*, 13(17), 9766. <https://doi.org/10.3390/app13179766>.
- [18] Susnjak, T. (2023). Applying BERT and ChatGPT for sentiment analysis of Lyme disease in scientific literature. *arXiv*. <https://arxiv.org/abs/2302.06474>.
- [19] Woźniak, S., & Kocoń, J. (2023, December). From Big to Small Without Losing It All: Text Augmentation with ChatGPT for Efficient Sentiment Analysis. In 2023 IEEE International Conference on Data Mining Workshops (ICDMW) (pp. 799-808). IEEE. <https://doi.org/10.1109/ICDMW60847.2023.00108>.
- [20] Zhang, Y., Wang, M., Rong, L., Yu, Y., Zhao, D., & Qin, J. (2024). Can ChatGPT be served as the sentiment expert? An evaluation of ChatGPT on sentiment and metaphor analysis. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 60(1). <https://doi.org/10.13209/j.0479-8023.2023.075>.