# **Análisis de Sentimiento de Comentarios sobre la Aplicación Threads**

## **1. Introducción**

En un mundo impulsado por la conectividad digital y las redes sociales, la toma de decisiones empresariales informadas se ha convertido en un pilar fundamental para el éxito de cualquier organización. La expansión estratégica hacia nuevas plataformas y canales de comunicación se ha vuelto esencial para llegar a audiencias diversificadas y participar en conversaciones significativas con los usuarios. En este contexto, la aplicación "Threads" se presenta como una potencial adición a la cartera de redes sociales de nuestra empresa.

"Threads" es una plataforma de comunicación que ha ganado notoriedad en el panorama de las redes sociales, especialmente entre las generaciones más jóvenes. Su enfoque en la comunicación privada y la compartición de momentos efímeros ha generado interés tanto en usuarios individuales como en empresas que buscan conectar con su audiencia de una manera auténtica y directa.

Sin embargo, antes de dar el paso hacia la asignación de recursos significativos en marketing y estrategias de promoción en Threads, es esencial comprender en profundidad la percepción y la experiencia de los usuarios actuales de la plataforma. Los comentarios y reseñas disponibles en la tienda Play Store proporcionan una ventana única para conocer la opinión de los usuarios sobre Threads.

Este perfil de investigación se enfoca en llevar a cabo un análisis de sentimiento exhaustivo de los comentarios y reseñas de Threads en la tienda Play Store. El objetivo principal es evaluar si Threads representa una plataforma prometedora para la inversión en marketing de nuestra empresa, identificando tanto los aspectos positivos que pueden potenciarse como las preocupaciones que deben abordarse.

A través del análisis de sentimiento, exploraremos la satisfacción general de los usuarios, las áreas de mejora potencial y las tendencias emergentes en las opiniones de los usuarios. Los resultados de esta investigación proporcionarán una base sólida para la toma de decisiones estratégicas relacionadas con Threads, permitiendo a nuestra empresa tomar medidas informadas para aprovechar el potencial de esta plataforma de redes sociales.

## **2. Antecedentes**

En el dinámico mundo de las redes sociales, la aparición de nuevas plataformas puede alterar significativamente el panorama competitivo. Meta, anteriormente conocida como Facebook, lanzó "Threads" con la intención de competir con plataformas establecidas, particularmente Twitter bajo la dirección de Elon Musk. Un análisis realizado en julio de 2023 sugiere que Threads tiene el potencial de aprovechar la base de usuarios de Facebook e Instagram y podría beneficiarse de los desafíos técnicos enfrentados por Twitter (Referencia: [URL del documento]). El estudio también subraya la influencia de los creadores de contenido en la retención y lealtad de los usuarios, y cómo Threads podría persuadir a estos creadores a cambiar de plataforma.

Para los inversores y las empresas, la percepción del usuario sobre una plataforma de redes sociales es un indicador crucial para determinar su viabilidad como canal de marketing y publicidad. Si bien los estudios académicos y de mercado pueden proporcionar una visión general, las opiniones y sentimientos directos de los usuarios, como los expresados en comentarios en tiendas de aplicaciones como Play Store, ofrecen una perspectiva más inmediata y auténtica.

Dado el potencial percibido de Threads y su impacto en el mercado, es esencial determinar si esta plataforma representa una oportunidad valiosa para atraer y retener clientes. La pregunta clave es: ¿Vale la pena invertir en Threads como un medio efectivo para el engagement y la adquisición de clientes, o su potencial es limitado?

## **3. Planteamiento del Problema**

En el contexto de la creciente competencia y los cambios significativos en el panorama de las redes sociales, la toma de decisiones estratégicas se vuelve aún más crucial para nuestra empresa. La reciente aparición de la aplicación "Threads" como una plataforma de comunicación y redes sociales plantea tanto oportunidades como desafíos. Threads ha surgido como un competidor directo en un espacio que ha sido dominado durante mucho tiempo por plataformas establecidas como Twitter.

La rivalidad entre Threads y Twitter es especialmente relevante debido a los cambios recientes introducidos por Twitter, que han ampliado su enfoque más allá de los mensajes cortos para centrarse en la comunicación efímera y la compartición de momentos, una característica distintiva de Threads. Esta competencia directa entre ambas plataformas nos obliga a considerar si asignar recursos significativos en marketing y estrategias de promoción en Threads es una medida estratégica viable.

El problema fundamental que abordaremos en esta investigación es el siguiente: ¿Cómo perciben y experimentan los usuarios actuales la plataforma Threads en comparación con otras opciones de redes sociales, en particular Twitter, en el contexto de los cambios recientes? La respuesta a esta pregunta es esencial para evaluar si Threads representa una inversión prometedora para nuestra empresa.

Para resolver este problema, llevaremos a cabo un análisis de sentimiento de los comentarios y reseñas de usuarios en la tienda Play Store relacionados con Threads. Este análisis proporcionará una visión profunda de la satisfacción de los usuarios, así como de las áreas en las que Threads puede destacar o necesitar mejoras. Además, nos permitirá identificar la relación entre Threads y Twitter en términos de la percepción de los usuarios.

Esta empresa se encuentra en una encrucijada estratégica en la que debe tomar decisiones informadas sobre la asignación de recursos en un entorno altamente competitivo de redes sociales. La investigación abordará este desafío al proporcionar información valiosa basada en datos concretos sobre la percepción de los usuarios y la competencia directa entre Threads y Twitter.

## **4. Descripción del objeto de estudio**

### 4.1 Objetivo general

* Ofrecer información clave para que la empresa pueda tomar decisiones informadas sobre la percepción y el compromiso con la plataforma Threads de los posibles clientes en función de datos objetivos y la retroalimentación de los usuarios.

### 4.2 Objetivos específicos

* Evaluar el sentimiento general de los comentarios y reseñas de los usuarios sobre la aplicación Threads en la tienda Play Store para comprender la percepción predominante hacia la plataforma.
* Identificar las principales áreas de fortaleza de Threads según las opiniones de los usuarios y destacar las características y experiencias que generan opiniones positivas.
* Proporcionar recomendaciones basadas en los resultados del análisis de sentimiento para guiar la toma de decisiones estratégicas relacionadas con la asignación de recursos de marketing y la participación en la plataforma Threads.

## **5. Justificación**

**5.1. Importancia Estratégica:** En un entorno empresarial altamente competitivo y en constante evolución, la toma de decisiones informadas es esencial para el éxito. La decisión de invertir recursos en una plataforma de redes sociales como Threads debe basarse en datos concretos y en la percepción real de los usuarios.

**5.2. Competencia en el Espacio de Redes Sociales:** La competencia directa entre Threads y Twitter, dos plataformas de comunicación con enfoques similares, plantea preguntas críticas sobre cómo posicionarse en el mercado. Entender cómo los usuarios perciben estas dos plataformas es vital para tomar decisiones estratégicas.

**5.3. Influencia en la Estrategia de Marketing:** La percepción de los usuarios puede tener un impacto significativo en la estrategia de marketing de una empresa. Si los usuarios tienen una experiencia positiva con Threads, es probable que estén más dispuestos a comprometerse y participar en la plataforma.

**5.4. Base de Datos Objetiva:** El análisis de sentimiento de comentarios en la tienda Play Store proporciona una fuente de datos objetiva y no sesgada sobre la opinión de los usuarios. Esto contrasta con las encuestas o estudios de mercado, que pueden estar sujetos a sesgos de respuesta.

**5.5. Identificación de Oportunidades y Desafíos:** Este proyecto permitirá identificar tanto las fortalezas como las áreas de mejora de Threads, lo que facilitará la toma de decisiones para abordar deficiencias y aprovechar oportunidades.

**5.6. Contribución al Conocimiento: E**l proyecto también contribuirá al conocimiento general sobre la percepción de los usuarios en el contexto de las redes sociales y la competencia entre plataformas similares.

**6. Metodología de Análisis de Sentimiento de Comentarios en Play Store**

**6.1. Recopilación de Datos:**

Acceso a la tienda Play Store: Utiliza técnicas de web scraping o APIs (si están disponibles) para recopilar los comentarios de los usuarios sobre la aplicación Threads en la tienda Play Store.

Almacena los datos recopilados en un formato adecuado para su procesamiento posterior.

**6.2. Limpieza de Datos:**

Elimina comentarios duplicados o irrelevantes, como respuestas automáticas o spam.

Normaliza los datos, incluyendo la estandarización de mayúsculas y minúsculas, la eliminación de caracteres especiales y la corrección de errores ortográficos si es necesario.

### 6.3. Procesamiento de Lenguaje Natural (NLP):

Tokenización: Divide los comentarios en palabras o frases (tokens) para su análisis individual.

Eliminación de stopwords: Elimina palabras comunes que no aportan información significativa, como "y", "en", "con", etc.

Stemming o lematización: Reduzca las palabras a sus formas base para agrupar variantes léxicas.

Análisis de sentimiento: Utiliza una librería de NLP (como NLTK o spaCy) o una API de análisis de sentimiento (como la de Google Cloud o Azure) para etiquetar los comentarios como positivos, negativos o neutrales.

### 6.4. Análisis de Datos:

Realiza un análisis exploratorio de los datos para obtener información sobre la distribución del sentimiento en los comentarios.

Calcula métricas estadísticas como la proporción de comentarios positivos, negativos y neutrales, así como el promedio de valoraciones de los usuarios.

### 6.5. Comparación con Twitter:

Si es relevante, compara la percepción de los usuarios hacia Threads con la de Twitter, especialmente en el contexto de los cambios recientes introducidos por Twitter. Esto podría implicar un análisis de palabras clave y temas recurrentes en los comentarios de ambas plataformas.

### 6.6. Visualización de Resultados:

Crea gráficos y visualizaciones que ayuden a representar de manera efectiva los resultados de tu análisis de sentimiento.

Utiliza gráficos de barras, nubes de palabras (word clouds) o diagramas de dispersión para resaltar tendencias y patrones.

**6.7. Interpretación de Resultados:**

Interpreta los resultados del análisis de sentimiento en función de los objetivos de tu investigación y la competencia con Twitter.

Identifica las principales áreas de fortaleza y debilidad de Threads según la percepción de los usuarios.

Proporciona una visión general de cómo los comentarios y valoraciones pueden influir en la decisión de asignar recursos de marketing a Threads.

**6.8. Conclusiones y Recomendaciones:**

Resuma las conclusiones clave de tu análisis de sentimiento y cómo se relacionan con los objetivos de tu investigación.

Proporciona recomendaciones claras sobre la estrategia de marketing y la asignación de recursos en función de los hallazgos.

**7. Marco Teórico**

### 7.1. Análisis de Sentimiento (Sentiment Analysis):

El análisis de sentimiento es una técnica de procesamiento de lenguaje natural (NLP) que se utiliza para determinar la actitud emocional en el texto de un usuario, ya sea positiva, negativa o neutral. Es ampliamente aplicable en la evaluación de opiniones y reseñas de productos o servicios en línea.

**7.2. Recopilación de Comentarios en Línea y Web Scraping:**

La recopilación de comentarios y reseñas en línea se ha convertido en una fuente valiosa de retroalimentación de los usuarios. Para obtener estos datos, se utilizan técnicas de web scraping para extraer información de la tienda Play Store y otros sitios web similares.

### 7.3. Percepción del Usuario en la Toma de Decisiones de Marketing

La percepción del usuario desempeña un papel fundamental en la toma de decisiones de marketing. Las opiniones de los usuarios pueden influir en la elección de los consumidores, la reputación de una marca y la asignación de recursos de marketing.

### 7.4. Competencia en el Espacio de Redes Sociales:

El entorno de las redes sociales es altamente competitivo. Plataformas como Threads compiten con gigantes establecidos como Twitter. La elección de una plataforma para la promoción y el marketing puede tener un impacto significativo en la estrategia de una empresa.

### 7.5. Cambios Recientes en Twitter:

Twitter, como plataforma de redes sociales, ha experimentado cambios significativos en su enfoque, incluyendo la incorporación de características de comunicación efímera y compartición de momentos. Esto ha impactado la dinámica de competencia en el espacio de las redes sociales.

### 7.6. Análisis de Sentimiento con Procesamiento del Lenguaje Natural (NLP):

El análisis de sentimiento se basa en técnicas de procesamiento del lenguaje natural (NLP) para comprender el contenido textual. Esto incluye la tokenización, la eliminación de stopwords y el etiquetado de texto como positivo, negativo o neutral.

### 7.7. Minería de Datos y Machine Learning en Comentarios de Usuarios:

La minería de datos y el machine learning se aplican para extraer conocimientos útiles de grandes conjuntos de datos, como los comentarios de los usuarios en la tienda Play Store. Esto incluye la identificación de patrones y tendencias en los datos de los comentarios.

### 7.8. Métodos de Minería de Datos y Machine Learning

En la minería de datos y el machine learning, se utilizan algoritmos y modelos para clasificar automáticamente los comentarios en categorías de sentimiento. Esto implica técnicas de aprendizaje supervisado y no supervisado.

### 7.9. Evaluación de Modelos de Machine Learning:

La evaluación de modelos de machine learning es esencial para determinar la eficacia del análisis de sentimiento. Se utilizan métricas como la precisión, la exhaustividad y la F1-score para medir el rendimiento de los modelos.

### 7.10. Aplicaciones Empresariales de NLP y Análisis de Sentimiento:

Las aplicaciones empresariales de NLP y análisis de sentimiento incluyen la toma de decisiones estratégicas basadas en datos, la personalización de experiencias de usuario y la mejora de la satisfacción del cliente.

## **8. Límites y alcances**

**8.1 Límites:**

**8.1.1. Fuente de Datos Limitada:** Este proyecto se basa en los comentarios y reseñas disponibles en la tienda Play Store. No incluirá opiniones de usuarios de otras plataformas o redes sociales, lo que podría limitar la visión general de la percepción de los usuarios sobre Threads.

**8.1.2. Idioma:** El análisis de sentimiento se centrará en comentarios y reseñas en un idioma español, lo que podría excluir a usuarios que se expresan en otros idiomas. Esto puede llevar a una visión parcial de la percepción de los usuarios.

**8.1.3. Contexto de la Tienda Play Store:** Los comentarios en la tienda Play Store pueden estar influenciados por factores como problemas técnicos o de rendimiento de la aplicación. Esto podría afectar la objetividad de las opiniones.

**8.1.4. Análisis de Sentimiento Automatizado:** El análisis de sentimiento se realizará utilizando herramientas de procesamiento de lenguaje natural y aprendizaje automático, lo que puede llevar a una cierta cantidad de errores en la clasificación de comentarios.

**8.2. Alcances:**

**8.2.1. Comentarios y Reseñas de Usuarios:** El análisis se centrará en evaluar la percepción de los usuarios a través de los comentarios y reseñas disponibles en la tienda Play Store. Esto proporcionará una visión específica de la experiencia de los usuarios en esta plataforma.

**8.2.2. Comparación con Twitter**: El proyecto incluirá una comparación entre la percepción de los usuarios hacia Threads y Twitter, especialmente en relación con los cambios recientes en Twitter. Esto permitirá entender mejor la competencia directa entre estas dos plataformas.

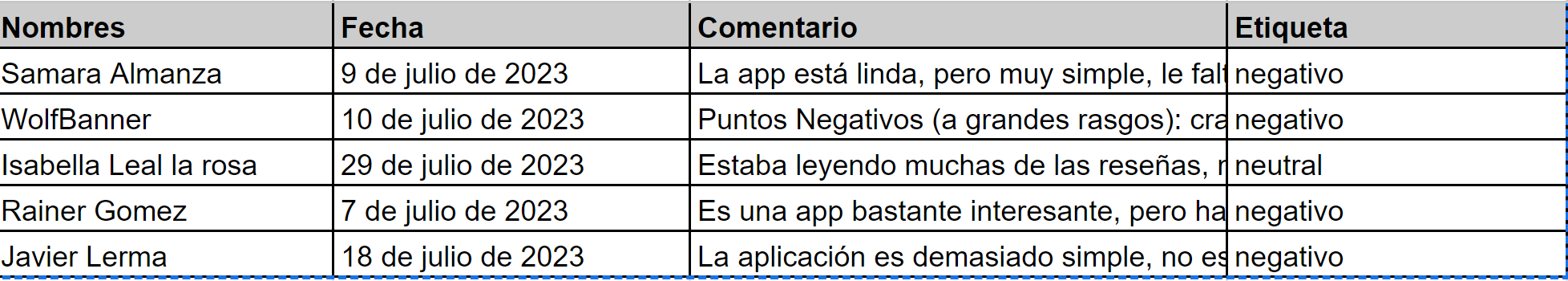
**8.2.3. Recomendaciones Estratégicas:** Se proporcionarán recomendaciones basadas en los resultados del análisis de sentimiento para guiar la toma de decisiones estratégicas relacionadas con la asignación de recursos de marketing y la participación en la plataforma Threads.

**8.2.4. Contribución al Conocimiento**: El proyecto contribuirá al conocimiento sobre la percepción de los usuarios en el contexto de las redes sociales y la competencia entre plataformas similares, enfocándose en Threads y Twitter.

## **9. Descripción del data set**

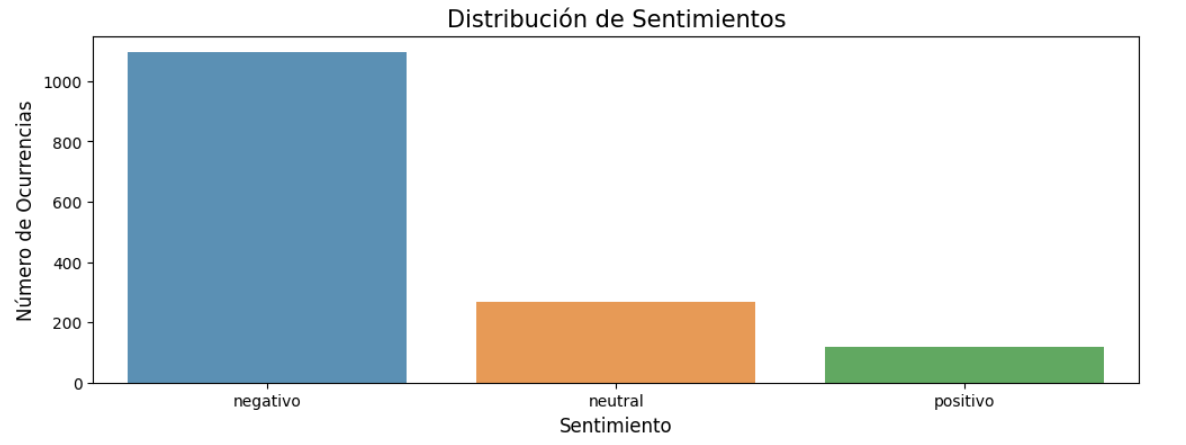
## Se cuenta con un data set de 1480 observaciones. Contiene comentarios relacionados con la opinión de la aplicación de Threads en PlayStore de diferentes usuarios. Los comentarios extraídos se realizaron entre julio y agosto de 2023. El data set contiene 1095 comentarios negativos, 267 neutrales y 118 positivos.

A continuación se observa una pequeña muestra del data set:



## **10. Análisis EDA**

**10.1 Distribución de sentimientos**



## Los comentarios en el Dataset se distribuyeron según el gráfico de barras precedente.

**10.1.1 Cantidad de comentarios etiquetados con el target positivo en el Dataset**



**10.1.2 Cantidad de comentarios etiquetados con el target neutral en el Dataset**

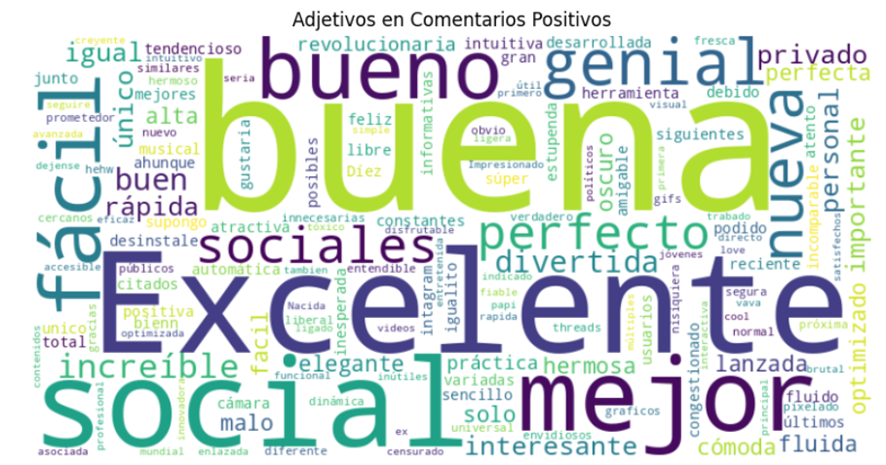


**10.1.3 Cantidad de comentarios etiquetados con el target negativo en el Dataset**

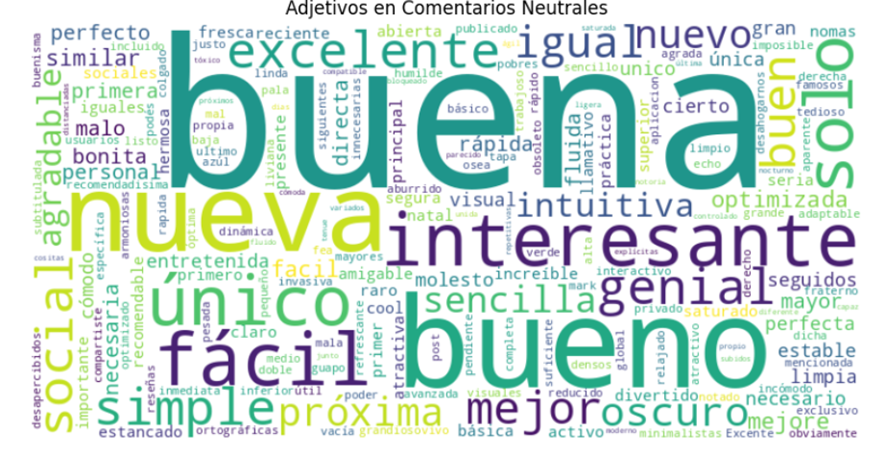


**10.2 Nubes de palabras según target de sentimiento**

**10.2.1 Positivo**

****

**10.2.2 Neutral**

****

**10.2.3 Negativo**

****

## 

## **11. Modelo**

Eliminación de stopwords:

* Las stopwords son palabras muy comunes en un idioma (como "el", "la", "en", "y", etc.) que a menudo no aportan mucho significado en el análisis de texto.
* La eliminación de stopwords es un paso de preprocesamiento común en el NLP para reducir el ruido en los datos y centrarse en las palabras que realmente importan.
* En español, las stopwords incluirían palabras como "de", "y", "en", "el", etc. Estas palabras se eliminan de los comentarios para simplificar el análisis y reducir la dimensionalidad del conjunto de datos.

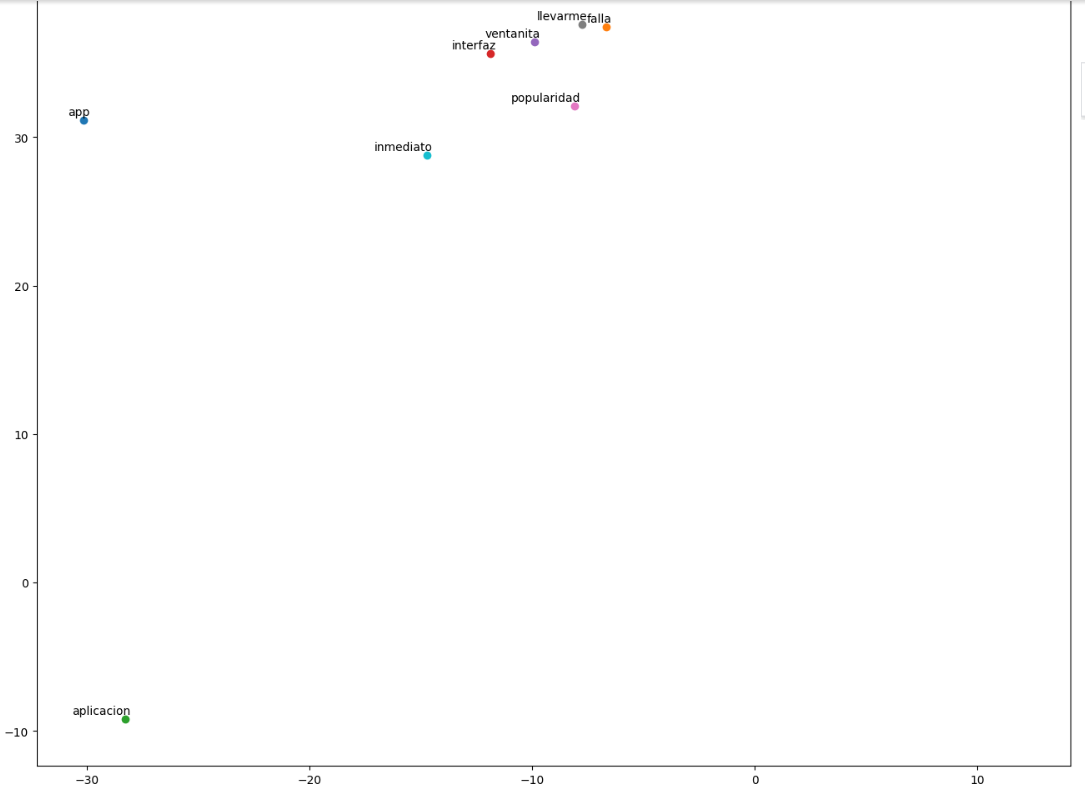
Tokenización de comentarios:

* La tokenización es el proceso de dividir un texto en unidades más pequeñas llamadas "tokens". En el contexto del procesamiento de texto, los tokens suelen ser palabras individuales o, en algunos casos, partes de palabras.
* Por ejemplo, la frase "El perro corre en el parque" se tokenizar en tokens individuales como ["El", "perro", "corre", "en", "el", "parque"].
* La tokenización es esencial para que el modelo pueda comprender y procesar el texto de manera adecuada, ya que trabaja a nivel de palabras o tokens individuales.

Vectorización a través de Doc2Vec:

* La vectorización es el proceso de convertir texto en números, ya que los modelos de aprendizaje automático, como las redes neuronales, solo pueden trabajar con datos numéricos.
* Doc2Vec es una técnica de vectorización que es una extensión de Word2Vec. Mientras que Word2Vec representa palabras como vectores, Doc2Vec extiende esto para representar documentos completos (como comentarios o documentos de texto) como vectores.
* Cada documento se asigna a un vector en un espacio de alta dimensionalidad, donde la posición de cada vector representa las características del documento.
* Esta representación vectorial de los comentarios es útil porque permite calcular similitudes entre documentos, agrupar documentos similares y alimentar estos vectores en modelos de aprendizaje automático para tareas específicas como clasificación de texto o recuperación de información.

Se ha graficado además un modelo TSNE, una técnica de reducción de dimensionalidad que se utiliza para visualizar datos complejos en un espacio de menor dimensión de manera que las relaciones y estructuras en los datos se conserven de manera efectiva. Es especialmente útil para la visualización y la exploración de datos en aplicaciones de aprendizaje automático y análisis de datos.



El objetivo del modelo es poder etiquetar nuevos comentarios sobre la aplicación Threads como positivos, neutros y negativos. Para esto se aplica el siguiente proceso:

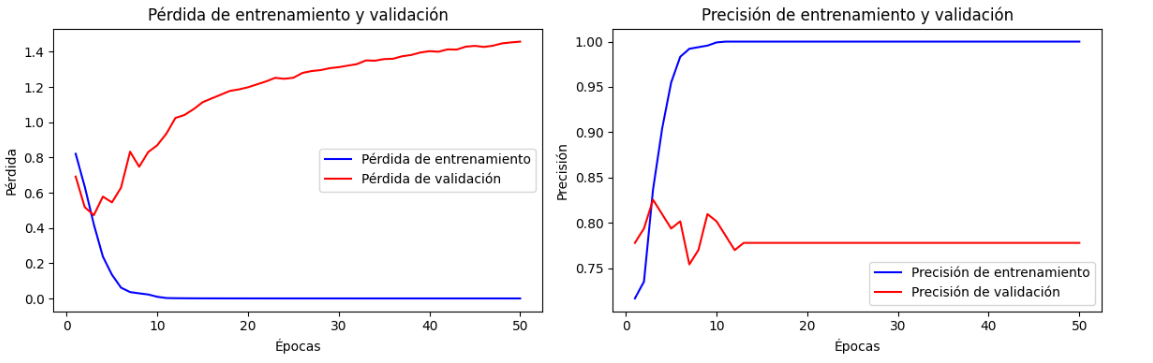
* Tokenización y secuenciación de texto donde se asegura que las secuencias tengan una misma longitud
* División de datos en conjuntos de entrenamiento y prueba. El 15% de los datos se utiliza como conjunto de prueba
* Creación del modelo LSTM donde se crea un modelo secuencial de Keras. Se agrega una capa de embedding que toma las secuencias numéricas y las convierte en vectores densos. Se agrega una capa LSTM con 128 unidades. También una capa densa con 3 unidades de salida y activación softmax, que es adecuada para la clasificación de 3 clases de sentimientos: negativo, neutral y positivo.
* Entrenamiento del modelo utilizando el optimizador 'adam' durante 50 épocas con un tamaño de lote de 32 ejemplos por lote.

## **12. Resultados Obtenidos:**

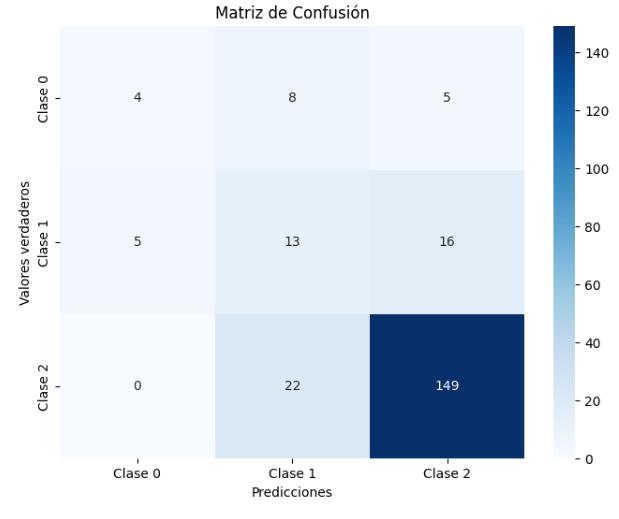
Se han obtenido además mediante la aplicación del modelo los siguientes resultados,



Se han creado dos gráficos en una sola figura: uno para la pérdida y otro para la precisión durante el entrenamiento y la validación de un modelo de aprendizaje automático. Estos gráficos son útiles para evaluar el rendimiento del modelo y detectar posibles problemas como el sobreajuste (overfitting) o el subajuste (underfitting). obteniendo los siguientes resultados,



Se ha obtenido además la matriz de confusión siguiente



Observando que se clasifican correctamente 4 comentarios positivos, 13 neutrales y 149 negativos, teniendo 13 falsos positivos, 5 neutrales y 38 negativos.

## **13. Bibliografía**

* [1] Artificial Inteligence, Machine Learning and Deep Learning by Mercury Learning and Informarion (Book)
* [2] Source Information of NLP :<https://monkeylearn.com/what-is-text-classification/>
* [3] Parameter doc2vec and LSTM inspired by : <https://dl.acm.org/doi/10.1145/3406601.3406624>
* [4] (Bing Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012.)