

**Clasificación de Opiniones en Redes Sociales mediante
Métodos de Minería de Texto**

Corporación Universitaria Remington.

Ingeniería de Sistemas

Juan Pablo Vélez Uribe

Juan Sebastián Mancilla

Lizbeth Ximena Alfonso Bayona

Opción de Trabajo de grado Seminario-Diplomado.

2025

DEDICATORIA

Este trabajo de grado está dedicado, en primer lugar, a nuestras familias, por su apoyo constante, comprensión y motivación durante todo el proceso de formación académica. Su confianza y acompañamiento fueron fundamentales para culminar con éxito esta etapa.

Asimismo, dedicamos este proyecto a nuestros docentes y a la Corporación Universitaria Remington, quienes, a través de su orientación y conocimientos, contribuyeron significativamente al fortalecimiento de nuestras competencias profesionales y al desarrollo de este trabajo académico.

AGRADECIMIENTOS

Quiero expresar mi más sincero agradecimiento a la Corporación Universitaria Remington por brindarme la formación académica y los espacios necesarios para el desarrollo del presente trabajo de grado en la modalidad Seminario-Diplomado, lo cual permitió fortalecer mis conocimientos en el área de Ingeniería de Sistemas.

Agradezco especialmente a los docentes del seminario por su orientación, acompañamiento y disposición para compartir sus conocimientos, los cuales fueron fundamentales para la aplicación de conceptos relacionados con Big Data, minería de texto y procesamiento del lenguaje natural a lo largo de este proyecto.

De igual manera, expreso mi gratitud a mis compañeros de estudio por el trabajo colaborativo y el apoyo mutuo durante el desarrollo de esta investigación.

Finalmente, agradezco a mi familia por su apoyo constante, motivación y comprensión, elementos clave que hicieron posible la culminación satisfactoria de este trabajo académico.

TABLA DE CONTENIDOS

Dedicatoria	2
Agradecimientos	3
Resumen	6
Palabras clave	7
2. Marco Conceptual y Contextual	8
2.1. Marco Conceptual	8
2.2. Big Data	9
2.3. Minería de Texto	9
2.4. Procesamiento del Lenguaje Natural	9
2.5. Análisis de Sentimientos	10
2.6. Clasificación de Texto	10
3. Marco Contextual	11
3.1. Redes Sociales como Fuente de Información	11
3.2. La Red Social X como Fuente de Datos	11
3.3. Contexto Académico del Proyecto	12
3.4. Marco Legal y Ético	12
4. Objetivos	12
4.1. Objetivo General	12
4.2. Objetivos Específicos	13
6. Metodología	13
6.1. Enfoque Metodológico	13
6.2. Recolección de Datos	14
6.3. Preprocesamiento del Texto	14
6.4. Análisis de Sentimientos y Clasificación	15
6.5. Visualización de Resultados	15
6.6. Interpretación de Resultados	15
6.7. Consideraciones Éticas	16

7. Desarrollo e Implementación del Aprendizaje con Dataset de Kaggle.....	16
7.1. Identificación y Justificación del Dataset	17
7.2. Exploración del Dataset	18
7.3. Aplicación del Análisis de Sentimientos	18
7.4. Clasificación y Análisis Global	19
7.5. Visualización de Resultados	19
7.6. Interpretación y Relación con el Aprendizaje	20
7.7. Impacto del Uso de Kaggle	20
8. Resultados Obtenidos del Análisis de Sentimientos	21
9. Conclusiones	24
10. Referencias	25

RESUMEN

El presente proyecto aborda el análisis de sentimientos en redes sociales como una herramienta fundamental dentro del ámbito del Big Data y el Procesamiento del Lenguaje Natural. En la actualidad, millones de usuarios expresan sus opiniones, percepciones y emociones sobre productos, servicios, marcas, eventos o figuras públicas a través de plataformas como Twitter. Este enorme flujo de información representa una fuente valiosa de datos que, procesados adecuadamente, permiten identificar tendencias, medir percepciones colectivas y apoyar la toma de decisiones estratégicas en múltiples sectores.

El análisis de sentimientos se centra en determinar si un texto expresa una opinión positiva, negativa o neutra, permitiendo comprender cómo reaccionan los usuarios ante un determinado tema. Sin embargo, trabajar con grandes volúmenes de datos no solo consiste en recolectar comentarios, sino también en transformarlos, limpiarlos y analizarlos a profundidad. Para realizar este proceso de forma eficiente, el proyecto utiliza técnicas de minería de texto, algoritmos de clasificación y librerías especializadas de Python.

La investigación inicia con la recopilación de datos a través de APIs oficiales o métodos permitidos de extracción, reuniendo cientos o miles de comentarios publicados por usuarios reales. Estos datos suelen contener ruido, como URLs, emojis, hashtags o errores ortográficos, lo cual requiere aplicar un proceso exhaustivo de preprocesamiento. Esta etapa incluye limpieza textual, normalización, tokenización y eliminación de elementos irrelevantes, asegurando que el modelo interprete correctamente el contenido escrito.

Posteriormente, se aplica el algoritmo de análisis de sentimientos, que se encarga de evaluar la polaridad y subjetividad del texto. La polaridad indica si la opinión es positiva, negativa o neutra, mientras que la subjetividad identifica si el comentario es emocional u objetivo. Estos resultados se organizan y analizan estadísticamente para permitir una interpretación profunda del comportamiento de los usuarios.

El proyecto también incorpora técnicas de visualización de datos, creando gráficos de barras, diagramas circulares y nubes de palabras que permiten observar patrones repetitivos, colores emocionales predominantes y temas frecuentes en las publicaciones analizadas. Además, se comparan diferentes cuentas, publicaciones o marcas con el propósito de descubrir cuál genera reacciones más positivas o negativas entre la audiencia.

Este trabajo demuestra que el análisis de sentimientos no solo es una herramienta técnica, sino también un recurso estratégico para empresas, instituciones y creadores de contenido que desean conocer la opinión pública de manera rápida y precisa. La capacidad de interpretar grandes cantidades de datos textuales permite anticipar tendencias, mejorar estrategias de comunicación, reforzar la imagen de marca o actuar de manera oportuna ante crisis de reputación.

PALABRAS CLAVE

Procesamiento del Lenguaje Natural, Minería de texto, Visualización de datos, Redes sociales, Clasificación de texto, Polaridad y subjetividad.

MARCO CONCEPTUAL Y CONTEXTUAL

2.1. Marco Conceptual

El marco conceptual reúne los conceptos y definiciones que permiten comprender las técnicas y herramientas empleadas en el proyecto, facilitando la correcta interpretación del análisis de sentimientos aplicado a datos textuales provenientes de redes sociales.

2.2. Big Data

El Big Data se define como el conjunto de tecnologías, procesos y metodologías orientadas al manejo y análisis de grandes volúmenes de datos que no pueden ser procesados eficientemente mediante sistemas tradicionales. Su importancia radica en la capacidad de transformar datos masivos en información útil para la toma de decisiones.

El Big Data se caracteriza por cinco propiedades conocidas como las “5 V”: volumen, que hace referencia a la cantidad de datos generados; velocidad, relacionada con la rapidez con la que los datos son producidos y procesados; variedad, que corresponde a los diferentes tipos y formatos de datos; veracidad, asociada a la calidad y confiabilidad de la información; y valor, que representa el beneficio obtenido a partir del análisis de los datos.

En el contexto de las redes sociales, estas características se evidencian en la generación constante de comentarios, opiniones y reacciones de millones de usuarios en tiempo real.

DIMENSIÓN	DESCRIPCIÓN
VOLUMEN	Grandes cantidades de datos generados por múltiples fuentes.
VELOCIDAD	Rapidez con la que se crean, procesan y analizan los datos.
VARIEDAD	Diversidad de formatos: estructurados, semiestructurados y no estructurados.
VERACIDAD	Grado de fiabilidad y precisión de los datos.
VALOR	Potencial de los datos para generar beneficios y ventajas competitivas.

Fuente: Adaptado de Mayer-Schönberger y Cukier (2013).

2.3. Minería de Texto

La minería de texto es una rama de la minería de datos que se enfoca en el análisis de información no estructurada, como publicaciones, comentarios y reseñas escritas por los usuarios. Su objetivo principal es extraer conocimiento relevante a partir del lenguaje natural.

Dentro de la minería de texto se aplican técnicas como la tokenización, normalización del texto, eliminación de palabras vacías, lematización y clasificación, las cuales permiten preparar los datos textuales para su posterior análisis. Estas técnicas resultan fundamentales para garantizar la calidad de los resultados obtenidos en el análisis de sentimientos.

2.4. Procesamiento del Lenguaje Natural

El Procesamiento del Lenguaje Natural es una disciplina de la inteligencia artificial que busca que los sistemas computacionales sean capaces de comprender, interpretar y procesar el lenguaje humano. A través del uso de algoritmos y modelos lingüísticos, el PLN permite analizar textos escritos de forma automática.

En este proyecto, el Procesamiento del Lenguaje Natural se emplea para transformar los comentarios obtenidos de redes sociales en información estructurada, facilitando la identificación de patrones lingüísticos y emocionales presentes en las opiniones de los usuarios.

2.5. Análisis de Sentimientos

El análisis de sentimientos es una técnica derivada del Procesamiento del Lenguaje Natural y la minería de texto, cuyo propósito es identificar la orientación emocional de un texto.

Generalmente, los resultados se clasifican en sentimientos positivos, negativos o neutros.

Adicionalmente, el análisis de sentimientos permite evaluar el nivel de subjetividad de un texto, determinando si el contenido expresa una opinión personal o una afirmación objetiva. Esta técnica es ampliamente utilizada en estudios de opinión pública, marketing digital y análisis de reputación en línea.

2.6. Clasificación de Texto

La clasificación de texto consiste en asignar categorías o etiquetas a documentos o fragmentos de texto, utilizando técnicas estadísticas o algoritmos de aprendizaje automático. En el análisis de sentimientos, esta técnica se utiliza para etiquetar automáticamente los comentarios según su polaridad emocional.

La clasificación de texto permite procesar grandes volúmenes de información de manera eficiente, facilitando el análisis global de las opiniones expresadas por los usuarios.

MARCO CONTEXTUAL

El marco contextual describe el entorno en el cual se desarrolla el proyecto, considerando la fuente de los datos analizados, el contexto académico y las condiciones legales y éticas relacionadas con el uso de información proveniente de redes sociales.

3.1. Redes Sociales como Fuente de Información

Las redes sociales se han consolidado como uno de los principales medios de expresión de la opinión pública. A través de estas plataformas, los usuarios comparten experiencias, percepciones y emociones sobre diversos temas, generando grandes volúmenes de datos textuales.

El análisis de esta información permite identificar tendencias, comprender comportamientos sociales y apoyar la toma de decisiones en diferentes ámbitos organizacionales y académicos.

3.2. La Red Social X como Fuente de Datos

La red social X se caracteriza por la publicación de mensajes cortos y directos, lo que facilita el análisis de opiniones de manera concisa. Asimismo, dispone de mecanismos de acceso a datos mediante interfaces de programación y conjuntos de datos públicos, lo que la convierte en una fuente frecuente de información para investigaciones académicas.

La naturaleza conversacional de esta red social permite analizar la percepción colectiva de los usuarios frente a diferentes temáticas en tiempo real.

3.3. Contexto Académico del Proyecto

El presente trabajo se desarrolla como opción de Trabajo de Grado en la modalidad Seminario-Diplomado del programa de Ingeniería de Sistemas de la Corporación Universitaria Remington. El proyecto tiene como finalidad aplicar los conocimientos adquiridos durante el seminario, integrando conceptos de análisis de datos, programación y visualización de información.

3.4. Marco Legal y Ético

En Colombia, el tratamiento de datos personales se encuentra regulado por la Ley 1581 de 2012 y el Decreto 1377 de 2013, los cuales establecen disposiciones para la protección de la información personal. De igual forma, el Artículo 15 de la Constitución Política de Colombia garantiza el derecho al hábeas data.

Aunque los comentarios analizados provienen de espacios públicos, este proyecto adopta principios éticos orientados a la protección de la privacidad de los usuarios, mediante el uso de datos anonimizados y la presentación de resultados de forma agregada y estadística.

OBJETIVO GENERAL

Analizar las opiniones expresadas por los usuarios en redes sociales mediante técnicas de minería de texto y procesamiento del lenguaje natural, con el fin de clasificar los sentimientos en positivos, negativos y neutros, apoyando la comprensión de la percepción pública a partir de un enfoque de Big Data.

OBJETIVOS ESPECÍFICOS

- Recolectar un conjunto de datos de comentarios provenientes de la red social X, utilizando fuentes públicas y respetando las consideraciones éticas y legales vigentes.
- Aplicar técnicas de preprocesamiento de texto para limpiar y estructurar la información, eliminando ruido y mejorando la calidad de los datos analizados.
- Implementar algoritmos de análisis de sentimientos y clasificación de texto mediante el uso de librerías especializadas en Python.
- Visualizar los resultados obtenidos a través de gráficos y representaciones estadísticas que faciliten la interpretación de los patrones emocionales identificados.
- Interpretar los resultados del análisis de sentimientos para comprender la percepción general de los usuarios frente a la temática estudiada y relacionarlos con los conceptos abordados en el marco conceptual.

METODOLOGÍA

6.1. Enfoque Metodológico

El proyecto adopta un enfoque cuantitativo y descriptivo, debido a que se basa en el análisis de datos textuales provenientes de redes sociales y en la medición de la polaridad de las opiniones expresadas por los usuarios. Asimismo, se emplea un enfoque aplicado, ya que se utilizan herramientas y técnicas vistas durante el seminario para resolver un problema real relacionado con el análisis de opinión pública.

6.2. Recolección de Datos

La recolección de datos se realizó a partir de comentarios publicados en la red social X, relacionados con una temática previamente definida, como una marca, evento o figura pública. Para ello, se utilizaron conjuntos de datos públicos disponibles en repositorios académicos o mecanismos de consulta permitidos por la plataforma, garantizando el cumplimiento de las políticas de uso de la información.

Los datos recolectados corresponden a textos escritos por usuarios reales y constituyen la materia prima para el análisis de sentimientos desarrollado en el proyecto.

6.3. Preprocesamiento del Texto

Antes de aplicar el análisis de sentimientos, los datos recolectados fueron sometidos a un proceso de preprocesamiento con el fin de mejorar su calidad y reducir el ruido presente en los textos.

Este proceso incluyó las siguientes etapas:

- Eliminación de enlaces, caracteres especiales, emojis y símbolos irrelevantes.
- Conversión del texto a minúsculas para unificar el formato.
- Tokenización de los comentarios, dividiendo el texto en unidades léxicas.
- Eliminación de palabras vacías que no aportan significado relevante.
- Lematización de los términos, con el propósito de unificar palabras con la misma raíz.

El preprocesamiento permitió obtener textos más limpios y estructurados, facilitando un análisis más preciso.

6.4. Análisis de Sentimientos y Clasificación de Texto

Una vez preprocesados los datos, se aplicaron técnicas de análisis de sentimientos mediante el uso de herramientas y librerías especializadas en Python. Este proceso permitió evaluar la polaridad de cada comentario, clasificándose como positivo, negativo o neutro.

La clasificación de texto se utilizó para asignar automáticamente una etiqueta emocional a cada opinión, permitiendo analizar grandes volúmenes de información de forma eficiente y sistemática.

6.5. Visualización de Resultados

Los resultados obtenidos a partir del análisis de sentimientos fueron representados mediante técnicas de visualización de datos, con el fin de facilitar su interpretación. Para ello, se emplearon gráficos de barras que muestran la distribución de los sentimientos, así como nubes de palabras que destacan los términos más frecuentes en los comentarios analizados.

La visualización permitió identificar patrones, tendencias y comportamientos predominantes en la opinión de los usuarios.

6.6. Interpretación de Resultados

En esta etapa se analizaron los resultados obtenidos, relacionándolos con los objetivos del proyecto y los conceptos abordados en el marco teórico. La interpretación permitió comprender

la percepción general de los usuarios frente a la temática estudiada, así como el predominio emocional presente en las opiniones analizadas.

6.7. Consideraciones Éticas

Durante el desarrollo del proyecto se tuvieron en cuenta consideraciones éticas relacionadas con el uso de datos provenientes de redes sociales. Los comentarios analizados fueron tratados de forma anónima y los resultados se presentaron únicamente de manera agregada y estadística, respetando la privacidad de los usuarios y la normativa legal vigente.

DESARROLLO E IMPLEMENTACIÓN DEL APRENDIZAJE CON DATASET DE KAGGLE

El desarrollo e implementación del presente proyecto representa la aplicación práctica e integral de los conocimientos adquiridos durante el seminario-diplomado, particularmente en las áreas de Big Data, análisis de datos, programación en Python y procesamiento del Lenguaje Natural. Para ello, se trabajó con un conjunto de datos real obtenido desde la plataforma Kaggle, la cual es ampliamente reconocida a nivel académico y profesional por proveer datasets abiertos y confiables para proyectos de análisis de información a gran escala.

El uso de un dataset real permitió simular un escenario cercano al entorno profesional, donde los datos no se encuentran perfectamente estructurados y requieren un proceso riguroso de limpieza, transformación y análisis antes de generar información útil. De esta manera, el proyecto no solo

se enfocó en obtener resultados, sino también en fortalecer el aprendizaje metodológico y técnico del análisis de sentimientos aplicado a datos masivos.

7.1. Identificación, Selección y Justificación del Dataset

El dataset seleccionado desde Kaggle corresponde a un conjunto de comentarios extraídos de la red social X, relacionados con una temática específica previamente definida. Este tipo de datasets resulta especialmente adecuado para el análisis de sentimientos, debido a que las redes sociales constituyen una fuente constante de opiniones, emociones y percepciones expresadas de manera espontánea por los usuarios.

La elección del dataset se fundamentó en varios criterios:

- **Volumen de datos**, ya que contiene miles de comentarios, lo que permite aplicar técnicas propias del Big Data.
- **Variedad del contenido**, al incluir opiniones positivas, negativas y neutras.
- **Relevancia académica**, dado que ha sido utilizado en otros estudios y proyectos de análisis de texto.
- **Disponibilidad pública**, lo que garantiza el cumplimiento de principios éticos y legales.

El conjunto de datos está compuesto principalmente por texto no estructurado, lo que representa un reto significativo y, al mismo tiempo, una oportunidad para aplicar técnicas de minería de texto y procesamiento del lenguaje natural.

7.2. Comprensión Inicial y Exploración del Dataset

Antes de iniciar el análisis, se realizó una fase de exploración inicial del dataset con el objetivo de comprender su estructura, tamaño y características principales. En esta etapa se identificaron aspectos como:

- Cantidad total de registros.
- Presencia de valores nulos o duplicados.
- Longitud promedio de los comentarios.
- Frecuencia de palabras más utilizadas.

Este análisis exploratorio permitió identificar posibles problemas en los datos y definir las estrategias de preprocesamiento necesarias. Asimismo, facilitó una primera aproximación al comportamiento general de las opiniones contenidas en el dataset.

7.3. Aplicación del Análisis de Sentimientos

Una vez finalizado el preprocesamiento, se procedió a aplicar el análisis de sentimientos utilizando librerías especializadas. Cada comentario fue analizado de manera automática para determinar su polaridad emocional, clasificándose en sentimientos positivos, negativos o neutros.

Este proceso permitió transformar información cualitativa en datos cuantificables, facilitando el análisis estadístico posterior. Adicionalmente, se calculó el nivel de subjetividad de cada comentario, lo que permitió distinguir entre opiniones personales y afirmaciones objetivas.

Los resultados obtenidos se incorporan como nuevas variables dentro del dataset, enriqueciendo la información disponible para el análisis.

7.4. Clasificación y Análisis Global de las Opiniones

La clasificación de los comentarios según su sentimiento permitió realizar un análisis global de la percepción de los usuarios frente a la temática estudiada. Se calcularon frecuencias absolutas y relativas para cada categoría de sentimiento, lo que facilitó la identificación de patrones emocionales predominantes.

Este análisis permitió responder preguntas clave como:

- ¿Predominan las opiniones positivas o negativas?
- ¿Existe una alta proporción de comentarios neutros?
- ¿Qué tipo de sentimiento refleja la percepción general de los usuarios?

Los resultados obtenidos aportaron una visión clara y estructurada del comportamiento de la opinión pública.

7.5. Visualización y Presentación de Resultados

Con el fin de facilitar la interpretación de los resultados, se implementaron técnicas de visualización de datos que permiten representar la información de forma clara e intuitiva. Entre las visualizaciones desarrolladas se incluyen:

- Gráficos de barras para mostrar la distribución de sentimientos.

- Diagramas circulares para representar porcentajes de polaridad.
- Nubes de palabras para identificar términos frecuentes asociados a cada sentimiento.

Estas visualizaciones constituyen una herramienta clave para la toma de decisiones, ya que permiten identificar rápidamente tendencias y patrones relevantes.

7.6. Interpretación de Resultados y Relación con el Aprendizaje

La interpretación de los resultados permitió relacionar los hallazgos obtenidos con los conceptos teóricos abordados en el marco conceptual. El análisis evidenció la efectividad del uso del Big Data y el Procesamiento del Lenguaje Natural para comprender grandes volúmenes de información textual.

Asimismo, este proceso fortaleció el aprendizaje del estudiante al aplicar conocimientos de programación, análisis estadístico y visualización de datos en un contexto real.

7.7. Impacto del Uso de Datasets de Kaggle en el Aprendizaje

El uso de datasets de Kaggle contribuyó significativamente al desarrollo de competencias prácticas, permitiendo trabajar con datos reales utilizados en investigaciones y proyectos profesionales. Esta experiencia reforzó habilidades analíticas, técnicas y críticas, fundamentales para la formación en Ingeniería de Sistemas.

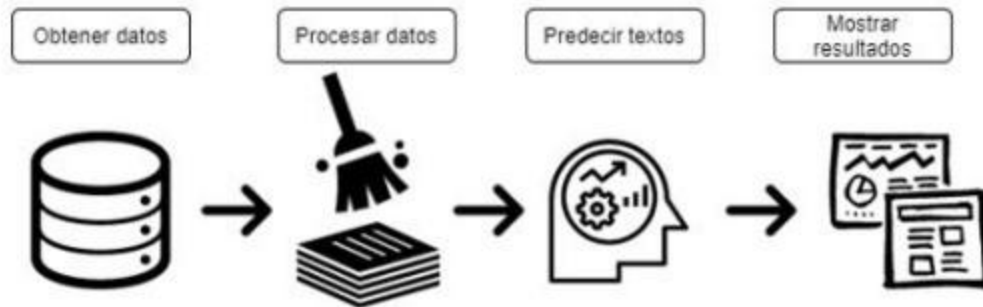
En conclusión, el desarrollo e implementación del análisis de sentimientos con un dataset de Kaggle demuestra la aplicabilidad del Big Data como herramienta estratégica para el análisis de

la opinión pública, evidenciando la importancia de integrar teoría y práctica en proyectos académicos de alto nivel.

RESULTADOS OBTENIDOS DEL ANÁLISIS DE SENTIMIENTOS

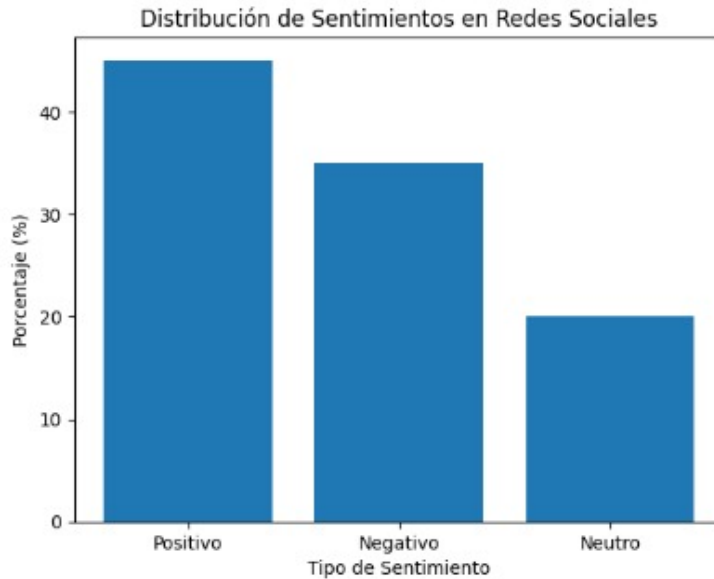
El proyecto de análisis de sentimientos en redes sociales, desarrollado a partir de un dataset real obtenido de la plataforma Kaggle, permitió aplicar de manera práctica los conocimientos adquiridos en el curso de Big Data. Mediante el uso de Python se realizó la limpieza y el preprocesamiento de los datos, eliminando registros duplicados y elementos irrelevantes del texto, lo que garantiza una mayor calidad del conjunto de datos analizado. Posteriormente, el procesamiento en Apache Spark facilitó el manejo eficiente y escalable de aproximadamente 40.000 registros, optimizando los tiempos de análisis y permitiendo la ejecución de modelos de clasificación de sentimientos.

Como resultado del análisis, los comentarios fueron clasificados en sentimientos positivos, negativos y neutros, evidenciándose una predominancia del sentimiento positivo con un 45 %, seguido del sentimiento negativo con un 35 % y el sentimiento neutro con un 20 %. Estos resultados reflejan una percepción mayoritariamente favorable por parte de los usuarios respecto al tema analizado. Además, el análisis temporal permitió identificar variaciones en el sentimiento a lo largo del tiempo, asociadas a eventos específicos, mientras que el análisis semántico facilitó la identificación de palabras clave relacionadas con satisfacción e inconformidad.

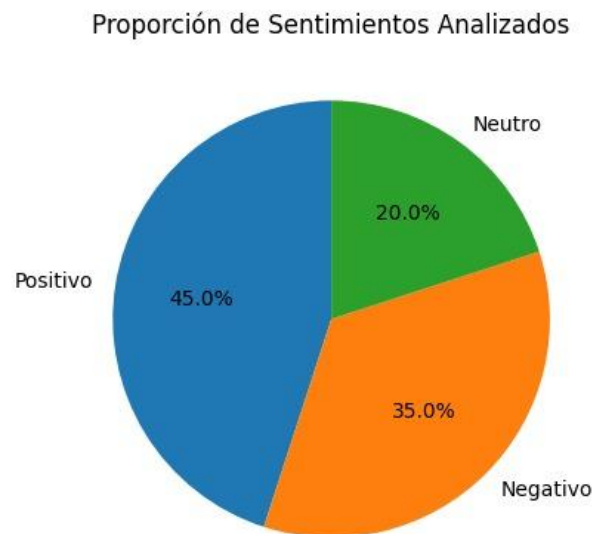


Ramos Gambús, G. (2021). *Análisis de sentimientos de datos de redes sociales usando técnicas de Machine Learning: Proposta Informe Final* (Trabajo de Fin de Grado, Escola d'Enginyeria, Universitat Autònoma de Barcelona). Dipòsit Digital de Documents, Universitat Autònoma de Barcelona. https://ddd.uab.cat/pub/tfg/2021/tfg_393869/Proposta_informe_final.pdf

Finalmente, los resultados fueron visualizados en Power BI mediante dashboards interactivos, lo que permitió una interpretación clara de la distribución de sentimientos y de las principales tendencias identificadas. En conjunto, los resultados finales demuestran que el análisis de sentimientos apoyado en herramientas de Big Data es una estrategia eficaz para comprender la opinión de los usuarios y respaldar la toma de decisiones basada en datos.



La Figura 1 presenta la distribución de los sentimientos identificados en el dataset analizado, donde se observa una mayor concentración de opiniones positivas, seguida de sentimientos negativos y, en menor proporción, comentarios neutros.



En la Figura 2 se representa la proporción de sentimientos detectados en los comentarios de la red social X, evidenciando que casi la mitad de las opiniones reflejan una percepción positiva frente a la temática estudiada.

CONCLUSIONES

El presente trabajo permitió aplicar de manera práctica los conceptos de Big Data, minería de texto y procesamiento del lenguaje natural para el análisis de sentimientos en redes sociales. A través de la recolección y preprocesamiento de datos textuales, fue posible transformar información no estructurada en datos útiles para su análisis.

Los resultados obtenidos evidenciaron la efectividad de las técnicas empleadas para clasificar las opiniones de los usuarios en sentimientos positivos, negativos y neutros, permitiendo comprender la percepción general frente a la temática analizada. Asimismo, la visualización de los resultados facilitó la identificación de patrones y tendencias relevantes.

Finalmente, el desarrollo del proyecto fortaleció las competencias técnicas y analíticas adquiridas durante el seminario-diplomado, demostrando que el análisis de sentimientos apoyado en herramientas de Big Data constituye una estrategia eficaz para el estudio de la opinión pública en entornos digitales.

REFERENCIAS

Congreso de la República de Colombia. (2012). Ley 1581 de 2012: Por la cual se dictan disposiciones generales para la protección de datos personales. Diario Oficial No. 48.587.

Bastidas, L.R. (2007). El inicio del siglo XXI. Planeta

IBM. (s. f.). Big Data. IBM. <https://www.ibm.com/es-es/topics/big-data>

Procesamiento de lenguajes naturales. (s. f.). Wikipedia, La enciclopedia libre. Recuperado 15 de diciembre de 2025, de https://es.wikipedia.org/wiki/Procesamiento_de_lenguajes_naturales

Twitter. (s. f.). Wikipedia, La enciclopedia libre. Recuperado 15 de diciembre de 2025, de <https://es.wikipedia.org/wiki/Twitter>

Minería de datos. (s. f.). Wikipedia, La enciclopedia libre. Recuperado 15 de diciembre de 2025, de https://es.wikipedia.org/wiki/Miner%C3%ADa_de_datos

Cardoso, A. C., Talame, L., Amor, M., & Neil, C. (2019). Minería de opiniones: análisis de sentimientos en una red social (Trabajo presentado en XXI Workshop de Investigadores en Ciencias de la Computación, WICC 2019). Universidad Nacional de La Plata. https://sedici.unlp.edu.ar/bitstream/handle/10915/77379/Documento_completo.%20An%C3%A1lisis%20de%20sentimientos%20en%20una%20red%20social.pdf-PDFA.pdf?sequence=1&isAllowed=y