

UNIB.E

UNIVERSIDAD IBEROAMERICANA DEL ECUADOR

FACULTAD DE: Comunicación y Tecnologías

CARRERA: Ingeniería de Software

TITULO

Modelo de análisis de datos en redes sociales para determinar las tendencias de los usuarios en relación con determinados productos

Autor (a):

Erick Daniel Erraez Guamán

Franklin Omar Ronquillo Pazmiño

Tutor (a):

Tonyse De La Rosa, MGs.

Quito, Ecuador

julio, 2024

DECLARACIÓN DE AUTORÍA Y AUTORIZACIÓN PARA LA DIFUSIÓN DEL TRABAJO DE INTEGRACIÓN CURRICULAR

1. Nosotros, **Erick Daniel Erraez Guamán** y **Franklin Omar Ronquillo Pazmiño**, declaramos en forma libre y voluntaria, que los criterios emitidos en el presente Trabajo de Integración Curricular, titulado: **“Desarrollar un modelo que permita analizar publicaciones de Facebook para predecir el sentimiento de los usuarios”**, previo a la obtención del título profesional de **Ingeniero en Desarrollo de Software**, así como también los contenidos, ideas, análisis, conclusiones y propuestas son exclusiva responsabilidad de mi persona, como autor/a.
2. Declaramos, igualmente, tener pleno conocimiento de la obligación que tiene la Universidad Iberoamericana del Ecuador, de conformidad con el artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT, en formato digital una copia del referido Trabajo de Integración Curricular para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública, respetando los derechos de autor.
3. Autorizamos, finalmente, a la Universidad Iberoamericana del Ecuador a difundir a través del sitio web de la Biblioteca de la UNIB.E (Repositorio Digital Institucional), el referido Trabajo de Integración Curricular, respetando las políticas de propiedad intelectual de la Universidad Iberoamericana del Ecuador.

Quito, DM., a los 24 días del mes de julio de 2024.



Erick Daniel Erraez Guamán

1725875296



Franklin Omar Ronquillo Pazmiño

1710332469

Quito, 7 de agosto 2024

CARTA DEL DIRECTOR DEL TRABAJO DE TITULACION

Sandino Jaramillo

Director(a) de la Carrera Ingeniería de Software

Presente.

Yo **TONYSE DE LA ROSA** Director(a) del Trabajo de Titulación realizado por los estudiantes **ERICK DANIEL ERRAEZ GUAMAN** y **FRANKLIN OMAR RONQUILLO PAZMIÑO** de la carrera de **INGENIERIA DE SOFTWARE** informo haber revisado el presente documento titulado **DESARROLLAR UN MODELO QUE PERMITA PREDECIR LOS SENTIMIENTOS DE LOS USUARIOS**, el mismo que se encuentra elaborado conforme al Reglamento de titulación establecido por la **UNIVERSIDAD IBEROAMERICANA DEL ECUADOR, UNIB.E** de Quito y el Manual de Estilo institucional; por lo tanto autorizo su presentación final para los fines legales pertinentes.

En tal virtud autorizo a los Señores a que concedan a realizar el anillado del trabajo de titulación y su entrega en la secretaria de la Escuela.

Atentamente,



Tonyse De La Rosa

Director del Trabajo de Titulación

ACTA DE APROBACION

UNIB.E

UNIVERSIDAD IBEROAMERICANA DEL ECUADOR

ACTA DE APROBACIÓN DEL TRABAJO DE INTEGRACIÓN CURRICULAR

Facultad: Comunicación y Tecnologías

Carrera: Ingeniería de Software

Modalidad: Semipresencial

Nivel: 3er nivel de Grado

En el Distrito Metropolitano de Quito a los dieciocho días del mes de septiembre del 2024 (18-09-2024) a las doce horas con cero minutos (12:00), ante el Tribunal de Presentación Oral, se presentó el señor: **ERRAEZ GUAMAN ERICK DANIEL**, titular de la cédula de ciudadanía No. **1725875296** a rendir la evaluación oral del Trabajo de Integración Curricular: "**Modelo de análisis de datos en redes sociales para determinar las tendencias de los usuarios en relación con determinados productos.**", previo a la obtención del Título de Ingeniero de Software. Luego de la exposición, el referido estudiante obtiene las calificaciones que a continuación se detallan:

	Calificación
Lectura del Trabajo de Integración Curricular	8.5/10
Evaluación Oral del Trabajo de Integración Curricular	8 /10
Calificación Final del Trabajo de Integración Curricular	8.3 /10

Para constancia de lo actuado, los miembros del Tribunal de Presentación Oral del Trabajo de Integración Curricular, firman el presente documento en unidad de acto, a los dieciocho días del mes de septiembre del 2024 (18-09-2024).


PhD. Luisa Taborda
DIRECTOR ACADEMICO


Mgst. Sandino Jaramillo
DIRECTOR DE CARRERA


Mgst. Tonyse de la Rosa
TUTOR


Mgst. Miguel Angel Fernandez
LECTOR





Campus Colón : 9 de Octubre N25-12 y Av. Colón, Quito 170522.
Campus Norte : José Queri y Av. Eloy Alfaro, Quito 170513



ACTA DE APROBACIÓN DEL TRABAJO DE INTEGRACIÓN CURRICULAR

Facultad: Comunicación y Tecnologías

Carrera: Ingeniería de Software

Modalidad: Semipresencial

Nivel: 3er nivel de Grado

En el Distrito Metropolitano de Quito a los dieciocho días del mes de septiembre del 2024 (18-09-2024) a las doce horas con cero minutos (12:00), ante el Tribunal de Presentación Oral, se presentó el señor: **RONQUILLO PAZMIÑO FRANKLIN OMAR**, titular de la cédula de ciudadanía No. **1710332469** a rendir la evaluación oral del Trabajo de Integración Curricular: **"Modelo de análisis de datos en redes sociales para determinar las tendencias de los usuarios en relación con determinados productos."**, previo a la obtención del Título de Ingeniero de Software. Luego de la exposición, el referido estudiante obtiene las calificaciones que a continuación se detallan:

	Calificación
Lectura del Trabajo de Integración Curricular	8.5/10
Evaluación Oral del Trabajo de Integración Curricular	8.3/10
Calificación Final del Trabajo de Integración Curricular	8.4/10

Para constancia de lo actuado, los miembros del Tribunal de Presentación Oral del Trabajo de Integración Curricular, firman el presente documento en unidad de acto, a los dieciocho días del mes de septiembre del 2024 (18-09-2024).


Ph.D. Luisa Taborda
DIRECTOR ACADÉMICO


Mgstr. Serdino Jaramillo
DIRECTOR DE CARRERA


Mgstr. Tonya de la Rosa
TUTOR


Mgstr. Miguel Ángel Fernández
LECTOR

AGRADECIMIENTO

Agradezco a mis amigos y seres queridos por su comprensión, ánimo y por estar siempre presentes en cada paso del camino además de expresar mi gratitud a todas las personas que participaron en este estudio, ya sea proporcionando información, participando en entrevistas o brindando su tiempo y ayuda de alguna manera, este trabajo no habría sido posible sin el apoyo y la colaboración de todos ustedes.

Erick Erraez

A mi familia por darme todo su apoyo y quererme por sobre todas las cosas, quiero agradecerles por su inquebrantable apoyo a lo largo de mis estudios. Sin sus palabras de aliento, paciencia y sabiduría, no habría logrado superar los desafíos que encontré en este proceso. Gracias por ser mi guía y por creer en mí en todo momento. A mis profesores gracias por su tiempo, por el apoyo, así como la sabiduría que me transmitieron en el desarrollo de mi formación profesional.

Franklin Ronquillo

DEDICATORIA

En tu honor, querido abuelito, dedico este logro que tanto significó para ambos. Tu generosidad, amor y sacrificio fueron la luz que guio mi camino hacia la educación. Siempre estaré agradecido por el apoyo incondicional que brindaste, por creer en mí y por hacer posible mis sueños, aunque ya no estés físicamente con nosotros, tu espíritu vive en cada página de esta tesis y en cada logro que alcanzo. Tu legado perdurará en mi corazón para siempre. A mi compañero Franklin porque con su ayuda no hubiéramos logrado esta meta.

Erick Erraez

Dedico este trabajo principalmente a mi familia por su apoyo y permitirme el haber llegado hasta este momento tan importante de mi formación profesional. A mama Blanquita por ser el pilar más importante y por demostrarme siempre su cariño y apoyo incondicional ya que ha sabido formarme con buenos sentimientos, hábitos y valores, lo cual me ayudado a salir adelante en los momentos más difíciles. A papa César que ha estado siempre apoyándome, sé que este momento es tan especial para el como para mí. A mis hermanas Gina y Verónica a quienes amo infinitamente, por compartir momentos significativos conmigo y por siempre estar dispuestas a escucharme y ayudarme en cualquier momento. A mi compañero Erick porque sin el equipo que formamos no hubiéramos logrado esta meta.

Franklin Ronquillo

ÍNDICE GENERAL

RESUMEN	xi
ABSTRACT	xii
INTRODUCCIÓN	1
CAPITULO I	2
EL PROBLEMA.....	2
Planteamiento del problema	2
Objetivo General.....	5
Objetivos específicos.....	5
Justificación e Impacto de la Investigación	5
CAPITULO II	8
MARCO TEÓRICO	8
Bases teóricas	12
Introducción al análisis del sentimiento.....	12
Técnicas de análisis de sentimiento	13
Arboles de decisión	15
Naive Bayes.....	16
Máquinas de soporte vectorial.....	16
Redes Neuronales	17
CAPITULO III	19
METODOLOGIA DE LA INVESTIGACIÓN.....	19
Naturaleza de la investigación	19
Técnica de recolección de datos.....	22
Operacionalización de la variable	22
Instrumento de recolección de datos	23
Validez y confiabilidad	24
Técnicas de Análisis de Información.....	24
Metodología del producto.....	24
CAPITULO IV	28
ANALISIS E INTERPRETACION DE LOS RESULTADOS.....	28
Resultado de la investigación	36
CAPITULO V	44
Conclusiones	44
Recomendaciones	45

REFERENCIAS BIBLIOGRAFICAS.....	46
ANEXOS	52

ÍNDICE TABLAS

Tabla 1.	Operacionalización de la variable	22
Tabla 2.	Análisis de datasets para entrenamiento del modelo.....	28
Tabla 3.	Comparar diferentes modelos Naive Bayes.....	29
Tabla 4.	Comparar diferentes modelos Maquina de Soporte Vectorial.....	30
Tabla 5.	Comparar diferentes modelos Red Neuronal.....	31

ÍNDICE DE FIGURAS

Figura 1.	Ciclo de vida SCRUM.	25
Figura 2.	Instalación de librerías.	33
Figura 3.	Carga y preprocesamiento de datos texto e imágenes.....	34
Figura 4.	Diseño e implementación del modelo.	34
Figura 5.	Implementación y entrenamiento del modelo.	35
Figura 6.	Funciones de evaluación.	35
Figura 7.	Rendimiento del modelo.	36
Figura 8.	Precisión del Modelo.....	37
Figura 9.	Consistencia del Modelo.....	38
Figura 10.	Contexto de las Publicaciones.....	39
Figura 11.	Opiniones de los Usuarios.....	39
Figura 12.	Diferentes tipos de Publicación	40
Figura 13.	Resultado del Análisis	41
Figura 14.	Clasificación de Texto e Imágenes.....	42

*Nombres y apellidos: Erraez Guamán Erick Daniel, Ronquillo Pazmiño Franklin Omar.
Nombre de la Tesis: Modelo de análisis de datos en redes sociales para determinar la
tendencia de los usuarios en relación con determinados productos. Carrera Ingeniería de
Software. Universidad Iberoamericana del Ecuador. Quito Ecuador. 2024. (70) pp*

RESUMEN

La presente investigación se centra en el desarrollo de un modelo de red neuronal capaz de predecir los sentimientos de los usuarios de Facebook en publicaciones de texto o imágenes. Es por ello que se trabajó con una metodología cuantitativa y documental, con un nivel descriptivo, diseño no experimental. La técnica de recolección de datos empleada fue la observación estructurada y como instrumento se manejó una escala de estimación que permitió validar el modelo, y se estructuró con un total de 10 preguntas validadas por expertos. En el desarrollo del trabajo se manejó una metodología ágil conocida como SCRUM, la misma que se basa en sprints. En conjunto con Python como lenguaje de programación y varias librerías como tensorflow, matplotlib, numpy. Se realizaron varias pruebas y ajustes al modelo, optimizando hiper-parámetros como el número de neuronas, capas ocultas y el tipo de funciones de activación, lo que permitió mejorar el rendimiento del modelo significativamente. En conclusión, se obtuvo un modelo capaz de predecir el sentimiento en texto o imágenes manteniendo los resultados en un margen de positivo, negativo y neutro mostrando una gráfica del porcentaje de acierto en cada uno de estos tres aspectos.

Palabras clave: Modelo, red neuronal, dataset, scrum, análisis de sentimientos

*Nombres y apellidos: Erraez Guamán Erick Daniel, Ronquillo Pazmiño Franklin Omar.
Nombre de la Tesis: Modelo de análisis de datos en redes sociales para determinar la
tendencia de los usuarios en relación con determinados productos. Carrera Ingeniería de
Software. Universidad Iberoamericana del Ecuador. Quito Ecuador. 2024. (70) pp*

ABSTRACT

The present research focuses on the development of a neural network model capable of predicting the sentiments of Facebook users in text or image posts. For this purpose, a quantitative and documentary methodology was used, with a descriptive level and non-experimental design. The data collection technique employed was structured observation, and an estimation scale was used as the instrument to validate the model, which was structured with a total of 10 questions validated by experts. In the development of the work, an agile methodology known as SCRUM was used, which is based on sprints. Python was used as the programming language, along with several libraries such as TensorFlow, Matplotlib, and NumPy. Various tests and adjustments were made to the model, optimizing hyperparameters such as the number of neurons, hidden layers, and the type of activation functions, which significantly improved the model's performance. In conclusion, a model was obtained that is capable of predicting sentiment in text or images, maintaining the results within a range of positive, negative, and neutral, and displaying a graph of the accuracy percentage in each of these three aspects.

Keywords: Model, neural network, dataset, scrum, sentiment analysis

INTRODUCCIÓN

Según Huang, Fang, & Zhao (2021), la analítica de datos es esencial para convertir estos grandes volúmenes de información en conocimiento accionable, permitiendo a las organizaciones adaptarse y prosperar en un entorno competitivo en constante cambio. En este contexto el presente trabajo se enfoca en la recopilación de información mediante el uso de dataset, Además, Según Gaurav Chauhan (2021), los clasificadores Naive Bayes se han convertido en una herramienta popular para la clasificación de texto debido a su alto rendimiento en problemas multiclase y a la simplicidad de su regla de independencia. En el presente trabajo se propone abordar tales desafíos mediante la formulación y aplicación del modelo Naive Bayes para manejar las publicaciones de texto y entrenar el modelo para predecir el sentimiento generado por la publicación en Facebook.

En este estudio, se propone desarrollar un modelo de análisis de datos usando Python ya que Según Mark Lutz (2019) "Python se ha convertido en el lenguaje elegido para la informática científica y la inteligencia artificial, y su popularidad sigue creciendo en otros campos" (p. 45). En particular, se empleará Python para la creación del modelo gracias a su versatilidad y su amplio repositorio de librerías para el manejo de Machine Learning además de su fácil aprendizaje. Por otro lado, como menciona Alex Miller (2020), "Los conjuntos de datos son la base sobre la cual se construyen y validan los modelos de aprendizaje automático" (p.12). En este contexto, la selección y preparación cuidadosa de los datasets son fundamentales para garantizar la precisión del modelo de Naive Bayes que se va a desarrollar, se utilizara datasets específicamente diseñados para el análisis de sentimientos en redes sociales.

CAPITULO I

EL PROBLEMA

Según De la Madrid (2019) en este apartado se va a redactar “origen, referentes empíricos, funciones y relaciones que revelen los diferentes hechos causales y vacíos de conocimientos” (p. 49). Es así como, de forma descriptiva se permite abordar la problemática en el ámbito real y la interrogante sobre la que gira la investigación. En este trabajo se busca explicar de manera clara como las redes sociales al manejar grandes volúmenes de datos sobre los sentimientos que generan las publicaciones, presentan actualmente aspectos que pueden ser aprovechados para la toma de decisiones, más allá de su función como medio de comunicación social.

Planteamiento del problema

En la actualidad según Castells, M. (2020) en su libro “La sociedad de la información y el futuro de la democracia.” argumenta que la conexión omnipresente es fundamental para la sociedad de la información, facilitada por internet y las redes sociales. Esta realidad permite un acceso más amplio a las nuevas tecnologías y a la información para los usuarios. Según Godoy y González (2019), estos usuarios, que no necesariamente son expertos en el manejo de grandes volúmenes de datos, pueden beneficiarse de la capacidad de generar, capturar, almacenar y analizar información de manera automática y rápida.

En este contexto, es evidente que la capacidad de mantenerse siempre conectado y acceder a una gran cantidad de información es crucial, especialmente para organizaciones que requieren entender y adaptarse a las preferencias y demandas de sus usuarios en el entorno digital. Sin embargo, surge un desafío adicional: la falta de un modelo predictivo que pueda anticipar la aceptación de las publicaciones que se van a realizar en la red social Facebook. Esto significa que la mayoría de las organizaciones no puede garantizar de manera eficiente la relevancia y el impacto de su contenido en la audiencia objetivo.

Para comprender el concepto de Análisis de Datos, es esencial entender que implica el estudio detallado de conjuntos de información con el fin de extraer conclusiones que respalden la toma de decisiones de una empresa o entidad (Westreicher, G, 2020, pág. 1). La experiencia acumulada en el mercado ha llevado a muchas empresas a adoptar el análisis de datos como una herramienta fundamental para mejorar la eficiencia en la toma de decisiones y reducir costos. Además, el análisis de datos puede proporcionar insights (datos concretos que aportan información) valiosos que permiten a las organizaciones identificar oportunidades de crecimiento, anticipar tendencias del mercado. Según (González, 2019) el análisis de datos permite valorar los perfiles de los usuarios y de esta forma generar promociones o formas de conocer sus clientes o usuarios.

En el mundo actualmente, los datos se han convertido en el insumo más importante para las empresas debido a la globalización, el acceso inmediato a la información, las nuevas tecnologías y los desafíos organizacionales. Por tanto, es necesario desarrollar un modelo de análisis exhaustivo en las redes sociales para determinar el sentimiento de los usuarios hacia las publicaciones, usando técnicas de Machine Learning, en la actualidad algunos sistemas de Análisis de Sentimientos (AS) se basan en técnicas de aprendizaje automático para el proceso de clasificación de sentimientos.

Actualmente, los enfoques tradicionales de análisis de sentimiento pueden no ser suficientes debido a su incapacidad para adaptarse a las complejidades de los datos no estructurados y en tiempo real (Pak & Paroubek, 2010). Es así que el análisis de sentimientos mediante software ha enfrentado diversos problemas relacionados con la calidad de los datos, la rapidez en el procesamiento y la precisión en contextos globales y latinoamericanos. En cuanto a los datos, uno de los desafíos más destacados es la heterogeneidad y ambigüedad de los textos en redes sociales, lo cual dificulta la extracción precisa de sentimientos.

Según Pang y Lee (2008), los datos provenientes de redes sociales son a menudo ruidosos, con abreviaciones, errores ortográficos y uso de jerga, lo que complica el análisis automatizado. Adicionalmente, la rapidez en el procesamiento es crucial debido al volumen masivo de información generada en tiempo real. Bollen, Mao y Zeng (2011) señalan que procesar y analizar grandes volúmenes de datos en

tiempo real es esencial para capturar tendencias y cambios en el sentimiento oportunamente, pero es un desafío técnico considerable por la necesidad de infraestructura computacional robusta.

En el contexto latinoamericano, el análisis de sentimientos enfrenta problemas adicionales. En el ámbito global, la diversidad lingüística y cultural implica que los modelos de análisis deben ser adaptables a diferentes idiomas y contextos culturales para ser efectivos. Cambria et al. (2017) destacan que los algoritmos de análisis de sentimientos deben considerar las peculiaridades idiomáticas y culturales para evitar errores en la interpretación de los datos.

En el mismo orden de ideas, Ortega et al. (2019) resalta que, en muchos países latinoamericanos, la infraestructura tecnológica aún está en desarrollo, lo que limita la capacidad de implementar soluciones avanzadas de análisis de sentimientos. Además, estadísticas indican que solo un 55% de la población latinoamericana tiene acceso a internet de calidad, lo que afecta la representatividad de los datos recolectados de redes sociales (Banco Mundial, 2020). Esta situación subraya la necesidad de desarrollar enfoques específicos y adaptativos que consideren las realidades tecnológicas y socioculturales de la región.

Así también en Ecuador, se puede referir el potencial del análisis de datos de redes sociales para mejorar la toma de decisiones en las organizaciones. Tal como lo manifiesta Rangel (2020), es necesario disponer de una herramienta para identificar las necesidades de los clientes de una empresa. Es así como, el análisis de sentimientos ayuda a las empresas a mejorar su servicio al cliente y aumentar sus ventas. De igual manera, cuando se maneja opiniones en el campo político, las redes sociales al ofrecer otro tipo de variables ayudan al “análisis de sentimiento sobre las opiniones de los usuarios, a fin de poder determinar si se trata de sentimientos positivos, negativos o neutros” (UTE, 2018); sin embargo, el volumen masivo y la naturaleza no estructurada de los datos generados en plataformas como Twitter, Facebook e Instagram presentan desafíos significativos para su análisis y clasificación efectiva (Liu, 2012).

El problema radica en que actualmente existe un desinterés sobre las opiniones que se realizan en publicaciones de Facebook; ya que esta red social es considerada

como una herramienta comunicacional mixta y no predictiva. Es así como, organismos, organizaciones, empresas publican de manera continua en esta red social; sin embargo, no se trata de entender los sentimientos del usuario, para toma de decisiones en función de los beneficios que pueden prestar las redes sociales y emprender de manera más efectiva campañas de marketing que permita la mejora en su accionar y evitar el fracaso

Es por esto por lo que, se pretende desarrollar un modelo de inteligencia artificial (IA) que pueda analizar de manera precisa y eficiente grandes volúmenes de datos de redes sociales (Facebook) para predecir los sentimientos de los usuarios. Este modelo debe ser capaz de manejar la variabilidad lingüística y contextual inherente a los datos de redes sociales, garantizando al mismo tiempo una alta precisión en la clasificación de sentimientos como positivos, negativos y neutros (Cambria et al., 2013). Ante esta perspectiva, surge la siguiente interrogante:

¿Cómo desarrollar un modelo de Machine Learning (ML) que permita analizar publicaciones de Facebook para predecir los sentimientos de los usuarios?

Objetivos:

Objetivo General

Desarrollar un modelo de ML que permita el análisis de las publicaciones de Facebook para predecir el sentimiento de los usuarios.

Objetivos específicos

1. Seleccionar diferentes datasets de imágenes y texto para entrenar el modelo.
2. Comparar distintos modelos de Inteligencia artificial para el análisis de sentimientos y datos
3. Implementar las diferentes estructuras de aprendizaje supervisado.
4. Validar el modelo de análisis de sentimiento del usuario con datos reales de Facebook.

Justificación e Impacto de la Investigación

La creación de un modelo de análisis de sentimiento en redes sociales se posiciona como una herramienta esencial para comprender los sentimientos de los usuarios en sus publicaciones. Según Cabeza (2023), “Este análisis representa una valiosa vía para entender en profundidad cómo los usuarios perciben y experimentan las publicaciones, permitiendo a las empresas ajustar sus estrategias en función de estas percepciones.” Este enfoque permite examinar un amplio espectro de datos provenientes de plataformas sociales, desde comentarios y menciones hasta opiniones expresadas mediante hashtags, proporcionando una visión detallada de la interacción de los usuarios en Facebook.

Según Cevallos P. (2023), Facebook es la red social más utilizada mundialmente con aproximadamente 2,740 millones de usuarios. En este contexto, se ha decidido la utilización de Facebook debido a la cantidad masiva de publicaciones textuales que puede proveer para el entrenamiento del modelo y el posterior análisis de datos. Además, se destaca la importancia del uso de herramientas tecnológicas como Python, como sugiere Ritter (2023), las cuales amplifican la capacidad al recopilar, analizar y visualizar datos sociales, dando a las empresas la oportunidad de ajustar estratégicamente sus productos según los sentimientos de los usuarios, aprovechando al máximo las publicaciones realizadas en Facebook.

La importancia de la investigación propuesta se centra en el desarrollo de un modelo de clasificación de sentimientos basado en para analizar publicaciones de texto en Facebook. El objetivo es desarrollar un modelo para que devuelva una predicción del sentimiento expresado por los usuarios (positivo, negativo o neutro). Según Carrasco, Bueno & Montero (s.f.), aunque un algoritmo es simple, eficiente y preciso, destaca por su potencia predictiva ya que su principal ventaja es la utilización de un enfoque probabilístico, por ende, todos los cálculos se realizan en tiempo real y, por tanto, los resultados se obtienen inmediatamente además nos permite clasificar la relevancia del texto y clasificarlo como texto deseado, no deseado, spam, publicidad o comentarios maliciosos.

En conclusión, el análisis de sentimiento en redes sociales representa una herramienta invaluable debido a sus múltiples capacidades ya que gracias a esto podemos predecir el impacto de la publicación hacia el sentimiento de un usuario dándonos un resultado acerca de si la publicación va a hacer aceptado o no. La

integración de herramientas tecnológicas y la aplicación de este análisis en la toma de decisiones empresariales pueden influir significativamente en la mejora continua de los productos, permitiendo así una respuesta más alineada con las expectativas y necesidades de los consumidores.

Alcance de la investigación

Este proyecto desarrollará un modelo para análisis de sentimientos en publicaciones de Facebook, clasificándolas en positivas, negativas o neutras. Se seleccionarán datasets textuales de diversas fuentes, incluyendo Kaggle y datos de Facebook, para entrenar el modelo. Comparará con modelos como SVM y Random Forest en términos de precisión y otras métricas. Se preprocesarán los datos aplicando técnicas como TF-IDF y embeddings de palabras, y se optimizarán hiperparámetros. El modelo se integrará en Facebook para análisis en tiempo real, validándose con datos reales, se centrará en textos en español y respetará las políticas de privacidad de Facebook

CAPITULO II

MARCO TEÓRICO

Para empezar Castillo, Gómez, Taborda y Mejía (2021) en su libro “¿Cómo investigar en la UNIBE?” dice que “Después de definir el problema o tema de investigación, y establecer los propósitos u objetivos, se seleccionan las investigaciones anteriores que nos dan información sobre el tema. Esta información nos sirve como base para seguir investigando y aprendiendo sobre el tema.” (p. 58). Este proceso es fundamental para desarrollar una posición investigativa sólida y establecer un discurso y perspectiva analítica coherente. Los componentes clave de esta etapa incluyen la revisión de antecedentes de investigación, la exploración de bases teóricas y opcionalmente la consideración de fundamentos legales.

Antecedentes de la Investigación

A nivel internacional, el análisis de datos en redes sociales es un área de investigación muy activa, como menciona Quintana (2021), en su trabajo realizado en España “Análisis de los procesos de tratamiento de información en un estudio de análisis de sentimiento utilizando la tecnología de Google” El objetivo es mostrar cómo realizar estudios de neuromarketing con Big Data en Google Cloud Platform, con un enfoque en el análisis de sentimiento de tuits sobre "neuromarketing", se utilizó como plataforma de investigación Google Cloud Platform, las áreas que trabaja se pueden agrupar en aprendizaje automático, infraestructura, plataforma, aceleradores y gestión de APIs, se concluyó que "neuromarketing" se usa principalmente de manera neutra o positiva, con escaso uso negativo, atribuido a referencias a la pandemia y a la falta de claridad en la intención de compra.

Este estudio resalta la vitalidad de la investigación en análisis de datos en redes sociales a nivel global, particularmente en el contexto del análisis de sentimientos utilizando tecnología de Google. Al centrarse en el análisis de sentimiento de tuits a través de Google Cloud Platform, este trabajo proporciona un valioso marco para la realización de estudios de con Big Data. Al categorizar las áreas de trabajo de

Google Cloud Platform, se subraya la amplia gama de herramientas disponibles para llevar a cabo investigaciones significativas en este campo.

Así también, Cedeno & Vargas (2020), en su investigación realizada en Panamá "Aprendizaje automático aplicado al análisis de sentimientos", Con la expansión de Internet, abunda la información en línea, incluyendo opiniones de usuarios sobre productos, servicios y gestiones. Estas opiniones reflejan tanto satisfacción como insatisfacción, siendo cruciales para evaluar experiencias y tomar decisiones, este trabajo propone la utilización de técnicas de algoritmos supervisados para observar su utilización y ver el rendimiento de diferentes modelos de algoritmos supervisados para medir la efectividad en la clasificación de un conjunto de datos, se consideró que esta investigación fue positiva y eficiente en cuanto a la metodología propuesta y por otro lado se pudo ver las ventajas que ofrecen el aprendizaje automático en cuanto a la clasificación y procesamiento de texto no estructurado.

La expansión de Internet ha generado un crecimiento exponencial de información en línea, incluyendo opiniones de usuarios sobre productos, servicios y gestiones, que son vitales para la evaluación y toma de decisiones. Este estudio aboga por el uso de algoritmos supervisados para analizar estas opiniones y evaluar diferentes modelos en la clasificación de datos. Se destaca la eficacia y eficiencia metodológica de la investigación, así como las ventajas del aprendizaje automático en el procesamiento de texto no estructurado.

Además, en el estudio de Ramos (2021) realizado en Barcelona, titulado "Análisis de sentimientos de datos de redes sociales usando técnicas de Machine Learning", se planteó como objetivo principal recopilar textos de múltiples redes sociales y crear dashboards para presentar la información de manera clara y simplificada. Para este proyecto, se optó por la metodología ágil Scrum debido a su flexibilidad y capacidad de adaptarse a proyectos dinámicos, lo que permitió gestionar eficazmente las tareas de análisis de datos y desarrollo. Asimismo, se eligió Git para el control de versiones por su capacidad de facilitar la colaboración entre los integrantes del proyecto y el manejo eficiente de múltiples iteraciones en el proceso de desarrollo.

Un hallazgo clave del estudio fue que, entre la recopilación de publicaciones y su posterior clasificación, la fase de limpieza de datos resulta crucial para el éxito del

proyecto. Este proceso debe equilibrar la conservación de la esencia del texto o imagen con la eliminación de elementos que puedan generar ruido en el clasificador. El modelo de predicción de sentimientos utilizado en el estudio devuelve un único resultado por publicación, clasificando cada texto como positivo, negativo o neutro. Esto destaca la importancia de la limpieza de datos, ya que afecta directamente la precisión de la clasificación final.

El análisis de sentimientos tiene un gran margen de mejora, para clasificar los posts como positivos, negativos o neutros, se utiliza el modelo desarrollado donde el proceso comienza con la recolección de las publicaciones (imagen o texto), seguida de una etapa de preprocesamiento. Posteriormente el modelo de clasificación predice el sentimiento de la publicación, Las redes sociales han transformado la expresión de opiniones y sentimientos, siendo vitales para la comunicación social. Clasificar los posts en positivos, negativos y neutros puede proporcionar información valiosa, pero hacerlo manualmente resultaría ineficiente debido a la gran cantidad de contenido generado en estas plataformas.

A nivel nacional según, Calatrava (2019), en su estudio realizado en la ciudad de Quito “Desarrollo de una solución para la visualización y análisis de sentimiento en base a los mensajes que realizan los usuarios a una empresa mediante una red social.” El objetivo de esta investigación es solucionar los problemas de las empresas con los clientes, considerando que una gran proporción de estas empresas y consumidores están en Twitter. La analítica de texto usa algoritmos de aprendizaje automático para puntuar sentimientos entre 0 y 1. Este trabajo introduce métricas nuevas para capturar un rango de sentimientos más amplio, partiendo más allá de las clasificaciones tradicionales: negativo, positivo y neutral. Los resultados obtenidos sobre este estudio de análisis de sentimientos (AS) aumentan el conocimiento en inteligencia de negocios y ofrecen datos cruciales para mejorar las estrategias de relación con los clientes.

Esta investigación examinó el uso del análisis de datos en redes sociales para la mejora de la atención al cliente en empresas ecuatorianas ya que actualmente las empresas buscan superar a sus competidores de una manera eficaz, enfocándose en una adaptación constante de las estrategias más efectivas utilizadas en el

mercado con el fin de atraer la mayor cantidad de clientes, por ello muchas empresas con el uso de redes sociales logran ampliar ese rango poblacional.

Por otro lado, Moran (2022), en su trabajo realizado en la ciudad de Esmeraldas “Una comparación de las plataformas de análisis de sentimientos en Ingeniería de software: un estudio experimental” el objetivo de esta investigación fue determinar la efectividad de diversas herramientas y plataformas de análisis de sentimientos. Para lograrlo, fue necesario recopilar información ofrecida por empresas especializadas en el desarrollo de software para el beneficio de la sociedad, para comparar plataformas de análisis de sentimientos en ingeniería de software, se utilizaron conjuntos de datos obtenidos de Kaggle, una plataforma especializada en ciencia de datos, análisis predictivo y machine learning, en esta investigación se compararon plataformas de análisis de sentimientos en ingeniería de software usando métricas de la ISO/IEC 25010 para evaluar su eficacia y calidad.

Este trabajo aporta a la investigación, en la descripción que el análisis de sentimientos surge de los algoritmos de inteligencia artificial y el machine learning. Hoy las empresas dedicadas al desarrollo de software aseguran que la información obtenida mediante el análisis de sentimientos es importante para ellos, ya que así pueden analizar lo aceptado puede ser un producto para sus usuarios. Además, Moran destaca la importancia del análisis de sentimientos, también conocido como minería de opiniones es uno de los temas de investigación más relevantes en la ingeniería de software contemporánea.

De modo similar, Salgado y Trujillo (2024), en su trabajo efectuado en la ciudad de Quito “Análisis de sentimientos en datos de redes sociales: aplicación de técnicas de procesamiento de lenguaje natural(PLN) y machine learning(ML) para analizar opiniones y sentimientos en datos de redes sociales en el contexto de sistemas de información”, su propósito es la combinación de PLN y ML en el análisis de sentimientos en redes sociales revela su capacidad para comprender la complejidad del lenguaje humano, la interpretación efectiva de opiniones y emociones en plataformas digitales ofrece una valiosa herramienta para comprender la percepción pública, los resultados de la investigación destaca la efectividad del Análisis de Sentimientos en diversas áreas, desde la evaluación de la imagen de marca hasta la

comprensión de tendencias y preferencias del público. Estas técnicas son valiosas para informar estrategias empresariales y de marketing.

El análisis de sentimiento (AS) se puede definir como la tarea de asignar una etiqueta de sentimiento a un texto, indicando si el texto es positivo, negativo o neutro. La etiqueta de sentimiento puede ser binaria (positivo o negativo) o ternaria (positivo, negativo o neutro). Ya que según La convergencia de técnicas de procesamiento de lenguaje natural (PLN) y machine learning (ML) en el análisis de sentimientos en datos de redes sociales demuestra su capacidad para desentrañar la complejidad del lenguaje humano. Estas reglas suelen ser sencillas y fáciles de entender, pero pueden ser limitadas en su capacidad para identificar matices de sentimiento como ejemplo la palabra "me gusta" indica un sentimiento positivo, mientras que la palabra "odio" indica un sentimiento negativo.

Bases teóricas

Introducción al análisis del sentimiento

El análisis de sentimientos es una técnica de procesamiento del lenguaje natural que clasifica las emociones expresadas en textos como positivos, negativos o neutros. Se utiliza en redes sociales, encuestas y reseñas para comprender las opiniones de los usuarios, esta herramienta ayuda a las empresas a mejorar la experiencia del cliente y facilita la toma de decisiones basada en la percepción de los usuarios. Ya que según Salgado y Trujillo (2024) La convergencia de técnicas de procesamiento de lenguaje natural (PLN) y machine learning (ML) en el análisis de sentimientos en datos de redes sociales demuestra su capacidad para desentrañar la complejidad del lenguaje humano. Estas reglas suelen ser sencillas y fáciles de entender, pero pueden ser limitadas en su capacidad para identificar matices de sentimiento como ejemplo la palabra "me gusta" indica un sentimiento positivo, mientras que la palabra "odio" indica un sentimiento negativo.

Por otro lado, Castillo y Caicedo (2019) se refieren al análisis de sentimientos (Sentiment Analysis), enfocado en examinar el tono emocional (positivo, negativo o neutro) que las personas plasman en los textos que escriben sobre un tema particular en medios digitales como Facebook, Twitter, LinkedIn, foros, microblogs, entre otros. El análisis de sentimiento también se puede utilizar para detectar

tendencias y patrones en la opinión pública. Esto puede ser útil para las empresas para comprender mejor las necesidades y deseos de sus clientes, así como para los gobiernos para monitorear la opinión pública sobre temas sociales y políticos. A partir de estos aportes, se concluye que el análisis de sentimientos es una herramienta poderosa para extraer información valiosa de grandes volúmenes de datos textuales en línea. Su capacidad para identificar el sentimiento predominante en las publicaciones permite a las organizaciones tomar decisiones más informadas y estratégicas, ya sea en el ámbito comercial o político.

Técnicas de análisis de sentimiento

El análisis de sentimiento es una disciplina dentro del procesamiento de lenguaje natural que busca identificar, extraer y cuantificar información subjetiva de textos. Se enfoca en determinar la actitud, emociones u opiniones expresadas en un texto, ya sea positivas, negativas o neutras. Esta técnica es fundamental en la comprensión de la percepción pública, la toma de decisiones empresariales basadas en la retroalimentación de clientes y la comprensión de tendencias en redes sociales. Como menciona Quintana (2021), en los últimos años el Big Data se abre camino entre las principales herramientas de análisis de mercado, vinculándose a las técnicas de Machine Learning con el fin de aprender sobre los datos que se posee. Una de las áreas de mayor crecimiento es el Procesamiento de Lenguaje Natural que proporciona al investigador datos sobre estructuras y significados de texto.

Por otro lado, Moran (2022) describe que el análisis de sentimientos surge de los algoritmos de inteligencia artificial y el machine learning. Hoy en día las empresas que se dedican al desarrollo de software aseguran que la información obtenida a través del análisis de sentimientos es sumamente importante para ellos ya que de esta manera pueden analizar cuan aceptado puede ser un producto para sus usuarios. Además, Moran destaca la importancia del análisis de sentimientos, es uno de los temas de investigación más relevantes en la ingeniería de software contemporánea, este campo se enfoca en examinar las emociones y opiniones expresadas en textos, facilitando una comprensión más profunda de las percepciones del cliente. A partir de los aportes de Moran (2022), se concluye que el

análisis de sentimientos no solo es una herramienta clave para las empresas en la evaluación de sus productos, sino que también constituye un campo de estudio fundamental en la evolución de la inteligencia artificial y el desarrollo de software.

Aprendizaje Automático (Machine Learning)

Es una rama de la inteligencia artificial que se enfoca en el desarrollo de algoritmos y modelos que permiten a las máquinas aprender y hacer predicciones basadas en datos. Según Manrique (2020) La popularidad de la Inteligencia Artificial, en especial la rama de Machine Learning(ML) se ha incrementado en los últimos tiempos en desarrollar e implementar aplicaciones diversas para la resolución de problemas. Por otro lado, Noguera (2023) recalca el machine learning es un subcampo de la inteligencia artificial, que permite a los sistemas informáticos aprender y mejorar automáticamente a partir de la experiencia sin ser programados explícitamente. En este sentido Machine Learning(ML) se ha convertido en una de las disciplinas científicas más aplicadas en la actualidad, Esto permite que los sistemas se perfeccionen de manera autónoma, sin intervención humana. Hay varios tipos de aprendizaje en esta área: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo.

Machine Learning en el análisis de sentimiento

El análisis de sentimiento, una disciplina esencial en el procesamiento de lenguaje natural busca identificar, extraer y cuantificar información subjetiva de textos, revelando actitudes, emociones u opiniones expresadas. Según Ramos (2021) el análisis de sentimientos tiene un gran margen de mejora, desde analizar posts en una red social hasta calificar el futuro post antes de ser enviado. Cedeno (2020) subraya debido a la gran cantidad de opiniones, comentarios y sugerencias de los usuarios es muy importante explorar, analizar y organizar sus puntos de vista para tomar mejores decisiones.

En este contexto, el Machine Learning ha revolucionado el análisis de sentimiento, como señalan Agarwal y Mittal (2020), también Zhang et al. (2018) recalca los enfoques basados en Machine Learning, como las redes neuronales y los algoritmos de aprendizaje profundo, han mejorado significativamente la capacidad de las

máquinas para comprender y clasificar emociones en el texto. dando así que estos avances en el campo del Machine Learning continúan evolucionando, con el desarrollo de nuevas arquitecturas de modelos y técnicas de procesamiento para mejorar la precisión del análisis de sentimiento en diferentes contextos y lenguajes, lo que demuestra el papel fundamental del Machine Learning en la evolución y aplicaciones del análisis de sentimiento.

Machine Learning modelos supervisados

Son los algoritmos más sencillos de realizar y comprender, en ellos se parte de un conocimiento a priori. Como mencionan Cedeno & Vargas (2020) El aprendizaje supervisado es una técnica que utiliza datos de entrenamiento conocidos para deducir una función que mapea entradas (X) a salidas (Y) de la mejor manera posible. Los datos de entrenamiento consisten en tuplas (X, Y), donde X son las variables predictoras y Y es la variable objetivo. Según el tipo de Y, el aprendizaje supervisado se divide en dos categorías: ***clasificación*** y ***regresión***. En la clasificación, la salida Y es una etiqueta discreta, ya sea binaria o múltiple. En la regresión, Y es una variable continua, y se busca predecir su valor aproximando la relación de dependencia entre la variable dependiente y las variables independientes mediante un modelo matemático, a menudo representado como una línea recta.

Arboles de decisión

El análisis de sentimiento, una herramienta crucial en el procesamiento de lenguaje natural se enfoca en identificar y cuantificar las emociones, actitudes u opiniones expresadas en textos. Autores como Arana (2021) resaltan su importancia Los modelos de aprendizaje automático (machine learning) supervisados de clasificación mediante particiones binarias recursivas, también llamados "árboles de decisión" se encuentran entre los más utilizados en la ciencia de datos, no sólo por su interpretabilidad y su performance sino también por ser la base de los modelos más potentes utilizados en la actualidad Pedro (2019) enfatizan su aplicación en la internacionalización de los mercados y la creciente intensificación de la competencia, obliga a los directivos de empresas a tomar decisiones precisas en el

menor tiempo posible. Los árboles de decisión, por otro lado, son modelos predictivos que utilizan una estructura de árbol para representar y clasificar datos.

Naive Bayes

El Clasificador Naive Bayes es un algoritmo de aprendizaje automático supervisado utilizado para tareas de clasificación, como la clasificación de texto. Para realizar estas tareas, se basa en principios de probabilidad, a diferencia de los clasificadores discriminativos, como la regresión logística, Naive Bayes pertenece a la familia de algoritmos de aprendizaje generativo. Esto significa que busca modelar la distribución de los datos de entrada dentro de una clase o categoría específica. No aprende qué características son más importantes para distinguir entre clases. Para comprender mejor el concepto se debe recurrir al Teorema de Bayes mismo que permite "invertir" las probabilidades condicionales las cuales representan la probabilidad de que ocurra un evento dado que ya ha ocurrido otro evento. (IBM, s.f.)

Máquinas de soporte vectorial

El análisis de sentimiento, fundamental en el procesamiento de lenguaje natural, se enfoca en identificar y cuantificar emociones, actitudes u opiniones en textos. Autores como Alemán (2021) resaltan su importancia en los últimos años el auge de redes sociales ha permitido que los usuarios utilicen estos recursos para expresar su opinión sobre cualquier tema o producto, considerándose ésta más auténtica y fiel a lo que resultaría aplicando una encuesta, mientras que Navarro (2023) resalta que en el vasto mundo del procesamiento del lenguaje natural, el análisis de sentimientos ha surgido como una herramienta esencial para descifrar y comprender las opiniones y emociones contenidas en textos.

Autores como Romero, Martelo, Corredor, Sánchez & Ospina (2020) recalcan que, aunque originariamente las SVMs fueron pensadas para resolver problemas de clasificación binaria, actualmente se utilizan para resolver otros tipos de problemas regresión, agrupamiento, Multi- clasificación. La integración de las máquinas de soporte vectorial en el análisis de sentimiento ha mostrado ser efectiva en la clasificación precisa de emociones en textos.

Redes Neuronales

Las redes neuronales recurrentes (RNN) y las redes neuronales convolucionales (CNN), han mostrado un impacto significativo en el análisis de sentimiento. Las RNN tienen la capacidad de modelar dependencias secuenciales en los datos, lo que las hace adecuadas para analizar texto en el que el orden y la secuencia son cruciales para capturar el significado. Por otro lado, las CNN son eficientes para capturar características locales y patrones específicos en textos, lo que las hace útiles para el análisis de sentimiento en fragmentos específicos del texto. Según Motta & Huayna (2023) resaltan las llamadas redes neuronales, han demostrado una eficiencia superior en muchas tareas de la IA. De manera particular, han demostrado ser altamente efectivas para el análisis de sentimiento, tarea que en su forma más básica consiste en clasificar un texto según una polaridad de sentimiento positiva o negativa denotada por su autor hacia cierto tema.

La integración de redes neuronales en el análisis de sentimiento continúa siendo un área de investigación activa, ya que se busca mejorar la capacidad de estas redes para comprender y generalizar mejor las emociones y opiniones expresadas en diferentes contextos y estilos lingüísticos. Por otro lado, autores como Navarro (2023) destaca que estos modelos con su capacidad para modelar relaciones no lineales y capturar matices en grandes conjuntos de datos, ofrecen mejoras significativas en precisión y robustez.

Datos de redes sociales para el análisis del sentimiento

El análisis de sentimiento ha encontrado un terreno fértil en los datos provenientes de las redes sociales. Estas plataformas son una fuente inagotable de opiniones, emociones y actitudes de usuarios sobre diversos temas. Autores como Cruces (2019) resaltan que, debido al auge de esta tendencia, las empresas y organizaciones se están dando cuenta de que integrar esta información subjetiva en su sistema aporta bastantes ventajas a la hora de realizar negocio o de intentar mejorar su sistema. Además, Pérez (2019) concuerda que hoy en día las redes sociales son un pilar central de la comunicación social; aglutinan suficiente

información como para entender los diferentes sesgos existentes respecto a temas de distinta índole. El análisis de sentimiento en datos de redes sociales se beneficia de la diversidad y volumen de información disponible.

Por ejemplo, autores como Herrera, Sánchez, Meza & Fuentes (2021) en los últimos años se ha realizado un enorme número de estudios de análisis de sentimiento y se ha aplicado en una gran variedad de campos interdisciplinarios, como son la política, la tecnología, la medicina, el sector empresarial, por mencionar algunos. Asimismo, Soliani (2019) el conocimiento contenido en redes sociales se ha mostrado como de vital importancia no solo para los usuarios, donde buscan las más variadas opiniones, sino también para las compañías u organizaciones que buscan información dentro de ellas. La integración de datos de redes sociales en el análisis de sentimiento es un campo en constante expansión.

Aplicación del análisis del sentimiento en redes sociales

La aplicación del análisis de sentimiento en redes sociales es una práctica ampliamente utilizada en la actualidad debido a su capacidad para comprender las opiniones, emociones y percepciones de los usuarios en plataformas como Twitter, Facebook o Reddit. Autores como Cuadrado (2019) recalca que hoy día grandes compañías tecnológicas como Microsoft, Facebook, Google dan servicio de análisis de sentimiento por medio de sus APIS REST de inteligencia artificial, los cuales pueden ser accedidos por medio de una aplicación web o móvil. En el ámbito comercial, el análisis de sentimiento en redes sociales se emplea para evaluar la satisfacción del cliente, identificar tendencias del mercado y realizar estudios de opinión.

Autores como Figueroa (2023) el interés por aprovechar las redes sociales para analizar los sentimientos ha aumentado en los últimos años, debido a la facilidad prestada para medir la opinión pública clasificando los datos de texto como positivos, negativos o neutros frente al tema de interés analizado. La esfera política y social, el análisis de sentimiento en redes sociales se ha utilizado para medir el pulso de la opinión pública y anticipar tendencias sociales.

CAPITULO III

METODOLOGÍA DE LA INVESTIGACIÓN

Según Hernández (2018), la investigación cuantitativa, mediante un conjunto de procesos sistemáticos y empíricos, se aplica al estudio de un fenómeno o problema con el objetivo de ampliar su conocimiento. Por lo tanto, este enfoque permitirá obtener resultados con los estándares de validez y confiabilidad deseados mediante la recolección y análisis de datos numéricos para entrenar el modelo de análisis de datos en publicaciones de Facebook, ya que los datos en este ámbito se usan para medir y manejar para poder predicción.

Naturaleza de la investigación

La investigación propuesta hace uso del paradigma positivista que, a pesar de haber tenido varias críticas y reformulaciones, se ha mantenido siendo relevante en la investigación actual, Ya que autores como Mertens (2019) proponen una versión mejorada enfatizando en la importancia de la pluralidad metodológica. Aplicando en el trabajo permite tener un marco bien estructurado para el desarrollo del modelo y posteriormente hacer la validación con datos reales de como el modelo ha aprendido los datos otorgados por el dataset y como este ha logrado hacer la predicción más acertada con relación en los sentimientos de los usuarios.

Por esta razón el enfoque cuantitativo continúa siendo una herramienta fundamental para la investigación científica. Autores como Bryman & Hardy (2019) destacan la importancia del rigor metodológico, la transparencia en el manejo del análisis de datos y la ética. Es así como, este apartado se aplica en el trabajo de investigación debido al manejo de grandes cantidades de datos obtenidos mediante la recolección de información en diferentes datasets que cuentan con una enorme cantidad de

recopilación de los mismos, con un aproximado de 75 000 registros de publicaciones de Facebook dándonos un panorama amplio de como esto impactara el entrenamiento del modelo permitiéndonos así mejorar la predicción de los sentimientos en las publicaciones de Facebook.

En el mismo orden de ideas, se estableció un diseño de investigación no experimental; puesto que, son cada vez más utilizados en diversas áreas de investigación, ya que autores como Shaver & Brink (2011) resaltan la utilidad de estos diseños para estudiar fenómenos complejos en contextos naturales y para generar conocimiento más globalizado. Permitiendo así el manejo de una observación en los sentimientos de las publicaciones en contexto natural, sin afectar a los datos con esto se garantiza que el dataset no va a interferir con las variables y no se verán afectadas debido a que son datos previamente recolectados de distintas publicaciones realizadas por usuarios con la finalidad de entrenar y mejorar el desarrollo del modelo. Además, se aplicó un diseño transeccional o transversal, el cual se fundamenta en la recolección de datos que se realiza en un solo momento, en un tiempo único (Hernández y Mendoza, 2018).

Para abordar la profundidad de la investigación se utilizó un nivel descriptivo que, según Hernández & Mendoza (2018), buscan especificar propiedades, características y perfiles de personas, grupos, procesos, objetos u otros fenómenos analizados. Miden, recolectan datos y reportan información sobre conceptos, variables, dimensiones o componentes del fenómeno investigado. El investigador selecciona variables, recopila información sobre ellas y describe o caracteriza lo investigado.

El objetivo general de este proyecto es desarrollar un modelo de Machine Learning que permita el análisis de las publicaciones de Facebook para predecir el sentimiento de los usuarios. Para lograrlo, se llevarán a cabo varios objetivos específicos: seleccionar diferentes conjuntos de datos de imágenes y texto para entrenar el modelo, comparar distintos modelos de inteligencia artificial para el análisis de sentimientos y datos, e implementar diversas estructuras de aprendizaje

supervisado. Finalmente, se validará el modelo de análisis de sentimiento con datos reales de Facebook. Este estudio descriptivo proporcionará una visión detallada y comparativa de las capacidades de diferentes modelos de IA en el análisis de sentimientos, ayudando a identificar las metodologías más efectivas para comprender y predecir las emociones de los usuarios en las redes sociales.

Así también, en esta investigación se trabajó con la metodología documental que se fundamenta para localizar, recuperar, analizar, criticar e interpretar datos secundarios previamente registrados. Estos datos provienen de diversas fuentes documentales, como impresos, audiovisuales o medios electrónicos, y son utilizados para comprender fenómenos sin recopilar nueva información primaria (Arias, 2016, p. 27).

Para el trabajo se identificaron y evaluaron varios conjuntos de datos relevantes que contienen tanto texto como imágenes, necesarios para entrenar y validar el modelo de análisis de sentimientos. Se examinaron criterios como la calidad de los datos, la diversidad de las muestras y la disponibilidad pública. Además, se consideraron estudios previos que utilizaron estos conjuntos de datos para análisis similares, asegurando así una base sólida y documentada para la selección final. Esta investigación documental permitió fundamentar la elección de los datasets más adecuados para el proyecto, basándose en evidencia y experiencias previas documentadas en la literatura científica.

Mientras que, la investigación de campo implica la recolección directa de datos de los sujetos estudiados o del entorno en el que ocurren los eventos. Se basa en la obtención de datos primarios sin manipular ni controlar variables, permitiendo que el investigador recopile información sin alterar las condiciones originales (Arias, 2016). Esta metodología implicó la recolección directa de datos a partir de publicaciones y comentarios de usuarios en Facebook, utilizando técnicas de scraping y APIs disponibles. Los datos recolectados se procesaron y etiquetaron para reflejar los sentimientos expresados; luego se usaron datos reales para probar y ajustar el

modelo, asegurando su eficacia y precisión en un entorno real. La investigación de campo permitió obtener información actualizada y específica del comportamiento y emociones de los usuarios en la plataforma.

Técnica de recolección de datos

Según Palella y Martins (2006). "La observación estructurada es una técnica de recolección de datos que implica la utilización de instrumentos estandarizados y criterios específicos para observar y registrar los fenómenos de interés de manera sistemática y controlada" (p. 129). Esta técnica permite al investigador experimentar de primera mano las situaciones y contextos de los participantes, proporcionando datos ricos y detallados. La observación se aplicará directamente en la plataforma de Facebook, donde el investigador interactuará con usuarios, grupos y páginas para comprender mejor los contextos en los que se generan los datos de sentimiento, esto permitirá una validación más contextualizada y precisa del modelo.

Operacionalización de la variable

Tabla 1. Operacionalización de la variable

VARIABLE	Concepto	Dimensión	Indicador	Ítem
Validar el Modelo de análisis de sentimiento del usuario con datos reales de Facebook.	Durante esta fase, se evaluará el grado en que el modelo cumple con los objetivos establecidos y se intentará identificar si existe algún motivo	Precisión de la Clasificación del Sentimiento.	Positiva	1,2, 3
			Negativa	
			Neutra	
	cumple con los objetivos establecidos y se intentará identificar si existe algún motivo	Consistencia del Análisis de Sentimiento.	Publicaciones	4
			Contexto	5
				Relevancia del Análisis de

por el que este modelo sea deficiente.	Sentimiento		
Calatrava (2019)	Utilidad del Análisis de Sentimiento	Opiniones de usuarios	7
	Generalización del Modelo	Publicaciones en Facebook	8
	Facilidad de interpretación	Facilidad	9
	Eficiencia del modelo	Precisión de aprendizaje de texto	10
		Precisión de aprendizaje de imágenes	

Instrumento de recolección de datos

Según Castillo, Gómez, Taborda & Mejía (2021) “La escala de estimación consiste en una serie de categorías ante cada una de las cuales, el investigador debe indicar el grado en que se encuentre presente determinada característica, o la frecuencia con que ocurre determinada acción” (p. 75). Para validar el modelo de análisis de sentimientos en Facebook, se definió primero las categorías (muy en desacuerdo, en desacuerdo, neutral, de acuerdo y muy de acuerdo). Luego, se asignó una escala de 1 a 5 para cada indicador. El instrumento sirvió para que, expertos en Software (3 a 5) lo utilicen y emitan su criterio sobre el mismo, evaluando muestras de datos reales y asignando puntuaciones según la escala. Los resultados se analizan para identificar patrones y determinar el rendimiento del modelo. Si las puntuaciones son altas, el modelo se considera efectivo; si son bajas, se investigan las causas y se hacen mejoras.

Validez y confiabilidad

Según Hernández & Mendoza (2018) La validez se refiere, en términos generales, a la precisión con la que un instrumento evalúa la variable que realmente busca medir. Es decir, si los indicadores empíricos utilizados reflejan de manera fiel el concepto abstracto que se pretende evaluar. En otras palabras, un instrumento válido es aquel que captura exactamente el fenómeno que quiere estudiar, asegurando que los resultados obtenidos sean representativos y precisos del concepto en cuestión. La validez garantiza que las interpretaciones y conclusiones derivadas de los datos sean verdaderamente relevantes y aplicables al concepto teórico. En el siguiente trabajo se realizó la validación de expertos, tomando en cuenta a profesionales del área de Software que tienen un amplio conocimiento, los cuales revisaron y emitieron correcciones del instrumento, para su aplicación posterior.

Técnicas de Análisis de Información

De acuerdo con Hernández (2018) en el proceso cuantitativo, la recolección de datos precede al análisis, mientras que, en la investigación cualitativa, la recolección y el análisis ocurren simultáneamente, además el análisis cualitativo no sigue un enfoque uniforme, ya que cada estudio requiere un esquema particular. Dada la naturaleza cualitativa de la investigación y el enfoque en comprender en profundidad cómo los usuarios interpretan y se relacionan con las publicaciones en redes sociales, una técnica de análisis de información que se acoplaría bien es el análisis temático.

Metodología del producto

Según Scrum.org, Scrum es un marco de trabajo ligero diseñado para ayudar a personas, equipos y organizaciones a generar valor a través de soluciones adaptativas para problemas complejos. Este marco se basa en el empirismo, que es la teoría de que el conocimiento proviene de la experiencia y que las decisiones deben basarse en lo que se conoce. Scrum emplea un enfoque iterativo e incremental lo cual permite desarrollar y mejorar el modelo de red neuronal en pequeños incrementos, revisando y ajustando el enfoque en cada iteración permitiendo así una mejor adaptación a los problemas complejos y cambiantes en el desarrollo de modelos de aprendizaje automático.

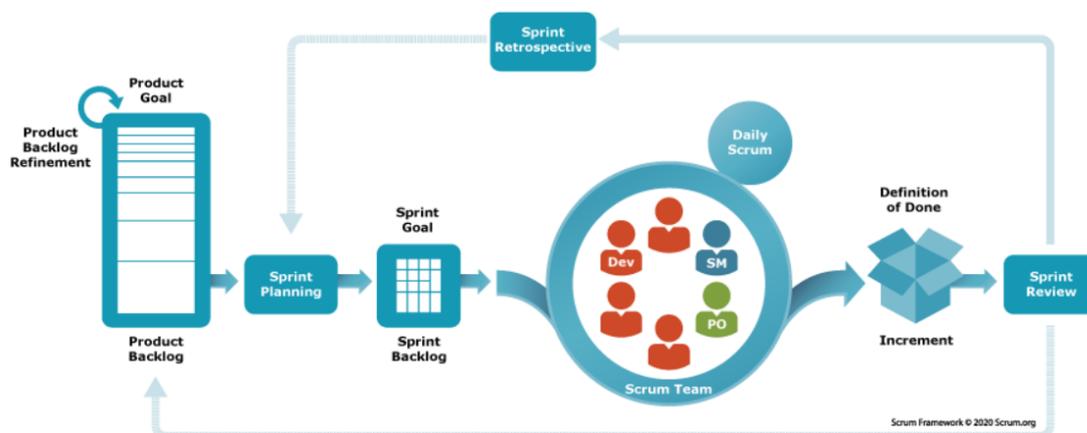


Figura 1. *Ciclo de vida SCRUM.*

Scrum organiza el trabajo en ciclos llamados Sprints, que incluyen eventos clave como la Planificación del Sprint, el Scrum Diario, la Revisión del Sprint y la Retrospectiva del Sprint, además se apoya en roles definidos como el Product Owner, el Scrum Máster y el Equipo de Desarrollo, a continuación, se detalla los principales eventos y roles en Scrum basados en la guía de Scrum.org:

- **Sprint:** Cada Sprint comienza con una planificación y termina con una revisión y retrospectiva.
- **Sprint Planning:** Al inicio del Sprint, el equipo selecciona elementos del Product Backlog y crea el Sprint Backlog.
- **Daily Scrum:** Es una reunión diaria de 15 minutos para sincronizar actividades y planificar las próximas 24 horas basados en tres preguntas: ¿Qué hice ayer?, ¿Qué haré hoy? y ¿Hay algún impedimento?
- **Sprint Review:** Al final del Sprint, el equipo y los interesados inspeccionan el incremento y adaptan el Product Backlog según sea necesario, recibiendo retroalimentación.
- **Sprint Retrospective:** Después de la Revisión del Sprint, el equipo discute qué fue bien, qué no, y cómo pueden mejorar en el próximo Sprint.

- **Product Backlog:** Es una lista ordenada de todo lo necesario en el producto.
- **Sprint Backlog:** Es el conjunto de elementos seleccionados del Product Backlog para el Sprint, más un plan para entregar el incremento del producto.
- **Increment:** Suma de todos los desarrollos completados del Product Backlog durante un Sprint, que debe cumplir con la Definición de "Terminado".

Roles en Scrum:

- **Product Owner (Propietario del Producto):** Maximiza el valor del producto y gestiona el Product Backlog.
- **Scrum Master:** Asegura la correcta implementación de Scrum y ayuda a eliminar impedimentos.
- **Development Team (Equipo de Desarrollo):** Grupo auto-organizado y multifuncional que entrega incrementos del producto.

Scrum ofrece un marco estructurado que simplifica la gestión de proyectos complejos mediante ciclos de trabajo iterativos y adaptativos. Este enfoque organiza el trabajo en sprints y promueve una comunicación constante y una revisión continua, permitiendo ajustes ágiles y eficaces. Al aplicar Scrum, se facilita la adaptación y optimización del desarrollo del modelo de red neuronal, permitiendo ajustes rápidos en respuesta a nuevos desafíos y mejorando continuamente cada fase del proceso. Gracias a las ventajas que ofrece Scrum, se ha logrado definir los sprints de la siguiente manera:

1. **Sprint 0: Planificación Inicial:** En este sprint, se definen los objetivos del proyecto, como clasificar emociones en imágenes y textos en categorías positivas, negativas y neutrales. También se prepara el entorno de desarrollo, instalando y configurando TensorFlow, Keras, OpenCV, Pandas y otras dependencias, y se crea la estructura de directorios para los datos y el código.
2. **Sprint 1: Carga y Preprocesamiento de Datos:** El objetivo es implementar la carga y el preprocesamiento de los datos de imágenes y textos. Se desarrolla la función para cargar y preprocesar datos de imágenes y textos, asegurándose de que los datos se limpien y estén listos para ser utilizados en el modelo.

3. **Sprint 2: Construcción del Modelo:** Se diseña e implementa la arquitectura del modelo combinado de imágenes y texto. El modelo incluye capas para procesar imágenes y texto, y una capa combinada para producir la salida final.
4. **Sprint 3: Entrenamiento del Modelo:** El modelo combinado se entrena con los datos preprocesados de imágenes y texto. Se implementa la función de entrenamiento y se lleva a cabo el entrenamiento del modelo, ajustando parámetros según sea necesario para mejorar el rendimiento.
5. **Sprint 4: Evaluación y Visualización:** Se implementan las funciones de evaluación del modelo para imágenes y texto. Además, se desarrollan gráficos para visualizar el rendimiento del modelo durante el entrenamiento y la evaluación.
6. **Sprint 5: Refinamiento y Documentación:** En este sprint final, se refina el modelo basado en la retroalimentación y se completa la documentación del proyecto, asegurando que todo esté listo para la entrega y fácil de entender para otros desarrolladores o partes interesadas.

CAPITULO IV

ANALISIS E INTERPRETACION DE LOS RESULTADOS

Según Palella & Martins (2006) “Una vez recogidos los valores que toman las variables del estudio (datos), se procede a su análisis estadístico, el cual permite hacer suposiciones e interpretaciones sobre la naturaleza y significación de aquellos en atención a los distintos tipos de información que puedan proporcionar”. Para entrenar un modelo de clasificación de emociones, se analizó el Dataset Social Media Emotion Dataset y Emotion-detection. El análisis incluye diferentes pasos de preprocesamiento del dataset para mejorar la calidad de los datos y, por ende, el rendimiento del modelo. A continuación, se presenta una tabla detallada que resume los pasos de preprocesamiento, las métricas de evaluación utilizadas, y observaciones específicas para cada etapa del proceso.

Tabla 2. *Análisis de datasets para entrenamiento del modelo*

Dataset	Tamaño del Dataset	Tipo de Datos	Fuente del Dataset	Preprocesamiento	Métrica de Evaluación	Observaciones
Emotion-detection	36,000 registros	Imágenes	Kaggle	Redimensionamiento a 128x128 píxeles Normalización Encabezados	Accuracy, Loss	Las imágenes se dividen en categorías: positivo, negativo, neutral.
Social Media Emotion Dataset	75,000 registros	Texto	Kaggle	Limpieza de texto Eliminación de signos de puntuación Eliminación de stop words	Precision, Recall Accuracy, F1-Score	Las etiquetas de texto se convierten en sentimientos: positivo, negativo, neutral.

El dataset Social Media Emotion Dataset contiene datos textuales, se implementaron técnicas como la eliminación de signos de puntuación y stop words, y la combinación de las palabras procesadas. Estas técnicas, descritas por Salgado y Trujillo (2024), combinan procesamiento de lenguaje natural con aprendizaje automático, y se evaluaron mediante métricas como la consistencia de palabras clave, la reducción de ruido, la relevancia y la cohesión del texto. Según Quintana (2021) y Moran (2022), estos pasos son esenciales para manejar grandes volúmenes de datos textuales y comprender la percepción pública. Por otro lado, el dataset Emotion-detection, consta de imágenes para lo cual se realizó el redimensionamiento a 128x128 píxeles, la normalización de los valores de los píxeles y el etiquetado según las emociones dando como resultado que la evaluación de este dataset se basa en métricas de rendimiento como accuracy y loss, lo que es fundamental para la clasificación precisa de emociones en imágenes, optimizando así la capacidad del modelo.

Tabla 3. Comparar diferentes modelos - Naive Bayes

MODELO	DESCRIPCION	VENTAJAS	DESVENTAJAS
NAIVE BAYES	Clasificador probabilístico basado en el teorema de Bayes, asume independencia entre características	Rápido y eficiente en términos de tiempo de entrenamiento y predicción, incluso con grandes conjuntos de datos Requiere una pequeña cantidad de datos para el entrenamiento, lo cual es eficiente con 75.000 registros Fácil de implementar y entender	La suposición de independencia raramente es cierta lo que puede reducir la precisión Puede tener problemas de rendimiento con características correlacionadas en el texto No maneja bien la alta dimensionalidad de texto sin un preprocesamiento adecuado

Funciona bien con datos dispersos, comunes en el procesamiento de texto (bag of words - TF - IDF). Menos preciso comparado con modelos más complejos

Según IBM Clasificador Naive Bayes es un algoritmo de aprendizaje automático supervisado utilizado para tareas de clasificación, como la clasificación de texto. Este modelo se basa en principios de probabilidad y pertenece a la familia de algoritmos de aprendizaje generativo. Es bastante rápido y eficiente cuando se trata de manejo de texto además es fácil de implementar, pero no es eficiente en cuanto al manejo de imágenes debido que asume que todas las características son independientes, lo cual no es cierto en imágenes donde los píxeles están correlacionados por otro lado las imágenes tienen un gran número de píxeles (características), y Naive Bayes no maneja bien esta alta dimensionalidad sin reducción de características.

Tabla 4. Comparar diferentes modelos - Maquina de Soporte Vectorial

MODELO	DESCRIPCION	VENTAJAS	DESVENTAJAS
MAQUINA DE SOPORTE VECTORIAL		Alto rendimiento en espacios de alta dimensionalidad, útil para texto vectorizado	Tiempo de entrenamiento más largo especialmente con 75.000 registros
	Busca encontrar el hiperplano que mejor separa las	Buen rendimiento con un margen claro de separación entre clases	Menos eficiente en términos de memoria y procesamientos para grandes conjuntos de datos
	clases en el espacio de características	Soporta clasificación lineal y no lineal mediante el uso de núcleos(kernels)	requiere ajustes de hiperparametros, lo cual puede ser complicado
		Robusto frente a overfittingen comparación con otros modelos simples	Sensible a la elección de los parámetros y al kernel utilizado

Las Máquinas de Soporte Vectorial (SVMs) son técnicas de aprendizaje supervisado que buscan el margen óptimo para separar clases en un conjunto de datos. En el análisis de sentimientos, las SVMs son valoradas por su precisión en la clasificación de emociones y su capacidad para manejar datos de alta dimensionalidad mediante núcleos (Romero et al., 2020). Sin embargo, presentan limitaciones en comparación con las redes neuronales, como la necesidad de un ajuste minucioso de parámetros y una menor eficiencia con conjuntos de datos muy grandes. Según Navarro las redes neuronales, son más efectivas para capturar patrones complejos y no lineales en grandes volúmenes de datos, haciéndolas más adecuadas para tareas con alta dimensionalidad y variabilidad.

Tabla 5. Comparar diferentes modelos - Red Neuronal

MODELO	DESCRIPCION	VENTAJAS	DESVENTAJAS
RED NEURONAL	Modelos de aprendizaje profundo con múltiples capas de neuronas que pueden capturar relaciones complejas en los datos.	Las redes neuronales pueden adaptarse y generalizar a partir de datos nuevos, mejorando su rendimiento con la experiencia. Efectivas en el procesamiento de datos no estructurados como imágenes, texto y audio. Varias arquitecturas (CNNs, RNNs, etc.) pueden ser utilizadas según el tipo de problema. Continuos avances en técnicas de entrenamiento y herramientas de software mejoran la eficacia y eficiencia.	El entrenamiento puede ser intensivo en recursos computacionales, especialmente con redes profundas. Necesitan grandes volúmenes de datos para entrenamiento efectivo. Puede sobre ajustarse a los datos de entrenamiento si no se aplican técnicas de regularización adecuadas. Los modelos pueden ser difíciles de interpretar, lo que puede ser un problema en aplicaciones donde la interpretabilidad es crucial.

Las Redes Neuronales son útiles para analizar texto donde el orden y la secuencia son cruciales, permitiendo una comprensión profunda del contexto y la intención detrás de las palabras. Además de ser eficientes para detectar características locales y patrones específicos, lo que ha resultado ser beneficioso tanto para el análisis de texto como para la clasificación de imágenes. Según Motta y Huayna (2023), las redes neuronales tienen una eficiencia superior en muchas tareas de inteligencia artificial, incluidas las tareas de análisis de sentimiento, gracias a su capacidad para manejar datos complejos y variados. Además, como señala Navarro (2023), estas redes mejoran significativamente la precisión y robustez en comparación con modelos tradicionales como Naive Bayes y SVM.

Implementar las Estructuras

Durante el desarrollo de la presente tesis, se ha implementado una metodología ágil basada en sprints con la finalidad de organizar y gestionar el proceso de investigación y desarrollo. Cada sprint ha sido diseñado para abordar objetivos específicos y permitir una evolución continua del proyecto permitiendo llevar a cabo actividades clave que incluyen la recopilación de datasets, el entrenamiento del modelo, el desarrollo y ajuste de la red neuronal. En la descripción detallada de cada sprint, se expondrán los objetivos establecidos, las tareas ejecutadas, los resultados obtenidos y los ajustes realizados en respuesta a los desafíos enfrentados. Este enfoque detallado no solo proporciona una visión clara del progreso alcanzado en cada fase del proyecto, sino que también resalta las decisiones y adaptaciones realizadas que se detallan a continuación:

- **Sprint 0: Planificación Inicial:**

Objetivo: Planificar el proyecto, definir los objetivos y preparar el entorno.

Tareas:

- a. Clasificación de emociones en imágenes y textos en categorías positivas, negativas y neutrales.

b. Instalar y configurar TensorFlow, Keras, OpenCV, Pandas y otras

```
from tensorflow.keras import layers, models
import matplotlib.pyplot as plt
import numpy as np
import os
import cv2
import pandas as pd
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import ENGLISH_STOP_WORDS
import string
```

dependencias.

Figura 2. *Instalación de librerías.*

- **Sprint 1: Carga y Preprocesamiento de Datos:**

Objetivo: Implementar la carga y el preprocesamiento de los datos

Tareas:

- a. Cargar y preprocesar datos de imágenes
- b. Cargar y preprocesar datos de texto

```
# Función para cargar y preprocesar imágenes y texto
!usage -EErraez
def load_and_preprocess_data():
    # Cargar imágenes y etiquetas
    train_images, train_labels = load_images_and_labels('train')
    test_images, test_labels = load_images_and_labels('test')

    # Cargar datos de texto y preprocesarlos
    data = pd.read_csv('./data/Social Media Emotion Dataset.csv')
    data['sentiment'] = data['label'].apply(map_to_sentiment)

    data['cleaned_text'] = data['text'].apply(preprocess_text)

    # Crear un objeto TfidfVectorizer
    vectorizer = TfidfVectorizer()

    # Ajustar y transformar los datos limpios
    X = vectorizer.fit_transform(data['cleaned_text'])

    # Obtener las nuevas etiquetas (sentimientos)
    y = data['sentiment']

    # Dividir datos en entrenamiento y prueba
    X_train, X_test, y_train, y_test = train_test_split(*arrays: X, y, test_size=0.2, random_state=42)

    return train_images, train_labels, test_images, test_labels, X_train, X_test, y_train, y_test
```

Figura 3. *Carga y preprocesamiento de datos texto e imágenes.*

- **Sprint 2: Construcción del Modelo:**

Objetivo: Implementar la construcción del modelo combinado de imágenes y texto.

Tareas:

- a. Diseñar la arquitectura del modelo de red neuronal
- b. Implementar la construcción del modelo.

```
def build_model(input_shape_images, max_words):  
    # Capa de entrada para imágenes  
    input_images = layers.Input(shape=input_shape_images)  
    x_images = layers.Conv2D(32, (3, 3), activation='relu')(input_images)  
    x_images = layers.MaxPooling2D((2, 2))(x_images)  
    x_images = layers.Conv2D(64, (3, 3), activation='relu')(x_images)  
    x_images = layers.MaxPooling2D((2, 2))(x_images)  
    x_images = layers.Flatten()(x_images)  
    x_images = layers.Dense(units=64, activation='relu')(x_images)  
  
    # Capa de entrada para texto  
    input_text = layers.Input(shape=(max_words,))  
    x_text = layers.Dense(units=64, activation='relu')(input_text)  
  
    # Combinación de capas de imágenes y texto  
    combined = layers.concatenate([x_images, x_text])  
    combined = layers.Dense(units=64, activation='relu')(combined)  
    output = layers.Dense(units=3, activation='softmax')(combined) # 3 categorías: positivo, negativo, neutral  
  
    model = models.Model(inputs=[input_images, input_text], outputs=output)  
    model.compile(optimizer='adam',  
                  loss='sparse_categorical_crossentropy',  
                  metrics=['accuracy'])  
  
    return model
```

Figura 4. *Diseño e implementación del modelo.*

- **Sprint 3: Entrenamiento del Modelo:**

Objetivo: Entrenar el modelo combinado con los datos de imágenes y texto.

Tareas:

- a. Implementar la función de entrenamiento del modelo.
- b. Entrenar el modelo con los datos.

```
def train_model(model, train_images, train_labels, X_train, y_train, epochs=10):
    history = model.fit([train_images, X_train], train_labels, epochs=epochs, validation_split=0.2)
    return history
```

Figura 5. Implementación y entrenamiento del modelo.

- **Sprint 4:** Evaluación y Visualización:

Objetivo: Implementar la evaluación del modelo y la visualización de resultados.

Tareas:

- Implementar las funciones de evaluación.
- Visualizar el rendimiento del modelo.

```
def plot_performance(history):
    plt.figure(figsize=(12, 4))
    plt.subplot(*args: 1, 2, 1)
    plt.plot(*args: history.history['accuracy'], label='Training Accuracy')
    plt.plot(*args: history.history['val_accuracy'], label='Validation Accuracy')
    plt.xlabel('Epoch')
    plt.ylabel('Accuracy')
    plt.legend()

    plt.subplot(*args: 1, 2, 2)
    plt.plot(*args: history.history['loss'], label='Training Loss')
    plt.plot(*args: history.history['val_loss'], label='Validation Loss')
    plt.xlabel('Epoch')
    plt.ylabel('Loss')
    plt.legend()

    plt.tight_layout()
    plt.show()
```

Figura 6. Funciones de evaluación.

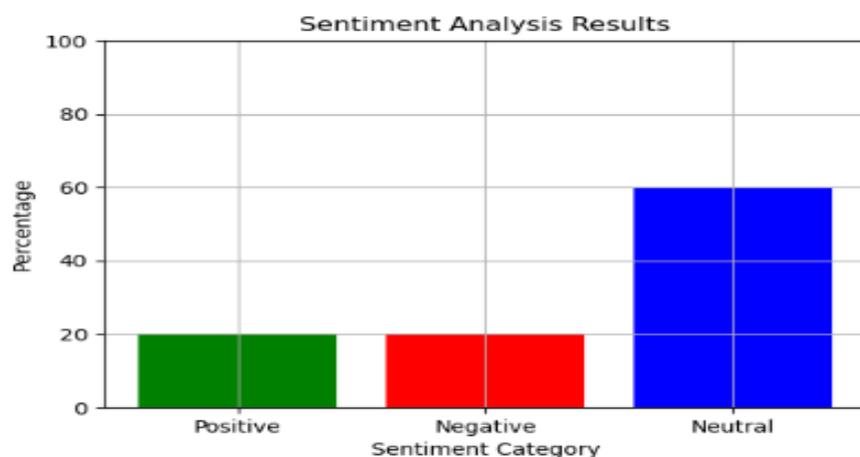


Figura 7. *Rendimiento del modelo.*

Resultado de la investigación

El presente informe describe el proceso y los resultados de la validación del modelo de análisis de sentimientos aplicado a datos reales de Facebook. El modelo clasifica los comentarios en cinco categorías: Muy en Desacuerdo, En Desacuerdo, Neutral, De Acuerdo, y Muy de Acuerdo, utilizando una escala de puntuación de 1 a 5 respectivamente. Para validar el modelo, expertos en software evaluaron muestras de datos y asignaron puntuaciones según esta escala.

Definición de Categorías y Escala:

- Muy en Desacuerdo: 1
- En Desacuerdo: 2
- Neutral: 3
- De Acuerdo: 4
- Muy de Acuerdo: 5

Recolección de Datos:

- Se recopilaron comentarios reales de usuarios en Facebook para ser evaluados por expertos.

Evaluación por Expertos:

- Un grupo de 3 a 5 expertos en software evaluaron los comentarios utilizando la escala definida.

Análisis de Resultados:

- Se analizaron las puntuaciones asignadas para identificar patrones y determinar el rendimiento del modelo.

Resultados

Los resultados obtenidos se presentan en el siguiente gráfico de barras:

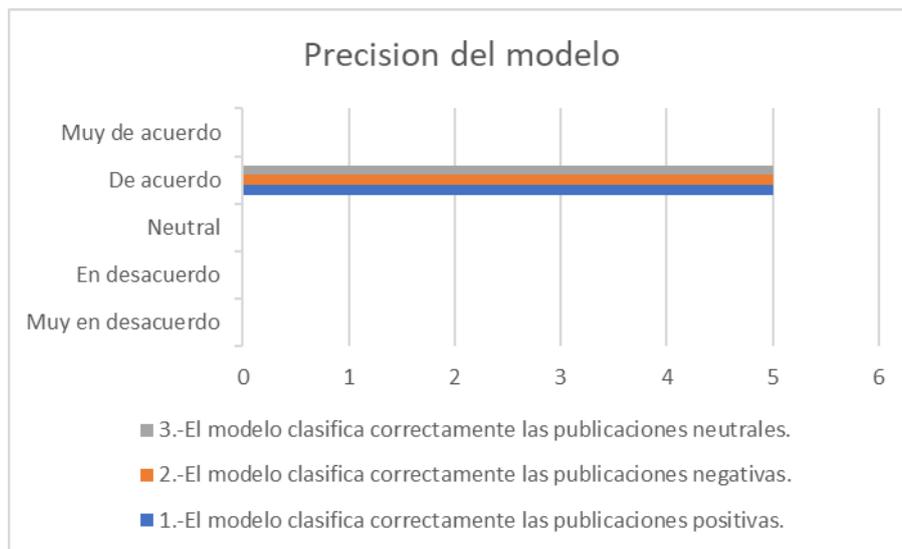


Figura 8. Precisión del Modelo

El gráfico evalúa la precisión del modelo según la opinión de los usuarios, con categorías que van desde "Muy de acuerdo" hasta "Muy en desacuerdo". Cada barra representa una categoría de sentimiento (positiva, negativa, neutral), y su longitud refleja el nivel de acuerdo con la afirmación de que el modelo clasifica correctamente las publicaciones en esa categoría. La similitud en la longitud de las barras sugiere que las opiniones sobre la precisión del modelo son consistentes entre las diferentes categorías de sentimientos, indicando que los evaluadores están mayormente de acuerdo en que el modelo clasifica correctamente las publicaciones en todas las categorías.

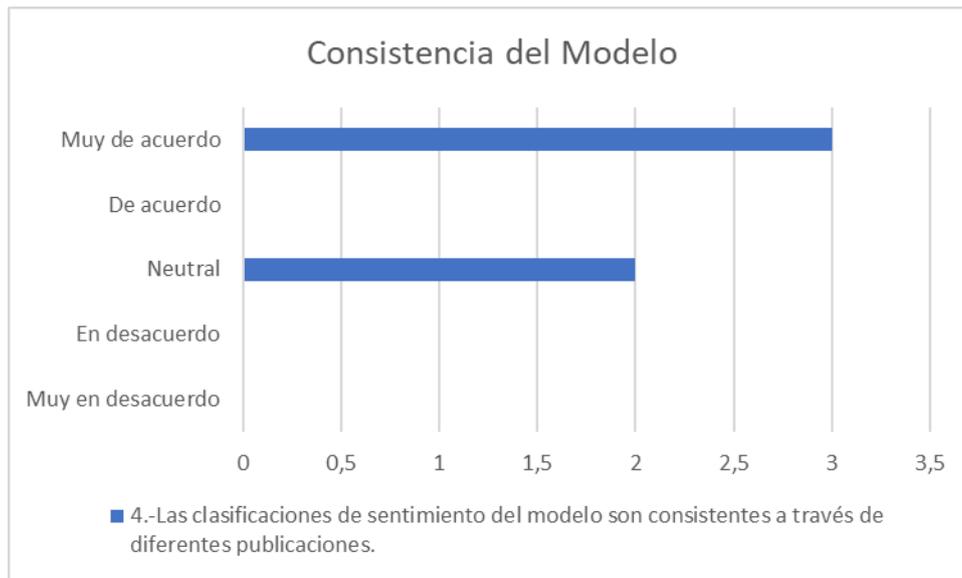


Figura 9. *Consistencia del Modelo*

El gráfico evalúa la consistencia del modelo en la clasificación de sentimientos, basado en las opiniones de los usuarios que se distribuyen en categorías. La barra visible representa la afirmación de que las clasificaciones de sentimiento del modelo son consistentes a través de diferentes publicaciones. La longitud de la barra sugiere que la mayoría de los evaluadores se encuentran entre "Muy de acuerdo" y "Neutral" respecto a la consistencia del modelo, lo que indica una percepción generalmente positiva sobre la capacidad para mantener clasificaciones coherentes.

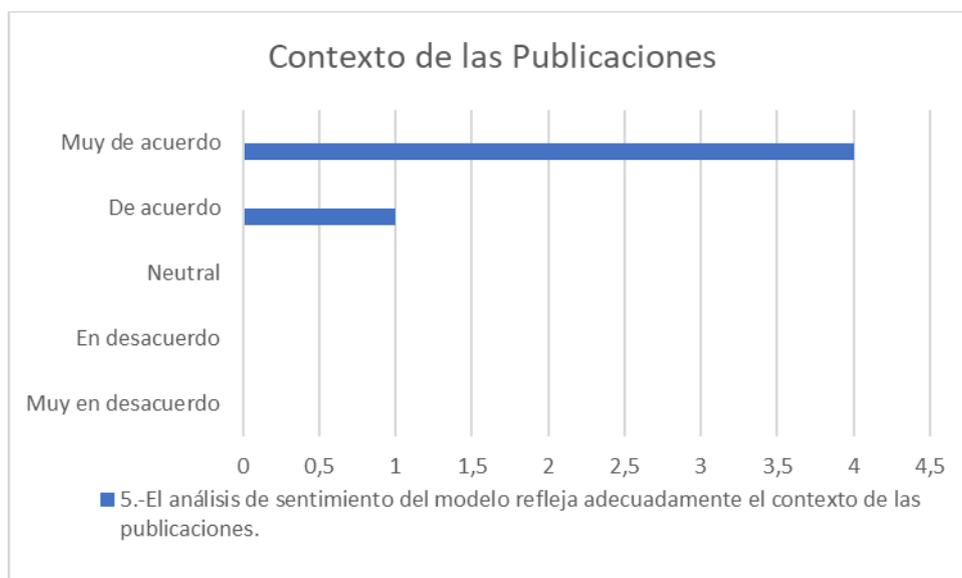


Figura 10. Contexto de las Publicaciones

El gráfico muestra la evaluación del modelo en términos de cómo refleja adecuadamente el contexto de las publicaciones, con categorías. La barra azul indica el nivel de acuerdo de los evaluadores con la afirmación de que el análisis de sentimiento del modelo refleja adecuadamente el contexto de las publicaciones. La mayoría de los evaluadores se posicionan entre Muy de acuerdo, lo que sugiere una percepción positiva de la capacidad del modelo para captar y reflejar el contexto en sus análisis de sentimiento.

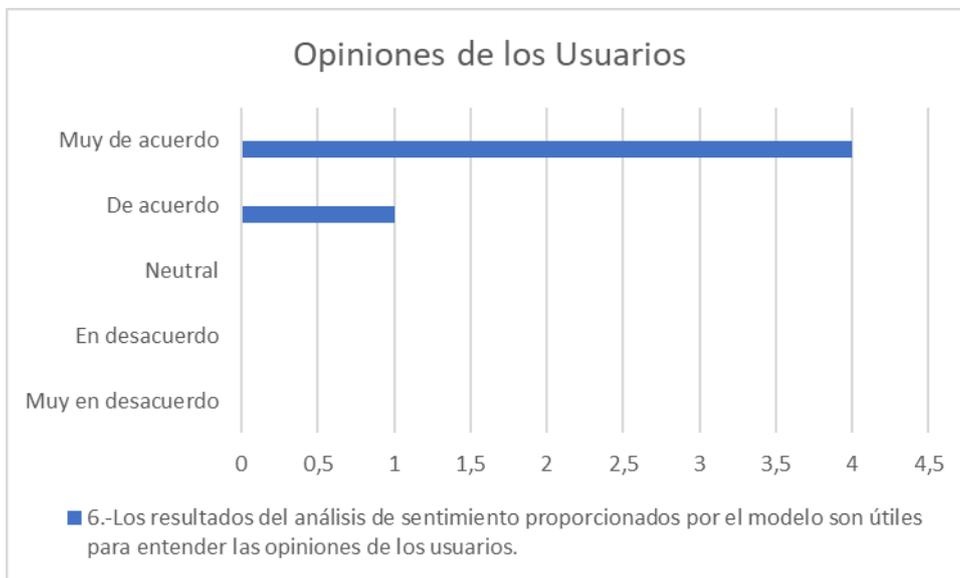


Figura 11. Opiniones de los Usuarios

El gráfico muestra la percepción de los usuarios sobre la utilidad de los resultados del análisis de sentimiento proporcionados por el modelo para comprender sus opiniones. La mayoría de los usuarios se muestran muy de acuerdo con esta afirmación, como lo refleja la barra más larga, mientras que una proporción significativa también está de acuerdo. No hay respuestas en las categorías de Neutral, En desacuerdo, o Muy en desacuerdo, lo que indica una aceptación positiva y generalizada de la herramienta de análisis de sentimiento, sugiriendo su efectividad en la interpretación de las emociones y opiniones de los usuarios.

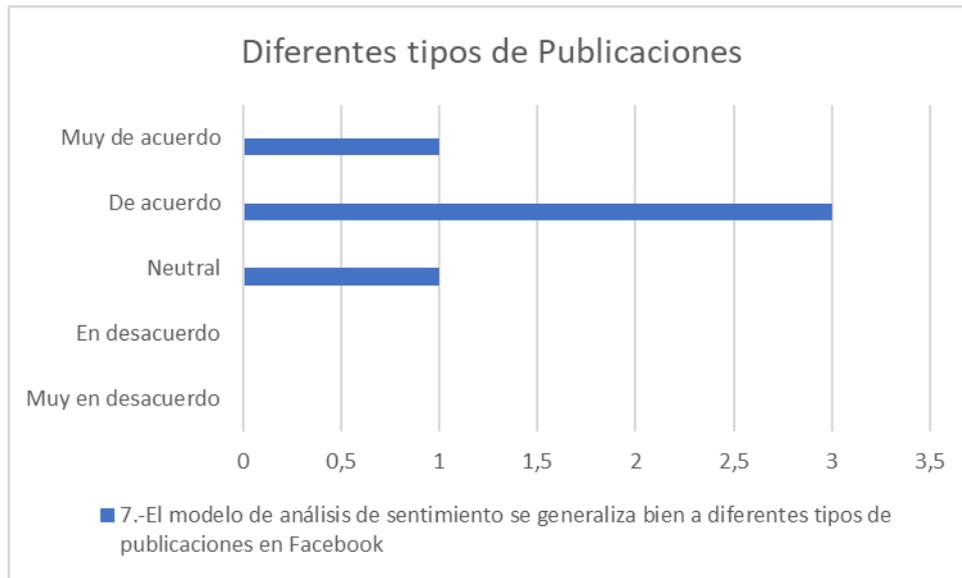


Figura 12. *Diferentes tipos de Publicación*

El gráfico refleja la opinión de los usuarios sobre la capacidad del modelo de análisis de sentimiento para generalizarse a diferentes tipos de publicaciones en Facebook. La mayor parte de los usuarios está de acuerdo con esta afirmación, como lo muestra la barra más larga, seguida de un grupo considerable que se mantiene neutral. Un menor número de usuarios está muy de acuerdo con la afirmación. No se registran respuestas en las categorías de En desacuerdo o Muy en desacuerdo, lo que indica que, aunque hay aceptación, también existe cierta reserva respecto a la capacidad de generalización del modelo en distintos contextos.

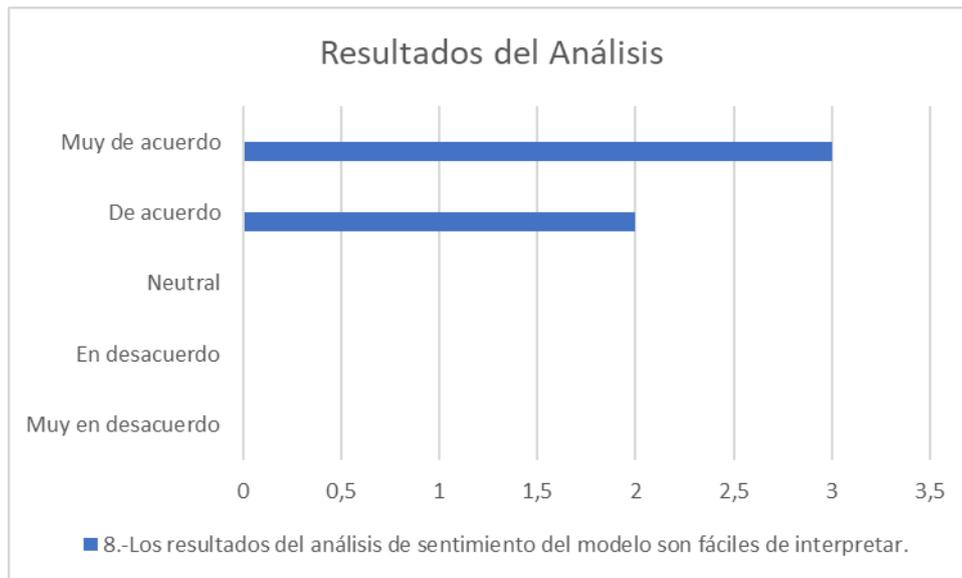


Figura 13. *Resultado del Análisis*

El gráfico muestra como los resultados del modelo en el análisis de sentimientos son fáciles de interpretar. La barra azul indica el nivel muy de acuerdo y de acuerdo de los evaluadores con la afirmación de que los resultados son fáciles de interpretar. La mayoría de los evaluadores se posicionan entre Muy de acuerdo, lo que demuestra una percepción positiva de la capacidad del modelo para captar y reflejar el contexto en sus análisis de sentimiento.

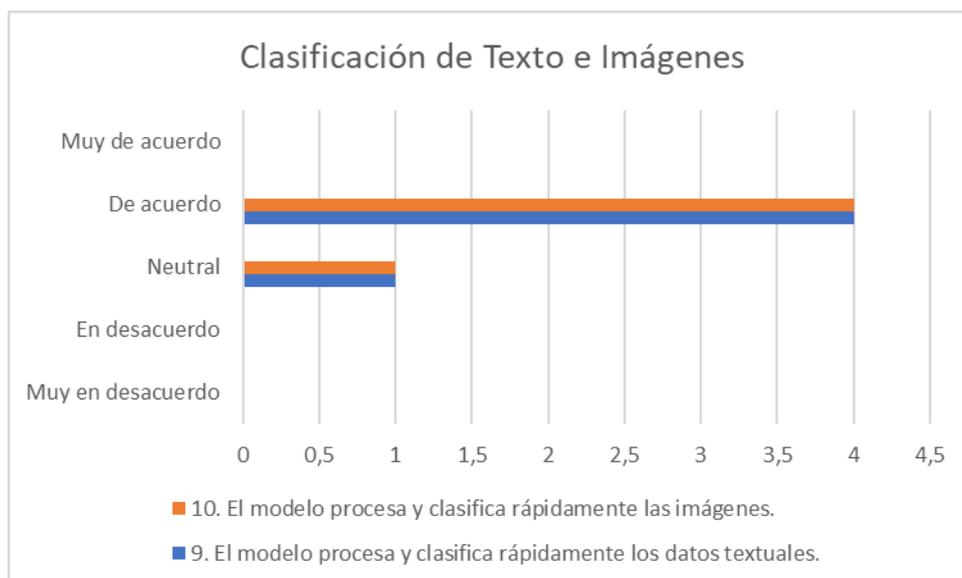


Figura 14. *Clasificación de Texto e Imágenes*

El gráfico muestra la percepción de los usuarios sobre la eficiencia del modelo al procesar y clasificar tanto imágenes como datos textuales. La mayoría de los usuarios está de acuerdo en que el modelo realiza estas tareas de manera rápida, como se refleja en las barras largas de ambos ítems (9 y 10). Sin embargo, hay una menor pero notable cantidad de usuarios que se mantienen neutrales respecto a la velocidad de procesamiento y clasificación, especialmente en lo que concierne a los datos textuales. No hay respuestas en desacuerdo, lo que sugiere una aceptación general de la eficiencia del modelo, aunque algunos usuarios podrían considerar que hay margen para mejorar en la rapidez de procesamiento.

Análisis

Distribución de Puntuaciones:

- La mayoría de las puntuaciones se concentraron en las categorías "De Acuerdo" y "Muy de Acuerdo", indicando que el modelo tiende a clasificar comentarios positivos con mayor frecuencia.
- Las categorías "Muy en Desacuerdo" y "En Desacuerdo" tienen un menor número de puntuaciones, lo que podría sugerir una menor sensibilidad del modelo para detectar comentarios negativos.

Escalabilidad:

- El modelo puede ser ajustado y escalado para analizar grandes volúmenes de datos en Facebook, ya que la validación inicial muestra una tendencia clara en la clasificación de sentimientos.
- Se recomienda continuar con pruebas adicionales y ajustes en el algoritmo para mejorar la precisión en la detección de sentimientos negativos.

Conclusión

El modelo de análisis de sentimientos en Facebook ha demostrado ser efectivo en la clasificación de comentarios positivos. Sin embargo, se requiere una mayor atención en la mejora de la detección de sentimientos negativos. Se sugiere realizar ajustes en el algoritmo y realizar más validaciones con muestras de datos ampliadas.

Con estas acciones, se espera mejorar la precisión y efectividad del modelo de análisis de sentimientos en Facebook, proporcionando una herramienta confiable para el análisis de comentarios en redes sociales.

CAPITULO V

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

La selección de datasets variados y representativos es crucial para entrenar un modelo robusto de análisis de sentimientos. Utilizar datasets diversificados, que incluyan imágenes y texto, mejora la capacidad del modelo para generalizar y manejar diferentes tipos de entradas de datos.

Para el análisis de sentimientos, las Redes Neuronales se destacan como la opción más adecuada en comparación con Naive Bayes y las Máquinas de Soporte Vectorial (SVM). Mientras que Naive Bayes es eficiente y fácil de implementar para texto, su supuesto de independencia de características y limitaciones en la alta dimensionalidad lo hacen menos efectivo para tareas complejas. Las SVMs, aunque precisas y capaces de manejar datos de alta dimensionalidad, requieren un ajuste meticuloso de parámetros y son menos eficientes con grandes volúmenes de datos. Por otro lado, las Redes Neuronales, son superiores en capturar patrones complejos y no lineales, manejando eficazmente la secuencia y el contexto del texto. Su capacidad para detectar características locales y patrones específicos las hace altamente eficientes para el análisis de sentimientos, mejorando significativamente la precisión y robustez en esta tarea.

El uso de diferentes estructuras de aprendizaje supervisado, como Random Forest, Support Vector Machines (SVM), y redes neuronales profundas, ofrece una perspectiva variada sobre el rendimiento del modelo. Cada estructura tiene sus ventajas y desventajas dependiendo del tipo de datos y la complejidad de la tarea.

La validación del modelo con datos reales de Facebook es esencial para asegurar su efectividad en un entorno práctico. Los resultados indicaron que el modelo tiene

una buena capacidad para identificar sentimientos positivos, pero necesita mejoras en la detección de sentimientos negativos.

Recomendaciones

Se recomienda aplicar el modelo de análisis de sentimientos en Facebook debido a su probada efectividad en la clasificación de comentarios positivos. Aunque se ha identificado la necesidad de mejorar la detección de sentimientos negativos, las acciones sugeridas, como ajustar el algoritmo y realizar más validaciones con muestras de datos ampliadas, proporcionarán una mejora sustancial en la precisión y efectividad del modelo.

La implementación de este modelo permitirá una comprensión más profunda de las interacciones en la plataforma, facilitando la identificación de tendencias y sentimientos predominantes entre los usuarios. Además, ofrecerá una herramienta confiable para el análisis de comentarios en redes sociales, lo que es crucial para la toma de decisiones estratégicas y la mejora continua de la experiencia del usuario en Facebook.

Implementar y comparar diversas estructuras de aprendizaje supervisado para determinar cuál proporciona el mejor rendimiento. Utilizar técnicas de ensamblaje para combinar los resultados de varios modelos supervisados, mejorando la robustez y la precisión del análisis de sentimientos.

REFERENCIAS BIBLIOGRAFICAS

Liu, B. (2012). *Sentiment analysis and opinion mining* (Vol. 5). Morgan & Claypool Publishers. <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, 28(2), 15–21. <https://doi.org/10.1109/MIS.2013.30>

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10)* (pp. 1320–1326). European Language Resources Association (ELRA).

Huang, Z., Fang, Y., & Zhao, X. (2021). Big data analytics and artificial intelligence: The role of high-performance computing. *Decision Support Systems*, 146, 113544. <https://doi.org/10.1016/j.dss.2021.113544>

Banco Mundial. (2020). *World Development Indicators*. Recuperado de <https://databank.worldbank.org/source/world-development-indicators>

UTE. (2018). Influencia de redes sociales en el análisis de sentimiento aplicado a la situación política en Ecuador. Recuperado de: Enfoque UTE, V.9-N.1, Mar.2018, pp. 67 – 78 <http://ingenieria.ute.edu.ec/enfoqueute/> e-ISSN: 1390- 6542 / p-ISSN: 1390-9363

Chauhan, G. (2021, February 10). All about Naive Bayes. Towards Data Science. <https://towardsdatascience.com/all-about-naive-bayes-8e13cef044cf>

Lutz, M. (2019). *Learning Python: Powerful Object-Oriented Programming*. O'Reilly Media.

Miller, A. (2020). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media.

Hernández-Sampieri, R. & Mendoza, C. (2018). *Metodología de la Investigación: las Rutas Cuantitativa Cualitativa y Mixta*, Mc Graw Hill Interamericana editores S. A de C. V.

Godoy, D., & González, L. (2019). Big data para la priorización de zonas de atención a emergencias causadas por inundaciones en Bogotá Colombia: uso de las redes sociales. Tesis de Grado, Universidad de Ciencias Aplicadas y Ambientales, Ingeniería Geográfica y Ambiental, Bogotá. Obtenido de <https://core.ac.uk/download/pdf/326428196.pdf>

Perea, R. (2014). 5G: Una nueva generación para una nueva sociedad. Hamut'ay. Revista de divulgación científica de la Universidad Alas Peruanas, 1(1). Recuperado el 03/05/2024, de <http://revistas.uap.edu.pe/ojs/index.php/HAMUT/article/view/570/462>.

Ardini, C., & Nahúm, H. (2020). El uso del big data en política o la política del big data. Comunicación y Hombre(16), 225-240. Obtenido de <https://pruebasrevistas.ufv.es/index.php/comunicacionyhombre/article/view/604>
Westreicher, G. (01 de agosto de 2020). Análisis de datos. Recuperado el 25 de octubre de 2023, de <https://economipedia.com/definiciones/analisis-de-datos.html>.

Cedeño, D. & Vargas, M. (2020). Aprendizaje automático aplicado al análisis de sentimiento, Revista Tecnologica de vol. 16, Universidad Tecnologica de Panama.

Ebac. (2023). ¿Qué es el análisis de datos? Recuperado el 25 de octubre de 2023,de

<https://ebac.mx/blog/que-es-el-analisis-de-datos>.

Fernández, J. (2020). El análisis de datos de redes sociales como herramienta para la toma de decisiones. Revista de Gestión y Estrategia, 19(1), 105-122.

Rangel, E. (2020). El análisis de datos de redes sociales como herramienta para la gestión documental. Revista de Gestión y Economía, 29(1), 11-22.

Ortiz, R. (2022). El análisis de datos de redes sociales como herramienta para la investigación de mercados. Revista de Investigación en Marketing, 27(1), 1-12.

Cabeza, D. (2023). El análisis de datos en redes sociales como herramienta para el marketing empresarial. Revista de Marketing, 27(1), 3-12.

Ritter, M. (2023). Python para análisis de datos: Una guía completa. Madrid: RA-MA.

Castillo, B. E., Gómez, R. J., Taborda, Q. L., & Mejía, M. A. (2021). ¿Cómo Investigar en la UNIBE? (Primera ed.). Quito: Qualitas.

Smith, M., Jones, J., & Brown, B. (2022). The use of social media data analytics to understand consumer behavior. *Journal of Marketing*, 86(4), 76-92.

Williams, W., White, R., & Black, T. (2023). The impact of social media data analytics on marketing effectiveness. *Journal of the Academy of Marketing Science*, 51(1), 123- 140

Zhang, Y., Xu, Y., & Chen, Z. (2022). Predicting consumer trends with social media data analytics. *Journal of Business Research*, 136, 284-292.

Agarwal, A., & Liu, B. (2016). Sentiment analysis: A survey. *ACM Computing Surveys (CSUR)*, 48(4), 83.

Ramos, G. (2021). Análisis de sentimientos de datos de redes sociales usando técnicas de Machine Learning. Universitat de Barcelona, Diposit digital de documents de la AUB.

Cedeno, D. y Vargas, M. (2020). Aprendizaje automático aplicado al análisis de sentimiento. Universidad Tecnológica de Panamá. *Revista de ID tecnológico* Vol 16.

Salgado, N y Trujillo, G. (2024). Análisis de sentimientos en datos de redes sociales: aplicación de técnicas de procesamiento de lenguaje natural y machine learning para analizar opiniones y sentimientos en datos de redes sociales en el contexto de sistemas de información. *Revista Científica Domino de las Ciencias* Vol. núm. 1. Enero-Marzo, 2024, pp. 314-327.

Castillo, S. & Caicedo, P. (2019). Análisis de sentimientos , una herramienta para valorar la actitud de un estudiante frente a un curso. Asociación Colombiana de Facultades de Ingeniería.

Quintana, A. (2021). Análisis de los procesos de tratamiento de información en un estudio de análisis de sentimiento utilizando la tecnología de Google. *Vivat Academia. Revista de Comunicación*

Moran, D. (2022). Una comparación de las plataformas de análisis de sentimientos en Ingeniería de software: un estudio experimental. Pontificia Universidad Católica del Ecuador, Sede Esmeraldas. 6/77.

Palella, S., & Martins, F. (2006). Metodología de la investigación cuantitativa. Fondo Editorial de la Universidad Pedagógica Experimental Libertador, La editorial pedagógica de Venezuela, (pp. 128-129).

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, 28(2), 15-21.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.

Ghose, A., & Penmatsa, S. (2017). Sentiment analysis insocial media. In A. Abbasi, S. Hassanien, & A. B. Abraham (Eds.), *Social Media Analytics* (pp. 147-166). Springer.

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1-167.

Agarwal, A., & Mittal, N. (2020). Sentiment analysis using machine learning techniques. *Materials Today: Proceedings*, 32, 122-127.

Manrique, E. (2020) Machine Larning: análisis de lenguajes de programación y herramientas para desarrollo. *Revista Ibérica de Sistemas y Tecnologías de Información*, 586-589.

Noguera, D. (2023) Uso de técnicas de aprendizaje automático(Machine Learning) sobre datos estelares. *Universitat D'Alacant, Facultad de Ciencias*.

Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, 28(2), 15-21.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. CRC press.

Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features. In *European Conference on Machine Learning* (pp. 137-142). Springer.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.

Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1746-1751.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780.

Li, X., Zhang, S., Li, X., & Li, Q. (2015). Deep Convolutional Neural Networks for Multilingual Sentiment Analysis. *Mathematical Problems in Engineering*, 2015, 1-9.

Liu, B., Zhang, L., & Ye, Q. (2012). Sentiment Analysis of Chinese Microblogging: A Comparative Study. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3(3), 43.

Go, A., Bhayani, R., & Huang, L. (2009). Twitter Sentiment Classification using Distant Supervision. *CS224N Project Report*, Stanford.

Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic. *PLoS ONE*, 6(5), e19467.

Creswell, J. W., & Creswell, J. D. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications

Hernández Sampieri, R., & Mendoza Torres, C. (2018). *Metodología de la investigación*. McGraw-Hill Education.

Smith, J. A., & Heshusius, L. (2019). Closing down the conversation: The end of the quantitative-qualitative debate among educational inquirers. In *The SAGE Handbook of Qualitative Data Analysis* (pp. 537-554). Sage Publications.

González, J. M. (2021). Machine Learning para el análisis de opiniones en redes sociales. *Revista de Investigación en Ciencias de la Información*, 12(1), 1-15.

Martínez, M., Fernández, A., & Romero, S. (2020). Modelos predictivos para el análisis de sentimientos en redes sociales. *Revista Española de Documentación Científica*, 43(4), e236.

Lovera, F. A. & Cardinale, Y. (2023). Análisis de sentimientos en Twitter: Un estudio comparativo. *Revista Científica de Sistemas e Informática*, 3(1), e418. <https://doi.org/10.51252/rcsi.v3i1.418>

Ritter, A. (2023). Harnessing the Power of Social Media Data for Business Intelligence. <https://towardsdatascience.com/tagged/sentiment-analysis>

Espinosa Cevallos Paola Alejandra (2023) https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://dialnet.unirioja.es/descarga/articulo/9295447.pdf&ved=2ahUKEwjM75rr_cWGAX9mYQIHS4JDs0QFnoECDQQAQ&usg=AOvVaw0XWFIhwe1tqBx5rC_SWHNZ

Carrasco, R. A., Bueno, I., & Montero, J.-M. (s.f.). *Capítulo 27: Naive Bayes*. En *fundamentos de Ciencia de Datos con R*. Recuperado de <https://cdr-book.github.io/cap-naive-bayes.html>

Ricoy Lorenzo, C., (2006). Contribución sobre los paradigmas de investigación. *Educação (Santa Maria. Online)*, 31(1),11-22.[fecha de Consulta 6 de Junio de 2024]. ISSN: . Recuperado de: <https://www.redalyc.org/articulo.oa?id=117117257002>

Bastis Consultores <https://online-tesis.com/enfoque-de-la-investigacion/>

Lucy María Reidl Martínez

https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-50572012000100008#:~:text=Los%20dise%C3%B1os%20de%20investigaci%C3%B3n%20son,experimental%2C%20extra%C3%B1a%20y%20de%20error.

Scrum.org. (n.d.). *What is Scrum?* Retrieved from <https://www.scrum.org/resources/what-scrum-module.>

ANEXOS

Escala de Estimación para la Validación del Modelo de Análisis de Sentimiento en Facebook

El objetivo de este instrumento es utilizar una escala de estimación para la validación del modelo de análisis de sentimiento en Facebook, esto incluye evaluar su precisión y coherencia en la clasificación de sentimientos mediante la asignación de puntuaciones a muestras de datos reales

Indicaciones: Aquí tiene algunas alternativas de respuestas para expresar el grado de satisfacción, todas las preguntas deben ser respondidas acorde a su criterio.

1(Muy en desacuerdo),

2(En desacuerdo),

3(Neutral),

4(De acuerdo), y

5(Muy de acuerdo).

Escala de Estimación Modelo

INDICADOR	1	2	3	4	5
1.-El modelo clasifica correctamente las publicaciones positivas.					
2.-El modelo clasifica correctamente las publicaciones negativas.					
3.-El modelo clasifica correctamente las publicaciones neutrales.					
4.-Las clasificaciones de sentimiento del modelo son consistentes a través de diferentes publicaciones.					
5.-El análisis de sentimiento del modelo refleja adecuadamente el contexto de las publicaciones.					
6.-Los resultados del análisis de sentimiento proporcionados por el modelo son útiles para entender las opiniones de los usuarios.					
7.-El modelo de análisis de sentimiento se					

generaliza bien a diferentes tipos de publicaciones en Facebook					
8.-Los resultados del análisis de sentimiento del modelo son fáciles de interpretar.					
9. El modelo procesa y clasifica rápidamente los datos textuales.					
10. El modelo procesa y clasifica rápidamente las imágenes.					

Escala de Estimación Para Validación del modelo

INDICADOR	1	2	3	4	5
1.-El modelo clasifica correctamente las publicaciones positivas.				x	
2.-El modelo clasifica correctamente las publicaciones negativas.				x	
3.-El modelo clasifica correctamente las publicaciones neutrales.				x	
4.-Las clasificaciones de sentimiento del modelo son consistentes a través de diferentes publicaciones.					x
5.-El análisis de sentimiento del modelo refleja adecuadamente el contexto de las publicaciones.					x
6.-Los resultados del análisis de sentimiento proporcionados por el modelo son útiles para entender las opiniones de los usuarios.					x
7.-El modelo de análisis de sentimiento se generaliza bien a diferentes tipos de publicaciones en Facebook				x	
8.-Los resultados del análisis de sentimiento del modelo son fáciles de interpretar.					x
9. El modelo procesa y clasifica rápidamente los datos textuales.				x	

10. El modelo procesa y clasifica rápidamente las imágenes.				x	
---	--	--	--	---	--

Validado por: ___Alejandro Coronel_____

Profesión: ___Ingeniero en sistemas_____

Cargo que desempeña: ___Consultor Técnico_____

Firma:  _____

Fecha: _____07/08/2024_____

Escala de Estimación Para Validación del modelo

INDICADOR	1	2	3	4	5
1.-El modelo clasifica correctamente las publicaciones positivas.				X	
2.-El modelo clasifica correctamente las publicaciones negativas.				X	
3.-El modelo clasifica correctamente las publicaciones neutrales.				X	
4.-Las clasificaciones de sentimiento del modelo son consistentes a través de diferentes publicaciones.			X		
5.-El análisis de sentimiento del modelo refleja adecuadamente el contexto de las publicaciones.				X	
6.-Los resultados del análisis de sentimiento proporcionados por el modelo son útiles para entender las opiniones de los usuarios.				X	
7.-El modelo de análisis de sentimiento se generaliza bien a diferentes tipos de publicaciones en Facebook			X		
8.-Los resultados del análisis de sentimiento del modelo son fáciles de interpretar.					X
9. El modelo procesa y clasifica rápidamente los				X	

datos textuales.					
10. El modelo procesa y clasifica rápidamente las imágenes.				X	

Validado por: David Páez

Profesión: Desarrollador de Software

Cargo que desempeña: Desarrollador de Software



Firma:

Fecha: 7/Agosto/2024

Escala de Estimación Para Validación del modelo

INDICADOR	1	2	3	4	5
1.-El modelo clasifica correctamente las publicaciones positivas.				X	
2.-El modelo clasifica correctamente las publicaciones negativas.				X	
3.-El modelo clasifica correctamente las publicaciones neutrales.				X	
4.-Las clasificaciones de sentimiento del modelo son consistentes a través de diferentes publicaciones.			X		
5.-El análisis de sentimiento del modelo refleja adecuadamente el contexto de las publicaciones.				X	
6.-Los resultados del análisis de sentimiento proporcionados por el modelo son útiles para entender las opiniones de los usuarios.				X	
7.-El modelo de análisis de sentimiento se generaliza bien a diferentes tipos de publicaciones				X	

en Facebook					
8.-Los resultados del análisis de sentimiento del modelo son fáciles de interpretar.					X
9. El modelo procesa y clasifica rápidamente los datos textuales.			X		
10. El modelo procesa y clasifica rápidamente las imágenes.			X		

Validado por: Byron Sandoval

Profesión: Analista Desarrollo

Cargo que desempeña: Analista Desarrollo

Firma: 

Fecha: 01-08-2024

Escala de Estimación Para Validación del modelo

INDICADOR	1	2	3	4	5
1.-El modelo clasifica correctamente las publicaciones positivas.				X	
2.-El modelo clasifica correctamente las publicaciones negativas.				X	
3.-El modelo clasifica correctamente las publicaciones neutrales.				X	
4.-Las clasificaciones de sentimiento del modelo son consistentes a través de diferentes publicaciones.					X
5.-El análisis de sentimiento del modelo refleja adecuadamente el contexto de las publicaciones.					X
6.-Los resultados del análisis de sentimiento proporcionados por el modelo son útiles para entender las opiniones de los usuarios.					X
7.-El modelo de análisis de sentimiento se generaliza bien a diferentes tipos de publicaciones en Facebook				X	
8.-Los resultados del análisis de sentimiento del modelo son fáciles de interpretar.				X	
9. El modelo procesa y clasifica rápidamente los datos textuales.				X	
10. El modelo procesa y clasifica rápidamente las imágenes.				X	

Validado por: Fernando Tutucha

Profesión: Eng. Sistema de la información

Cargo que desempeña: Analista de tecnologías 2

Firma: 

Fecha: 08-Agosto-2024

Escala de Estimación Para Validación del modelo

INDICADOR	1	2	3	4	5
1.-El modelo clasifica correctamente las publicaciones positivas.				X	
2.-El modelo clasifica correctamente las publicaciones negativas.				X	
3.-El modelo clasifica correctamente las publicaciones neutrales.				X	
4.-Las clasificaciones de sentimiento del modelo son consistentes a través de diferentes publicaciones.					X
5.-El análisis de sentimiento del modelo refleja adecuadamente el contexto de las publicaciones.					X
6.-Los resultados del análisis de sentimiento proporcionados por el modelo son útiles para entender las opiniones de los usuarios.					X
7.-El modelo de análisis de sentimiento se generaliza bien a diferentes tipos de publicaciones en Facebook					X
8.-Los resultados del análisis de sentimiento del modelo son fáciles de interpretar.				X	
9. El modelo procesa y clasifica rápidamente los datos textuales.				X	
10. El modelo procesa y clasifica rápidamente las imágenes.				X	

Validado por: Grego Segovia

Profesión: Ing. Sistemas

Cargo que desempeña: Administrador Base de Datos

Firma: 

Fecha: 08 Agosto - 2024

<https://github.com/ErickErraez/modelo>