

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO



**Modelo de Medición de Desempeño de Servicios Financieros en Perú a Través de
Análisis de Sentimiento Utilizando Métodos de Decisiones Multicriterio**

**TESIS PARA OBTENER EL GRADO DE MAGÍSTER EN
GERENCIA DE TECNOLOGÍAS DE INFORMACIÓN OTORGADO POR
LA PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ**

PRESENTADA POR

Araceli Yoselín Cueva Mendoza, DNI: 71452422

César Blademir Cueva Mendoza, DNI: 44125093

Judith Cathy Huacac Huañec, DNI: 41585382

Ananda Adelaida Junco Navarro, DNI: 10680728

ASESOR

Dr. Luis Alfredo Negrón Naldos, DNI: 10788917

ORCID 0000-0003-1328-0323

JURADO

Mag. Gonzalo Javier Cachay Silva

Dr. Marco Antonio Salcedo Huarcaya

Surco, noviembre 2022

Declaración Jurada de Autenticidad

Yo, Luis Alfredo Negrón Naldos, docente del Departamento Académico de Posgrado en Negocios de la Pontificia Universidad Católica del Perú, asesor(a) de la tesis/el trabajo de investigación titulado: "Modelo de Medición de Desempeño de Servicios Financieros en Perú a Través de Análisis de Sentimiento Utilizando Métodos de Decisiones Multicriterio".

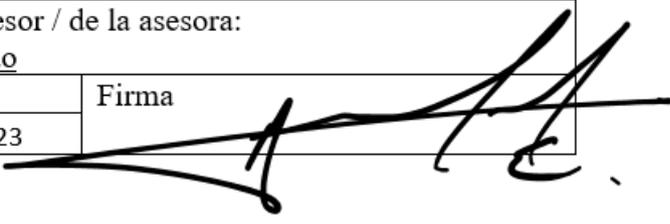
del/de la autor(a)/ de los(as) autores(as)

- Araceli Yoselín Cueva Mendoza, DNI: 71452422
- César Blademir Cueva Mendoza, DNI: 44125093
- Judith Cathy Huacac Huañec, DNI: 41585382
- Ananda Junco Navarro, DNI: 10680728

dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 17%. Así lo consigna el reporte de similitud emitido por el software *Turnitin* el 11/07/2023.
- He revisado con detalle dicho reporte y confirmo que cada una de las coincidencias detectadas no constituyen plagio alguno.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha:Lima, 11 de Julio del 2023.....

Apellidos y nombres del asesor / de la asesora: <u>Negron Naldos, Luis Alfredo</u>	
DNI:10788917	Firma 
ORCID: 0000-0003-1328-0323	

Agradecimientos

Un especial agradecimiento a nuestro asesor; Dr. Luis Negrón, que con su apoyo en sus tiempos y asesorías se logró desarrollar esta tesis con mucho esfuerzo, entusiasmo y dedicación. A todos los profesores de CENTRUM PUCP – MGTI III que nos demostraron calidad, aptitud y vocación de enseñanza, agradecer por su tiempo en compartir sus experiencias en cada clase los fines de semana y en tiempos difíciles del COVID.



Dedicatoria

A mis padres y hermanos por ser el principal impulso para encaminarme en mi vida profesional, al haberme acompañado y orientado en este proceso.

Araceli Yoselín Cueva Mendoza.

A mi hijo, mis padres y mis hermanas por la motivación constante , a mi novia por todo el apoyo incondicional y aliento para lograr mis objetivos

César Blademir Cueva Mendoza.

A mi esposo e hija por su paciencia, amor y soporte acompañándome en esta travesía, a mis padres y hermanos por el apoyo y motivación, a mi madre por su amor incondicional.

Judith Cathy Huacac Huañec.

A mi esposo, por apoyarme en poder realizar esta maestría, a mis hijos por la paciencia y comprensión en los tiempos de ausencia. A mi madre, por el gran esfuerzo y dedicación en mi educación.

Ananda Junco Navarro.

Resumen Ejecutivo

El presente trabajo de investigación tuvo como objetivo establecer un modelo de medición de enfoque cuantitativo, puesto que se busca resolver un problema práctico a partir de la teoría; para luego, brindar desempeño de servicios financieros haciendo uso de una de las técnicas de procesamiento de lenguaje natural (PLN) como es el análisis de sentimiento que, complementado con un método de decisiones multicriterio (MCDM), permite realizar una evaluación de productos o servicios ofrecidos por los cuatro principales bancos de Perú, analizando opiniones que expresan sentimientos positivos, negativos o neutrales en la red social Twitter. La investigación permitió identificar un conjunto de aspectos o criterios a partir de los cuales se evalúa la polaridad del sentimiento expresada en los tuits, generando un consolidado de tablas de puntuación, para finalmente evaluar el desempeño de cada entidad bancaria según los aspectos identificados. La investigación se estableció dentro de un marco de diseño como una investigación aplicada, bajo una fuente de información alternativa que soporte la toma de decisión de un cliente bancario al momento de contratar un servicio financiero, de tal manera que se pueda mejorar su satisfacción. Para la construcción del modelo se evaluaron 15,546 tuits de los que se escogieron 5,276 en los cuales se menciona al menos un aspecto. Dicho conjunto de datos se evaluó con métodos de polarización de sentimiento como Stanza, VADER, TextBlob y BETO (BERT) y; posteriormente, se utilizó el método MCDM-VIKOR para proporcionar información adicional que permita determinar la mejor opción de servicio a contratar. Los resultados obtenidos arrojan que el mejor método para determinar la polaridad de los tuits es BETO, logrando un nivel de exactitud de 88%, además de verificar que tanto BETO como VIKOR cumplen con las condiciones de aceptabilidad definidas por los métodos, logrando con ello identificar de manera adecuada la mejor alternativa de servicio.

Palabras clave: Procesamiento de lenguaje natural, análisis de sentimiento, satisfacción del cliente, MCDM, VIKOR

Abstract

This study aims to create a model for measuring the performance of financial services using Natural Language Processing (NLP) techniques like Sentiment Analysis (SA) and a Multicriteria Decision Method. (MCDM) enables the evaluation of products or services offered by Peru's four major banks by analyzing positive, negative, or neutral opinions expressed on the social network Twitter. The study identifies a set of aspects or criteria for scoring the sentiment expressed in tweets, generates a consolidated score table, and finally evaluates the performance of each banking entity based on the identified aspects. The research was established within a design framework as applied research, using a quantitative approach, since it seeks to solve a practical problem based on the theory and then provide an alternative source of information that supports a bank client's decision-making when hiring a financial service to improve satisfaction. For the model's construction, 15,546 tweets were evaluated, with only 5,276 chosen that mentioned at least one aspect. The data set was analyzed using sentiment polarization methods such as Stanza, VADER, TextBlob and BETO (BERT), and then the MCDM VIKOR provides additional information to determine the best service option to contract. The results obtained show us that the best method to determine the polarity of the tweets is BETO, achieving an accuracy level of 88%, in addition to verifying that both BETO and VIKOR meet the acceptability conditions defined by the methods; achieving this adequately resolve the best service alternative.

Keywords: Natural language processing, sentiment analysis, customer satisfaction, MCDM, VIKOR

Tabla de Contenidos

Lista de Tablas	x
Lista de Figuras.....	xii
Capítulo I: Introducción	1
Antecedentes.....	2
Problema de la Investigación.....	5
Propósito de la Investigación.....	6
Objetivo general	6
Objetivos específicos.....	6
Preguntas de la Investigación	6
Justificación de la Investigación.....	7
Marco Teórico Conceptual	7
Satisfacción del cliente.....	8
Análisis de sentimiento	9
Enfoques para la clasificación del análisis de sentimiento	9
Métodos de decisión multicriterio (MCDM)	11
Modelo de clasificación mediante MCDM y análisis de sentimiento.....	15
Clasificación utilizando MCDM	18
Definición de Términos del Estudio.....	18
Limitaciones	19
Delimitaciones	20
Resumen	20
Capítulo II: Revisión de la Literatura	22
Análisis de Sentimiento	22
Procesamiento de Lenguaje Natural (PLN).....	23

Preprocesamiento de Texto	25
Extracción.....	25
Limpieza.....	26
Tokenización	26
Lematización	27
Etiquetado morfosintáctico.....	27
Análisis de Decisiones con Múltiples Criterios (MCDM)	29
Redes Sociales	32
Facebook	32
Twitter	33
LinkedIn	33
Instagram	33
TikTok.....	34
Matriz de Confusión	34
Exactitud.....	34
Precisión	35
Exhaustividad o sensibilidad.....	35
Especificidad	35
Valor F1.....	35
Proceso del Método VIKOR.....	37
Resumen	38
Capítulo III: Metodología	40
Diseño de la Investigación.....	40
Justificación del Diseño.....	40
Población	41

Muestra	42
Consentimiento Informado	42
Procedimiento de Recolección de Datos	43
Instrumentos de Medición	43
Análisis e Interpretación de Datos	43
Recopilación de datos.....	43
Preprocesamiento	44
Determinación de la opinión orientada a aspectos	44
Puntuación del sentimiento	44
Clasificación utilizando MCDM	45
Validez y Confiabilidad.....	45
Resumen	45
Capítulo IV: Presentación y Análisis de Resultados.....	47
Recopilación de Datos (Etapa 1)	47
Proceso: Limpieza de datos y eliminación de palabras vacías.....	47
Proceso: Tokenización	49
Proceso: Determinación de frecuencia de ocurrencias.....	50
Proceso: Identificación de aspectos.....	50
Proceso: Filtrado de reseñas con al menos un aspecto.....	51
Preprocesamiento y Determinación de Criterios (Etapa 2)	51
Proceso: Tokenización y división de oraciones	52
Proceso: Expansión de tókenes de varias palabras.....	52
Proceso: Lematización	53
Proceso: Etiquetado morfosintáctico.....	53
Proceso: Análisis sintáctico de dependencias	53

Proceso: Reconocimiento de las entidades nombradas	54
Determinación de la Opinión Orientada a Aspectos (Etapa 3).....	54
Generación de gráficos de análisis	54
Proceso: Identificación de aspectos similares	55
Puntuación del Sentimiento (Etapa 4)	56
Proceso: Determinar la polaridad de sentimientos	56
Proceso: Comparar métodos de polaridad de sentimiento	57
Proceso: Generar tabla de puntuación	60
Clasificación Utilizando MCDM (Etapa 5).....	62
Proceso: Análisis jerárquico (AHP)	62
Proceso: Aplicación del método VIKOR	65
Discusión de Resultados	66
Resumen	71
Capítulo V: Conclusiones y Recomendaciones.....	72
Conclusiones.....	72
Recomendaciones	74
Referencias.....	75
Apéndice A Código Fuente de Limpieza y Modelos para Determinar la Polaridad	
de los Tuits.....	86

Lista de Tablas

Tabla 1	<i>Intensidad de Importancia</i>	12
Tabla 2	<i>Ejemplo de Salida para el Proceso de Tokenización (Sentencia 1)</i>	26
Tabla 3	<i>Ejemplo de Salida para el Proceso de Tokenización (Sentencia 2)</i>	26
Tabla 4	<i>Ejemplo de Salida para el Proceso de Lematización</i>	27
Tabla 5	<i>Conjunto de Etiquetas de Dependencias Universales</i>	28
Tabla 6	<i>Características de los Tuits</i>	41
Tabla 7	<i>Cuenta Twitter por Banco</i>	44
Tabla 8	<i>Datos en Archivos de Tipo CSV por cada Entidad Bancaria Extraídos del 1 de Enero al 1 de Septiembre de 2022</i>	47
Tabla 9	<i>Lista de Caracteres que Generan Ruido</i>	48
Tabla 10	<i>Tuits Procesados</i>	48
Tabla 11	<i>Tuit Tokenizado para el Banco D</i>	49
Tabla 12	<i>Palabras Identificadas por Frecuencia</i>	50
Tabla 13	<i>Tokenización</i>	52
Tabla 14	<i>Expansión de Tókenes de Varias Palabras</i>	52
Tabla 15	<i>Aspectos Identificados con Palabras Similares</i>	55
Tabla 16	<i>Resultados de Predicción del Método de Polaridad Stanza</i>	58
Tabla 17	<i>Resultados de Predicción del Método de Polaridad VADER</i>	58
Tabla 18	<i>Resultados de Predicción del Método de Polaridad TextBlob</i>	59
Tabla 19	<i>Resultados de Predicción del Método de Polaridad BETO</i>	60
Tabla 20	<i>Puntuación de Sentimiento para el Banco A</i>	60
Tabla 21	<i>Puntuación de Sentimiento para el Banco B</i>	61
Tabla 22	<i>Puntuación de Sentimiento para el Banco C</i>	61
Tabla 23	<i>Puntuación de Sentimiento para el Banco D</i>	61

Tabla 24	<i>La Matriz de Resultados para los Cuatro Bancos y los Respectivos Aspectos....</i>	62
Tabla 25	<i>Matriz de Comparación por Pares de los Aspectos.....</i>	63
Tabla 26	<i>Matriz Normalizada de Comparación de Aspectos</i>	63
Tabla 27	<i>Matriz Normalizada de Comparación de Aspectos</i>	64
Tabla 28	<i>Vector de Prioridades</i>	64
Tabla 29	<i>Función de Variables a Maximizar.....</i>	65
Tabla 30	<i>Cálculo de la Utilidad (S) y el Arrepentimiento (R) de cada Aspecto</i>	65
Tabla 31	<i>Cálculo el Índice Q de cada Alternativa.....</i>	65
Tabla 32	<i>Ordenamiento de las Alternativas de Acuerdo con el Índice Q.....</i>	66
Tabla 33	<i>Descripción de Resultados de Predicción por Modelo de Polarización</i>	69

Lista de Figuras

Figura 1	<i>Modelo de Clasificación Mediante MCDM y Análisis de Sentimiento</i>	16
Figura 2	<i>Métricas de la Matriz de Confusión (Precisión y Exactitud)</i>	36
Figura 3	<i>Distribución de Aspectos Identificados con Mayor Frecuencia</i>	51
Figura 4	<i>Resultado de Lematización</i>	53
Figura 5	<i>Etiquetado Morfosintáctico</i>	53
Figura 6	<i>Análisis Sintáctico de Dependencias</i>	54
Figura 7	<i>Generación de Gráficos de Análisis</i>	55
Figura 8	<i>Reseña del Total de Tuits por cada Entidad Bancaria</i>	56
Figura 9	<i>Matriz de Confusión Utilizando la Librería Stanza</i>	57
Figura 10	<i>Matriz de Confusión Utilizando la Librería VADER</i>	58
Figura 11	<i>Matriz de Confusión Utilizando la Librería TextBlob</i>	59
Figura 12	<i>Matriz de Confusión Utilizando la Librería BETO</i>	59

Capítulo I: Introducción

El sector financiero es de alta importancia para el crecimiento y la estabilidad económica de un país, puesto que permite el incremento del producto bruto interno (PBI) y ayuda en la reducción de la tasa de desempleo (Kumbirai & Webb, 2010, p. 31). Dentro de los entes que brindan soporte a este sector se encuentran los mercados de capital, las gestoras de fondos de inversión, los bancos, las entidades regulatorias, entre otros. En cuanto a los bancos, desde su aparición han ido evolucionando y adaptándose a las necesidades de sus clientes y, gracias a la tecnología desde inicios del siglo XXI se vive una transformación sin precedentes en la historia de la banca; todo ello marcado por una revolución de los servicios financieros a través de la tecnología puesto que con mayor frecuencia surgen nuevos competidores con base digital, que amenazan con quebrar el statu quo de la banca tradicional. Este escenario obliga a los bancos a: pensar en el usuario, cómo brindarle servicios a medida, ayudarlo a crecer y facilitarle la vida; y que todo lo indicado resulte siendo rentable, sostenible y perdurable en el tiempo (Scasserra, 2022, pp. 10-20).

La convergencia de la comunicación y del internet trae consigo el avance de la comunicación digital, cuyos medios cumplen un rol central y definitivo en esta revolución (Kaul, 2012, pp. 114-115). Con el auge de los medios digitales han cambiado las formas de relacionarse los clientes con las empresas (Assaad & Marx Gómez, 2011, pp. 13-22), permitiendo una comunicación bidireccional a través de medios digitales tales como redes sociales, blogs, sitios especializados, etc. Esta realidad no solo ha permitido mejorar los esfuerzos de las organizaciones para diseñar e implementar innovadores programas de mercadeo, sino también asumir el reto de entender y conocer los comportamientos, hábitos y percepciones de sus clientes que son expresados por estas nuevas plataformas tales como redes sociales, que en la actualidad se han convertido en un medio de comunicación directa entre las organizaciones y sus consumidores, lo que conlleva a que los usuarios conozcan más sobre sus productos y/o servicios dentro de estas plataformas. Es de esta manera que las

organizaciones aplican estrategias de vínculo emocional con el consumidor de manera que el usuario objetivo pueda percibir a la organización de manera positiva (Pretel Jiménez et al., 2018, p. 230).

Alrededor del 80% de los consumidores cambian su decisión de compra basada en opiniones positivas o negativas que encuentran dentro de las redes sociales (Kutchera et al., 2014); por lo que resulta importante brindarles la posibilidad de elegir el servicio más adecuado a su necesidad. En ese sentido y enfocado en el sector financiero, la presente investigación propone un modelo de medición de desempeño de servicios financieros mediante el análisis de sentimiento utilizando métodos de decisión multicriterio (MCDM, por sus siglas en inglés), que ayuden a comparar las opiniones en la red social Twitter de un conjunto de servicios para ayudar al cliente a elegir, evitando la tarea engorrosa de analizar toda la alta cantidad de comentarios y seleccionar el más apto. Para ello, se plantea analizar estas opiniones mediante técnicas de procesamiento del lenguaje natural para la clasificación de mensajes en positivos, negativos o neutrales para luego mediante un enfoque de MCDM seleccionar un banco que se ajuste a su necesidad específica.

Antecedentes

En las entidades financieras existe el interés creciente por el análisis de sentimientos puesto que el éxito de un banco depende de la relación de alta calidad con los clientes. Por lo tanto, se requiere medir continuamente los niveles de satisfacción, puesto que los clientes satisfechos son el verdadero activo para una entidad bancaria que garantiza su retención a largo plazo generando fidelidad y rentabilidad económica. Es por lo que a continuación se detallan diferentes estudios que han analizado la problemática desde distintos enfoques.

Al Karim y Chowdhury (2014) en su investigación “Satisfacción del Cliente sobre la Calidad del Servicio en el Sector de la Banca Comercial Privada en Banglades”, realizaron encuestas a una población pequeña en diferentes bancos donde los datos se analizaron utilizando el software Statistical Package for the Social Sciences (SPSS). El estudio demostró

que existen diversos factores combinados que determinan la calidad del servicio en la industria bancaria como son la tangibilidad, la confiabilidad, la capacidad de respuesta, la seguridad y la empatía; los cuales influyeron de manera significativa y positiva en las actitudes de los clientes y son factores clave que afectan la satisfacción de estos. Además, este estudio determinó que mejorar la calidad del servicio en los bancos es uno de los factores que influye en la satisfacción del cliente, incluyendo una lista de criterios como: interés, préstamo, cuenta a cargo, hipoteca, tasas, etc. Para la presente investigación se considera una lista de criterios identificados conformada por tarjeta, cuenta, cajero, sistema y agencia.

Greaves et al. (2014) en su artículo “Tweets sobre la Calidad Hospitalaria: Un Estudio de Métodos Mixtos”, midieron la calidad de atención hospitalaria hacia los pacientes del Servicio Nacional de Salud (NHS por sus siglas en inglés) de Inglaterra, obteniendo data desde Twitter. La investigación tuvo como finalidad extraer información cuantitativa con una muestra de 198,499 tuits enviados a hospitales ingleses que forman la red NHS por un periodo de un año; y a nivel cualitativo se analizó dicha información según el tipo de contenido que se tiene en esta red social, obteniendo de forma aleatoria 1,000 tuits relacionándolos con uno o más temas clave identificados como: calidad, información de salud, información práctica u organizacional sobre el hospital, mensajes promocionales y mensajes para pacientes que reciben cuidado, entre otros. Se aplicaron algoritmos de análisis de sentimiento con los tuits recopilados, obteniendo una puntuación de sentimiento para cada tuit, clasificando la polaridad en positivo con 47%, negativo con 6% y neutral con 47%. Esta investigación fue utilizada como base para el sustento en la aplicación de métodos y algoritmos con el uso del análisis de sentimiento teniendo como fuente de información las redes sociales.

Panico (2018) en su tesis de grado “La Eficacia del Análisis de Sentimientos para la Empresa: El Caso de Estudio Dell Technologies Inc.”, abordó la efectividad del análisis de sentimiento utilizado por la empresa de tecnología Dell, la cual viene transformándose como

organización, diversificando su cartera de productos y servicios. Esta diversificación es la razón del por qué era necesario para esta empresa entender a sus clientes por medio de las redes sociales, las cuales son utilizadas para ofrecer nuevos productos, monitoreando las opiniones y menciones sobre la marca en Twitter, clasificándolas en *dominios* como: envío/entrega, productos/soluciones, soporte técnico, página web, consultoría, instalación, etc. Para ello, Dell aplica técnicas y algoritmos utilizando un motor de procesamiento de lenguaje natural (PLN) patentado, basado en reglas y un diccionario que le permite segmentar y analizar las opiniones de los usuarios, clasificándolas como positivas, negativas o neutrales y, evaluadas por la escala de sentimiento definida por la métrica utilizada. La investigación de Panico ayuda a reconocer la importancia para las organizaciones de utilizar técnicas de análisis de sentimiento puesto que, con los resultados se puede tanto medir el desempeño de productos o servicios como el monitoreo y la retroalimentación de las opiniones.

Kumar (2018) en su investigación “Un Enfoque de Toma de Decisiones de Criterios Múltiples para Recomendar un Producto Utilizando Análisis de Sentimiento”, describió el propósito de recomendar un producto o servicio a partir de las opiniones de los usuarios que se encuentran en las plataformas digitales. Se observó que los usuarios antes de realizar una compra revisan las opiniones que se encuentran en estos medios y; para ello, utilizan métodos de aprendizaje automático (*machine learning*) que permiten clasificar los sentimientos en función de las opiniones, relacionándolas con las características de un producto o servicio determinado y considerando las múltiples combinaciones obtenidas. Sin embargo, los resultados no permitieron concluir en recomendar el mejor o peor producto; por ello, para facilitar la tarea al usuario se aplica el MCDM. Ahora bien, para obtener los resultados, se definieron los casos; en ese sentido, para la selección de la mejor cámara digital y el mejor teléfono inteligente, utilizaron un proceso de recopilación de reseñas desde Amazon y FlipKart para luego realizar la identificación de características, puntuación y clasificación de sentimiento y; por último, aplicar el enfoque de selección de multicriterio con la técnica de

ordenación de preferencias por similitud a la solución ideal (TOPSIS, por sus siglas en inglés). Finalmente, se obtuvo como resultado el mejor producto a partir de la asignación de pesos de acuerdo con la importancia de sus atributos, considerando que la alternativa seleccionada se aproxime a la solución ideal positiva. El aporte de este estudio es considerar la aplicación de la integración de ambas metodologías como es el caso del análisis de sentimiento y el enfoque de decisiones de múltiples criterios teniendo como finalidad la recomendación al usuario sobre el mejor producto o servicio a partir de las opiniones evaluadas.

Problema de la Investigación

Las personas toman decisiones todo el tiempo, siendo esta una tarea compleja, puesto que involucra a diferentes actores de opinión como: personas, grupos de personas, instituciones y organizaciones del Estado. En la última década la toma de decisiones está fuertemente influenciada por las redes sociales debido a que dentro de estas existe una diversidad de foros de discusión y sitios de comercio electrónico, donde los clientes pueden expresar experiencias acerca de los productos y/o servicios que contratan lo cual confirma que las redes sociales se han convertido en una fuente de datos única para obtener información de opiniones positivas, negativas o neutras y; recopilar esta información resulta una tarea compleja para el consumidor al momento de elegir el artículo más adecuado para su necesidad. De acuerdo con Sherlock Communications (2021), en Latinoamérica los peruanos son quienes más confían en las redes sociales (76%) antes de realizar una compra por Internet.

En el Perú el sector financiero ofrece una variedad de alternativas de servicios mediante las redes sociales u otros medios, donde se evidencia la compleja tarea para el cliente de seleccionar un servicio financiero que ayude a cubrir sus necesidades. Existen diferentes plataformas que muestran opciones de decisión sobre productos, pero no existe una que mida los sentimientos desde una red social Twitter, donde un usuario pueda tomar una

decisión de acuerdo con su necesidad, basada en los mensajes realizados por otros usuarios acerca de la calidad de un servicio o producto que ofrecen diversas entidades financieras.

Propósito de la Investigación

Para abordar el propósito de la investigación referida al sector financiero, se definieron tanto su objetivo general como sus objetivos específicos, los cuales se detallan a continuación.

Objetivo general

Desarrollar un modelo de medición de desempeño de servicios financieros en Perú, mediante el análisis de sentimiento y el enfoque de toma de decisiones de criterios múltiples de publicaciones sobre experiencias y percepciones que realizan los usuarios en la red social Twitter.

Objetivos específicos

Se definieron cuatro objetivos específicos, los cuales son:

1. Consolidar en una base de datos los comentarios registrados por los clientes en la plataforma Twitter acerca de los servicios financieros de los principales bancos en Perú.
2. Determinar los aspectos de evaluación para el modelo.
3. Determinar la polarización de cada tuit mediante métodos de PLN basados en un lexicón.
4. Establecer el método de enfoque decisión de criterios múltiple que proporcione información adicional para determinar la mejor opción de servicio a contratar.

Preguntas de la Investigación

La presente investigación propone responder las siguientes interrogantes:

- ¿Es la plataforma Twitter una fuente de datos para recolectar y consolidar los comentarios de usuarios sobre servicios financieros?
- ¿Cuáles son los principales aspectos de evaluación necesarios en el modelo?
- ¿Cómo determinar la polarización de tuits?

- ¿La aplicación de MCDM permite seleccionar, comparar y clasificar diferentes servicios financieros?

Justificación de la Investigación

Los beneficios esperados luego de la implementación de la investigación son:

- Como justificación social: lograr obtener un resumen de la polarización de las opiniones de los servicios financieros a partir de los comentarios en la red social Twitter, con el fin de que los clientes puedan identificar entre las principales entidades financieras el mejor servicio que se adapte a sus necesidades.
- Como justificación práctica: reducir la complejidad de analizar una significativa cantidad de datos estructurados y no estructurados; y
- Como justificación teórica: la presente investigación se realiza con el propósito de aportar al conocimiento existente sobre el uso del método de clasificación de decisión multicriterio, donde el proceso jerárquico de análisis AHP permite calcular pesos para cada aspecto o criterio definido. Además, permite aplicar el método VIKOR para lograr un mejor resultado a través de un ranking de alternativas de los productos o servicios que ofrecen los bancos.

Marco Teórico Conceptual

En esta sección se describen los principales conceptos teóricos que ayudan a entender el análisis de sentimiento de los clientes a través de las redes sociales, que de acuerdo con la calificación en positivo, negativo o neutral de sus mensajes, los clientes los consideran para tomar una mejor decisión al momento de elegir un producto o servicio. Se describe de forma breve la satisfacción del cliente y con mayor detalle se abordan los conceptos y aplicaciones del análisis de sentimiento, que combina técnicas de minería de texto y PLN, así como también los diferentes métodos de decisión multicriterio (MCDM). Por último, se describen los pasos necesarios del modelo de clasificación mediante estos métodos y aplicando análisis de sentimiento, los cuales son: recopilación de datos, reprocesamiento, determinación de la opinión orientada a aspectos, puntuación del sentimiento y clasificación utilizando MCDM.

Satisfacción del cliente

La satisfacción del cliente depende del desempeño percibido de un producto en relación con las expectativas del comprador; en ese sentido, si el desempeño del producto es inferior a las expectativas, el cliente queda insatisfecho; si el desempeño es igual a las expectativas, el cliente está satisfecho; y si el desempeño es superior a las expectativas, el cliente está muy satisfecho e incluso *encantado* (Kotler & Armstrong, 2017). Por otro lado, la mayoría de los estudios muestran que un nivel más alto de satisfacción del cliente lleva a mayor lealtad por parte de este, lo que a su vez provoca un mejor desempeño de la empresa. Las empresas inteligentes tienen como meta encantar a sus clientes prometiéndoles solamente lo que pueden entregar y entregando más de lo que prometieron. Los clientes encantados no solo harán compras repetidas, sino también tendrán la disposición de convertirse tanto en socios de promoción de la empresa como “clientes predicadores” que comuniquen a los demás sus buenas experiencias (Kotler & Armstrong, 2017).

Según Zeithaml et al. (2018, pp. 80-81), la satisfacción también puede estar relacionada con otros tipos de sentimiento, dependiendo del contexto o tipo de servicio. La satisfacción también puede estar asociada con sentimientos de placer por los servicios que hacen que el cliente se sienta bien o están asociados con una sensación de felicidad. En algunas situaciones, donde la eliminación de un negativo conduce a la satisfacción, el cliente puede asociar una sensación de alivio con satisfacción. Finalmente, la satisfacción puede estar asociada con sentimientos de ambivalencia cuando existe una mezcla de experiencias positivas y negativas asociadas con el producto o servicio. La satisfacción del cliente está influenciada por: (a) el producto específico o las características del servicio, (b) las percepciones de la calidad del producto y del servicio y (c) el precio. Además, factores personales como el estado de ánimo o el estado emocional del cliente y factores situacionales como las recomendaciones de los miembros de la familia influyen en la satisfacción (Zeithaml et al., 2018, p. 79).

Deloitte (2018) sostuvo que la recomendación tiene alto impacto positivo comercial en la organización, asegurando la lealtad de los clientes con la marca y considerando las exigencias actuales de los consumidores. Es más probable que los clientes recuerden la experiencia positiva con la marca que los descuentos y/o beneficios ofrecidos por la organización; por lo que se recomienda que las empresas hagan un esfuerzo constante trabajando en un vínculo emocional positivo a largo plazo.

Análisis de sentimiento

Análisis de sentimiento viene a ser un conjunto de métodos, técnicas y herramientas que permiten detectar y clasificar información subjetiva como opiniones y actitudes expresados en un texto respecto a un servicio, producto, persona o tema (Mäntylä et al., 2018, p. 17). Del mismo, Athar (2011) lo definió como la tarea de identificar opiniones, sentimientos, emociones y actitudes positivas, neutrales o negativas haciendo uso de características como n-gramas, presencia de adjetivos, adverbios, palabras negativas, relaciones gramaticales, etc., todo ello a partir de frases, palabras, oraciones y documentos.

Enfoques para la clasificación del análisis de sentimiento

El análisis de sentimiento es una herramienta que se usa para clasificar las opiniones como positivas, negativas o neutrales, que combina técnicas de minería de texto y PLN. Existen tres enfoques: (a) basado en léxico, (b) aprendizaje automático e (c) híbrido que combina los dos anteriores.

Enfoque basado en lexicones. Es un enfoque semántico donde se define un documento *lexicón* que contiene una serie de palabras polarizadas o clasificadas en una determinada categoría de emoción (Strapparava & Mihalcea, 2008). Para realizar el análisis, se definen determinadas reglas, por ejemplo, para el análisis de sentimiento, lo más básico es contar las palabras positivas, negativas o neutrales según el lexicón y; la polaridad que tenga mayor cantidad de palabras es la polaridad resultante del texto. Por otro lado, Segura

Navarrete et al. (2021) sostuvieron que la mayoría de los lexicones disponibles fueron creados para el idioma inglés; sin embargo, estos también se pueden extender para el español, lo que permite que sea sean utilizados para realizar análisis de subjetividad en texto en el idioma indicado. Finalmente, según Priyavrat y Sharma (2018), la puntuación de polaridad de las palabras positivas y negativas en una oración o documento se suman para obtener la puntuación de sentimiento total de esa oración o documento.

Enfoque basado en aprendizaje automático. El *aprendizaje automático* se define como el campo de estudio que le da a las computadoras la habilidad de aprender sin haber sido explícitamente programadas (Samuel, 1969). Para realizar el análisis de sentimiento o emociones, se utilizan distintos algoritmos de este campo que identifican tendencias o patrones en los datos y que permiten posteriormente estimar la polaridad o emoción que presenta un texto. Existen dos grandes categorías de aprendizaje automático que básicamente se diferencian en la forma de entrenar los algoritmos: (a) aprendizaje supervisado y (b) aprendizaje no supervisado.

Aprendizaje supervisado. Este se realiza a partir de un conjunto de datos; es decir, para lo que el valor del atributo objetivo o que interesa estimar (por ejemplo, polaridad) es conocido dentro de ese conjunto de datos de entrenamiento. A partir de ello, los algoritmos identifican patrones que permiten estimar el valor del atributo objetivo en los datos.

Aprendizaje no supervisado. A diferencia del aprendizaje supervisado, los datos de entrenamiento no tienen especificados los resultados; por lo tanto, el algoritmo va aprendiendo, ajustando los datos o maximizando una función objetivo. Dado que el clasificador de opiniones se entrena en los datos etiquetados de un dominio, frecuentemente no funciona con otro dominio. Para superar esta brecha, se recomiendan enfoques basados en lexicon (Bonta et al., 2019).

Enfoque híbrido. Este enfoque aprovecha la fuerza o las propiedades combinadas de los recursos léxicos con las técnicas de aprendizaje automático (Yusof et al., 2015). Una

ventaja del enfoque híbrido es que los resultados son en su mayoría precisos. La segunda ventaja es que es más útil en un entorno de estilo cruzado donde el entrenamiento se realiza en un conjunto de datos y la prueba se realiza en diferentes conjuntos de datos de prueba (Mudinas et al., 2012; Mumtaz & Ahuja, 2018). La tercera ventaja del enfoque híbrido es que también puede detectar y medir sentimientos a nivel de concepto (Mudinas et al., 2012). En contraste, una desventaja de usar un enfoque híbrido es que para una oración compleja o si las palabras irrelevantes y el ruido ocurren con frecuencia en el texto, no se llega a detectar ningún sentimiento y se asigna a esas oraciones una puntuación neutral (D'Andrea et al., 2015).

Métodos de decisión multicriterio (MCDM)

Los métodos de decisión multicriterio son enfoques apropiados para tomar decisiones en presencia de múltiples criterios que están generalmente en conflicto. Los MCDM se clasifican en dos categorías: (a) MCDM discreto o toma de decisiones multiatributo (MADM, por sus siglas en inglés) y (b) MCDM continuo o toma de decisiones con objetivos múltiples (MODM, por sus siglas en inglés) (Zavadskas et al., 2014). No obstante, en la presente investigación se utiliza el MCDM discreto puesto que se evalúa la toma de decisiones a partir de múltiples atributos de un servicio financiero.

Los MCDM discretos se subdividen en dos tipos: (a) nítidos, entre los que se encuentran: solución de compromiso y optimización de múltiples criterios (VIKOR, por sus siglas en bosnio), proceso analítico jerárquico (AHP, por sus siglas en inglés), técnica de ordenación de preferencias por similaridad a la solución ideal (TOPSIS, por sus siglas en inglés) y método de organización de clasificación de preferencia para evaluaciones de enriquecimiento (PROMETHEE, por sus siglas en inglés) II y (b) difusos, entre los que se encuentran a eliminación y elección traduciendo la realidad (ÉLECTRE, por sus siglas en francés) III y toma de decisiones de atributos múltiples difusos (FMADM, por sus siglas en inglés). Mientras que VIKOR y ÉLECTRE III se basan en el principio de máxima utilidad y

mínimo arrepentimiento individual, PROMETHEE II se basa solo en el principio de máxima utilidad grupal; por otro lado, VIKOR y TOPSIS son métodos basados en buscar la distancia más corta de la solución ideal positiva y más lejos de la solución ideal negativa (Ravi & Ravi, 2017). En esta investigación se aplica el método VIKOR, puesto que permite considerar en las publicaciones de los tuits la separación de los atributos insatisfactorios que afectan a la selección de un servicio financiero, sobre todo los que ofrecen los entes financieros, debido a que el cliente se enfrenta a la oferta de varios servicios similares en la misma categoría.

AHP. El proceso analítico jerárquico (AHP, por sus siglas en inglés) propuesto por Saaty (1980), presenta las siguientes características: (a) descompone un problema no estructurado y complejo, (b) construye un orden jerárquico, (c) valora cuantitativamente la importancia de las variables, (d) sintetiza los juicios o valoraciones y (e) determina qué variables tienen mayor prioridad.

Se utiliza AHP para determinar el peso de cada criterio o aspecto, lo cual se realiza mediante dos fases: (a) se elabora una matriz de comparación por pares para los aspectos o criterios seleccionados para que luego; estos aspectos puedan ser colocados en filas y columnas; y (b) todos los aspectos son comparados con el fin de obtener una importancia relativa en una escala del 1 al 9 (ver Tabla 1). La diagonal de la matriz representa la comparación entre los aspectos o criterios similares.

Tabla 1

Intensidad de Importancia

Intensidad de importancia	Definición
1	Igual importancia
3	Importancia moderada
5	Fuertemente importante
7	Muy importante
9	Extremadamente importante
2, 4, 6 y 8	Valores intermedios
1/2, 1/3, 1/4, ... 1/9	Si la intensidad de la importancia del criterio A sobre el criterio B es 5, entonces el recíproco (1/5) es la importancia del criterio B sobre el criterio A.

TOPSIS. La técnica de ordenación de preferencias por similitud a la solución ideal (TOPSIS, por sus siglas en inglés), fue propuesto por Hwang y Yoon (1981) y se basa en buscar la solución que es menos distante de la solución ideal y más distante de la solución negativa. En el proceso de TOPSIS una alternativa es considerada un vector y a cada una le pertenece un punto en el espacio n-dimensional, donde la mejor alternativa se enfoca en la elección que tiene la menor distancia a la solución ideal. Las ventajas que se identifican de este método son: (a) está basado en la distancia y necesita un cálculo simple y (b) no necesita ningún parámetro excepto pesos relativos de diferentes criterios. Por otro lado, algunas desventajas son: (a) no considera la importancia relativa de las distancias de las soluciones y (b) la normalización de vectores puede ser problemática debido a las diferentes unidades de criterio.

VIKOR. El método solución de compromiso y optimización de múltiples criterios (VIKOR, por sus siglas en bosnio) fue propuesto por Opricovic (1998), el cual considera la cercanía a la solución ideal para clasificar el método alternativo, donde el valor más cercano a cero corresponde a la mejor alternativa. VIKOR a comparación de otros MCDM, es útil puesto que se pueden establecer pesos de utilidad máximo del grupo y de manera diferente a las opiniones de arrepentimiento individual según diversas situaciones. Ello permite que la medición de satisfacción de cliente pueda ser usada por las personas que toman decisiones sobre un servicio completo.

Las ventajas que se identifican de este método son: (a) basado en la distancia, que sigue el principio de máxima utilidad grupal y mínimo arrepentimiento individual; (b) puede proporcionar clasificación, así como una solución de compromiso con una tasa de ventaja y (c) proporciona análisis de estabilidad de peso y análisis de compensaciones entre criterios. Por otro lado, algunas desventajas son: (a) la función agregada (es decir, las medidas Q) puede ser errónea debido a la comparación de cantidades potencialmente incomparables (criterios no conmensurables) y (b) no puede considerar una función de preferencia no lineal.

PROMETHEE II. El método de organización de clasificación de preferencia para evaluaciones de enriquecimiento (PROMETHEE, por sus siglas en inglés) II, propuesto por Brans et al. (1986), es utilizado para clasificar acciones o alternativas en función de grados de preferencia. Las ventajas que se identifican de este método son: (a) pertenece a la categoría de método de superación, que proporciona una rica función de preferencia como lineal y gaussiana que necesita solo un punto de inflexión para ser especificado por un tomador de decisiones, (b) la matriz de preferencias se puede utilizar para determinar “acciones locales incomparables” pero que pueden ser “indiferentes” en la clasificación final (Ishizaka & Nemery, 2013) y (c) puede representar el poder de discriminación de un criterio utilizando la representación del plano de análisis geométrico para ayuda interactiva (GAIA, por sus siglas en inglés). Por otro lado, algunas desventajas son: (a) funciona solo en el principio de máxima utilidad; (b) el tomador de decisiones no puede expresar una preferencia en una escala ordinal y (c) un ciclo puede ocurrir en comparaciones por pares de acciones, lo que provoca un fenómeno de inversión de rango.

ELECTRE III. El método de eliminación y elección traduciendo la realidad (ÉLECTRE, por sus siglas en francés) III fue propuesto por (Roy, 1978), y requiere dos parámetros adicionales que son: peso (w) y el umbral de veto (v) para poder clasificar un conjunto de alternativas. Las ventajas que se identifican de este método son: (a) pertenece a la categoría de método de clasificación superior y funciona con la máxima utilidad y el principio de arrepentimiento mínimo individual como VIKOR; y (b) no es propenso al proceso de normalización, por lo que evita la distorsión de los datos originales. En lo que se refiere a desventajas, se identifican las siguientes: (a) el tomador de decisiones necesita especificar distintos parámetros; (b) cuando ocurre un empate entre las alternativas, los parámetros α y β deben ajustarse para romper el empate; (c) el cálculo del grado superior es arbitrario y (d) puede sufrir un problema de inversión de rango.

FMADM. La toma de decisiones de atributos múltiples difusos (FMADM, por sus siglas en inglés) es un método que sigue el enfoque general MADM, el cual se ha desarrollado para resolver problemas en la evaluación de la importancia relativa de los atributos. MADM especifica cómo se hace la elección entre las alternativas a partir de la información de los atributos.

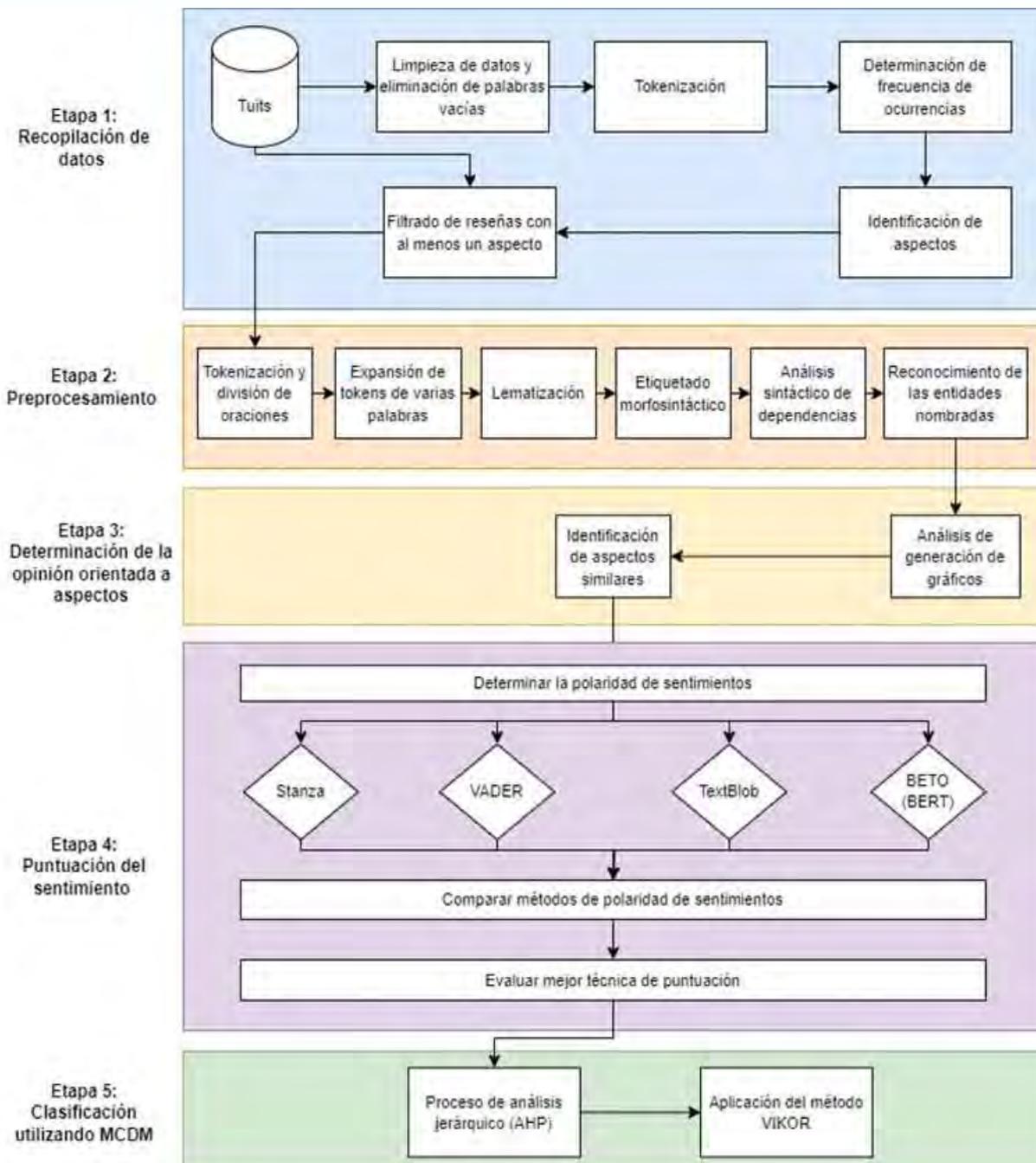
Modelo de clasificación mediante MCDM y análisis de sentimiento

Para la presente investigación se ha tomado como referencia el modelo de clasificación mediante MCDM y análisis de sentimiento. En ese sentido, Vyas et al. (2019) propusieron un modelo que se compone de cinco etapas: (a) recopilación de datos, (b) preprocesamiento y determinación de criterios, (c) determinación de la opinión orientada a aspectos, (d) puntuación del sentimiento y (e) clasificación utilizando MCDM. A partir del modelo referido, se realizaron adaptaciones de acuerdo con el contexto del estudio para asegurar el cumplimiento de los objetivos de la investigación, donde en la Figura 1 se muestra el flujo de cada una de las etapas del enfoque propuesto para esta.

Recopilación de datos. En esta etapa se recopilan los tuits desde la plataforma Twitter de los bancos principales en Perú que son: Banco A, Banco B, Banco C y Banco D. Luego de la recopilación se realiza la *tokenización* que consiste en el análisis de cada tuit eliminando los signos “.”, “?”, “!”, los sustantivos y los pronombres con el fin de reducir datos y realizar el preprocesamiento, ello teniendo cuidado de no alterar la subjetividad de cada tuit, para luego preparar un listado de unigramas y bigramas y obtener la frecuencia de ocurrencia en el conjunto de datos. Algunos ejemplos de unigramas y bigramas son “tasa de interés”, “cargo por servicio”, “cuota”, “préstamo”, etc. Seguidamente, se consideran los aspectos o características comunes en cada tuit, para luego convertir todos los bigramas en unigramas insertando el signo “-” entre palabras (Vyas et al., 2019).

Figura 1

Modelo de Clasificación Mediante MCDM y Análisis de Sentimiento



Nota. Adaptado de “Aspect-Based Approach to Measure Performance of Financial Services Using Voice of customer [Enfoque Basado en Aspectos para Medir el Desempeño de los Servicios Financieros Usando la Voz del Cliente]”, por V. Vyas, V. Uma, y K. Ravi, 2019, *Journal of King Saud University - Computer and Information Sciences*, 34(5), p. 2264.

<https://doi.org/10.1016/j.jksuci.2019.12.009>

Preprocesamiento. En esta fase se aplican técnicas como tokenización y eliminación de palabras vacías con el fin de reducir el ruido del conjunto de datos, para luego hacer el proceso de lematización que consiste en convertir cada palabra en su forma raíz. Por ejemplo, “aplicado” y “aplicando” son convertidos en “aplicar” (Vyas et al., 2019).

Determinación de la opinión orientada a aspectos. Esta etapa se subdivide en tres actividades: (a) análisis de generación de gráficos, y ocurre cuando dos palabras guardan relación y están conectadas por un borde, lo que se puede expresar gráficamente, donde cada palabra se puede representar con un nodo y su relación con un borde; (b) identificación de palabras vecinas, en este paso se crea una lista de palabras vecinas por cada aspecto a evaluar y (c) extracción de opinión para un aspecto o criterio y; para ello, se extraen los verbos, adverbios y adjetivos así como también palabras de opinión.

Puntuación del sentimiento. En la presente investigación se realiza la clasificación de sentimiento a través de cuatro herramientas basadas en el enfoque léxico: Stanza, razonador de sentimiento del diccionario consciente del valor (VADER, por sus siglas en inglés), TextBlob, y BETO (BERT). En la investigación de Mohan et al. (2022) se menciona a tres clasificaciones de sentimiento: TextBlob, VADER y Stanza.

TextBlob. Clasifica los sentimientos por puntuación de subjetividad y polaridad. La polaridad positiva indica un sentimiento positivo, la polaridad negativa indica un sentimiento negativo y una polaridad de cero indica un sentimiento neutral (Mohan et al., 2022).

VADER. Es una biblioteca de código abierto totalmente basada en reglas/léxico que utiliza un enfoque de “bolsa de palabras”. Devuelve una sola puntuación unidimensional con un rango de -1 a +1. Los valores positivos significan que el sentimiento de un texto determinado es positivo y los valores negativos significan que el sentimiento es negativo. Los valores de sentimiento neutral están entre -0.05 y +0.05 (Mohan et al., 2022).

Stanza. Es una biblioteca de programación neurolingüística creada por Stanford Core NLP Group para analizar texto en más de 70 idiomas. Utiliza el modelo clasificador de red

neuronal convolucional (CNN, por sus siglas en inglés) para clasificar el sentimiento de un conjunto de datos determinado, teniendo como resultado una puntuación de 0 a 2, donde 0 representa la clase negativa, 1 representa la clase neutral y 2 representa la clase positiva (Mohan et al., 2022).

Pérez et al. (2021) en su estudio “pysentimiento: Un Kit de Herramientas de Python para Análisis de Sentimiento y Tareas de SocialNLP”, abordó el uso del *pysentimiento*, el cual es un conjunto de herramientas de Python multilingüe para el análisis del sentimiento, donde se proporcionan modelos basados en transformaciones de vanguardia soportado en inglés y en español. En este estudio se menciona la utilización de la biblioteca Transformers, la cual brinda el modelo de lenguaje preentrenado como el de la representación de codificador bidireccional de transformadores (BERT, por sus siglas en inglés) para textos en inglés y el modelo BETO que es una versión de BERT en español, donde se realizaron diferentes experimentos con distintos modelos, teniendo como por ejemplo modelos en español como mBERT y DistilBERT, logrando mejores resultados con el modelo BETO.

Clasificación utilizando MCDM

Para esta parte, se aplica el MCDM que permite seleccionar, comparar y clasificar diferentes alternativas en este estudio, obteniendo la priorización mayor valorada para el servicio que tenga el mejor desempeño de los bancos evaluados. Por otro lado, se utiliza AHP para determinar el peso de cada criterio evaluado, representado en una matriz para comparar entre sí y colocar una importancia relativa de escala del 1 al 9, donde a mayor peso mejor rango o priorización. Además, también se aplica el enfoque VIKOR puesto que sirve para resolver problemas de decisión con criterios conflictivos siendo el más utilizado en diversos campos de aplicación.

Definición de Términos del Estudio

Los términos que más se utilizaron en el presente estudio fueron los siguientes:

- Procesamiento del lenguaje natural (PLN), “es una disciplina de la inteligencia artificial que se ocupa del tratamiento automático del lenguaje natural hablado o escrito” (Bansal et al., 2008, como se citó en Montoyo et al., 2012, p. 675).
- Análisis de sentimiento: “habilitar computadoras para reconocer y expresar emociones” (Picard, 1997, como se citó en Pang & Lee, 2008, p. 10).
- Métodos multicriterio de toma de decisiones (MCDM): “son métodos que consideran características y criterios sobre diferentes alternativas con el objetivo de seleccionar la mejor alternativa” (Hwang & Yoon, 1981, p. 62).
- VIKOR: “es un método solución de compromiso y optimización de múltiples criterios” (Opricovic, 1998, p. 18).
- Proceso analítico jerárquico (AHP): “permite determinar la prioridad de un conjunto de variables” (Saaty, 1980, p. 161).

Limitaciones

Para el desarrollo de la investigación se presentaron limitaciones para cumplir con los objetivos propuestos, las cuales son:

- La investigación se limitó al uso de referencias internacionales, debido a que en el Perú no existen precedentes de investigaciones aplicando análisis de sentimiento complementado con enfoque MCDM.
- La coyuntura internacional que se viene presentando debido a la guerra entre Rusia y Ucrania influye en el entorno económico del país. Según BBVA Research (2022), la aceleración de la inflación en Perú persistiría si se prolonga la guerra entre estos países, debido a que las cotizaciones internacionales de combustibles y algunos alimentos se están elevando como resultado del conflicto bélico, lo que provocaría que el descenso de la inflación sea más lento.
- En cuanto a la coyuntura nacional, la limitación consiste en la complejidad política que atraviesa el país; puesto que, pese a la buena gestión macroeconómica, no se tiene

un Estado fuerte y diligente y; los avances de crecimiento económico son limitados, debido a que las medidas diseñadas para contener los estragos generados por la pandemia de la COVID-19 no han sido satisfactorias (Banco Mundial, 2021).

- Por otro lado, se prevé que, la economía peruana crecería ligeramente por debajo del ritmo del 3% previo a la pandemia, debido al aumento del volumen de exportaciones. Por otro lado, la demanda interna se desaceleraría en un contexto de escasa confianza entre los consumidores y las empresas, generando menor crecimiento en la cantidad de socios comerciales y a la vez cierta volatilidad de precios en cuanto a energía. Además, se estima que la pobreza se mantendría por encima de los niveles anteriores a la pandemia en los próximos dos años, como consecuencia de la caída de la calidad media del empleo (Banco Mundial Perú, s.f.).

Delimitaciones

La investigación se aplica solo en las principales entidades bancarias del país que, por motivos de protección de la identidad de los bancos, se denotan como Banco A, Banco B, Banco C y Banco D.

Resumen

En el Capítulo I se desarrollaron los antecedentes donde se detallan las investigaciones previas acerca del uso del análisis de sentimiento en redes sociales, que en conjunto con los métodos de decisiones multicriterio permiten la medición de desempeño de productos y/o servicios en las organizaciones, los cuales es una necesidad en la actualidad debido a que los usuarios buscan recomendaciones en las plataformas digitales al momento de adquirir un producto o servicio.

Por otro lado, se detalló el marco teórico de la investigación, desarrollando cada uno de los conceptos utilizados en la presente investigación como: satisfacción del cliente, redes sociales, procesamiento de lenguaje natural, análisis de sentimiento y métodos de decisión

multicriterio. También se precisó el problema de investigación en donde se evidenció la necesidad de contar con una herramienta que permita a los clientes de las principales entidades financieras peruanas elegir los servicios de estas basándose en las evaluaciones de otros usuarios que hayan contratado el servicio previamente, realizando una evaluación de acuerdo con diferentes características o aspectos.



Capítulo II: Revisión de la Literatura

En el presente capítulo se desarrolla la revisión de la literatura, evaluando las diferentes teorías y estudios existentes que permitan profundizar acerca del problema planteado en la investigación. En ese sentido, se abordan los conceptos implicados en el estudio, como son: análisis de sentimiento, procesamiento de lenguaje natural, análisis de decisiones con múltiples criterios y redes sociales; profundizando en la definición de cada una de ellas y analizando referencias en estudios previos a nivel mundial, los cuales permitirán enriquecer el marco teórico de nuestra investigación y a la vez permitirán establecer un análisis GAP, que buscaremos cubrir con nuestra investigación.

Análisis de Sentimiento

Diversos resultados de estudios sobre el análisis de sentimiento de los datos de Twitter se adaptaron a las organizaciones de servicios en su búsqueda por examinar opiniones positivas, negativas y neutras sobre la satisfacción del producto y servicio (Algur & Patil, 2017; Botchway et al., 2019; Kouloumpis et al., 2011). Por otro lado, para que las empresas obtengan información relevante para la gestión de relaciones con sus clientes, la minería de texto se ha vuelto imperativa para estas, especialmente en el mercadeo digital del siglo XXI.

Un estudio realizado en Columbia University llamado “Análisis de Sentimiento de los Datos de Twitter” (Agarwal et al., 2011), tuvo como objetivo indagar en Twitter para tener una idea del sentimiento hacia un producto mediante la construcción de modelos que permitan determinar si un tuit expresa un sentimiento positivo, negativo o neutral. El análisis se basó en el desarrollo y validación de tres modelos: (a) basado en características, (b) basado en el árbol de Kernel y (c) el de unigrama. Para entrenar y validar el modelo se usaron un total de 10,000 tuits y, de los modelos desarrollados se encontró que el que tenía mejor desempeño era el de unigrama, que supera hasta en un 20% a los otros dos restantes. Además, el análisis también se realizó combinando modelos como, por ejemplo, modelo de unigrama con el basado en características y este último con el basado en el árbol de Kernel. Estas dos

combinaciones superan al modelo de unigrama hasta en un 4%. Como recursos se usaron un diccionario de emoticones y otro de acrónimos. Este último es una base de datos de léxico en inglés que permite agrupar sustantivos, verbos y adjetivos en conjuntos de sinónimos cognitivos los cuales finalmente permite determinar si el tuit expresa un sentimiento positivo, negativo o neutral.

Procesamiento de Lenguaje Natural (PLN)

Se entiende como *procesamiento de lenguaje natural* al área de la inteligencia artificial (IA) que estudia la comunicación entre las personas y los computadores utilizando lenguas naturales mostrándose en un texto y pasando por cuatro tipos de análisis: (a) morfológico, (b) sintáctico, (c) semántico y (d) pragmático, teniendo como resultado un texto procesado. El PLN utiliza diversas técnicas y enfoques basados en reglas y algoritmos para interpretar el lenguaje humano (Talamé et al., 2019).

En cuanto al uso del PLN para la puntuación en el análisis de sentimiento, en un estudio de Mohan et al. (2022), se utilizaron las opiniones de artículos, publicaciones de blog y noticias de virólogos, científicos y expertos en salud, relacionados con la tercera ola de la COVID-19 en la India. En esa línea, sus opiniones hacia la potencial tercera ola se registraron utilizando bibliotecas de PLN basado en un lexicón como TextBlob, VADER y Stanza para el análisis de sentimiento. Los resultados del PLN revelaron sentimientos negativos en la mayoría de los artículos y blogs con algunas excepciones.

Bonta et al. (2019) determinaron que una de las principales ventajas del enfoque PLN basado en lexicón es su independencia de dominio, y también que se puede ampliar y mejorar fácilmente. Compararon las herramientas de análisis de sentimiento como VADER (Hutto & Gilbert, 2014), TextBlob y el kit de herramientas de lenguaje natural (NLTK, por sus siglas en inglés), y la finalidad era clasificar críticas de películas descargadas de la página web <http://www.rottentomatoes.com> mediante las tres librerías mencionadas anteriormente. Los resultados experimentales confirmaron que VADER supera tanto a TextBlob y NLTK como

herramienta. Concluyeron en este estudio que, si el sentimiento es la única finalidad que se planea hacer para un corpus procedente de redes sociales o microblogueo y si necesita ser procesado rápidamente, entonces VADER es una mejor opción con un umbral de +/-0.05.

Zárate Calderón (2021) en su tesis de grado “Análisis de Sentimiento en Información de Medios Periodísticos y Redes Sociales Mediante Redes Neuronales Recurrentes”, enfatiza la importancia del análisis de sentimiento para el uso de diferentes aplicaciones computacionales debido al crecimiento de las redes sociales. Tener este tipo de información debidamente procesada puede ser de gran utilidad para las organizaciones. Respecto a la extracción de la data evaluada, esta provino de dos tipos de fuente, pero para efectos de la presente investigación se aborda solo la fuente de la red social Twitter, donde se aplicaron diferentes técnicas para determinar la polaridad de los sentimientos, teniendo como resultado de mayor relevancia el valor-F1 con 86.55% sobre un conjunto de tuits en inglés con la librería Transformers del modelo BERT obteniendo un buen rendimiento. Sin embargo, el modelo se tuvo que adaptar a una arquitectura basada en un conjunto de datos en español, por lo que se cambió de modelo (de BERT a BETO), teniendo las mismas características, pero con textos en español.

Zárate Calderón (2021), para evaluar su modelo con BETO, aplicó diversos ajustes en la etapa de preprocesamiento y en la optimización de este, obteniendo mejores resultados de entrenamiento, como el valor-F1 con 76.92%. Después, evaluando el modelo completo según la precisión de los aciertos de la polaridad de los positivos, negativos y neutrales, se obtuvo un valor de precisión del 76.42%.

Por otro lado, Rosenbrock et al. (2021) en una la investigación presenta la aplicación, búsqueda de hiperparámetros, comparación y análisis de resultados de distintas técnicas de aprendizaje automático utilizadas para el PLN como los métodos Word2vec+LSTM (donde Word2vec es una técnica de incrustación de palabras desarrollada en 2013 por Mikolov, que

utiliza como representación de palabras un vector multidimensional y; LSTM es un tipo de redes neuronales recurrentes que tienen la capacidad de recordar a muy largo plazo) y BERT (BETO). Rosenbrock et al. (2021) tomaron un conjunto de más de 50,000 comentarios en lenguaje español coloquial sobre películas (extraídos de la página web <https://www.cinesargentinos.com.ar/>). Los resultados de otras investigaciones utilizando el modelo BERT estuvieron entre un 85% y un 94% de exactitud (los textos se encontraban en inglés). Pero utilizando el modelo BETO (BERT) que ya cuenta con una incrustación preentrenada con palabras en español, el valor F1 obtenido fue de 88%, clasificando de forma incorrecta solo 1,815 del total de 10,462 comentarios.

Por último, este estudio es un referente para la presente investigación puesto que permite utilizar el modelo BETO al tener un rendimiento favorable en los resultados con textos en español y con métricas de evaluación aceptables.

Preprocesamiento de Texto

Las tareas básicas de PLN implican la simbolización y el análisis sintáctico, lematización/derivación, etiquetado morfológico, detección del lenguaje e identificación de relaciones semánticas. En general, las tareas de PLN dividen el lenguaje en trozos o piezas elementales más cortas o tókenes, con la idea de comprender las relaciones entre dichos tókenes y explorar cómo funcionan estas piezas juntas para crear significado (Yue et al., 2020). Dichas tareas implícitas se emplean frecuentemente en recursos PLN como: extracción, limpieza, tokenización, lematización y etiquetado morfosintáctico.

Extracción

Es el proceso de obtener información de sitios específicos, consolidándola en una colección de documentos. Este proceso se ejecuta normalmente con ayuda de herramientas o sistemas mostrando la información más fácil de analizar (Quispe Aguilar, 2015). El proceso de extracción también se puede dar de forma automática en datos estructurados de fuentes basadas en texto.

Limpieza

Se refiere a la limpieza básica donde se remueven caracteres extraños, espacios, signos de puntuación, direcciones web u otros caracteres extras en los documentos obtenidos en el proceso de extracción (Agudelo Restrepo, 2022).

Tokenización

Según Hasan Basari et al. (2013), la *tokenización* es el proceso de separar las cadenas de texto más largas en trozos más pequeños o tókenes. Los trozos de texto más grandes pueden dividirse en frases, las frases a su vez pueden dividirse en palabras, etc. A continuación, se detalla un ejemplo de cómo realizar tokenización y segmentación de oraciones a partir de una pieza de texto en idioma español, donde la sentencia evaluada es: “Esta es una sentencia de prueba para Tokenizar. Esta es otra sentencia.”

A partir de la sentencia indicada, en las Tablas 2 y 3 se muestra que el texto está segmentado en dos oraciones, donde cada una ellas contienen algunos tókenes.

Tabla 2

Ejemplo de Salida para el Proceso de Tokenización (Sentencia 1)

No. de token	Token
1	Esta
2	es
3	una
4	sentencia
5	de
6	prueba
7	para
8	tokenizar
9	.

Tabla 3

Ejemplo de Salida para el Proceso de Tokenización (Sentencia 2)

No. de token	Token
1	Esta
2	es
3	otra
4	sentencia
5	.

Lematización

Taboada et al. (2011) definieron la *lematización* como un proceso simple de eliminar las terminaciones de las palabras que no están en el diccionario; en otras palabras, es el proceso de reducir las diferentes formas de una palabra a una sola forma. A continuación, se presenta un ejemplo para observar la salida del proceso de lematización (ver Tabla 4), donde la sentencia evaluada es: “Juan construye una casa en Colombia”. En dicho ejemplo se observa que las palabras *construye*, *construyendo* o *construido* se transforman a su palabra base *construir*.

Tabla 4

Ejemplo de Salida para el Proceso de Lematización

Token	Lema
Juan	Juan
construye	construir
una	una
casa	casa
en	en
Colombia	Colombia

Etiquetado morfosintáctico

El *etiquetado morfosintáctico* es el proceso de asignar una parte del discurso o un marcador de clase léxica a cada palabra en una colección (Kumawat & Jain, 2015). Las clases léxicas son: (a) clase abierta, que son los adjetivos, adverbios, sustantivos, verbos, pronombres e interjecciones; (b) clase cerrada, conformada por preposiciones, los auxiliares y determinantes; y (c) otros, que son signos de puntuación, *emojis* y símbolos. Para realizar el etiquetado morfosintáctico, se elige un conjunto estándar de etiquetas de trabajo. Se podría elegir conjuntos de etiquetas muy gruesos como las propuestas por Nivre et al. (2016) que se muestran en la Tabla 5.

Tabla 5*Conjunto de Etiquetas de Dependencias Universales*

Tag	Descripción	Ejemplo
Clase abierta		
ADJ	Adjetivo: modifica al sustantivo describiendo sus propiedades.	Rojo, joven, impresionante, etc.
ADV	Adverbio: modifica al verbo en tiempo y lugar.	Muy, despacio, casa, ayer, etc.
NOUN	Sustantivo: palabra que describe personas animales, cosas, etc.	Algoritmo, gato, mango, bello, etc.
VERB	Verbo: palabra que describe acciones o procesos.	Dibujar, proporcionar, ir, etc.
PROPN	Nombre propio: de una persona organización, lugar, etc.	Regina, IBM, Colorado, etc.
INTJ	Interjección: exclamación, saludo, respuestas sí/no, etc.	Oh!, ummm!, Sí, Hola!, etc.
Clase cerrada		
ADP	Preposición/Posposición: marca la relación espacial, temporal, u otra relación con el sustantivo.	En, sobre, debajo, etc.
AUX	Auxiliar: verbo auxiliar que marca tiempo, aspecto, modo, etc.	Puede, debe, son, etc.
CCONJ	Conjunción coordinante: une dos frases o cláusulas.	Y, o, pero, etc.
DET	Determinante: marca las propiedades del sintagma nominal.	Este, esto, etc.
NUM	Numeral	Uno, dos, primero, segundo, etc.
PART	Participio: preposición que se usa junto a un verbo.	Arriba, abajo, en, apagado, sobre, etc.
PRON	Pronombre: abreviatura para referirse a una entidad o un evento.	Ella, él, ellos, etc.
SCONJ	Conjunción subordinada: une una oración principal con una oración subordinada como complemento oracional.	Tanto que, con que, etc.
Otros		
PUNCT	Puntuación	“,” “,” “()” “:” “.” “,” etc.
SYM	Símbolos como \$ o emojis	\$, %, etc.
X	Otro	Asdf, qwfg, etc.

Nota. Adaptado de “Universal Dependencies v1: A Multilingual Treebank Collection”

[Dependencias Universales v1: Una Colección de Treebank Multilingüe], por J. Nivre, M.-C. de Marneffe, F. Ginter, Y. Goldberg, J. Hajič, C. D. Manning, R. McDonald, S. Petrov, S. Pyysalo, N. Silveira, R. Tsarfaty, y D. Zeman, 2016. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (p. 1661)
(<https://aclanthology.org/L16-1262>)

Análisis de Decisiones con Múltiples Criterios (MCDM)

En cuanto a los métodos de decisiones multicriterio, Kang y Park (2014) en su artículo “Medición Basada en Revisión de la Satisfacción del Cliente en el Servicio Móvil: Análisis de Sentimiento y Enfoque VIKOR”, presentaron un marco de estudio con respecto a los servicios móviles, que consistió en extraer las opiniones de los clientes. Luego de aplicar análisis de sentimiento encontraron que no son suficientes los resultados que obtuvieron como medición para la satisfacción del cliente, por lo que fue necesario aplicar el método de clasificación VIKOR y determinaron que los atributos insatisfactorios afectan notablemente la selección de todo el servicio móvil, debido a que este tiene varios servicios similares en la misma categoría. Este estudio propone analizar la satisfacción del cliente con los servicios móviles, a través de la consideración simultánea que tienen los clientes con preferencias por múltiples y diversos atributos del servicio, en lugar de valoraciones generales de todo un servicio. El proceso de investigación se hizo en dos etapas. La primera consistió en la recolección y procesamiento de datos y la segunda en la medición de la satisfacción del cliente. En la primera etapa se utiliza la minería de textos para compilar los diccionarios de atributos y palabras de sentimiento. En la segunda etapa se utiliza análisis de sentimiento para colocar las puntuaciones de los atributos por cada servicio y; adicionalmente se utilizó el método de clasificación VIKOR. Este método de clasificación optimiza el proceso de respuestas múltiples y el índice de clasificación se deriva considerando tanto la máxima utilidad grupal como también el mínimo arrepentimiento individual del oponente; es decir, clasifica y selecciona de un conjunto de alternativas en presencia de criterios en conflicto basados en la cercanía a la solución ideal. Como resultado de esta investigación, se confirma que el enfoque MCDM es adecuado para respaldar la opinión del cliente con análisis de sentimiento porque en el servicio, múltiples atributos afectan la satisfacción del cliente simultáneamente.

Otro estudio relacionado con MCDM fue el de K. Ravi y Ravi (2017), cuyo título es “Clasificación de Productos de Marca Utilizando Análisis de Sentimiento Orientado a Aspectos y Toma de Decisiones de Criterios Múltiples Combinados”, donde se propuso un novedoso procedimiento para análisis de sentimiento orientado a aspectos, combinado con un conjunto de métodos de toma de decisiones de criterios múltiples como: TOPSIS, VIKOR, PROMETHEE II, ÉLECTRE III, y FMADM. El objetivo de estudio se basó en establecer un enfoque para la determinación de opinión orientado a aspectos, además de proponer un enfoque híbrido de cálculo de puntuación de sentimientos a nivel de aspecto, y de desarrollo de dominios para representar aspectos destacados. Por otro lado, se realizó una combinación de los temas anteriormente descritos con MCDM para obtener un estudio comparativo entre un conjunto de teléfonos móviles y cámaras digitales, empleando los comentarios de los clientes provenientes de Twitter y reseñas de las clientes dejadas por correo electrónico.

Ahora bien, la amplia disponibilidad de productos y servicios que se ofertan por internet genera un problema en el consumidor a la hora de elegir el artículo más adecuado de acuerdo con su necesidad, puesto que los consumidores se guían de comentarios que encuentran en redes sociales sobre un producto o servicio en particular y de acuerdo con sus características. El principal resultado de este estudio fue que el análisis de sentimiento con un enfoque a nivel de aspectos permite realizar la comparación de los comentarios sobre algún producto o servicio a un buen nivel de granularidad, combinado con métodos de toma de decisiones de criterios múltiples, realizando una comparación de diferentes productos en términos de aspectos y permitiendo tener un resultado que oriente al consumidor entre los aspectos importantes de un producto y/o servicio, además de un conjunto de otras alternativas (K. Ravi & Ravi, 2017).

Finalmente, se identificó un estudio relacionado a la medición de servicios financieros que tuvo el mismo propósito de la presente investigación fue el de Vyas et al. (2019), cuyo

título fue “Enfoque Basado en Aspectos para Medir el Desempeño de los Servicios Financieros Usando la Voz del Cliente”. En dicho estudio analizaron el problema que tienen las personas al momento de seleccionar una entidad bancaria. En ese sentido, dado que estas instituciones brindan toda una cartera de servicios con distintos criterios, el propósito del estudio fue trabajar en un modelo que permita medir el desempeño de los servicios financieros usando la metodología por decisión multicriterio (MCDM) para la clasificación de los bancos de acuerdo con la priorización obtenida. Este modelo trabaja en conjunto con el análisis de sentimiento; por ello, en esta investigación se plantearon las siguientes preguntas: ¿Cómo se mide el desempeño de los servicios financieros? y ¿El método multicriterio ayudará en la toma de decisiones de las personas para calificar a un banco? Para responder estas preguntas, se aplicó la metodología experimental analizando como muestra las reseñas de los bancos HDFC Bank, ICICI Bank, SBI y Axis Bank que se encuentran en la página web <https://www.bankbazaar.com/>, teniendo 510, 500, 520 y 500 reseñas, respectivamente, para luego mediante algoritmos de tokenización, lematización y de n-gramas, además de las herramientas del análisis de sentimiento, obtener una matriz donde se relaciona la puntuación de los sentimientos por cada criterio identificado y analizado por banco. Luego, con esta información cada banco es evaluado con los enfoques FMADM y VIKOR. La fusión de estos métodos y su comparación generan resultados determinando la priorización, criterio y pesos por orden de jerarquía, obteniendo a los bancos con mayor puntuación: SBI > ICICI Bank > HDFC Bank > Axis Bank. Los instrumentos que utilizaron en la metodología para el proceso de obtener la puntuación de los sentimientos fueron diferentes diccionarios como el cálculo de orientación semántica (SO-CAL, por sus siglas en inglés), SenticNet y SentiWordNet, además del analizador de los gráficos Stanford. Los resultados de esta investigación con el modelo propuesto permitirían a las personas tener información para comparar, clasificar y elegir la entidad financiera según sus necesidades.

Redes Sociales

Las redes sociales representan un medio de comunicación digital entre personas que comparten contenidos, expresando sus sentimientos hacia algún tema de interés, así como también comparten experiencias sobre los productos y/o servicios que brindan las organizaciones públicas o privadas. La evolución de las plataformas digitales y de la conectividad ha hecho que la sociedad cambie dentro de sus hábitos cotidianos la forma de comunicarse entre las personas. Parte de la vida humana ha sido trasladada a las plataformas digitales como lo son las redes sociales, relacionándose en comunidad participativa en la parte social, política, cultural, profesional, entretenimiento, entre otros (Van Dijck, 2016, pp. 11-12).

En los últimos años se ha observado un incremento en el uso de las redes sociales en el ámbito mundial, donde las personas usan estas plataformas para emitir opiniones, comentarios y referencias acerca de la calidad de un servicio o producto. Galeano (2022) realizó un análisis del informe digital del año 2022 elaborado por We Are Social y Hootsuite, en el cual se indica el porcentaje de preferencia por las redes sociales de parte de los usuarios entre las edades de 16 a 64 años. En esa línea, el uso de redes sociales en el ámbito mundial tiene la siguiente distribución: Facebook (14.5%), Twitter (3.3%), TikTok (4.3%), WhatsApp (15.7%) e Instagram (14.8%). En el Perú para el año 2020, según Ipsos Perú (2020), se estimó que 13.2 millones de usuarios se encontraban activos en las redes sociales, representando el 78% de la población urbana entre 18 y 70 años, siendo las más utilizadas la red social Facebook (94%), WhatsApp (86%), YouTube (62%), Instagram (60%) y Twitter (29%).

Facebook

La red social Facebook fue fundada por Mark Zuckerberg en el año 2004, donde se le ofrece al usuario alojar contenidos como videos, fotos, imágenes, texto, etc., permitiendo compartir y comentar sobre las publicaciones con una red de contacto. Las organizaciones no son ajenas a tener esta interacción con los usuarios puesto que permite ofrecer sus productos

y/o servicios por medio de esta red social (Gurevich, 2016, pp. 217-238). Según We Are Social y Hootsuite (como se citó en Galeano, 2022), en el año 2022 Facebook registró 2,910 millones de usuarios conectados y se ubicó en el puesto 1 en la clasificación mundial para esta categoría.

Twitter

Twitter es una red social y un blog a la vez, puesto que permite que su contenido sea compartido y visible por todos los usuarios, organizaciones, comunidades, etc., comentando de forma abierta distintas opiniones sobre alguna publicación de algún tema de interés (Carballar, 2011, p. 4). Esta red social es una forma de microblogueo, lo que significa que las actualizaciones de Twitter (tuits) pueden contener solo 280 caracteres. Los tuits permiten a las empresas crear comunidades, ofrecer retroalimentación o asistencia inmediata y promocionar sus productos/servicios (Baruah, 2012, p. 1). Según We Are Social y Hootsuite (como se citó en Galeano, 2022), en el 2022 Twitter registra un crecimiento del 23%, alcanzando los 436 millones de usuarios conectados en el ámbito mundial.

LinkedIn

La misión de LinkedIn es conectar a los profesionales del mundo para que sean más productivos y exitosos. Al unirse a LinkedIn, los usuarios obtienen acceso a personas, trabajos, noticias, actualizaciones y conocimientos que mejoraría el conocimiento y el rendimiento de los usuarios; mientras que para las organizaciones es de utilidad captar el talento por medio de esta plataforma. LinkedIn es una red social donde principalmente los usuarios conectados son los profesionales y las organizaciones, los cuales aprovechan este canal para explotar su marca personal (Costa-Sánchez & Corbacho-Valencia, 2015).

Instagram

Instagram es una red social para compartir fotos y videos, donde se permite agregar diversos filtros permitiendo la creatividad de los usuarios y ofreciendo oportunidades en mercadeo y publicidad gratuita para los profesionales, negocios y marcas (Ramos, 2013) Según We Are Social y Hootsuite (como se citó en Galeano, 2022), es una de las redes

sociales más utilizadas dentro de las plataformas digitales teniendo como dueño a Mark Zuckerberg y; además, en el año 2022 esta red contaba con 1,478 millones de usuarios conectados en el ámbito mundial.

TikTok

TikTok es una de las redes sociales con mayor tendencia de crecimiento. Para julio del 2018 los usuarios la utilizaban por un promedio de 52 minutos al día, y con ella se pueden crear videos cortos que fácilmente se pueden viralizar. Asimismo, esta red permite que las marcas conecten a los usuarios con las organizaciones (Ramos, 2020). Según We Are Social y Hootsuite (como se citó en Galeano, 2022), en el año 2022 TikTok contaba con 1,000 millones de usuarios conectados en el ámbito mundial.

Matriz de Confusión

Según Barrios (2019), la herramienta *matriz de confusión* es muy utilizada dentro del campo de la inteligencia artificial permitiendo visualizar el desempeño o rendimiento de un algoritmo de aprendizaje supervisado aplicando a modelos por clases o de clasificación. Cada columna de la matriz representa la cantidad de predicciones de cada clase; mientras que cada fila representa la cantidad de la clase real, identificando los falsos positivos y los falsos negativos en una matriz binaria. Más aún, esta herramienta permite medir el desempeño de los modelos mediante métricas a obtener, las cuales son: exactitud, precisión, exhaustividad o sensibilidad, especificidad y valor-F1.

Exactitud

Es la aproximación del resultado de una medición al valor verdadero; es decir, define la exactitud en la cantidad de predicciones positivas que fueron correctas (ver Fórmula 1).

$$Exactitud = \frac{(VP + VN)}{(VP + FP + FN + VN)} \quad (1)$$

Donde: VP = Verdadero positivo; VN = Verdadero negativo; FP = Falso positivo; y FN = Falso negativo.

Precisión

Se refiere a la dispersión del conjunto de valores obtenidos a partir de mediciones repetidas de una magnitud. Cuanto menor es la dispersión mayor es la precisión y se muestra como el porcentaje de casos positivos detectados (ver Fórmula 2).

$$\text{Precisión} = \frac{VP}{(VP + FP)} \quad (2)$$

Donde: VP = Verdadero positivo y FP = Falso positivo.

Exhaustividad o sensibilidad

Esta métrica tiene la capacidad de discriminar los casos positivos de los negativos. La exhaustividad o sensibilidad representa la fracción de verdaderos positivos (ver Fórmula 3).

$$\text{Exhaustividad} = \frac{VP}{(VP + FN)} \quad (3)$$

Donde: VP = Verdadero positivo y FN = Falso negativo.

Especificidad

Esta métrica representa la fracción de los verdaderos negativos y trata de los casos negativos que el algoritmo ha clasificado correctamente (ver Fórmula 4).

$$\text{Especificidad} = \frac{VN}{(VN + FP)} \quad (4)$$

Donde VN = Verdadero negativo y FP = Falso positivo.

Valor F1

Esta métrica resume la precisión y sensibilidad en una sola y es muy utilizada cuando la distribución de las clases es desigual (ver Fórmula 5).

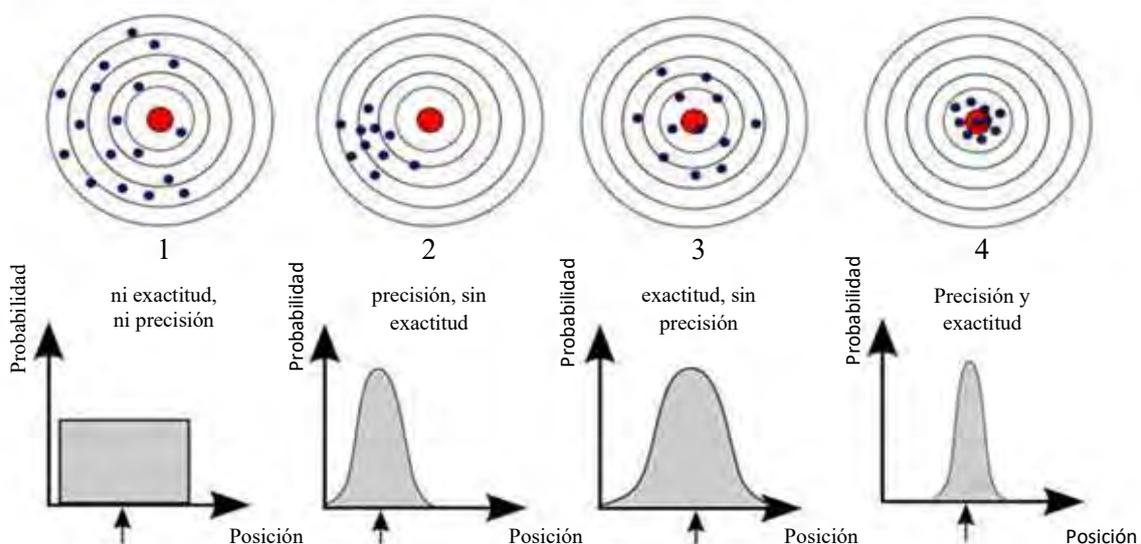
$$\text{Valor F1} = 2 \times \frac{\text{Precisión} \times \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}} \quad (5)$$

Alemán Viteri (2021) en su investigación “Análisis de Sentimientos para Twitter con Vader y TextBlob” profundizó en el análisis de sentimiento que se encuentra en las redes

sociales dado que es un tema relevante para el público en general. En este estudio se explica el proceso del análisis de sentimiento utilizando dos librerías (VADER y TextBlob), donde se enfatiza la importancia de medir el rendimiento de los algoritmos comparando con el análisis de la validación manual para determinar el mejor resultado. Para ello, se aplica el método de matriz de confusión, el cual tiene como métricas la precisión, exactitud, sensibilidad y especificidad. En este estudio en particular se indica que las métricas más importantes en relación con los datos son exactitud y precisión, donde la métrica de exactitud se refiere a qué tan cerca el valor calculado se aproxima a la proporción de verdaderos ya sean en negativos o positivos, mientras que la métrica de precisión hace referencia a la dispersión de los valores; es decir, si un conjunto de datos posee exactitud y precisión se tiene un conjunto de datos concentrado en torno al objetivo (ver Figura 2). Para este estudio se concluye que el mejor rendimiento en sus resultados es el del modelo TextBlob presentando en general sus métricas con mayor exactitud, precisión y sensibilidad.

Figura 2

Métricas de la Matriz de Confusión (Precisión y Exactitud)



Nota. Tomado de “La Matriz de Confusión y Sus Métricas”, por J. I. Barrios Arce, 2019, Health Big Data (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>). En el dominio público.

El estudio de Alemán Viteri (2021) ofrece a la presente investigación un soporte para la utilización del método de matriz de confusión, el cual tiene la finalidad de validar el rendimiento y desempeño de los modelos BETO, Stanza, VADER y TextBlob utilizando las métricas precisión, exactitud, sensibilidad y el valor F1, los cuales tanto los modelos como las métricas se usan en la investigación en los Capítulos IV y V,

Proceso del Método VIKOR

Los pasos que debe seguir el algoritmo VIKOR para determinar una clasificación son los siguientes (Opricovic & Tzeng, 2007; Tzeng & Huang, 2011):

- Se calculan los mejores f_i^* y peores f_i^- de cada uno de los criterios a evaluar. Si la función i representa un beneficio, los valores se calculan mediante $f_i^* = \max_i f_{ij}$ y $f_i^- = \min_i f_{ij}$ y; para el caso donde la función i representa un costo, los valores se calculan mediante $f_i^* = \min_i f_{ij}$ y $f_i^- = \max_i f_{ij}$
- Posteriormente, se calculan los valores de la utilidad (S_j), arrepentimiento (R_j) y Q_j de cada una de las variables mediante la aplicación de las siguientes formulas:

$$S_j = \sum_{i=1}^n w_i \frac{f_i^* - f_{ij}}{f_i^* - f_i^-} \quad (6)$$

$$R_j = \max_i \left[w_i \frac{f_i^* - f_{ij}}{f_i^* - f_i^-} \right] \quad (7)$$

$$Q_j = v \frac{S_j - S^*}{S^- - S^*} + (1 - v) \frac{R_j - R^*}{R^- - R^*} \quad (8)$$

Donde: $S^* = \min_j S_j$; $S^- = \max_j S_j$; $R^* = \min_j R_j$; $R^- = \max_j R_j$; y v es ingresado como un peso de la estrategia de máxima utilidad de grupo, mientras que $(1 - v)$ es el peso de la oposición individual.

- Luego se ordenan las alternativas según los valores S, R, y Q en orden decreciente.

- Finalmente, para determinar la solución compromiso a' , que es la solución con mejor clasificación, se selecciona aquella con el menor valor Q , siempre y cuando satisfagan dos condiciones: (a) ventaja aceptable, que es determinada mediante la fórmula $Q(a'') - Q(a') \geq DQ$, donde a'' es la alternativa en la segunda posición y $DQ = \frac{1}{J-1}$, donde J es el número de alternativas y (b) estabilidad aceptable en el proceso de decisión, la cual valida si la alternativa a' también es la mejor clasificada según el listado de valores S y/o R . Finalmente, si estas dos condiciones no se cumplen, se considera una solución cercana mas no una solución ideal.

Resumen

En el presente capítulo se revisó la literatura de los conceptos necesarios para entender y fundamentar el problema planteado en la investigación. En ese sentido, se pudo presentar investigaciones acerca del análisis de sentimiento, donde estas describen la necesidad de contar con herramientas que permitan a las organizaciones determinar factores para medir el interés de los clientes en cuanto a sus productos o servicios; clasificando opiniones como positivas negativas o neutrales. A partir de análisis de dichas investigaciones, se pudo evidenciar que luego de determinar la polaridad de las opiniones extraídas, no se establecen procedimientos adecuados para corroborar si la aplicación de los diferentes métodos es precisa para determinar la polaridad de un sentimiento en negativo, positivo o neutral; de tal modo que se pueda garantizar un buen análisis de los datos.

Por otro lado, también se pudo determinar que la aplicación de técnicas de PLN necesarias para el análisis de sentimiento no cubre por completo la problemática planteada; por lo cual, es importante incluir otro tipo de técnicas como MCDM, para poder evaluar los productos o servicios de acuerdo con múltiples atributos que afectan la satisfacción del cliente.

Finalmente, también se encontró que las personas buscan evaluaciones dentro de las distintas redes sociales para poder decidir que producto contratar; donde dicho proceso no se

soporta en ninguna herramienta que pueda simplificar esta actividad para los clientes de los diferentes sectores económicos. Por ende, también se concluye que es importante incluir una herramienta que permita hacer este proceso simplificado, principalmente en el sector financiero, que es uno de los sectores más importantes de la economía en el ámbito mundial.



Capítulo III: Metodología

En el presente capítulo se especifica el diseño de la metodología de investigación que se aplica en el Capítulo V. Se presenta también el detalle de la población y la muestra tomada para realizar el estudio. Por otro lado, se detalla el procedimiento utilizado para la recolección de los datos y los instrumentos a aplicar que permiten la medición de un modelo propuesto. Finalmente, se establece el procedimiento para realizar el análisis e interpretación de los datos y la evaluación de la validez y confiabilidad del modelo de medición de desempeño de servicios financieros.

Diseño de la Investigación

La investigación que se desarrolló fue de tipo *aplicada* dada la orientación de resolver problemas prácticos a partir de la teoría, que, para efectos de la presente tesis, se aplicaron los métodos del análisis de sentimiento y el modelo de decisión multicriterio. Además, la investigación sigue el método deductivo y según su enfoque es cuantitativa que, con la ayuda de metodologías y herramientas permite identificar y dimensionar variables a partir del análisis, tratamiento y procesamiento de los datos, permitiendo cuantificar por medio de puntuaciones la clasificación de la polaridad de los sentimientos ya sea en positivo, negativo o neutro. Esta información previa sirve para aplicar al modelo de decisión multicriterio que soporta la clasificación y medición de los servicios financieros con mejor desempeño. Se debe indicar que, la investigación según su nivel es una investigación descriptiva dado que se definen y miden variables con los datos recolectados en Twitter dentro de un contexto (Hernández Sampieri et al., 2014; Álvarez Risco, 2020).

Justificación del Diseño

En la última década, las personas expresan sus emociones y opiniones de las marcas que consumen por medio de las redes sociales, lo cual puede influir a los consumidores al momento de elegir un servicio o producto teniendo como referencia este medio de

comunicación. Los usuarios de las entidades bancarias no son ajenos a esta participación en las redes sociales; por ello, estas se consideran una fuente importante de información para las organizaciones y los clientes (Kharde & Sonawane, 2016). En Perú en el año 2020 se estimó que 13.2 millones de los usuarios se encuentran activos en las redes sociales, representando para la red social Twitter un 29% (Ipsos Perú, 2020). Para el modelo que se propone, se utiliza Twitter, red social que tiene una significativa cantidad de datos expresados en sentimientos en forma de tuits. Además, al tener las opiniones las características de: (a) no poseer estructura, (b) ser heterogéneas y (c) frecuentemente tener carga emocional; este escenario ha generado el interés de diversos investigadores de realizar estudios en la aplicación del análisis de sentimiento con los datos de esta red social (Kharde & Sonawane, 2016). Más aún, el modelo propuesto se apoya en la metodología de decisión multicriterio (MCDM), puesto que a los clientes se les permitiría evaluar los bancos de acuerdo con el servicio de mayor valoración (Vyas et al., 2019).

Población

La población de estudio para esta investigación son los tuits que se encuentran en la red social Twitter y que hacen mención a las cuentas oficiales de las principales entidades financieras del Perú que se denominan para la presente tesis como: Banco A, Banco B, Banco C y Banco D. En ese sentido, se seleccionaron 15,546 tuits registrados entre el periodo del 1 de enero al 1 de septiembre del año 2022. A partir de cada tuit, se extraen los campos fecha, usuario y tuit, cuyas características se describen en la Tabla 6.

Tabla 6

Características de los Tuits

Atributo	Descripción
Fecha	Fecha de creación del tuit
Usuario	Usuario que realizó la publicación del tuit
Tuit	Texto donde se encuentra la opinión de los usuarios

El periodo fue seleccionado de acuerdo con el restablecimiento de las actividades financieras, las cuales se fueron reanudando progresivamente a partir de enero del 2022 luego de un periodo de inactividad a causa de la pandemia de la COVID-19. Desde la fecha indicada se lograron restablecer los servicios de los bancos con normalidad, los cuales son: agentes, cajeros, oficinas, etc.

Muestra

Según Morales Vallejo (2012, pp. 187-203), para calcular el tamaño de la muestra, se debe hacer uso de la Formula 9:

$$n = \frac{Z^2 pqN}{NE^2 + Z^2 pq} \quad (9)$$

Donde: n es el tamaño de la muestra, Z es el nivel de confianza (95%) con 1.96 de número de sigmas, p es la variabilidad positiva (éxito) (0.5), q es la variabilidad negativa (fracaso), N es el tamaño de la población (15,546) y E es la precisión o error (5% o 0.05). Considerando dichos valores, se realiza el cálculo de la Fórmula 9 y se obtiene un resultado de 375; sin embargo, al tratarse de una investigación aplicada, se hace uso de una muestra por conveniencia y; para ello, se seleccionaron 5,276 tuits, que son los que refieren al menos un aspecto. Se debe indicar que esta muestra es significativamente más grande que el resultado calculado. Para la selección de aspectos, se realizó un análisis de frecuencias sobre todos los tuits seleccionados para los cuatro bancos mediante librerías de Phyton, de los cuales se obtuvieron las palabras con mayor frecuencia, relacionadas con los principales criterios del dominio bancario. Con los datos mencionados, se construye una lista de criterios principales para todos los bancos, donde los aspectos identificados con mayor frecuencia para el modelo propuesto son: tarjeta, cuenta, cajero, agencia y sistema.

Consentimiento Informado

Para la presente investigación no ha sido necesario contar con un consentimiento informado, puesto que la información con la que se ha trabajado es de acceso público y accedido sin ninguna restricción.

Procedimiento de Recolección de Datos

Para la presente investigación el método utilizado para la recopilación de las opiniones de los usuarios de los bancos se ejecuta por medio de la librería Snsrape de Python, la cual permite extraer los tuits de esta red social en un periodo de tiempo determinado. Con los datos recolectados, se procede a la realización del análisis e interpretación de los datos con la finalidad de validar las preguntas planteadas al inicio de la investigación.

Instrumentos de Medición

Con la finalidad de analizar el modelo que se propone en esta investigación para el proceso de polarización de los sentimientos, se utilizan diferentes modelos como Stanza, VADER, TextBlob y BETO; y posteriormente para la evaluación del desempeño de los servicios financieros se complementa con el enfoque MCDM con el método VIKOR.

Análisis e Interpretación de Datos

Para realizar el análisis e interpretación de los datos obtenidos para el modelo propuesto, se definió una serie de procesos la cual se compone de cinco etapas: (a) recopilación de datos, (b) preprocesamiento, (c) determinación de la opinión orientada a aspectos, (d) puntuación del sentimiento y (e) clasificación utilizando MCDM.

Recopilación de datos

El procedimiento para la recopilación de las opiniones de los usuarios en Twitter utiliza la librería Snsrape de Python, la cual permite extraer los tuits de esta red social de los principales bancos del Perú en un periodo del 01 de enero al 30 de septiembre del año 2022 con las cuentas de usuario de Twitter de cada entidad bancaria (ver Tabla 7).

Después de la recopilación de los datos, se aplican técnicas de limpieza para normalizar el texto eliminando caracteres extraños que no aportan valor. Con este paso previo, se aplican técnicas de tokenización para dividir todo el texto en una lista de palabras,

oraciones, caracteres, reseñas, entre otros. Con ello, se identifican los principales criterios del dominio bancario y de acuerdo con la concurrencia de ellos se construye una lista de criterios identificados para la investigación, conformada por *tarjeta, cuenta, cajero, sistema y agencia*.

Tabla 7

Cuenta Twitter por Banco

Banco	Cuenta Twitter
Banco A	@bancoA
Banco B	@bancoB
Banco C	@bancoC
Banco D	@bancoD

Preprocesamiento

En esta etapa se aplican técnicas de preprocesamiento de lenguaje natural como: (a) tokenización, para dividir los tuits en tókenes y oraciones; (b) lematización, para determinar la raíz de cada palabra; (c) etiquetado morfosintáctico, para añadir una etiqueta morfológica a cada palabra del tuit; y (d) análisis sintáctico de dependencias, para establecer el orden jerárquico de cada palabra del tuit.

Determinación de la opinión orientada a aspectos

Con la utilización de las técnicas y herramientas ya indicadas, se procesa el texto y se extraen todas las características o criterios comunes identificados en la etapa anterior para usarlos en el modelo. Esta etapa se subdivide en tres actividades: (a) análisis de generación de gráficos, donde se utilizan las palabras que guardan relación, expresándolas mediante gráficos; (b) identificación de palabras similares, donde se elabora una lista de palabras vecinas por cada criterio a evaluar; y (c) extracción de opinión para el criterio evaluado, donde se seleccionan los tuits en los que refieren a al menos un aspecto.

Puntuación del sentimiento

Para realizar el análisis de sentimiento se aplica el método bajo el enfoque basado en lexicón con los modelos Stanza, VADER, TextBlob y BETO.

Clasificación utilizando MCDM

Se aplica MCDM, lo cual permite seleccionar, comparar y clasificar diferentes alternativas en la presente investigación, obteniendo la priorización mayor valorada para el servicio que tenga el mejor desempeño de los bancos evaluados.

Validez y Confiabilidad

Para determinar la confiabilidad de los modelos Stanza, VADER, TextBlob y BETO que se utilizan en el análisis de sentimiento, se realiza una comparación entre la polaridad determinada por los modelos versus la polaridad determinada a juicio humano sobre 528 tuits que fueron seleccionados de forma aleatoria de la muestra usada en el estudio, correspondiendo al 10% de la misma, logrando determinar las métricas de exactitud, precisión, exhaustividad y valor F1, producto de la aplicación de la matriz de confusión, obteniendo como resultado una exactitud de 65%, 45%, 31% y 88% sobre los modelos Stanza, VADER, TextBlob y BETO, respectivamente, de tal manera que se seleccionó a BETO como el modelo a utilizar en la presente investigación.

Por otro lado, para determinar la confiabilidad del modelo VIKOR, validado mediante las dos condiciones de aceptabilidad propias del modelo que son *ventaja aceptable* ($0.55 - 0 \geq 0.33$), siendo 0.55 el resultado de la alternativa en la posición 2 y 0.33 el valor de DQ; y la condición de *estabilidad aceptable en el proceso de decisión*, donde S y R tienen un valor de 0.00 para la mejor alternativa.

Resumen

En el presente capítulo se definió la tesis como una investigación aplicada, con un enfoque cuantitativo y de alcance descriptivo debido a que esta se basa en un modelo teórico para resolver un problema planteado y se apoya en herramientas que permiten definir y medir variables de los datos recolectados, haciendo énfasis en la influencia de las redes sociales al momento de elegir un producto o servicios bancario. El tamaño de la población estudiada fue de 15,546 tuits con una muestra de 5,276, los cuales hacían referencia al menos uno de los

aspectos del estudio. Por otro lado, se definieron los modelos PNL como Stanza, VADER, TextBlob y BETO como instrumentos de medición para determinar la polaridad de sentimientos de los comentarios de Twitter y; también se definió el modelo VIKOR como instrumento de medición para evaluar el desempeño de los bancos. Además, para la recolección de datos se establece el uso de la librería de Python Snsrape, siguiendo las cinco fases establecidas en el marco teórico. Para evaluar la validez y confiabilidad, se establece el uso de juicio humano para evaluar la exactitud de los modelos PNL y las condiciones de aceptabilidad propias del modelo VIKOR.



Capítulo IV: Presentación y Análisis de Resultados

En este capítulo se revisan los resultados obtenidos luego de aplicar toda la metodología planteada en el Capítulo IV. Para realizar el análisis y la interpretación de los datos obtenidos para el modelo propuesto se definió una serie de procesos que se encuentran incluidos en cinco etapas.

Recopilación de Datos (Etapa 1)

Para el procedimiento de la recopilación de las opiniones de los usuarios en la red social Twitter, se utilizó la librería Snsrape de Python, la cual permite extraer los tuits de esta red social de los principales bancos del país, los cuales siguen la siguiente notación para sus nombres: Banco A, Banco B, Banco C y Banco D, para un periodo del 1 de enero al 1 de septiembre del año 2022, aplicando la cuenta de usuario oficial de cada entidad bancaria en Twitter (mencionada en la parte de la metodología), teniendo como resultado la extracción de un total de 15,546 tuits que representan la totalidad para los cuatro bancos. Dicha información se extrajo en archivos de tipo de valores separados por comas (CSV, por sus siglas en inglés), como se muestra en la Tabla 8.

Tabla 8

Datos en Archivos de Tipo CSV por cada Entidad Bancaria Extraídos del 1 de Enero al 1 de Septiembre de 2022

Banco	Tamaño (kB)
BancoA.csv	444
BancoB.csv	1,400
BancoC.csv	550
BancoD.csv	185

Después de la recopilación de los datos por cada banco, se aplican diversas técnicas que se mencionan a continuación.

Proceso: Limpieza de datos y eliminación de palabras vacías

En este proceso se aplican técnicas de limpieza para normalizar el texto, eliminando caracteres extraños y signos que no aportan valor por cada archivo de tipo CSV trabajado en

el proceso anterior. En la Tabla 9 se muestran los diferentes casos de caracteres que se depuraron.

Tabla 9

Lista de Caracteres que Generan Ruido

Ítem	Característica
URL	Ejemplos: https://t.co/JttS6oYtSm y https://t.co/tTlzVZSMxM
Signos de puntuación	“.”, “?”, “(”, “)”, “_”, “!”, “:”, “;”, “:”, “/”, “*”, “#”, “:” y “@”
Emojis	🙄, 😊, 🤖, 🧑, etc.
Palabras vacías	Palabras cuyos tamaños son menores que dos letras.

Como resultado del proceso de eliminación, en la Tabla 10 se muestran los tuits procesados.

Tabla 10

Tuits Procesados

Tuit	
Recopilado	Procesado (luego de su limpieza)
@bancoA que fue mano no hay naaaa :(@ClaroPeru justo entré a pagarte weee. #sad #bancoA https://t.co/h3rYQGct1U	@bancoA que fue mano hay naaaa @claroperu justo entre pagarte weee sad bancoA
@bancoA Su app funciona ??? No puedo ingresar a la app	@bancoA app funciona puedo ingresar app
@bancoA La App está caída y no puedo hacer transferencias	@bancoA app esta caida puedo hacer transferencias
@bancoA Buenas el app del bancoA (Perú) está fallando no me permite acceder. https://t.co/1eM0oEaHmi	@bancoA buenas app del bancoA esta fallando permite acceder
@bancoA Su aplicativo no está informando nada. Revisenlo urgente.	@bancoA aplicativo esta informando nada revisenlo urgente
@bancoA Por favor señores del bancoA la aplicación no funciona, últimamente ya es recurrente este tipo de fallas con su app hagan algo por favor	@bancoA por favor senores del bancoA aplicacion funciona ultimamente recurrente este tipo fallas con app hagan algo por favor
@bancoA Me respondieron que es mi responsabilidad sobre lo sucedido, detectaron que era un movimiento inusual y responden de esa manera, mal.	@bancoA respondieron que responsabilidad sobre sucedido detectaron que era movimiento inusual responden esa manera mal

Proceso: Tokenización

Este proceso permite dividir cada tuit en una lista de tókenes. Para el presente modelo se tiene como ejemplo la extracción del tuit del Banco D: “decepcionado de su servicio que desde ayer no funcionan las transferencias y ni por ser cuenta sueldo tengo beneficios ya que su app esta caída hace meses”.

Utilizando la librería Stanza de Python se ejecuta el proceso de tokenización teniendo como resultado el tuit tokenizado, como se muestra en la Tabla 11.

Tabla 11

Tuit Tokenizado para el Banco D

Clave	Texto
(1,)	decepcionado
(2,)	de
(3,)	su
(4,)	servicio
(5,)	que
(6,)	desde
(7,)	ayer
(8,)	no
(9,)	funcionan
(10,)	las
(11,)	transferencias
(12,)	y
(13,)	ni
(14,)	por
(15,)	ser
(16,)	cuenta
(17,)	sueldo
(18,)	tengo
(19,)	beneficios
(20,)	ya
(21,)	que
(22,)	su
(23,)	app
(24,)	esta
(25,)	caída
(26,)	hace
(27,)	meses

Proceso: Determinación de frecuencia de ocurrencias

Este proceso permite identificar palabras ocurrentes dentro de un total de palabras. Para efectos del modelo, se aplica la librería nltk.FreqDist de Python, la cual representa la distribución de los tókenes para obtener los términos más comunes. Los resultados de todos los documentos analizados para los cuatro bancos son los listados de palabras y sus respectivas frecuencias. En la Tabla 12 se muestran las 20 palabras con las mayores frecuencias para el conjunto de datos.

Tabla 12

Palabras Identificadas por Frecuencia

Token	Cantidad
para	2,668
http	2,645
banco	1,588
cuenta	1,251
pero	1,127
todo	1,050
tarjeta	1,037
hacer	1,015
como	980
dinero	924
problema	919
gracia	862
desd	836
cuando	825
puedo	817
tengo	805
transferencia	738
servicio	583
esta	572
cajero	552

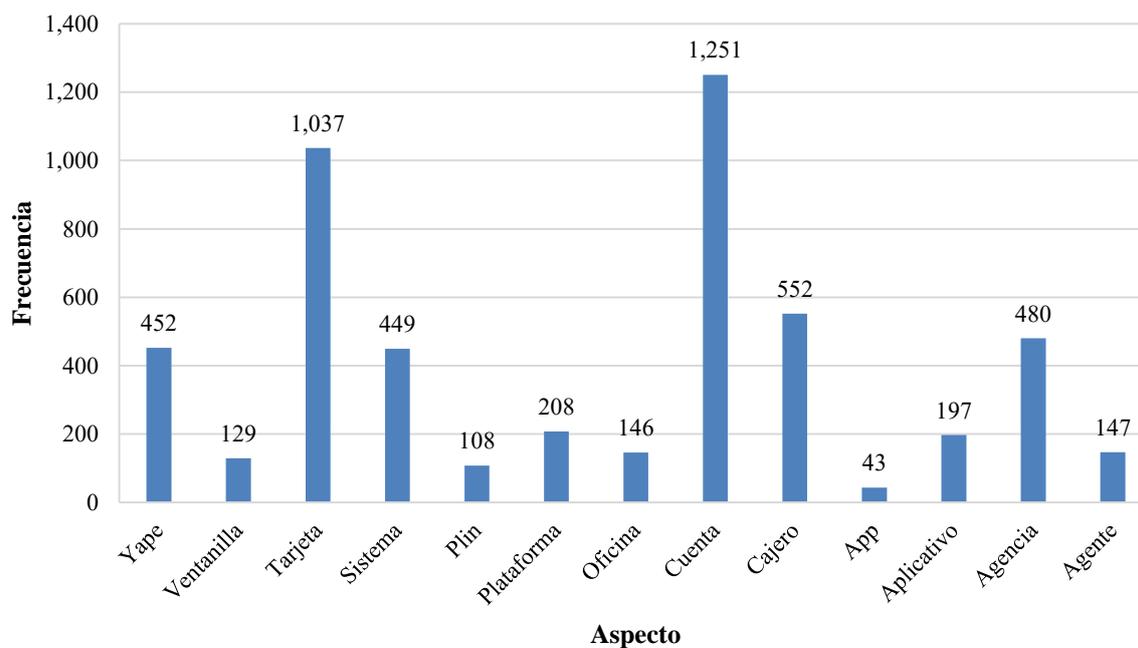
Proceso: Identificación de aspectos

Se procede a la identificación de los aspectos relacionados con los principales criterios del dominio bancario y según la concurrencia de ellos se construye una lista de

criterios principales para los bancos definidos: tarjeta, cuenta, cajero, agencia y sistema. En la Figura 3 se muestran los aspectos que poseen mayor frecuencia para el modelo propuesto.

Figura 3

Distribución de Aspectos Identificados con Mayor Frecuencia



Proceso: Filtrado de reseñas con al menos un aspecto

Se seleccionaron 5,276 tuits que son los que se refieren a al menos un aspecto que fueron determinados en el acápite anterior. Al finalizar esta etapa, se cumple con el objetivo específico 1: consolidar en una base de datos los comentarios registrados por los clientes en la plataforma Twitter acerca de los servicios financieros de los principales bancos en Perú.

Preprocesamiento y Determinación de Criterios (Etapa 2)

Desde esta etapa se incluyó un paquete de procesamiento de lenguaje natural (PLN) de Python, llamado BETO, que se utiliza para convertir una cadena que contiene texto en lenguaje humano en listas de oraciones y palabras. En ese sentido, se define un ejemplo de tuit original “@bancoD que pésima su atención de banca empresa”, el cual debe pasar por diversos procesos que se detallan a continuación.

Proceso: Tokenización y división de oraciones

Se dividen los tuits de entrada en tókenes y oraciones. En la Tabla 13 se muestra este accionar para el ejemplo indicado anteriormente y como consecuencia el tuit de ejemplo se divide o se segmenta en ocho tókenes.

Tabla 13

Tokenización

Clave	Texto
(1,)	@bancoD
(2,)	que
(3,)	pésima
(4,)	su
(5,)	atención
(6,)	de
(7,)	banca
(8,)	empresa

Proceso: Expansión de tókenes de varias palabras

En este proceso se puede expandir un token sin procesar en varias palabras sintácticas, lo que facilita la realización de análisis de dependencias universales en algunos idiomas. El resultado de este proceso genera que cada token del tuit de ejemplo se divida o se segmente en ocho palabras respectivamente (ver Tabla 14).

Tabla 14

Expansión de Tókenes de Varias Palabras

Token	Palabra
@bancoD	@bancoD
que	que
pésima	pésima
su	su
atención	atención
de	de
banca	banca
empresa	empresa

Proceso: Lematización

En este proceso se recupera la forma del lema para cada palabra de entrada. En la Figura 4 se muestra cómo la palabra *atención* se lematiza en *atender*.

Figura 4

Resultado de Lematización

El diagrama muestra la lematización de las palabras de un tuit. Las palabras originales están en una fila superior, cada una dentro de un recuadro con un borde negro. Debajo de ellas, se muestran sus formas de lema correspondientes. Las palabras originales son: @bancod, que, pésima, su, atender, de, banca, empresa. Las formas de lema son: @BancoD, que, pésima, su, atender, de, banca, empresa. Las palabras originales están conectadas a sus formas de lema por líneas curvas que se extienden desde el borde inferior de los recuadros superiores hacia las palabras de lema.

Proceso: Etiquetado morfosintáctico

Se etiqueta cada parte del tuit identificando cada palabra según sus características morfológicas. En la Figura 5 se muestra cómo se clasifican las palabras del tuit ejemplo según sus características morfológicas que se han definido en el Capítulo II.

Figura 5

Etiquetado Morfosintáctico

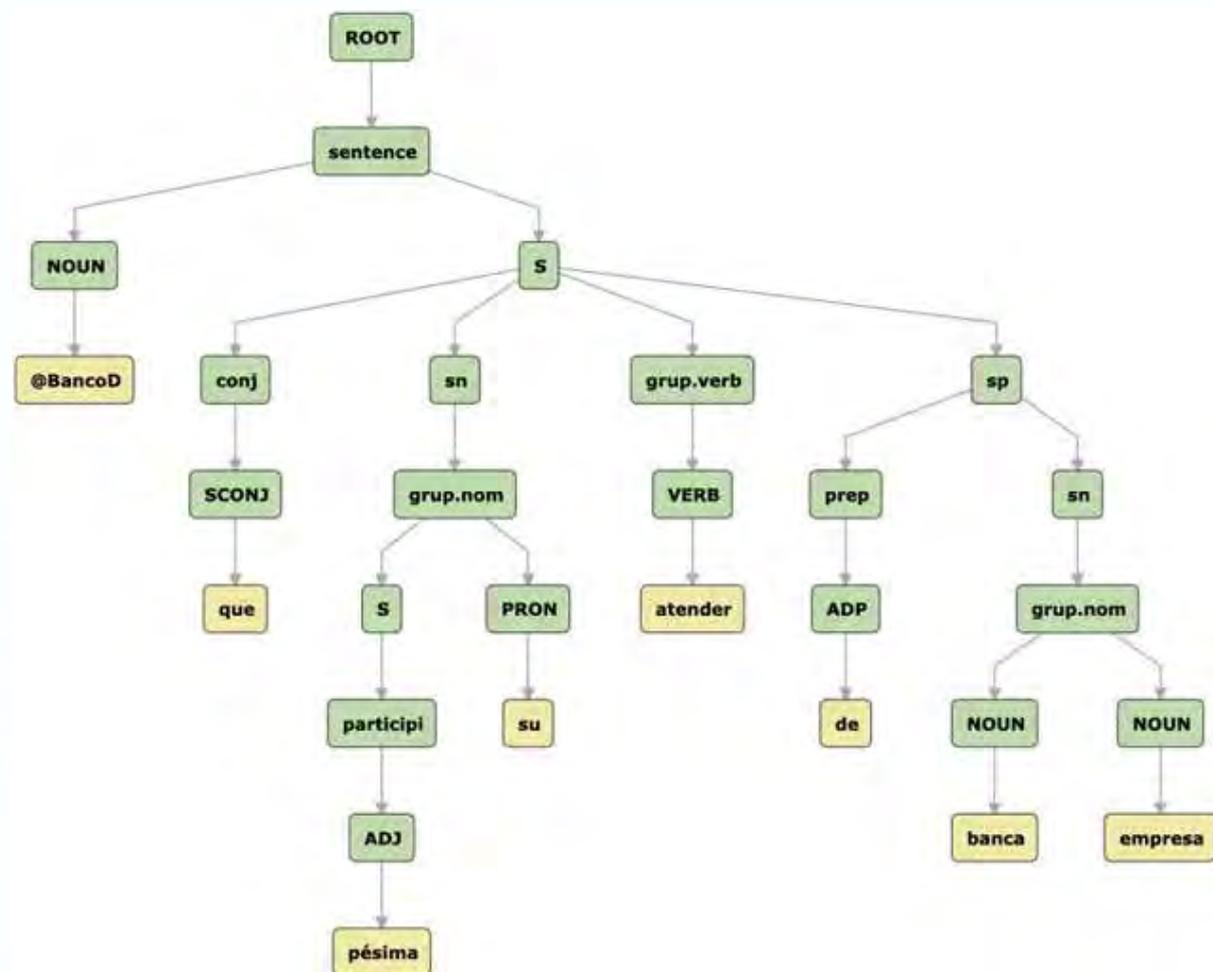
El diagrama muestra el etiquetado morfosintáctico de las palabras de un tuit. Las palabras originales están en una fila superior, cada una dentro de un recuadro con un borde negro. Debajo de ellas, se muestran sus etiquetas morfológicas correspondientes. Las palabras originales son: @BancoD, que, pésima, su, atender, de, banca, empresa. Las etiquetas morfológicas son: NOUN, SCONJ, ADJ, PRON, VERB, ADP, NOUN, NOUN. Las palabras originales están conectadas a sus etiquetas morfológicas por líneas curvas que se extienden desde el borde inferior de los recuadros superiores hacia las etiquetas morfológicas.

Proceso: Análisis sintáctico de dependencias

Se generan los gráficos para la extracción de conceptos de sentido común, como se muestra en la Figura 6. Es decir, se organizan los textos de manera jerárquica, estableciendo una estructura de árbol de palabras a partir de la oración de entrada; que representa las relaciones de dependencia sintáctica entre las palabras. Las representaciones de árbol resultantes, que siguen el formalismo de dependencias universales son útiles en distintas aplicaciones posteriores.

Figura 6

Análisis Sintáctico de Dependencias



Proceso: Reconocimiento de las entidades nombradas

En este proceso se reconocen lapsos de mención de un tipo de entidad en particular (por ejemplo, persona u organización) en la oración de entrada. Para el ejemplo anterior, el resultado sería: “BancoD que pésima su atender de banca empresa”.

Determinación de la Opinión Orientada a Aspectos (Etapa 3)

Esta etapa se compone de las siguientes actividades: (a) generación de gráficos de análisis e (b) identificación de aspectos similares.

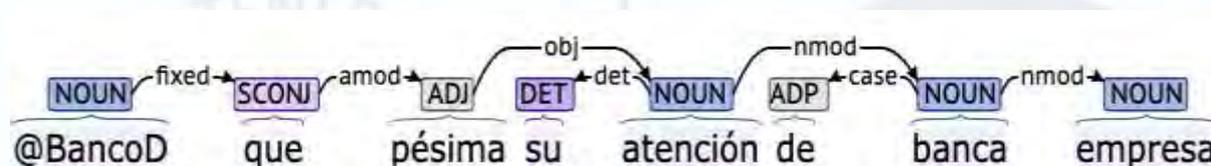
Generación de gráficos de análisis

En la etapa anterior correspondiente al proceso de análisis sintáctico de dependencias, se obtuvo el diagrama de árbol del ejemplo, donde se determinó el encabezado sintáctico de

cada palabra en la oración y la relación de dependencia entre las palabras, generando luego un diagrama que permite analizar la estructura gramatical de una oración, estableciendo relaciones entre las palabras principales y las que modifican esos encabezados, tal como se muestra en la Figura 7, donde por ejemplo a la palabra *que* se le asigna la etiqueta “amod” que describe la naturaleza exacta de la dependencia y significa que es un modificador adjetival y que afecta al adjetivo *pésima*.

Figura 7

Generación de Gráficos de Análisis



Proceso: Identificación de aspectos similares

A partir de la lista de aspectos identificados, se agrupan aquellos que tengan un contexto similar dentro de la industria financiera. En ese sentido, en Tabla 15 se muestra la agrupación de los aspectos, quedando finalmente cinco de ellos para el desarrollo de la investigación.

Tabla 15

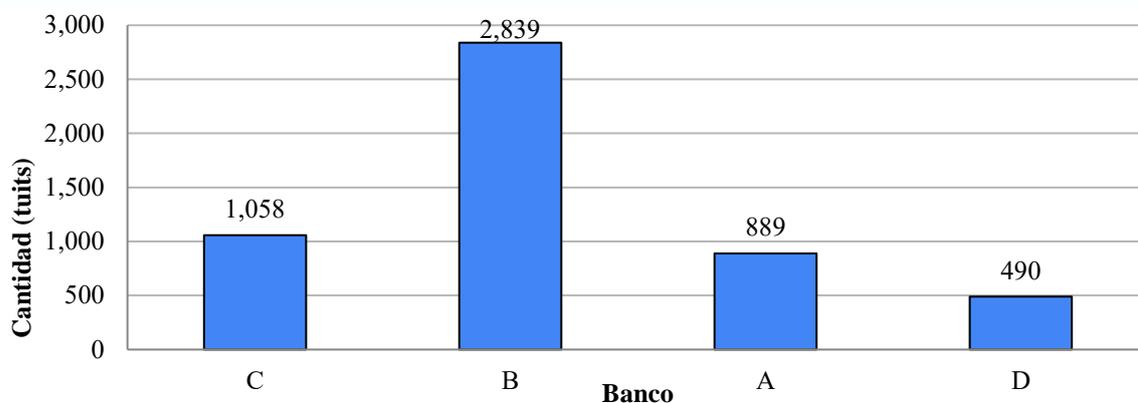
Aspectos Identificados con Palabras Similares

Aspecto	Palabras similares
Cuenta	
Tarjeta	
Cajero	
Agencia	Oficina, agente, plataforma y ventanilla
Sistema	Aplicativo, Yape, Plin y app

Después de identificar los aspectos y la relación con las palabras similares, se realiza el procedimiento de filtrado de las reseñas con los aspectos que se trabajan para el modelo por cada entidad bancaria, teniendo la cantidad de tuits por cada una de estas (ver Figura 8).

Figura 8

Reseña del Total de Tuits por cada Entidad Bancaria



Nota. Se representa el total de tuits por los aspectos identificados por cada entidad bancaria que se trabajan en el modelo.

Puntuación del Sentimiento (Etapa 4)

En esta etapa se realiza la identificación de la polaridad del tuit por cada aspecto según cada banco y; para ello, se toma el archivo con todos los datos preprocesados según la Etapa 2.

Proceso: Determinar la polaridad de sentimientos

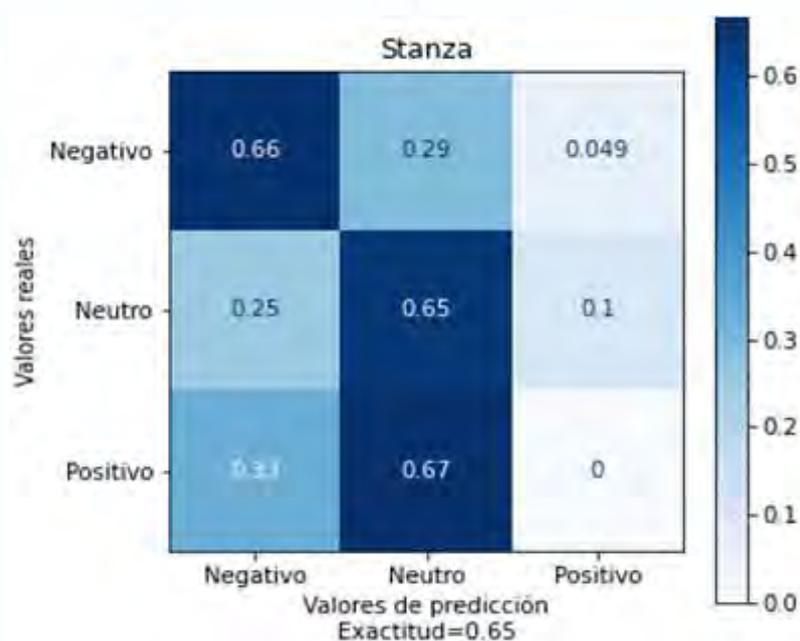
Para el proceso de obtener la puntuación de los sentimientos, se utilizan cuatro librerías de Python para PLN: (a) Stanza, basada en The Stanford CoreNLP, la que asigna a cada tuit una puntuación de sentimiento; donde para una polaridad positiva se asigna el valor 2, para una negativa el valor 0 y para una neutral el valor 1; (b) TextBlob, que clasifica los sentimientos por puntuación de subjetividad y polaridad, donde el valor 1 indica un sentimiento positivo, el valor -1 uno negativo y el valor 0 uno neutral; (c) VADER, que se basa en reglas/lexicón y que utiliza un enfoque de bolsa de palabras, asigna una sola puntuación unidimensional con un rango de -1 a +1, donde el valor positivo significa que el sentimiento de un texto determinado es positivo y el valor negativo significa que el sentimiento es negativo; mientras que, si el sentimiento es neutral, el valor se encuentra entre -0.05 y + 0.05; y (d) BETO, la cual es una versión en español del modelo original BERT, que ha sido preentrenada mediante aprendizaje no supervisado.

Proceso: Comparar métodos de polaridad de sentimiento

Para seleccionar el mejor método de polaridad de sentimiento, se puede realizar una comparación entre: Stanza, VADER, TextBlob y BETO; pero para realizar un análisis real, se tiene que hacer una evaluación manual de los mensajes extraídos de Twitter, para así saber si el sentimiento es correcto. Entonces, se selecciona una muestra aleatoria de 528 tuits, que corresponde al 10% de nuestra muestra. Para toda esta muestra, se identifica manualmente la opinión expresada implícitamente y se califica como positiva, negativa o neutral. Luego, ya se puede comparar con la polaridad de sentimiento calculado, utilizando las cuatro librerías descritas anteriormente (ver Figuras 9, 10, 11 y 12). Para obtener la precisión de un método de polaridad de sentimiento en comparación con los tuits analizados manualmente, se utiliza la matriz de confusión, donde en la línea vertical, se ubican los valores reales (manual) y; en la línea horizontal, los valores de predicción (según el modelo de polaridad). Los resultados de predicción de los métodos de polaridad para las cuatro librerías indicadas se muestran en las Tablas 16, 17, 18 y 19.

Figura 9

Matriz de Confusión Utilizando la Librería Stanza



Nota. La matriz de confusión y sus métricas realizada con librerías de Python.

Tabla 16

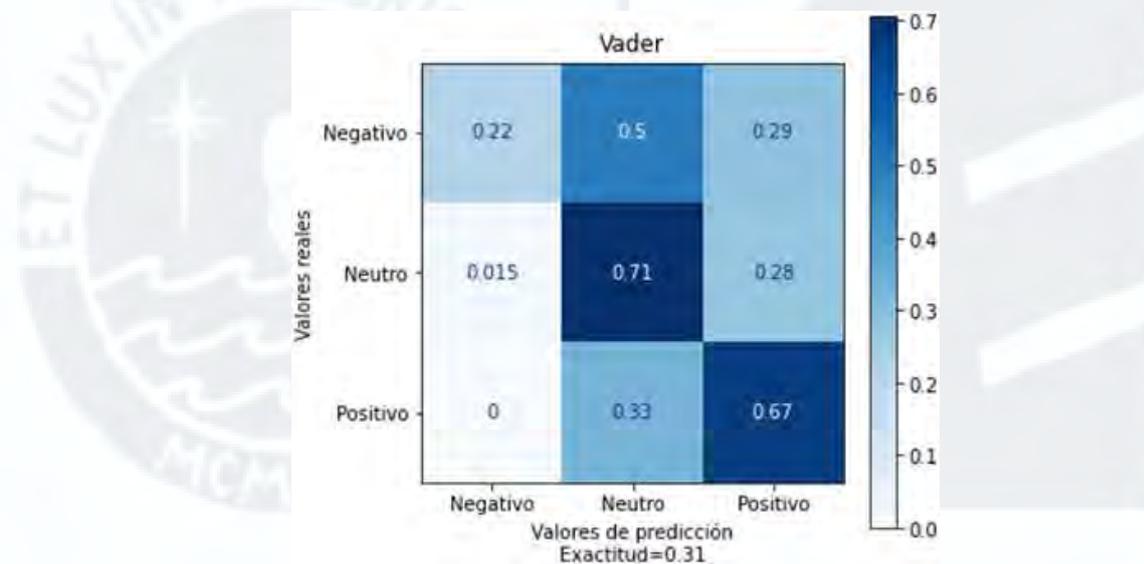
Resultados de Predicción del Método de Polaridad Stanza

Predicción	Métrica		Valor-F1
	Precisión	Exhaustividad	
Negativo	.91	.66	.77
Neutral	.34	.65	.44
Positivo	.00	.00	.00
Exactitud			.65

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Phyton.

Figura 10

Matriz de Confusión Utilizando la Librería VADER



Nota. La matriz de confusión y sus métricas realizada con librerías de Phyton.

Tabla 17

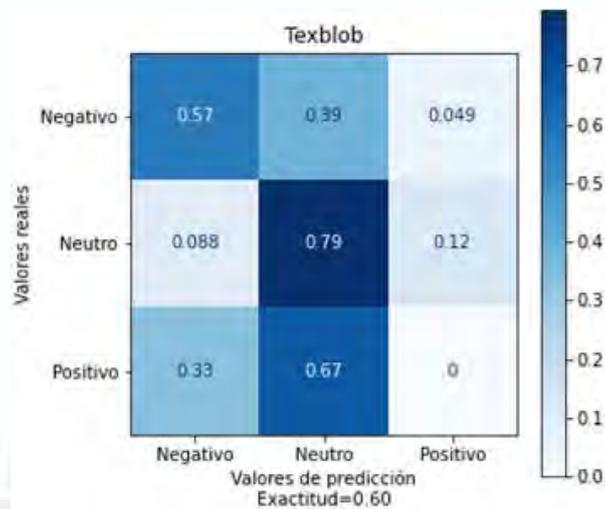
Resultados de Predicción del Método de Polaridad VADER

Predicción	Métrica		Valor-F1
	Precisión	Exhaustividad	
Negativo	.98	.22	.35
Neutral	.25	.71	.37
Positivo	.02	.67	.04
Exactitud			.31

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Phyton.

Figura 11

Matriz de Confusión Utilizando la Librería TextBlob



Nota. La matriz de confusión y sus métricas realizada con librerías de Phyton.

Tabla 18

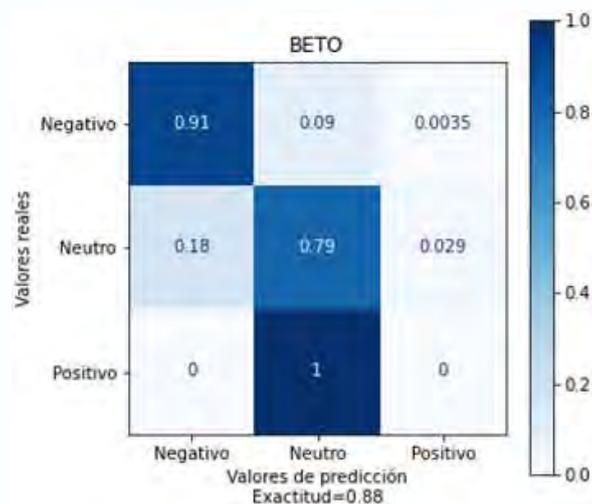
Resultados de Predicción del Método de Polaridad TextBlob

Predicción	Métrica		Valor-F1
	Precisión	Exhaustividad	
Negativo	.96	.57	.71
Neutral	.32	.79	.46
Positivo	.00	.00	.00
Exactitud			.60

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Phyton.

Figura 12

Matriz de Confusión Utilizando la Librería BETO



Nota. La matriz de confusión y sus métricas realizada con librerías de Phyton.

Tabla 19*Resultados de Predicción del Método de Polaridad BETO*

Predicción	Métrica		Valor-F1
	Precisión	Exhaustividad	
Negativo	.96	.91	.93
Neutral	.65	.79	.72
Positivo	.00	.00	.00
Exactitud			.88

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Phyton.

Finalmente, para la presente investigación, se elige el método de clasificación de polaridad de sentimiento BETO, puesto que comparando los cuatro resultados obtenidos con la cantidad de tuits calificados utilizando las cuatro librerías de Phyton, se observa que con BETO se tiene el más alto valor de exactitud (88%); por lo que resulta convincente la detección de la polaridad de sentimiento con respecto a la calificación manual. Además, se observa una detección de polaridad negativa de alto valor siendo el valor-F1 igual a 93%.

Proceso: Generar tabla de puntuación

Para este proceso, se obtiene un archivo con la puntuación de polaridad e información del aspecto como etiqueta por cada tuit que se obtuvo con el método de clasificación de polaridad de sentimiento BETO. Después, se procede a construir una matriz $n \times m$, donde n son los aspectos y m representa la polaridad, y con esta matriz se obtiene la cantidad total de tuits por cada polaridad según el aspecto considerado en la investigación. En las Tablas 20, 21, 22 y 23 se muestran las puntuaciones de sentimiento para cada banco.

Tabla 20*Puntuación de Sentimiento para el Banco A*

Aspecto	Polaridad			Total
	Negativo = 0	Neutro = 1	Positivo = 2	
Cuenta	165	44	2	211
Tarjeta	145	37	1	183
Cajero	59	7	0	66
Agencia	84	19	1	104
Sistema	276	45	3	324

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Phyton.

Tabla 21*Puntuación de Sentimiento para el Banco B*

Aspecto	Polaridad			Total
	Negativo = 0	Neutro = 1	Positivo = 2	
Cuenta	353	169	1	523
Tarjeta	218	83	1	302
Cajero	165	30		195
Agencia	556	76	1	633
Sistema	1,310	188	11	1,509

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Phyton.

Tabla 22*Puntuación de Sentimiento para el Banco C*

Aspecto	Polaridad			Total
	Negativo = 0	Neutro = 1	Positivo = 2	
Cuenta	208	68	5	281
Tarjeta	159	76	13	248
Cajero	66	15		81
Agencia	106	32	2	140
Sistema	208	64	8	280

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Phyton.

Tabla 23*Puntuación de Sentimiento para el Banco D*

Aspecto	Polaridad			Total
	Negativo = 0	Neutro = 1	Positivo = 2	
Cuenta	67	40	1	108
Tarjeta	71	27	0	98
Cajero	34	9	0	43
Agencia	45	13	0	58
Sistema	132	29	0	161

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Phyton.

Según K. Ravi y Ravi (2017), para calcular el puntaje total de sentimiento, se usa la

Fórmula 6:

$$SCT_{ij} = \sum_{p=1}^{P_i} (SCT)_{j_p}^{POW} - \sum_{n=1}^{N_i} (SCT)_{j_n}^{NOW} \quad (6)$$

En la Tabla 24 se muestra la matriz de resultados de puntaje de sentimiento (SCT), que es la aplicación de la Fórmula 6, obteniendo el resultado de la suma de los tuits calificados con polaridad positiva menos la suma de los tuits calificados con la polaridad negativa.

Tabla 24

La Matriz de Resultados para los Cuatro Bancos y los Respectivos Aspectos

Banco	Aspecto				
	Cuenta	Tarjeta	Cajero	Agencia	Sistema
A	-163	-144	-59	-83	-273
B	-352	-217	-165	-555	-1,299
C	-203	-146	-66	-104	-200
D	-66	-71	-34	-45	-132
Total	-784	-578	-324	-787	-1,904

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Python.

Clasificación Utilizando MCDM (Etapa 5)

Para esta etapa se aplica MCDM lo cual permite seleccionar, comparar y clasificar diferentes alternativas, obteniendo la priorización mayor valorada para el servicio que tenga el mejor desempeño de los bancos evaluados.

Proceso: Análisis jerárquico (AHP)

En este proceso se determinan los pesos de cada criterio o aspecto y; para ello, se aplica el método de decisión AHP, cuyos pasos se describen.

Construcción de la matriz de comparación por pares. Se realiza una valoración de cada criterio respecto del otro para validar cuánto más o cuánto menos aporta cada criterio para tomar una decisión. Para realizar la valoración se construye una matriz de orden (5 x 5), en la cual se valora de acuerdo con los valores descritos en la Tabla 25, donde las celdas de la diagonal principal registran el valor de 1 puesto que tienen igual importancia. Para el resto de

los aspectos, se debe considerar que existe una relación recíproca; de tal manera que, si el aspecto *tarjeta* es de “importancia moderada” sobre el aspecto *cajero*, entonces tiene un valor de 3 y; recíprocamente para *cajero* sobre *tarjeta* tiene un valor de $1/3 = 0.33$. En la Tabla 25 se muestra la matriz completa de comparación para cada uno de los aspectos.

Tabla 25

Matriz de Comparación por Pares de los Aspectos

	Aspecto				
	Tarjeta	Cuenta	Cajero	Sistema	Agencia
Tarjeta	1.00	2.00	3.00	9.00	2.00
Cuenta	0.50	1.00	2.00	2.00	5.00
Cajero	0.33	0.50	1.00	2.00	2.00
Sistema	0.11	0.50	0.50	1.00	2.00
Agencia	0.50	0.20	0.50	0.50	1.00

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Python.

Después de la construcción de la matriz de comparación por pares, se estima el vector de prioridades y ; para ello, se normaliza la matriz anterior utilizando un procedimiento denominado “proporción total”, donde se suman los valores de cada columna (ver Tabla 26).

Tabla 26

Matriz Normalizada de Comparación de Aspectos

	Aspecto				
	Tarjeta	Cuenta	Cajero	Sistema	Agencia
Tarjeta	1.00	2.00	3.00	9.00	2.00
Cuenta	0.50	1.00	2.00	2.00	5.00
Cajero	0.33	0.50	1.00	2.00	2.00
Sistema	0.11	0.50	0.50	1.00	2.00
Agencia	0.50	0.20	0.50	0.50	1.00
Total	2.44	4.20	7.00	14.50	12.00

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Python.

Ahora bien, para construir la matriz normalizada, se divide cada valor de las celdas entre el valor total que se encuentra en la última fila de sus respectivas columnas (ver Tabla 27).

Tabla 27

Matriz Normalizada de Comparación de Aspectos

	Aspecto				
	Tarjeta	Cuenta	Cajero	Sistema	Agencia
Tarjeta	0.41	0.48	0.43	0.62	0.17
Cuenta	0.20	0.24	0.29	0.14	0.42
Cajero	0.14	0.12	0.14	0.14	0.17
Sistema	0.05	0.12	0.07	0.07	0.17
Agencia	0.20	0.05	0.07	0.03	0.08

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Phyton.

Para obtener el vector de prioridades, se calcula el promedio de los valores de cada fila (ver Tabla 28).

Tabla 28

Vector de Prioridades

	Aspecto					Vector de prioridades
	Tarjeta	Cuenta	Cajero	Sistema	Agencia	
Tarjeta	0.41	0.48	0.43	0.62	0.17	0.42
Cuenta	0.20	0.24	0.29	0.14	0.42	0.26
Cajero	0.14	0.12	0.14	0.14	0.17	0.14
Sistema	0.05	0.12	0.07	0.07	0.17	0.09
Agencia	0.20	0.05	0.07	0.03	0.08	0.09

Nota. Obtenido de la matriz de confusión y sus métricas realizada con librerías de Phyton.

A partir de la Tabla 28, se puede concluir que el aspecto *tarjeta* es el que más contribuye al objetivo, seguido de *cuenta*, *cajero*, *sistema* y *agencia*, donde el ratio de consistencia de la matriz que se obtiene es 0.098, que indica que es consistente ($RC < 0.10$), por lo que es admisible el vector.

Proceso: Aplicación del método VIKOR

Paso 1. Primero se debe determinar la función de cada criterio o aspecto. Para la presente investigación, todos los aspectos tienen la función de maximizar la cantidad de opiniones positivas (ver Tabla 29).

Tabla 29

Función de Variables a Maximizar

Función	Aspecto				
	Tarjeta	Cuenta	Cajero	Sistema	Agencia
	Max	Max	Max	Max	Max

Paso 2. Se calcula la medida de utilidad (S) y de arrepentimiento (R) de cada alternativa (ver Tabla 30) y el índice de orden de selección (Q) (ver Tabla 31).

Tabla 30

Cálculo de la Utilidad (S) y el Arrepentimiento (R) de cada Aspecto

Banco	Aspecto					S	R
	Tarjeta	Cuenta	Cajero	Sistema	Agencia		
A	0.25	0.21	0.18	0.14	0.04	0.65	0.21
B	0.38	0.45	0.51	0.68	0.29	0.00	0.00
C	0.25	0.26	0.20	0.11	0.05	0.61	0.21
D	0.12	0.08	0.10	0.07	0.02	1.00	0.42

Nota. A partir de la matriz de resultados (ver Tabla 24), se realiza la normalización y se procede al cálculo de S = Medida de utilidad y R = Medida de arrepentimiento.

Tabla 31

Cálculo el Índice Q de cada Alternativa

Banco	S	R	Q
A	0.65	0.21	0.58
B	0.00	0.00	0.00
C	0.61	0.21	0.56
D	1.00	0.42	1.00

Paso 3. Se procede con el ordenamiento de las alternativas de acuerdo con el índice Q , donde la mejor solución es la que registra menor índice Q (ver Tabla 32).

Tabla 32

Ordenamiento de las Alternativas de Acuerdo con el Índice Q

Banco	S	R	Q	Jerarquía
A	0.71	0.26	0.66	3
B	0.00	0.00	0.00	1
C	0.62	0.20	0.55	2
D	1.00	0.42	1.00	4

Paso 4. Se verifica que se cumplan las condiciones de aceptabilidad.

Condición de ventaja aceptable ($Q(a'') - Q(a') \geq DQ$). Se obtuvieron valores $Q(a'') = 0.55$ (alternativa en la segunda posición), $Q(a') = 0.00$ (alternativa en la primera posición), $DQ = 0.33$ (DQ es igual a $\frac{1}{J-1}$ donde J es igual a cuatro [alternativas]). Luego, reemplazando valores se tiene $0.55 - 0.00 \geq 0.33$; lo cual cumple la condición.

Estabilidad aceptable en el proceso de decisión. En esta parte se valida que para a' su utilidad $S' = 0.0$ y su arrepentimiento $R' = 0.0$, que en ambos casos es más bajo que las demás alternativas; por lo tanto, esta condición también se cumple.

Finalmente, se puede afirmar que la presente investigación proporciona información adicional para determinar la mejor opción de servicio a contratar, siendo el Banco B el que ocupa el primer lugar, seguido del Banco C, Banco A y Banco D, en ese orden, tomando en consideración los aspectos *tarjeta, cuenta, cajero, sistema y agencia*.

Discusión de Resultados

Esta investigación se centra en dos aspectos principales: (a) seleccionar el mejor modelo para determinar la polaridad del sentimiento y (b) aplicar el método MCDM-VIKOR.

Sobre el primer aspecto, se realizó la validación de los resultados de polaridad de los cuatro modelos: Stanza, VADER, TextBlob y BETO, comparándolos con los obtenidos por juicio humano para, finalmente, determinar la *exactitud* de dichos modelos, puesto que según Barrios Arce (2019), este indicador permite lograr ello. En la Tabla 33 se muestra la

precisión, la exhaustividad, el valor-F1 de cada modelo con su respectiva polaridad y exactitud.

En la Tabla 33 se observa que, según los resultados de la métrica *exactitud* de cuatro modelos, el que presenta mayor proporción de resultados acertados es el modelo BETO, obteniendo 88%, logrando un alto porcentaje de predicciones acertadas con respecto a la calificación manual; por lo tanto, se puede afirmar que el modelo es adecuado para determinar la polaridad de sentimientos de comentarios en Twitter para las principales entidades financieras peruanas. Esto también fue evidenciado en otras investigaciones como la de Rosenbrock et al. (2021), donde el resultado de la métrica exactitud sobre el modelo BETO fue 83%; es decir, se obtuvo 5% de diferencia, lo cual puede tener diversas causas, como diferencias propias del lenguaje, diferencias del tipo de industria y preentrenamiento con un conjunto de datos de menor tamaño.

La diferencia entre la precisión de nuestros resultados y los presentados por Rosenbrock pueden deberse a varias razones. Uno de los factores más importantes es el conjunto de datos utilizado para la presente investigación. Los conjuntos de datos pueden variar significativamente en su tamaño, calidad y representatividad del lenguaje natural. Por ejemplo, algunos conjuntos de datos pueden contener principalmente texto informal, mientras que otros pueden contener principalmente texto formal. Esto puede afectar la precisión del modelo, ya que algunos modelos pueden estar mejor adaptados a ciertos tipos de texto que a otros. Adicional a ello, el modelo BETO presenta diversas limitaciones; según (Rosenbrock et al., 2021), estas pueden clasificarse en varios factores: (a) incapacidad para manejar incongruencias de etiquetas, en donde BETO puede tener dificultades para clasificar correctamente los comentarios cuando la connotación del texto no coincide con la etiqueta proporcionada, como por ejemplo, un comentario positivo etiquetado como negativo; (b) dificultad para tratar con comentarios ambiguos, ya que BETO puede tener problemas para

clasificar comentarios que contienen una mezcla de sentimientos positivos y negativos; (c) dificultad para interpretar lenguaje figurado, en la que el modelo puede no ser capaz de entender y procesar correctamente las frases con sentido figurado; (d) problemas con la negación, en donde BETO puede tener dificultades para interpretar correctamente las frases que contienen negaciones, especialmente cuando se usan múltiples negaciones en una misma frase; (e) un rendimiento comparativamente más bajo en idiomas que no sean inglés, ya que BETO ha obtenido resultados ligeramente inferiores en español en comparación con resultados publicados de experimentos similares en inglés; (f) la necesidad de ajuste fino, pues el modelo BETO puede requerir ajustes finos específicos para mejorar su rendimiento, lo cual puede añadir complejidad al proceso de implementación del modelo; y (g) la falta de comprensión del significado real del texto, ya que, al igual que otros modelos de procesamiento de lenguaje natural, BETO no comprende realmente el significado del texto, sino que se basa en las relaciones de co-ocurrencia que encuentra entre las palabras. No obstante, estas limitaciones no son exclusivas del modelo BETO y representan grandes desafíos comunes en el campo del procesamiento del lenguaje natural.

El modelo BETO a comparación de distintas técnicas de aprendizaje automático utilizadas como los modelos Naïve Bayes, Random Forest, Regresión logística, SVM, LSTM + Word2Vec; es superior con la aplicación de búsqueda de hiperparámetros, para el Procesamiento del Lenguaje Natural. El caso de estudio fue un conjunto de más de 50.000 comentarios en lenguaje español coloquial sobre películas, extraídos del sitio web (<https://www.cinesargentinos.com.ar/>). Aunque los porcentajes de acierto obtenidos en ese estudio son menores que los publicados sobre casos similares en los cuales los textos se encuentran en idioma inglés. Esta diferencia puede tener varias causas: diferencias propias del lenguaje, preentrenamiento con un corpus de menor tamaño, diferencias en el nivel de informalidad del lenguaje coloquial utilizado, o incluso, mejor ajuste de algunos hiperparámetros.

Tabla 33*Descripción de Resultados de Predicción por Modelo de Polarización*

Predicción	Precisión	Exhaustividad	Valor-F1
Stanza (Exactitud = .65)			
Negativo	.91	.66	.77
Neutral	.34	.65	.44
Positivo	.00	.00	.00
Vader (Exactitud = .31)			
Negativo	.98	.22	.35
Neutral	.25	.71	.37
Positivo	.02	.67	.04
TextBlob (Exactitud = .60)			
Negativo	.96	.57	.71
Neutral	.32	.79	.46
Positivo	.00	.00	.00
BETO (Exactitud = .88)			
Negativo	.96	.91	.93
Neutral	.65	.79	.72
Positivo	.00	.00	.00

Sobre el segundo aspecto, esto es, la aplicación del método MCDM-VIKOR; ello permitió determinar que la mejor alternativa es Banco B, seguido por Banco C, Banco A y Banco D, en ese orden, con respecto a los aspectos evaluados: tarjeta, cuenta, cajero, sistema y agencia. La aplicación de este método fue validado por el cumplimiento de sus condiciones de aceptabilidad que son: (a) condición de ventaja aceptable ($Q(a'') - Q(a') \geq DQ$), cuyos valores de cada variables son $Q(a'') = 0.55$, $Q(a') = 0.00$ y $DQ = 0.33$ y; realizando los cálculos se obtuvo $0.55 \geq 0.33$, con lo que la aplicación del modelo cumple con esta condición; y (b) estabilidad aceptable en la toma de decisiones, que se cumple cuando la utilidad (S) y el arrepentimiento (R) es el valor más bajo en comparación con las otras alternativas (bancos). Este método también fue utilizado por Vyas et al. (2019) para clasificar un conjunto de instituciones bancarias en función de la opinión de las personas.

Finalmente, a partir de estos resultados, se puede afirmar que la aplicación del modelo propuesto es de utilidad para proporcionar información adicional al cliente, esto es, que lo ayude a tomar una decisión antes de contratar los servicios ofrecidos por las principales entidades bancarias peruanas.

Por otro lado, en la presente investigación se ha logrado consolidar una base de datos de tuits sobre las opiniones de clientes acerca de los servicios ofrecidos por los principales bancos del Perú dentro de la plataforma Twitter, sobre la cuales se aplicó una serie de algoritmos para realizar la limpieza de dichos datos. Con ello, se logró responder de manera afirmativa a la primera pregunta de investigación referente a la utilidad de la red social Twitter como fuente para recolectar y consolidar opiniones de los usuarios de servicios financieros. Además, es posible extraer información desde Twitter para ampliar el análisis no solo para el sector financiero sino también para otros sectores, ampliando el volumen de datos.

En cuanto a la segunda pregunta vinculada a los principales aspectos de evaluación necesarios para el modelo propuesto, para responderla, se ha realizado un análisis de frecuencias de los datos obtenidos desde Twitter, donde se identificó que los aspectos a estudiar dentro de este modelo serían tarjeta, cuenta, cajero, sistema y agencia, puesto que fueron los predominantes en dicho análisis.

Con respecto a la tercera pregunta de investigación, se ha determinado la polaridad de los tuits evaluando distintos modelos de PLN desarrollados y preentrenados como: Stanza, VADER, TextBlob y BETO, de los cuales, para determinar la exactitud de los modelos, se compararon contra el etiquetado manual, logrando determinar que el mejor modelo para determinar la polaridad de los tuits extraídos es BETO.

Con respecto a la cuarta pregunta de investigación, si la aplicación del modelo MCDM permite seleccionar, comparar y clasificar diferentes servicios financieros, efectivamente en la presente investigación el método MCDM-VIKOR permitió realizar la

clasificación de los bancos A, B, C y D utilizando una clasificación con el enfoque de votación por mayoría, cuyo resultado determinó que el Banco B fue la mejor alternativa a elegir con respecto a los aspectos tarjeta, cuenta, cajero, sistema y agencia.

Resumen

En el presente capítulo se presentaron los resultados del modelo propuesto que permitió la medición del desempeño de servicios financieros mediante el análisis de sentimiento y métodos de decisiones multicriterio, logrando demostrar la viabilidad de la aplicación del modelo. Además, se analizaron cuatro métodos para determinar la polaridad del sentimiento (Stanza, VADER, TextBlob y BETO) y se profundizó en el detalle del modelo propuesto que utilizó el método BETO, determinándose que sí podría funcionar como modelo base puesto que se obtuvo una exactitud de 88% y; además, se ha validado la aceptabilidad del método VIKOR como herramienta que permite realizar una clasificación de alternativas.

Capítulo V: Conclusiones y Recomendaciones

Conclusiones

1. Hoy en día las personas buscan información sobre la percepción de calidad acerca de productos o servicios que ofrecen las empresas dentro de las diferentes redes sociales antes de tomar la decisión de adquirirlos; pero debido a la alta cantidad de comentarios positivos, negativos o neutrales, termina siendo una tarea engorrosa sintetizar dicha información y tomar la mejor alternativa. En este contexto, en la presente investigación se propuso un modelo de medición de servicios financieros que sería de utilidad para los clientes como complemento que ayude a sintetizar los diversos comentarios mediante una tabla de puntuación; para luego ordenar las alternativas a partir de los pesos de los aspectos seleccionados, dando como resultado una lista de opciones para el cliente. De los resultados obtenidos, se puede concluir que el modelo propuesto es útil para apoyar a los clientes en la selección del producto o servicio a contratar, tomando como referencia los comentarios dejados por usuarios en Twitter.
2. Para el desarrollo del modelo propuesto, se logró consolidar un conjunto de datos de 15,546 tuits sobre las opiniones de clientes acerca de los servicios ofrecidos por las principales entidades bancarias del Perú dentro de la plataforma Twitter. Estos resultados obtenidos permitieron señalar a esta red social como una fuente valiosa de datos para realizar análisis diversos tanto para el estudio propuesto como para estudios en otros sectores.
3. Además, se logró determinar los aspectos de evaluación mediante un análisis de frecuencias sobre el conjunto de datos extraído. Los aspectos tarjeta, cuenta, cajero, sistema y agencia se registraron como los más destacados. No obstante, estos pueden ser definidos por otro tipo de métodos como juicio de expertos en el caso de que el modelo sea extrapolado a otro sector o realidad.

4. Para determinar la polarización de cada tuit, se utilizó al modelo BETO, debido a que la métrica exactitud fue superior que el resto de los modelos analizados, logrando clasificar en positivos, negativos o neutrales a los 5,276 tuits de la muestra. Aplicando el modelo BETO, se obtiene 88% de precisión; y se observa una detección de polaridad negativa de alto valor siendo el valor-F1 igual a 93%, en comparación con otras investigaciones como la de Rosenbrock et al. (2021), donde el resultado de la métrica exactitud sobre el modelo BETO fue 83%. Por lo tanto, se puede concluir que el modelo es adecuado para determinar la polaridad de sentimientos de comentarios en Twitter para las principales entidades financieras peruanas.
5. Se estableció el método de enfoque de decisión de criterios múltiple MCDM-VIKOR, que permitió realizar una clasificación de los bancos a través de los aspectos o criterios, utilizando el enfoque AHP para elegir el más preferible a través de un conjunto de aspectos. De esta manera, se logró ordenar a las instituciones financieras en función de las opiniones de las personas, siendo la secuencia Banco B, Banco C, Banco A y Banco D (en ese orden), tomando en consideración los aspectos *tarjeta, cuenta, cajero, sistema y agencia*, cumpliendo con las condiciones de aceptabilidad establecida por el modelo, que son: (a) condición de ventaja aceptable ($Q(a'') - Q(a') \geq DQ$), cuyos valores de cada variables son $Q(a'') = 0.55$, $Q(a') = 0.00$ y $DQ = 0.33$ y; realizando los cálculos se obtuvo $0.55 \geq 0.33$ y (b) estabilidad aceptable en la toma de decisiones, que se cumple cuando la utilidad (S) y el arrepentimiento (R) es el valor más bajo en comparación con las otras alternativas (bancos).
6. De acuerdo a los resultados obtenidos con la validación de lo explicado vemos que el modelo ayudará a los usuarios a tomar una mejor decisión , así mismo permitirá a las entidades financieras peruanas a identificar y analizar la información de manera más específica, en el sentido que podrán evaluar qué es lo que más les aqueja a sus clientes en cuanto a sus productos y/o servicios a diferencia de las herramientas tradicionales que de

acuerdo a diversos estudios (Debnath, 2023; Solca, 2020; Vidya et al., 2015) coincidieron que estas herramientas como es el caso del indicador NPS solo mide la lealtad y el nivel de satisfacción del cliente de forma limitada no teniendo un análisis de forma granular hacia el verdadero problema del cliente, por ello en la actualidad las organizaciones complementan las herramientas tradicionales con herramientas de análisis de sentimiento, por lo cual concluimos que nuestro modelo de, análisis de sentimiento evaluando los aspectos para obtener con ello el mejor desempeño de los bancos, permitirá a complementar estas a herramientas tradicionales para un mejor análisis organizacional.

Recomendaciones

1. A futuro se podría incorporar un nuevo diccionario en dominio financiero dentro de los modelos empleados para así mejorar la detección de sentimientos.
2. Se recomienda ampliar el estudio, incorporando conceptos como ambigüedad de palabras, errores ortográficos, etc.
3. Se recomienda aplicar el presente modelo de medición de desempeño por periodos de tiempo, puesto que la opinión y la preferencia cambia en el tiempo.
4. El modelo planteado puede ser usado en diferentes industrias para apoyarse a mejorar sus productos o servicios gradualmente y; de esta manera, atraer o fidelizar clientes.
5. Se recomienda mejorar el entrenamiento de los modelos para oraciones positivas, puesto que, según los resultados, los cuatro modelos presentaron alta precisión cuando la oración era negativa; sin embargo, la precisión se reduce drásticamente cuando se trata de oraciones positivas.

Referencias

- Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011). Sentiment analysis of Twitter data [Análisis de sentimiento de los datos de Twitter]. En *Proceedings of the Workshop on Language in Social Media (LSM 2011)* (pp. 30-38).
<https://aclanthology.org/W11-0705.pdf>
- Agudelo Restrepo, L. (2022). *Identificación de factores a mejorar para aumentar la recomendación, análisis de NPS para clientes neutros a partir de procesamiento de lenguaje natural* [Tesis de maestría, Universidad Nacional de Colombia].
<https://repositorio.unal.edu.co/handle/unal/81608>
- Al Karim, R., & Chowdhury, T. (2014). Customer satisfaction on service quality in private commercial banking sector in Bangladesh [Satisfacción del cliente sobre la calidad del servicio en el sector de la banca comercial privada en Bangladesh]. *British Journal of Marketing Studies*, 2(2), 1-11.
- Alemán Viteri, S. B. (2021). Análisis de sentimientos para Twitter con Vader y TextBlob. *ODIGOS*, 2(3), 1-17. <https://doi.org/10.35290/ro.v2n3.2021.494>
- Algur, S. P., & Patil, R. H. (2017). Sentiment analysis by identifying the speaker's polarity in Twitter data [Análisis de sentimiento mediante la identificación de la polaridad del hablante en los datos de Twitter]. *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)* (pp. 1-5). IEEE.
- Álvarez Risco, A. (2020). *Clasificación de las investigaciones* [Nota académica]. Universidad de Lima.
- Assaad, W., & Marx Gómez, J. (2011). Social network in marketing (social media marketing): Opportunities and risks [Red social en marketing (social media marketing): Oportunidades y riesgos]. *International Journal of Managing Public Sector Information and Communication Technologies*, 2(1). 13-22.
<https://doi.org/10.5121/ijmpict.2011.2102>

Athar, A. (20 de junio de 2011). *Sentiment analysis of citations using sentence structure-based features* [Análisis de opinión de las citas utilizando características basadas en la estructura de la oración]. Proceedings of the ACL 2011 Student Session (pp. 81-87), Portland, OR, Estados Unidos. <https://aclanthology.org/P11-3015>

Banco Mundial. (s.f.). *Perú Panorama general*.

<https://www.bancomundial.org/es/country/peru/publication/repensar-el-futuro-del-per-apuntes-de-pol-tica-para-transformar-al-estado-en-un-gestor-del-bienestar-y-el-desarrollo>

Banco Mundial. (25 de agosto de 2021). *Repensar el futuro del Perú: Notas de política para transformar al Estado en un gestor de bienestar y desarrollo*.

<https://www.bancomundial.org/es/country/peru/publication/repensar-el-futuro-del-per-apuntes-de-pol-tica-para-transformar-al-estado-en-un-gestor-del-bienestar-y-el-desarrollo>

Barrios Arce, J. I. (26 de julio de 2019). *La matriz de confusión y sus métricas (consultores estratégicos en ciencia de datos)*. Health Big Data. <https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>

Baruah, T. D. (2012). Effectiveness of social media as a tool of communication and its potential for technology enabled connections: A micro-level study [Eficacia de las redes sociales como herramienta de comunicación y su potencial para conexiones habilitadas por tecnología: Un estudio a nivel micro]. *International Journal of Scientific and Research Publications*, 2(5), 1-9.

BBVA Research. (4 de marzo de 2022). *El conflicto entre Rusia y Ucrania y el impacto en la inflación en Perú*. <https://www.bbva.com/es/pe/conflicto-entre-rusia-y-ucrania-y-el-impacto-en-la-inflacion-en-peru/>

- Bonta, V., Kumaresh, N., & Janardhan, N. (2019). A comprehensive study on lexicon based approaches for sentiment analysis [Un estudio exhaustivo sobre los enfoques basados en el léxico para el análisis de sentimientos]. *Asian Journal of Computer Science and Technology*, 8(S2), 1-6. <https://doi.org/10.51983/ajcst-2019.8.s2.2037>
- Botchway, R. K., Jibril, A. B., Kwarteng, M. A., Chovancova, M., & Oplatková, Z. K. (2019). A review of social media posts from UniCredit bank in Europe: A sentiment analysis approach [Una revisión de las publicaciones en las redes sociales del banco UniCredit en Europa: un enfoque de análisis de sentimiento]. *Proceedings of the 3rd International Conference on Business and Information Management* (pp. 74-79). ACM.
- Brans, J. P., Vincke, Ph., & Mareschal, B. (1986). How to select and how to rank projects: The Promethee method. *European Journal of Operational Research*, 24(2), 228-238. [https://doi.org/10.1016/0377-2217\(86\)90044-5](https://doi.org/10.1016/0377-2217(86)90044-5)
- Carballar Falcón, J. A. (2011). *Twitter: Marketing personal y profesional*. RC Libros.
- Costa-Sánchez, C., & Corbacho-Valencia, J. M. (2015). LinkedIn para seleccionar y captar talento: Análisis del uso de LinkedIn como herramienta de employer branding. *Prisma Social*, 14, 187-221.
- D'Andrea, A., Ferri, F., Grifoni, P., & Guzzo, T. (2015). Approaches, tools and applications for sentiment analysis implementation [Enfoques, herramientas y aplicaciones para la implementación del análisis de sentimientos]. *International Journal of Computer Applications*, 125(3). <https://doi.org/10.5120/ijca2015905866>
- Deloitte. (2018). *NPS sector bancario: Net promoter score para sector en Colombia (Informe ejemplo: Datos ficticios)*. <https://www2.deloitte.com/content/dam/Deloitte/co/Documents/strategy/DELOITTE%20-%20Estudio%20ejemplo%20sector%20Bancario.docx.pdf>

- Galeano, S. (28 de enero de 2022). *Cuáles son las redes sociales con más usuarios del mundo* (2022). Marketing 4 Ecommerce. <https://marketing4ecommerce.net/cuales-redes-sociales-con-mas-usuarios-mundo-ranking/>
- Greaves, F., Laverty, A. A., Ramírez Cano, D., Moilanen, K., Pulman, S., Darzi, A., & Millett, C. (2014). Tweets about hospital quality: A mixed methods study [Tuits sobre la calidad hospitalaria: Un estudio de métodos mixtos]. *BMJQS*, 23(10), 838-846. <https://doi.org/10.1136/bmjqs-2014-002875>
- Gurevich, A. (2016). El tiempo todo en Facebook. *Aposta*, 69, 217-238. <https://www.redalyc.org/pdf/4959/495952431008.pdf>
- Hasan Basari, A. S., Hussin, B., Ananta, G. P., & Zeniarja, J. (2013). Opinion mining of movie review using hybrid method of support vector machine and particle swarm optimization [Extracción de opiniones de reseñas de películas utilizando un método híbrido de máquina de vectores de soporte y optimización de enjambre de partículas]. *Procedia Engineering*, 53, 453-462. <https://doi.org/10.1016/j.proeng.2013.02.059>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. (2014). *Metodología de la investigación* (6a ed.). McGraw-Hill.
- Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text [Un modelo parsimonioso basado en reglas para el análisis de sentimientos del texto de las redes sociales]. *Eighth International AAI Conference on Weblogs and Social Media*, 8(1).216-225. <https://doi.org/10.1609/icwsm.v8i1.14550>
- Hwang, C. L., & Yoon, K. (1981). Methods for multiple attribute decision making [Métodos para la toma de decisiones de atributos múltiples]. En M. Beckmann, & H. P. Kunzi (Eds.), *Multiple attribute decision making. Lecture notes in economics and mathematical systems* (Vol. 186). Springer. https://doi.org/10.1007/978-3-642-48318-9_3

Ipsos Perú. (6 de agosto de 2020). *Uso de redes sociales entre peruanos conectados 2020*.

<https://www.ipsos.com/es-pe/uso-de-redes-sociales-peruanos-conectados-2020>

Ishizaka, A., & Nemery, P. (2013). *Multi-criteria decision analysis: Methods and software*

[Análisis de decisión multicriterio: Métodos y software]. Wiley.

<http://dx.doi.org/10.1002/9781118644898>

Kang, D., & Park, Y. (2014). Review-based measurement of customer satisfaction in mobile

service: Sentiment analysis and VIKOR approach [Medición basada en revisión de la satisfacción del cliente en el servicio móvil: Análisis de sentimiento y enfoque

VIKOR]. *Expert Systems with Applications*, 41(4), 1041-1050.

<https://doi.org/10.1016/j.eswa.2013.07.101>

Kaul, V. (2012). The digital communications revolution [La revolución de las

comunicaciones digitales]. *Online Journal of Communication and Media*

Technologies, 2(3), 113-130. <https://doi.org/10.29333/OJCMT%2F2392>

Kharde, V. A., & Sonawane, S. (2016). *Sentiment analysis of Twitter data: A survey of*

techniques [Análisis de sentimiento de los datos de Twitter: Una encuesta de

técnicas]. <https://doi.org/10.48550/arXiv.1601.06971>

Kotler, P., & Armstrong, G. (2017). *Fundamentos de marketing* (L. A. Araujo & L. E. Pineda

Ayala, Trad.; 13a ed.). Pearson Educación de México.

Kouloumpis, E., Wilson, T., & Moore, J. (2011). Twitter sentiment analysis: The good the

bad and the omg! [Análisis de sentimiento de Twitter: ¡Lo bueno, lo malo y lo

bueno!]. *Fifth International AAI Conference on Weblogs and Social Media*.

Catalonia, Spain.

Kumar, G. (2018, 29-31 mayo). A multi-criteria decision making approach for

recommending a product using sentiment analysis [Un enfoque de toma de decisiones

de criterios múltiples para recomendar un producto utilizando análisis de

sentimiento]. *2018 12th International Conference on Research Challenges in*

Information Science, Nantes, France. <http://dx.doi.org/10.1109/RCIS.8406679>

- Kumawat, D., & Jain, V. (2015). POS tagging approaches: A comparison [Enfoques de etiquetado de POS: Una comparación]. *International Journal of Computer Applications*, 118(6), 32-38. <https://doi.org/10.5120/20752-3148>
- Kumbirai, M., & Webb, R. (2010). A financial ratio analysis of commercial bank performance in South Africa [Un análisis de razones financieras del desempeño de los bancos comerciales en Sudáfrica]. *African Review of Economics and Finance*, 2(1), 30-53.
- Kutchera, J., Hilda, G., & Fernández, A. (2014). *E-X-I-T-O: Su estrategia de marketing digital en 5 pasos*. Grupo Editorial Patria.
<https://www.editorialpatria.com.mx/mobile/pdf/files/9786074386387.pdf>
- Linoff, G. S., & Berry, M. J. A. (2011). *Data mining techniques: For marketing, sales, and customer relationship management* [Técnicas de minería de datos: Para marketing, ventas y gestión de relaciones con los clientes] (3a ed.). Wiley.
- Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis: A review of research topics, venues, and top cited papers [La evolución del análisis de sentimientos: una revisión de los temas de investigación, los lugares y los artículos más citados]. *Computer Science Review*, 27, 16-32.
<https://doi.org/10.1016/j.cosrev.2017.10.002>
- Mohan, S., Solanki, A. K., Taluja, H. K., Anuradha, & Singh, A. (2022). Predicting the impact of the third wave of COVID-19 in India using hybrid statistical machine learning models: A time series forecasting and sentiment analysis approach [Predecir el impacto de la tercera ola de COVID-19 en India utilizando modelos híbridos de aprendizaje automático estadístico: un enfoque de pronóstico de series temporales y análisis de sentimiento]. *Computers in Biology and Medicine*, 144, 1-13.
<https://doi.org/10.1016/j.combiomed.2022.105354>

- Montoyo, A., Martínez-Barco, P., & Balahur, A. (2012). Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments [Análisis de subjetividad y sentimiento: Una descripción general del estado actual del área y desarrollos previstos]. *Decision Support Systems*, 53(4), 675-679.
<https://doi.org/10.1016/j.dss.2012.05.022>
- Morales Vallejo, P. (2012). *Estadística aplicada a las Ciencias Sociales: Tamaño necesario de la muestra: ¿Cuántos sujetos necesitamos?*
<http://data.evalua.cdmx.gob.mx/docs/gral/taller2015/S0202EAC.pdf>
- Mudinas, A., Zhang, D., & Levene, M. (2012). Combining lexicon and learning based approaches for concept-level sentiment analysis [Combinación de enfoques basados en el léxico y el aprendizaje para el análisis de sentimientos a nivel de concepto]. *Proceedings of the first international workshop on issues of sentiment discovery and opinion mining* (p. 5). ACM.
- Mumtaz, D., & Ahuja, B. (2018). *A lexical and machine learning-based hybrid system for sentiment análisis* (Vol. 713) [Un sistema híbrido basado en aprendizaje automático y léxico para el análisis de sentimientos]. Springer.
- Nivre, J., de Marneffe, M.-C., Ginter, F., Goldberg, Y., Hajič, J., Manning, C. D., McDonald, R., Petrov, S., Pyysalo, S., Silveira, N., Tsarfaty, R., & Zeman, D. (2016). Universal dependencies v1: A multilingual treebank collection [Dependencias universales v1: Una colección de treebank multilingüe]. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 1659-1666).
<https://aclanthology.org/L16-1262>
- Opricovic, S. C. (1998) *Multicriteria optimization of civil engineering systems* [Optimización multicriterio de sistemas de ingeniería civil; [Tesis doctoral, Facultad de Ingeniería Civil, Universidad de Belgrado].

- Opricovic, S. C., & Tzeng, G. H. (2007). Extended VIKOR method in comparison with outranking methods [Método VIKOR extendido en comparación con los métodos de clasificación superior]. *European Journal of Operational Research*, 178(2), 514-529. <https://doi.org/10.1016/j.ejor.2006.01.020>
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis [Minería de opinión y análisis de sentimiento]. *Foundations and Trends® in Information Retrieval*, 2(1-2), 1-135. <http://dx.doi.org/10.1561/15000000011>
- Panico, C. (2018). *La eficacia del análisis de sentimientos para la empresa: El caso de estudio Dell Technologies Inc.* [Trabajo de fin de grado, Universidad Complutense de Madrid]. https://www.ucm.es/data/cont/docs/758-2019-01-04-TFG_Panico_Chicara_TFG.pdf
- Pérez, J. M., Giudici, J. C., & Luque, F. (2021). *pysentimiento: A Python toolkit for sentiment analysis and SocialNLP tasks* [pysentimiento: un kit de herramientas de Python para análisis de sentimiento y tareas de SocialNLP]. <https://www.arxiv-vanity.com/papers/2106.09462/>
- Pizarro, J. (2000). El análisis de estudio cualitativo. *Atención Primaria*, 25(1), 42-46.
- Pretel Jiménez, M., De Frutos Torres, B., & Sánchez Valle, M. (2018). Relación entre marca y consumidor en las redes sociales: Estudio del vínculo afectivo de los jóvenes con dos marcas tecnológicas. *Revista de Comunicación*, 17(2), 229–245. <https://doi.org/10.26441/RC17.2-2018-A10>
- Priyavrat, C., & Sharma, N. (2018). Sentiment Analysis using tidytext package in R [Análisis de sentimiento utilizando el paquete tidytext en R]. *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*. <http://dx.doi.org/10.1109/ICSCCC.2018.8703296>
- Quispe Aguilar, J. S. (2015). *Modelo de extracción y análisis de información de la red social Twitter* [Tesis de grado, Universidad Mayor de San Andrés]. <http://repositorio.umsa.bo/xmlui/handle/123456789/8718>

- Ramos, J. (2013). *Instagram para empresas*. Juanjo Ramos.
- Ramos, J. (2020). *TikTok para empresas*. XinXii.
- Ravi, K., & Ravi, V. (2017). Ranking of branded products using aspect-oriented sentiment analysis and ensembled multiple criteria decision-making [Clasificación de productos de marca utilizando análisis de sentimientos orientados a aspectos y toma de decisiones de criterios múltiples combinados]. *International Journal of Knowledge Management in Tourism and Hospitality*, 1(3), 317-359.
<https://doi.org/10.1504/IJKMTH.2017.086816>
- Rosenbrock, G., Trossero, S., & Pascal, A. (2021). Técnicas de análisis de sentimientos aplicadas a la valoración de opiniones en el lenguaje español. *XXVII Congreso Argentino de Ciencias de la Computación (CACIIC 2021)* (pp. 291-300). Universidad Nacional de Salta. <https://cacic2021.unsa.edu.ar/wp-content/uploads/2021/11/LIBRO-DE-ACTAS-CACIC-2021-SALTA.pdf>
- Roy, B. (1978). ELECTRE III: Algorithme de classement base sur une representation floue des preferences en presence de criteres multiples [ELECTRE III: Algoritmo de clasificación basado en una representación variable de preferencias en presencia de criterios múltiples]. *Cahiers Du CERO*, 20(1), 3-24.
- Saaty, T. L. (1980). *The analytic hierarchy process* [El proceso de la jerarquía analítica]. McGraw-Hill.
- Samuel, A. L. (1969). *Some studies in machine learning using the game of checkers: II-recent progress* [Algunos estudios sobre aprendizaje automático utilizando el juego de damas: II-avance reciente]. Pergamon Press.
- Scasserra, A. T. (2022). *Challenging: Cómo hacer producto en un banco digital y construir las finanzas del mañana..* Corgi Business Publishing.
<https://www.amazon.com/Challenging-producto-digital-construir-finanzas-ebook/dp/B09Y73HQRB>

- Segura Navarrete, A., Martínez-Araneda, C., & Vidal-Castro, C., & Rubio-Manzano, C. (2021). A novel approach to the creation of a labelling lexicon for improving emotion analysis in text [Un enfoque novedoso para la creación de un léxico de etiquetado para mejorar el análisis de emociones en el texto]. *The Electronic Library*.
<http://dx.doi.org/10.1108/EL-04-2020-0110>
- Sherlock Communications. (2021). *El estado del SEO en América Latina*.
<https://www.sherlockcomms.com/es/informe-seo-latam/>
- Strapparava, C., & Mihalcea, R. (2008). Learning to identify emotions in text [Aprendiendo a identificar emociones en el texto]. *SAC '08: Proceedings of the 2008 ACM Symposium on Applied Computing* (pp. 1556-1560).
<https://doi.org/10.1145/1363686.1364052>
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis [Métodos basados en léxico para el análisis de sentimientos]. *Computational Linguistics*, 37(2), 267-307. https://doi.org/10.1162/COLI_a_00049
- Talamé, L., Cardoso, A., & Amor, M. (2019). Comparación de herramientas de procesamiento de textos en español extraídos de una red social para Python. XX *Simposio Argentino de Inteligencia Artificial (ASAI 2019) - JAIIO 48* (pp. 53-67).
- Tzeng, G., & Huang, J. (2011). *Multiple attribute decision making methods and applications* [Métodos y aplicaciones de toma de decisiones de atributos múltiples]. CRC Press, Taylor and Francis Group, A Chapman & Hall Book.
- Van Dijck, J. (2016). *La cultura de la conectividad*. Siglo XXI.
- Vyas, V., Uma, V., & Ravi, K. (2019). Aspect-based approach to measure performance of financial services using voice of customer [Enfoque basado en aspectos para medir el desempeño de los servicios financieros usando la voz del cliente]. *Journal of King Saud University - Computer and Information Sciences*, 34(5), 2262-2270.
<https://doi.org/10.1016/j.jksuci.2019.12.009>

- Yue, K., Zhao, C., Chee-Wee, T., Qian, H., & Hefu, L. (2020). Natural language processing (NLP) in management research: A literature review [Procesamiento del lenguaje natural (PNL) en la investigación gerencial: Una revisión de la literatura]. *Journal of Management Analytics*, 7(2), 139-172.
<https://doi.org/10.1080/23270012.2020.1756939>
- Yusuf, S. M., Muazu, M. B., & Akinsanmi, O. (2015). A novel hybrid fuzzy time series approach with applications to enrollments and car road accidents [Un novedoso enfoque híbrido de series temporales difusas con aplicaciones para inscripciones y accidentes de tráfico]. *International Journal of Computer Applications*, 129(2), 37-44.
- Zárate Calderón, G. H. (2021). *Análisis de sentimiento en información de medios periodísticos y redes sociales mediante redes neuronales recurrentes* [Tesis de grado, Pontificia Universidad Católica del Perú]. <http://hdl.handle.net/20.500.12404/21525>
- Zavadskas, E. K., Turskis, Z., & Kildienė, S. (2014). State of art surveys of overviews on mcdm/madm methods [Encuestas de estado del arte de resúmenes sobre métodos mcdm/madm]. *Technological and Economic Development of Economy*, 20(1), 165-179. <https://doi.org/10.3846/20294913.2014.892037>
- Zeithaml, V. A., Bitner, M. J., & Gremler, D. D. (2018). *Services marketing: Integrating customer focus across the firm* [Marketing de servicios: Integración del enfoque en el cliente en toda la empresa]. Mc Graw Hill.

Apéndice A

Código Fuente de Limpieza y Modelos para Determinar la Polaridad de los Tuits

RunnerFiles.py

```

from tqdm import tqdm #tqdm para barra de progreso

from TweetSentimentAnalysisWithStanza import TweetSentimentAnalysisWithStanza
from TweetSentimentAnalysisWithTextBlob import TweetSentimentAnalysisWithTextBlob
from TweetSentimentAnalysisWithVader import TweetSentimentAnalysisWithVader

from TweetSentimentAnalysisWithStanzaTranslate import
TweetSentimentAnalysisWithStanzaTranslate

from TweetSentimentAnalysisWithBETO import TweetSentimentAnalysisWithBETO

from PreprocesamientoServiceV4 import PreprocesamientoServiceV4

SPREADSHEET_ID_SHEET_DATA_FINAL =
'1r5pDe2xvCb6UwB052hMTCuqC_pceNFEkoQsrpzhLhBw'

LIST_FILES_BANKS =['bancoA', 'bancoB', 'bancoC', 'bancoD']

if __name__ == "__main__":
    print("inciando")

    for bank in tqdm(LIST_FILES_BANKS):

PreprocesamientoServiceV4(SPREADSHEET_ID_SHEET_DATA_FINAL,bank).main()

PreprocesamientoServiceV4(SPREADSHEET_ID_SHEET_DATA_FINAL,bank).loadDataT
oGoogleSheet()

    for bank in LIST_FILES_BANKS:

TweetSentimentAnalysisWithStanza(SPREADSHEET_ID_SHEET_DATA_FINAL,bank).m
ain()

    TweetSentimentAnalysisWithVader(SPREADSHEET_ID_SHEET_DATA_FINAL,b
ank).main()
    TweetSentimentAnalysisWithTextBlob(SPREADSHEET_ID_SHEET_DATA_FINA
L,bank).main()

    TweetSentimentAnalysisWithStanzaTranslate(SPREADSHEET_ID_SHEET_DATA
_FINAL,bank).main()

```

```
TweetSentimentAnalysisWithBETO(SPREADSHEET_ID_SHEET_DATA_FINAL,bank).main()
```

PreprocesamientoServiceV4.py

```
# coding=utf-8
#!/usr/bin/python
import re
import pandas as pd
from cleantext import clean
from tqdm import tqdm #tqdm para barra de progreso

from ServiceUtils import ServiceUtils
from RepositoryGoogleSheet import RepositoryGoogle
from AppModel import Tweet
FOLDER_TEMP= 'temps/'
class PreprocesamientoServiceV4():
    def __init__(self, SPREADSHEET_ID_SHEET_DATA_FINAL, bank):
        print("PreprocesamientoService")
        self.serviceUtils = ServiceUtils()
        self.repositoryGoogleSheet= RepositoryGoogle()
        self.SPREADSHEET_ID_SHEET_DATA_FINAL =
        SPREADSHEET_ID_SHEET_DATA_FINAL
        self.bank = bank
    def readCSV(self,fileName):
        print('readCSV')
        df = pd.read_csv('%s%s' % (FOLDER_TEMP,fileName))
        print("tamaño", fileName, df.index)
        return df
    def remove_pattern(self, input_txt, pattern):
        r = re.findall(pattern, input_txt)
```

```

for i in r:
    input_txt = re.sub(i, "", input_txt)

return input_txt

def main(self):
    list_aspects=['cuenta','tarjeta', 'cajero','agencia','sistema', 'aplicativo', 'oficina', 'agente',
'yape','plin','app','plataforma','ventanilla','producto']

    NAME_FILE_TWEETS_RAW = self.bank

    temp_df = self.readCSV("%s.csv" % (NAME_FILE_TWEETS_RAW))

    frames = [temp_df]

    tweets_df = pd.concat(frames, keys=[NAME_FILE_TWEETS_RAW])

    print('Tamaño datos:::',len(tweets_df.index))

    print("ANALIZANDO %s" % NAME_FILE_TWEETS_RAW)

    tweets_with_polarity= []

    for index, row in tqdm(tweets_df.iterrows()):

        tweetObj = Tweet()

        tweetObj.created_at =row['Date']

        tweetObj.author_id = row['User']

        tweetObj.text = row['Tweet']

        text_clean = re.sub(r'https?:/\S+', "", row['Tweet'])

        text_clean =
clean(str(text_clean).replace("?",").replace("!",").replace(':',").replace(',').replace('(',").replace(')',').replace('/',").replace('*',").replace('\"',").replace('#',"),no_emoji=True) ##eliminacion de .
? ! y emojis

        text_clean = ".join([w for w in text_clean if not w.isdigit()]) #numbers

        text_clean = '.join([w for w in text_clean.split() if len(w)>2]) #Removing stops
Words

        text_clean = re.sub('@[A-Za-z0-9]+', "", text_clean) # Eliminación menciones

        tweetObj.text_clean = text_clean

        tweetObj.bank = str(NAME_FILE_TWEETS_RAW)

        aspec_final = ""

        for aspec in list_aspects:

```

```

        if aspect in text_clean:
            aspect_final = aspect
            break

        tweetObj.aspec= aspect_final

        tweets_with_polority.append(tweetObj.toJSON())

    FINAL_FILE_NAME_TAGS = "%s_POLARITY" % (NAME_FILE_TWEETS_RAW)

self.serviceUtils.writeFileDataJSON(tweets_with_polority,FINAL_FILE_NAME_TAGS)

def loadDataToGoogleSheet(self):
    print("loadDataToGoogleSheet")
    NAME_FILE_TWEETS_RAW = self.bank
    FINAL_FILE_NAME_TAGS = "%s_POLARITY" % (NAME_FILE_TWEETS_RAW)
    data = self.serviceUtils.openJSONFile(FINAL_FILE_NAME_TAGS)
    RANGE = "!A:H2"
    SHET_NAME = FINAL_FILE_NAME_TAGS
    DATA_TO_SHEET=[]
    for item in data:
        item_json= item
        new_item = []
        new_item.append(item_json['aspec'])
        new_item.append(item_json['author_id'])
        new_item.append(item_json['bank'])
        new_item.append(item_json['created_at'])
        new_item.append(item_json['idTweet'])
        new_item.append(item_json['text'])
        new_item.append(item_json['text_clean'])
        DATA_TO_SHEET.append(new_item)

self.repositoryGoogleSheet.updateDataOnExistedGoogleSheet(self.SPREADSHEET_ID_SHEET_DATA_FINAL,SHET_NAME, RANGE, DATA_TO_SHEET)

if __name__ == "__main__":

```

```
print("inciando")
```

```
PreprocesamientoServiceV4('1r5pDe2xvCb6UwB052hMTCuqC_pceNFEkoQsrpzhLhBw','p
rueba').main()
```

```
PreprocesamientoServiceV4('1r5pDe2xvCb6UwB052hMTCuqC_pceNFEkoQsrpzhLhBw','p
rueba').loadDataToGoogleSheet()
```

TweetSentimentAnalysisWithStanza.py

```
import json
```

```
from turtle import pen
```

```
from tqdm import tqdm #tqdm para bara de progreso
```

```
#LIBRERIA STANZA
```

```
import stanza
```

```
stanza.download('es') # download Spanish model
```

```
nlp = stanza.Pipeline(lang='es', processors='tokenize, mwt, lemma, pos,sentiment',
tokenize_no_sspllit=True)
```

```
from ServiceUtils import ServiceUtils
```

```
from RepositoryGoogleSheet import RepositoryGoogle
```

```
from AppModel import Tweet
```

```
class TweetSentimentAnalysisWithStanza:
```

```
    def __init__(self, SPREADSHEET_ID_SHEET_DATA_FINAL, bank):
```

```
        print("TweetSentimentAnalysisWithStanza %s -- %s" %
(SPREADSHEET_ID_SHEET_DATA_FINAL, bank))
```

```
        self.repositoryGoogleSheet = RepositoryGoogle()
```

```
        self.SPREADSHEET_ID_SHEET_DATA_FINAL =
SPREADSHEET_ID_SHEET_DATA_FINAL
```

```
        self.bank = bank
```

```
    def main(self):
```

```
        print("main")
```

```
        NAME_FILE_TWEETS_RAW = self.bank
```

```
        FINAL_FILE_NAME = "%s_POLARITY" % (NAME_FILE_TWEETS_RAW)
```

```
        RANGE = "!A:G2"
```

```

list_tweets_raw =
self.repositoryGoogleSheet.ReadDataFromSheet(self.SPREADSHEET_ID_SHEET_DATA_
FINAL,FINAL_FILE_NAME,RANGE)

list_tweets_with_sentiment = []

for item in tqdm(list_tweets_raw):

    tweet = Tweet()

    tweet.text = item[5]

    if len(item) > 6:

        tweet.text_clean = item[6]

    else:

        tweet.text_clean = ""

    tweet.score_polarity_stanza = self.sentiment_scores(tweet.text_clean)

    list_tweets_with_sentiment.append(tweet.toJSON())

self.writeDataOnGoogleSheet(list_tweets_with_sentiment, FINAL_FILE_NAME)

def sentiment_scores(self,sentence):

    if sentence != "":

        doc = nlp(sentence) # run annotation over a sentence

        score_by_tweet= -1

        for i, sentence in enumerate(doc.sentences):

            score_by_tweet = sentence.sentiment

            #print("%d -> %d" % (i, sentence))

        if score_by_tweet == 2:

            return "Positive"

        elif score_by_tweet == 0:

            return "Negative"

        else :

            return "Neutral"

    else:

        return 'Neutral'

def writeDataOnGoogleSheet(self, data,NAME_SHEET):

```

```

print("writeDataOnGoogleSheet")
RANGE = "!H:H2"
DATA_TO_SHEET=[]
for item in data:
    item_json= json.loads(item)
    new_item = []
    new_item.append(item_json['score_polarity_stanza'])
    DATA_TO_SHEET.append(new_item)

self.repositoryGoogleSheet.updateDataOnExistedGoogleSheet(self.SPREADSHEET_ID_SHEET_DATA_FINAL,NAME_SHEET, RANGE, DATA_TO_SHEET)
if __name__ == "__main__":
    print("inciando")

TweetSentimentAnalysisWithStanza('1r5pDe2xvCb6UwB052hMTCuqC_pceNFEkoQsrpzhLhBw','prueba').main()
TweetSentimentAnalysisWithVader.py
import json
from turtle import pen
from tqdm import tqdm #tqdm para bara de progreso
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
from ServiceUtils import ServiceUtils
from RepositoryGoogleSheet import RepositoryGoogle
from AppModel import Tweet
class TweetSentimentAnalysisWithVader:
    def __init__(self,SPREADSHEET_ID_SHEET_DATA_FINAL, bank):
        print("TweetSentimentAnalysisWithVader %s -- %s" %
(SPREADSHEET_ID_SHEET_DATA_FINAL, bank))
        self.repositoryGoogleSheet = RepositoryGoogle()
        self.SPREADSHEET_ID_SHEET_DATA_FINAL =
SPREADSHEET_ID_SHEET_DATA_FINAL
        self.bank = bank

```

```

def main(self):

    print("TweetSentimentAnalysisWithVader")

    NAME_FILE_TWEETS_RAW = self.bank

    FINAL_FILE_NAME = "%s_POLARITY" % (NAME_FILE_TWEETS_RAW)

    RANGE = "!A:G2"

    list_tweets_raw =
self.repositoryGoogleSheet.ReadDataFromSheet(self.SPREADSHEET_ID_SHEET_DATA_
FINAL,FINAL_FILE_NAME,RANGE)

    list_tweets_with_sentiment = []

    for item in tqdm(list_tweets_raw):

        tweet = Tweet()

        tweet.text = item[5]

        if len(item) > 6:

            tweet.text_clean = item[6]

        else:

            tweet.text_clean = "

        tweet.score_polarity_vader = self.sentiment_scores(tweet.text_clean)

        list_tweets_with_sentiment.append(tweet.toJSON())

    self.writeDataOnGoogleSheet(list_tweets_with_sentiment, FINAL_FILE_NAME)

def sentiment_scores(self,sentence):

    print(sentence)

    if sentence !="":

        try:

            sid_obj = SentimentIntensityAnalyzer()

            sentiment_dict = sid_obj.polarity_scores(sentence)

            if sentiment_dict['compound'] >= 0.05 :

                return "Positive"

            elif sentiment_dict['compound'] <= - 0.05 :

                return "Negative"

        else :

```

```

        return "Neutral"
    except:
        return 'Neutral'
    else:
        return 'Neutral'

def writeDataOnGoogleSheet(self, data,NAME_SHEET):
    print("writeDataOnGoogleSheet")
    RANGE = "!!I:I2"
    DATA_TO_SHEET=[]
    for item in data:
        item_json= json.loads(item)
        new_item = []
        new_item.append(item_json['score_polarity_vader'])
        DATA_TO_SHEET.append(new_item)

self.repositoryGoogleSheet.updateDataOnExistedGoogleSheet(self.SPREADSHEET_ID_SHEET_DATA_FINAL,NAME_SHEET, RANGE, DATA_TO_SHEET)

if __name__ == "__main__":
    print("inciando")

```

```

TweetSentimentAnalysisWithVader('1r5pDe2xvCb6UwB052hMTCuqC_pceNFEkoQsrpzhLhBw','prueba').main()

```

TweetSentimentAnalysisWithTextBlob.py

```

import json
from turtle import pen
from tqdm import tqdm #tqdm para bara de progreso
from ast import AnnAssign
from textblob import TextBlob
import textblob.exceptions
from ServiceUtils import ServiceUtils
from RepositoryGoogleSheet import RepositoryGoogle

```

```

from AppModel import Tweet

class TweetSentimentAnalysisWithTextBlob:

    def __init__(self,SPREADSHEET_ID_SHEET_DATA_FINAL, bank):

        print("TweetSentimentAnalysisWithTextBlob %s -- %s" %
              (SPREADSHEET_ID_SHEET_DATA_FINAL, bank))

        self.repositoryGoogleSheet = RepositoryGoogle()

        self.SPREADSHEET_ID_SHEET_DATA_FINAL =
        SPREADSHEET_ID_SHEET_DATA_FINAL

        self.bank = bank

    def main(self):

        print("main")

        NAME_FILE_TWEETS_RAW = self.bank

        FINAL_FILE_NAME = "%s_POLARITY" % (NAME_FILE_TWEETS_RAW)

        RANGE = "!A:G2"

        list_tweets_raw =
        self.repositoryGoogleSheet.ReadDataFromSheet(self.SPREADSHEET_ID_SHEET_DATA_
        FINAL,FINAL_FILE_NAME,RANGE)

        list_tweets_with_sentiment = []

        for item in tqdm(list_tweets_raw):

            tweet = Tweet()

            tweet.text = item[5]

            if len(item) > 6:

                tweet.text_clean = item[6]

            else:

                tweet.text_clean = "

            tweet.score_polarity_textblob = self.sentiment_scores(tweet.text_clean)

            list_tweets_with_sentiment.append(tweet.toJSON())

        self.writeDataOnGoogleSheet(list_tweets_with_sentiment, FINAL_FILE_NAME)

    def sentiment_scores(self,sentence):

        print("t--",len(sentence))

        if sentence !=" and len(sentence)>2:

```

```

print(sentence)
analysis = TextBlob(sentence)
try:
    analysis_ready = analysis.translate(from_lang="es", to='en')
    print(analysis_ready)
    if analysis_ready.sentiment.polarity > 0:
        return 'Positive'
    elif analysis_ready.sentiment.polarity == 0:
        return 'Neutral'
    else:
        return 'Negative'
except:
    return 'Neutral'
else:
    return 'Neutral'
def writeDataOnGoogleSheet(self, data,NAME_SHEET):
    print("writeDataOnGoogleSheet")
    RANGE = "!J:J2"
    DATA_TO_SHEET=[]
    for item in data:
        item_json= json.loads(item)
        new_item = []
        new_item.append(item_json['score_polarity_textblob'])
        DATA_TO_SHEET.append(new_item)

self.repositoryGoogleSheet.updateDataOnExistedGoogleSheet(self.SPREADSHEET_ID_SHEET_DATA_FINAL,NAME_SHEET, RANGE, DATA_TO_SHEET)
if __name__ == "__main__":
    print("inciando")

```

```
TweetSentimentAnalysisWithTextBlob('1r5pDe2xvCb6UwB052hMTCuqC_pceNFEkoQsrp
zhLhBw','prueba').main()
```

```
TweetSentimentAnalysisWithVader.py
```

```
import json
```

```
from turtle import pen
```

```
from tqdm import tqdm #tqdm para bara de progreso
```

```
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
```

```
from ServiceUtils import ServiceUtils
```

```
from RepositoryGoogleSheet import RepositoryGoogle
```

```
from AppModel import Tweet
```

```
class TweetSentimentAnalysisWithVader:
```

```
    def __init__(self,SPREADSHEET_ID_SHEET_DATA_FINAL, bank):
```

```
        print("TweetSentimentAnalysisWithVader %s -- %s" %
(SPREADSHEET_ID_SHEET_DATA_FINAL, bank))
```

```
        self.repositoryGoogleSheet = RepositoryGoogle()
```

```
        self.SPREADSHEET_ID_SHEET_DATA_FINAL =
SPREADSHEET_ID_SHEET_DATA_FINAL
```

```
        self.bank = bank
```

```
    def main(self):
```

```
        print("TweetSentimentAnalysisWithVader")
```

```
        NAME_FILE_TWEETS_RAW = self.bank
```

```
        FINAL_FILE_NAME = "%s_POLARITY" % (NAME_FILE_TWEETS_RAW)
```

```
        RANGE = "!A:G2"
```

```
        list_tweets_raw =
self.repositoryGoogleSheet.ReadDataFromSheet(self.SPREADSHEET_ID_SHEET_DATA_
FINAL,FINAL_FILE_NAME,RANGE)
```

```
        list_tweets_with_sentiment = []
```

```
        for item in tqdm(list_tweets_raw):
```

```
            tweet = Tweet()
```

```
            tweet.text = item[5]
```

```
            if len(item) > 6:
```

```

        tweet.text_clean = item[6]
    else:
        tweet.text_clean = ""
    tweet.score_polarity_vader = self.sentiment_scores(tweet.text_clean)
    list_tweets_with_sentiment.append(tweet.toJSON())
self.writeDataOnGoogleSheet(list_tweets_with_sentiment, FINAL_FILE_NAME)
def sentiment_scores(self,sentence):
    print(sentence)
    if sentence != "":
        try:
            sid_obj = SentimentIntensityAnalyzer()
            sentiment_dict = sid_obj.polarity_scores(sentence)
            if sentiment_dict['compound'] >= 0.05 :
                return "Positive"
            elif sentiment_dict['compound'] <= - 0.05 :
                return "Negative"
            else :
                return "Neutral"
        except:
            return 'Neutral'
    else:
        return 'Neutral'
def writeDataOnGoogleSheet(self, data,NAME_SHEET):
    print("writeDataOnGoogleSheet")
    RANGE = "!I:I2"
    DATA_TO_SHEET=[]
    for item in data:
        item_json= json.loads(item)
        new_item = []

```

```
new_item.append(item_json['score_polarity_vader'])  
DATA_TO_SHEET.append(new_item)  
  
self.repositoryGoogleSheet.updateDataOnExistedGoogleSheet(self.SPREADSHEET_ID_SHEET_DATA_FINAL,NAME_SHEET, RANGE, DATA_TO_SHEET)  
  
if __name__ == "__main__":  
    print("inciando")  
  
TweetSentimentAnalysisWithVader('1r5pDe2xvCb6UwB052hMTCuqC_pceNFEkoQsrpzhLhBw','prueba').main()
```

