

Ruso Alexander Morales Gonzales

MORALES GONZALES, RUSSO ALEXANDER-ALGORITMOS DE BUSINESS INTELLIGENCE EN LA ANALÍTICA DE DATOS EMO...

 Quick Submit Quick Submit Universidad Politécnica del Perú

Detalles del documento

Identificador de la entrega

trn:oid:::1:3165237659

Fecha de entrega

24 feb 2025, 8:10 p.m. GMT-5

Fecha de descarga

24 feb 2025, 8:19 p.m. GMT-5

Nombre de archivo

E_EN_LA_ANAL_TICA_DE_DATOS_EMOCIONALES_DE_TWEETS_PUBLICADOS.docx

Tamaño de archivo

1.9 MB

187 Páginas

85,180 Palabras

261,784 Caracteres

3% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Texto citado
- Coincidencias menores (menos de 20 palabras)

Fuentes principales

- 2%  Fuentes de Internet
- 0%  Publicaciones
- 2%  Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

Fuentes principales

- 2% Fuentes de Internet
- 0% Publicaciones
- 2% Trabajos entregados (trabajos del estudiante)

Fuentes principales

Las fuentes con el mayor número de coincidencias dentro de la entrega. Las fuentes superpuestas no se mostrarán.

1	Trabajos del estudiante Universidad Alas Peruanas	1%
2	Internet hdl.handle.net	<1%
3	Internet renati.sunedu.gob.pe	<1%
4	Internet repositorio.uap.edu.pe	<1%
5	Internet repositorio.unu.edu.pe	<1%
6	Internet pt.scribd.com	<1%
7	Trabajos del estudiante Universidad Politécnica del Perú	<1%
8	Trabajos del estudiante Corporación Universitaria Minuto de Dios, UNIMINUTO	<1%
9	Trabajos del estudiante Universidad TecMilenio	<1%
10	Internet openaccess.uoc.edu	<1%
11	Internet www.roforum.net	<1%

12	Trabajos del estudiante	Universidad Andina Nestor Caceres Velasquez	<1%
13	Internet	repositorio.undac.edu.pe	<1%
14	Trabajos del estudiante	Universidad de Navarra	<1%
15	Trabajos del estudiante	Hogeschool Utrecht - Tii	<1%
16	Internet	repositorio.unc.edu.pe	<1%



**UNIVERSIDAD ALAS PERUANAS
VICERRECTORADO ACADÉMICO
ESCUELA DE POSGRADO**

**ALGORITMOS DE BUSINESS INTELLIGENCE EN LA
ANALÍTICA DE DATOS EMOCIONALES DE TWEETS
PUBLICADOS POR CLIENTES DE MCDONALD'S Y KFC, LIMA
2022**

**PARA OPTAR EL GRADO ACADÉMICO DE:
DOCTOR EN INGENIERÍA DE SISTEMAS**

**PRESENTADO POR:
Mg. RUSO ALEXANDER MORALES GONZALES
CÓDIGO ORCID: <https://orcid.org/0000-0003-4077-7271>**

**Asesor : Dr. ELMER MARCIAL LIMACHE SANDOVAL
Codigo orcid: <https://orcid.org/0000-0003-4852-1916>**

**LIMA – PERÚ
2025**

DEDICATORIA

Con mucho amor esta tesis está dedicada a mi familia conformada por mi padre Dante, y mi madre María Antonieta, y también a la familia de mi hermano Dante, su esposa María y su hija Daneth.

AGRADECIMIENTO

Estoy profundamente agradecido al Dr, Limache, mi asesor de tesis, y al Dr. Jorge Luis Bringas Salvador mi revisor de tesis, ambos han sido seres humanos valiosos para este trabajo.

1

RECONOCIMIENTO

Doy un reconocimiento a toda la comunidad de cibernautas de la especialidad que con sus aportes en foros y sitios web ayudan expandiendo el conocimiento de las ciencias informáticas.

ÍNDICE

CARÁTULA.....	i
DEDICATORIA	iii
AGRADECIMIENTO	iv
RECONOCIMIENTO	v
RESUMEN	xi
ABSTRACT.....	xii
RESUMO.....	xiii
INTRODUCCIÓN.....	xiv
CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA	16
1.1 DESCRIPCIÓN DE LA REALIDAD PROBLEMÁTICA	16
1.2 DELIMITACIÓN DE LA INVESTIGACIÓN.....	20
1.2.1 DELIMITACIÓN ESPACIAL.....	20
1.2.2 DELIMITACIÓN SOCIAL	20
1.2.3 DELIMITACIÓN TEMPORAL	21
1.2.4 DELIMITACIÓN CONCEPTUAL	22
1.3 PROBLEMAS DE INVESTIGACIÓN	22
1.3.1 PROBLEMA PRINCIPAL.....	22
1.3.2 PROBLEMAS ESPECÍFICOS	22
1.4 OBJETIVOS DE LA INVESTIGACIÓN	23
1.4.1 OBJETIVO GENERAL	23
1.4.2 OBJETIVOS ESPECÍFICOS	23
1.5 JUSTIFICACIÓN E IMPORTANCIA DE LA INVESTIGACIÓN	23

1.5.1 JUSTIFICACIÓN.....	23
1.5.2 IMPORTANCIA	26
1.6 FACTIBILIDAD DE LA INVESTIGACIÓN.....	27
1.7 LIMITACIONES DEL ESTUDIO	27
CAPÍTULO II: MARCO FILOSÓFICO	29
2.1 FUNDAMENTACIÓN ONTOLÓGICA	29
2.1.1 LA NATURALEZA ONTOLÓGICA DE LOS DATOS EN BUSINESS INTELLIGENCE	30
2.1.2 ONTOLOGÍA DE LA ANALÍTICA DE DATOS EMOCIONALES..	32
2.2 FUNDAMENTACIÓN EPISTEMOLÓGICA	34
2.2.1 EPISTEMOLOGÍA DE LOS ALGORITMOS DE BUSINESS INTELLIGENCE	35
2.2.2 EPISTEMOLOGÍA DE LA ANALÍTICA DE DATOS EMOCIONALES	37
CAPÍTULO III: MARCO TEÓRICO CONCEPTUAL	40
3.1 ANTECEDENTES DEL PROBLEMA.....	40
3.2 BASES TEÓRICAS O CIENTÍFICAS	50
3.2.1 INTRODUCCIÓN.....	50
3.2.2 ALGORITMOS DE BUSINESS INTELLIGENCE	51
3.2.2.1 DIMENSIONES DEL BUSINESS INTELLIGENCE	52
3.2.3 ANALÍTICA DE DATOS EMOCIONALES	55
3.2.3.1 DIMENSIONES DE LA ANALÍTICA DE DATOS EMOCIONALES	57
3.3 DEFINICIÓN DE TÉRMINOS BÁSICOS	59
CAPÍTULO IV: HIPÓTESIS Y VARIABLES.....	62

4

4.1 HIPÓTESIS GENERAL.....	62
4.2 HIPÓTESIS ESPECÍFICA.....	62
4.3 DEFINICIÓN CONCEPTUAL Y OPERACIONAL DE LAS VARIABLES..	63
4.4 CUADRO DE OPERACIONALIZACIÓN DE VARIABLES.....	65
CAPÍTULO V: METODOLOGÍA DE LA INVESTIGACIÓN.....	66
5.1 ENFOQUE DE INVESTIGACIÓN.....	66
5.2 TIPO Y NIVEL DE INVESTIGACIÓN.....	66
5.2.1 TIPO DE INVESTIGACIÓN.....	66
5.2.2 NIVEL DE INVESTIGACIÓN.....	67
5.3 MÉTODOS Y DISEÑO DE INVESTIGACIÓN.....	67
5.3.1 MÉTODOS DE INVESTIGACIÓN.....	67
5.3.2 DISEÑO DE LA INVESTIGACIÓN.....	67
5.4 POBLACIÓN Y MUESTRA DE LA INVESTIGACIÓN.....	87
5.4.1 POBLACIÓN.....	87
5.4.2 MUESTRA.....	88
5.5 TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS.....	90
5.5.1 TÉCNICAS.....	90
5.5.2 INSTRUMENTOS.....	90
5.6 VALIDEZ Y CONFIABILIDAD.....	90
5.7 PROCESAMIENTO Y ANÁLISIS DE DATOS.....	93
5.8 ÉTICA EN LA INVESTIGACIÓN.....	95
CAPÍTULO VI: RESULTADOS.....	97
6.1 ANÁLISIS DESCRIPTIVO.....	97
6.1.1 RESULTADOS DE LOS ALGORITMOS.....	97

vii

6.1.2 RESULTADOS DE LA VARIABLE DEPENDIENTE.....	100
6.1.3 RESULTADOS DE LA VARIABLE INDEPENDIENTE.....	102
6.2 ANÁLISIS INFERENCIAL.....	104
6.2.1 PRUEBA DE NORMALIDAD.....	104
6.2.2 PRUEBA DE HIPÓTESIS.....	106
6.2.2.1 PRUEBA DE HIPÓTESIS GENERAL.....	106
6.2.2.2 PRUEBA DE HIPÓTESIS ESPECÍFICAS.....	109
CAPÍTULO VII: DISCUSIÓN DE LOS RESULTADOS.....	116
CONCLUSIONES.....	122
RECOMENDACIONES.....	123
FUENTES DE INFORMACIÓN.....	125
ANEXO 1. MATRIZ DE CONSISTENCIA.....	134
ANEXO 2. INSTRUMENTOS DE RECOLECCIÓN DE DATOS.....	135
ANEXO 3. FICHAS DE VALIDACIÓN DEL INSTRUMENTO.....	139
ANEXO 4. COPIA DE LOS DATOS PROCESADOS.....	149
ANEXO 5. AUTORIZACIÓN DE LA ENTIDAD.....	185
ANEXO 6. CONSENTIMIENTO INFORMADO.....	187
ANEXO 7. DECLARATORIA DE AUTENTICIDAD DE LA TESIS.....	188

ÍNDICE DE TABLAS

Tabla 1. Cuadro de operacionalización de variables	65
Tabla 2. Validez de los instrumentos, según los expertos consultados	90
Tabla 3. Valoración de la confiabilidad según el coeficiente KR ₂₀	91
Tabla 4. Valoración de la confiabilidad de los instrumentos.....	92
Tabla 5. Métricas de evaluación de los algoritmos.....	98
Tabla 6. Resultados de la pre-prueba y pos-prueba de la variable dependiente ...	100
Tabla 7. Resultados de la pre-prueba y pos-prueba de la variable independiente	102
Tabla 8. Resultados de la prueba de hipótesis general.....	109
Tabla 9. Resultados de la prueba de hipótesis específica 1	112
Tabla 10. Resultados de la prueba de hipótesis específica 2	114
Tabla 11. Resultados de las pruebas de hipótesis específicas.....	115

ÍNDICE DE FIGURAS

Figura 1. Fórmula matemática de la función sigmoide del algoritmo de regresión logística.....	53
Figura 2. Diseño experimental, con pre y pos prueba y grupo control.....	67
Figura 3. Diagrama de bloques del sistema	72
Figura 4. Fórmula del tamaño de la muestra	89
Figura 5. Fórmula 20	91
Figura 6. Figura de los datos de la tabla 6	101
Figura 7. Figura de los datos de la tabla 7	103

RESUMEN

Esta investigación empleó un enfoque cuantitativo, y su tipología fue aplicada, se utilizó el método de investigación hipotético-deductivo. El objetivo general consistió en determinar la influencia de los algoritmos más usados en business intelligence en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, en Lima en el último trimestre del año 2022. El nivel de investigación fue descriptivo-explicativo. Se adoptó un diseño experimental puro que incluyó prepruebas y pospruebas, así como un grupo de control. Se recopilaron tweets aleatorios que mencionaban a McDonald's y KFC en Lima durante el periodo mencionado, con una población total de 270,058 tweets. La muestra, obtenida mediante muestreo aleatorio simple, consistió en 384 tweets y se dividió en dos grupos: experimental y control. Para el análisis, se aplicaron distintos algoritmos en cada grupo. El grupo experimental se analizó utilizando algoritmos de business intelligence, específicamente regresión logística y árboles de decisión, mientras que al grupo control se le aplicó un algoritmo de Deep Learning llamado BERT. Los resultados obtenidos se compararon para evaluar la eficacia de los distintos algoritmos. Se diseñaron dos fichas de observación con preguntas cerradas de tipo dicotómico relacionadas con las variables de investigación. Los datos fueron analizados utilizando técnicas estadísticas descriptivas e inferenciales. La prueba de *McNemar* se utilizó para la prueba de hipótesis, y el análisis de datos se realizó con el lenguaje de programación Python. Las etapas del análisis incluyeron recolección de datos, definición de grupos, fase de pretest, fase de postest, evaluación y comparación, y conclusiones finales. La variable dependiente fue evaluada en dos etapas: pretest y postest. En el postest ambos grupos (experimental y control) mostraron una mejora en la cantidad de resultados. Sin embargo, el grupo experimental mostró una mejora significativamente mayor en comparación con el grupo control (Deep Learning). Los resultados iniciales indicaron que los algoritmos de business intelligence podrían ser una alternativa viable y efectiva para la analítica de datos emocionales en los mensajes de tweets. Las hipótesis específicas fueron evaluadas, con la hipótesis 1 (H_1) obteniendo un p-valor de $8.16e-20$ y la hipótesis 2 (H_2) un p-valor de $8.63e-22$. La prueba de *McNemar* proporcionó evidencia suficiente para rechazar las hipótesis nulas específicas, lo que llevó a la aceptación de ambas hipótesis específicas. La hipótesis general también fue aceptada, con un p-valor de $1.51e-37$ en comparación con el valor de significancia estándar de 0.05. Esto confirma que los algoritmos de business intelligence influyen en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

Palabras claves: Analítica de datos emocionales, Aprendizaje automático, Aprendizaje profundo, Inteligencia de Negocios.

X

ABSTRACT

This research used a quantitative approach, and its typology was applied, the hypothetical-deductive research method was used. The general objective was to determine the influence of the most used algorithms in business intelligence in the analysis of emotional data of tweets published by McDonald's and KFC customers in Lima in the last quarter of 2022. The level of research was descriptive-explanatory. A pure experimental design was adopted that included pre-tests and post-tests, as well as a control group. Random tweets mentioning McDonald's and KFC in Lima were collected during the mentioned period, with a total population of 270,058 tweets. The sample, obtained by simple random sampling, consisted of 384 tweets and was divided into two groups: experimental and control. For the analysis, different algorithms were applied in each group. The experimental group was analyzed using business intelligence algorithms, specifically logistic regression and decision trees, while the control group was applied a Deep Learning algorithm called BERT. The results obtained were compared to evaluate the effectiveness of the different algorithms. Two observation sheets were designed with closed dichotomous questions related to the research variables. The data were analyzed using descriptive and inferential statistical techniques. The *McNemar* test was used for hypothesis testing, and data analysis was performed using the Python programming language. The stages of the analysis included data collection, group definition, pretest phase, posttest phase, evaluation and comparison, and final conclusions. The dependent variable was evaluated in two stages: pretest and posttest. In the posttest, both groups (experimental and control) showed an improvement in the number of results. However, the experimental group showed a significantly greater improvement compared to the control group (Deep Learning). Initial results indicated that business intelligence algorithms could be a viable and effective alternative for emotional data analytics on tweets. Specific hypotheses were tested, with hypothesis 1 (H_1) obtaining a p-value of $8.16e-20$ and hypothesis 2 (H_2) a p-value of $8.63e-22$. *McNemar's* test provided sufficient evidence to reject the specific null hypotheses, leading to the acceptance of both specific hypotheses. The general hypothesis was also accepted, with a p-value of $1.51e-37$ compared to the standard significance value of 0.05. This confirms that business intelligence algorithms influence the emotional data analytics of tweets posted by McDonald's and KFC customers, Lima 2022.

Keywords: Emotional data analytics, Machine learning, Deep learning, Business intelligence.

RESUMO

Esta pesquisa utilizou uma abordagem quantitativa, e como tipologia foi aplicado o método de pesquisa hipotético-dedutivo. O objetivo geral foi determinar a influência dos algoritmos mais utilizados em inteligência de negócios na análise de dados emocionais de tweets publicados por clientes do McDonald's e KFC em Lima no último trimestre de 2022. O nível de pesquisa foi descritivo.-explicativo. Foi adotado um delineamento experimental puro que incluiu pré-testes e pós-testes, bem como um grupo de controle. Tweets aleatórios mencionando McDonald's e KFC em Lima foram coletados durante o período mencionado, com uma população total de 270.058 tweets. A amostra, obtida por amostragem aleatória simples, foi composta por 384 tweets e foi dividida em dois grupos: experimental e controle. Para a análise, diferentes algoritmos foram aplicados a cada grupo. O grupo experimental foi analisado usando algoritmos de inteligência de negócios, especificamente regressão logística e árvores de decisão, enquanto o grupo de controle foi submetido a um algoritmo de Deep Learning chamado BERT. Os resultados obtidos foram comparados para avaliar a eficácia dos diferentes algoritmos. Foram elaborados dois formulários de observação com questões dicotômicas fechadas relacionadas às variáveis da pesquisa. Os dados foram analisados por meio de técnicas estatísticas descritivas e inferenciais. O teste de *McNemar* foi usado para testes de hipóteses, e a análise de dados foi realizada usando a linguagem de programação Python. As etapas de análise incluíram coleta de dados, definição do grupo, fase de pré-teste, fase de pós-teste, avaliação e comparação e conclusões finais. A variável dependente foi avaliada em duas etapas: pré-teste e pós-teste. No pós-teste, ambos os grupos (experimental e controle) apresentaram melhora no número de resultados. Entretanto, o grupo experimental apresentou melhora significativamente maior em comparação ao grupo controle (Deep Learning). Os resultados iniciais indicaram que algoritmos de inteligência de negócios podem ser uma alternativa viável e eficaz para analisar dados emocionais em mensagens de tweets. As hipóteses específicas foram avaliadas, com a hipótese 1 (H_1) obtendo um valor de p de $8,16e-20$ e a hipótese 2 (H_2) um valor de p de $8,63e-22$. O teste de *McNemar* forneceu evidências suficientes para rejeitar as hipóteses nulas específicas, levando à aceitação de ambas as hipóteses específicas. A hipótese geral também foi aceita, com um valor de p de $1,51e-37$ comparado ao valor de significância padrão de 0,05. Isso confirma que algoritmos de inteligência de negócios influenciam a análise de dados emocionais de tweets postados por clientes do McDonald's e KFC, Lima 2022.

Palavras-chave: Análise de dados emocionais, Aprendizado de máquina, Aprendizado profundo, Inteligência de negócios.

INTRODUCCIÓN

El contexto actual está marcado por una interacción constante en las redes sociales digitales, donde las empresas buscan comprender las percepciones de los usuarios para optimizar sus estrategias y reforzar su posicionamiento en el mercado. En este escenario, el presente estudio tiene como objetivo evaluar el uso de algoritmos de inteligencia de negocios (del inglés *business intelligence* o simplemente BI) para analizar las percepciones de los usuarios hacia McDonald's y KFC en Lima, a partir de datos emocionales extraídos de los mensajes en Twitter durante el año 2022. La investigación sigue un enfoque cuantitativo, de tipo descriptivo-explicativo, con la finalidad de determinar la influencia de los algoritmos de BI en la analítica de datos emocionales obtenidos de los tweets de los clientes de McDonald's y KFC en Lima, durante el año mencionado. Para ello, se adoptó un diseño experimental puro, con pretest y posttest, en el que se establecieron dos grupos: uno experimental, en el que se analizaron los tweets utilizando algoritmos de BI (regresión logística y árboles de decisión), y otro de control activo, en el que se empleó un algoritmo de Deep Learning.

Los datos fueron recolectados mediante un muestreo aleatorio simple, obteniendo una muestra de 384 tweets, los cuales fueron analizados utilizando técnicas estadísticas descriptivas e inferenciales. Se incluyó el test de *McNemar* para la prueba de hipótesis. Los resultados finales revelaron que tanto los algoritmos de BI como los de Deep Learning son herramientas eficaces para analizar el contexto emocional en los mensajes de los usuarios en redes sociales. Sin embargo, tras la optimización del modelo, el grupo experimental que utilizó los algoritmos de BI mostró una mejora significativa en la detección y clasificación de emociones en los tweets analizados. Esto sugiere que los algoritmos de BI podrían constituir una alternativa viable y efectiva para el análisis de datos emocionales en Twitter, actualmente llamado X, e incluso superar el rendimiento de los algoritmos de Deep Learning.

Este estudio proporciona evidencia relevante sobre la utilidad de los algoritmos de BI para comprender las percepciones de los usuarios en redes sociales, lo cual resulta de gran interés para las empresas que buscan perfeccionar sus estrategias de marketing y fortalecer su presencia en el mercado. La investigación se organiza en siete capítulos que detallan la metodología, los resultados y las conclusiones de forma sistemática.

- **Capítulo I: Planteamiento del problema:** Introduce el problema de investigación, describe la realidad problemática y expone los objetivos y la justificación del estudio.
- **Capítulo II: Marco filosófico:** Establece el marco filosófico que sustenta la investigación, abordando aspectos ontológicos, epistemológicos y axiológicos.
- **Capítulo III: Marco teórico conceptual:** Presenta los antecedentes, las bases teóricas y científicas del estudio, así como la definición de los términos clave.
- **Capítulo IV: Hipótesis y variables:** Formula las hipótesis generales y específicas de la investigación, y define las variables tanto conceptualmente como operacionalmente.
- **Capítulo V: Metodología de la investigación:** Describe el tipo y nivel de investigación, los métodos y el diseño investigativo empleado a lo largo del desarrollo del trabajo de campo, también se cuenta con un apartado sobre la población y muestra del estudio, las técnicas de recolección de datos, así como el procesamiento y análisis de ellos.
- **Capítulo VI: Resultados:** Presenta los resultados del estudio, incluyendo el análisis descriptivo e inferencial de las variables dependiente e independiente.
- **Capítulo VII: Discusión de los resultados:** Discute los resultados obtenidos, los vincula con el marco teórico y presenta las implicaciones de los hallazgos.
- **Conclusiones:** Manifiesta las conclusiones derivadas de la investigación.
- **Recomendaciones:** Propone recomendaciones tanto para futuras investigaciones como para la aplicación práctica de los resultados en el ámbito empresarial.
- **Fuentes de información:** Lista las fuentes consultadas para la elaboración del estudio.
- **Anexos:** Incluye los instrumentos de recolección de datos, las fichas de validación del instrumento, los datos procesados, la autorización de la entidad y la declaración de autenticidad de la tesis.

CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA

1.1 DESCRIPCIÓN DE LA REALIDAD PROBLEMÁTICA

En el contexto global contemporáneo, la digitalización ha provocado una transformación profunda en la interacción entre las empresas y los consumidores, así como en las metodologías utilizadas para la recopilación, procesamiento y análisis de datos. Las redes sociales, particularmente Twitter (ahora conocido como la red social digital X), se han consolidado como plataformas fundamentales donde los usuarios comparten sus opiniones, ideas y emociones en tiempo real, lo que genera un flujo continuo de datos emocionales. Este volumen masivo de información representa una oportunidad invaluable para las empresas, permitiéndoles entender y predecir los comportamientos de consumo, lo que a su vez facilita la toma de decisiones estratégicas fundamentadas. A nivel global, los algoritmos más ampliamente utilizados y estudiados en el campo de la **Inteligencia de Negocios (BI)**, como la **regresión logística** y los **árboles de decisión**, han sido herramientas clave en el análisis de diversos aspectos empresariales, tales como la segmentación de clientes, la evaluación de rentabilidad y la predicción de resultados financieros. No obstante, con los avances en el campo del **Deep Learning** (traducido como aprendizaje profundo), especialmente tecnologías como las **Redes Neuronales Recurrentes (RNN)**, las **Redes Neuronales Convolucionales (CNN)** y BERT (Bidirectional Encoder Representations from Transformers), inicialmente aplicadas en

áreas como el análisis de secuencias de datos y el procesamiento de imágenes, han mostrado un gran potencial también en el ámbito empresarial. Estos métodos, a través del **procesamiento de lenguaje natural** (PLN), han facilitado avances en áreas como la clasificación de documentos, la predicción de demanda y ventas, y la evaluación de fluctuaciones en los mercados financieros. En América Latina, las redes sociales, especialmente la red **X**, desempeñan un papel crucial en la comunicación y el marketing. La presente investigación toma como caso de estudio a dos marcas globales de comida rápida, **McDonald's** y **KFC**, cuyos clientes en Lima, Perú, dejan mensajes diarios en redes sociales que abarcan una gran diversidad lingüística y cultural, lo cual presenta desafíos únicos en la analítica de datos emocionales. A nivel técnico, los algoritmos tradicionales empleados en **Machine Learning** (ML) y heredados por **Inteligencia de Negocios (BI)** han sido principalmente enfocados en la predicción del comportamiento del consumidor, el análisis de riesgos y la segmentación de mercados. Sin embargo, estos algoritmos no han sido diseñados específicamente para tareas relacionadas al análisis de sentimientos o, en términos más concretos, para la analítica de datos emocionales presentes en los textos; por tanto los algoritmos de BI quedan cortos en capacidad para analizar el lenguaje natural, por consiguiente algoritmos específicos como la regresión logística y los árboles de decisiones están limitados en su aplicabilidad en el **procesamiento de lenguaje natural humano**, lo que dificulta a su vez, su uso en el análisis de emociones en redes sociales (o llamado también análisis de sentimientos). Por otro lado, las técnicas avanzadas de **Deep Learning** ofrecen una capacidad superior para abordar el análisis de sentimientos, un campo donde amerita volverlo a recalcarlo, las metodologías de **BI** aún presentan limitaciones. La creciente adopción de algoritmos avanzados para la identificación y análisis de opiniones y emociones expresadas en los textos está ganando relevancia, particularmente en el contexto de las redes sociales digitales. En Lima, las interacciones en redes sociales entre consumidores y marcas representan una fuente rica de datos emocionales. En los casos de estudio de McDonald's y KFC, los comentarios de los clientes no sólo reflejan opiniones sobre los productos, sino inclusive también sobre el servicio al cliente, la calidad percibida, la relación calidad-precio, preocupaciones sobre la salud y la nutrición, preferencias alimentarias y opiniones sobre nuevas tendencias en el menú. Sin embargo, estas interacciones presentan desafíos significativos para procesar eficientemente tanto el volumen como la diversidad de los

datos. Volviendo a lo técnico los algoritmos tradicionales de **Machine Learning** utilizados en **BI**, aunque útiles por su simplicidad y rapidez, no son capaces de capturar la complejidad y la sutileza de las emociones humanas en documentos escritos, especialmente cuando los textos contienen sarcasmo, ironía o están enmarcados dentro de contextos culturales específicos. En contraposición, las técnicas de **Deep Learning** tienen un gran potencial para analizar datos no estructurados, como los textos libres; no obstante, estas metodologías enfrentan desafíos debido a sus elevados requerimientos técnicos, alta complejidad y largos tiempos de entrenamiento. Un ejemplo claro de la necesidad de analizar los datos emocionales de las interacciones en redes sociales es el caso de **McDonald's**, donde los clientes a menudo expresan insatisfacción con la inconsistencia en la calidad de productos y servicios entre distintas sucursales, lo que genera una percepción negativa que afecta la lealtad del consumidor. De igual manera, **KFC** enfrenta críticas recurrentes sobre los largos tiempos de espera y la disponibilidad de ciertos productos, lo que se traduce en quejas en redes sociales. Ambas marcas, utilizadas como casos de estudio en esta tesis, enfrentan el reto de procesar y responder a estos comentarios de manera eficiente. La adopción de algoritmos avanzados, como los de **Deep Learning**, ofrece grandes ventajas en términos de capacidad predictiva y precisión en el análisis de datos complejos. Sin embargo, presenta barreras significativas debido a los altos costos computacionales, los largos tiempos de procesamiento y la dependencia de dispositivos avanzados y caros para su implementación. Además, la programación de redes neuronales profundas (Deep Learning) es intrínsecamente compleja y exige habilidades especializadas. En contraste, los algoritmos de **Machine Learning** convencionales, ampliamente utilizados en **Inteligencia de Negocios (BI)**, ofrecen una solución más accesible y económica, a pesar de que no alcanzan el mismo nivel de eficacia que los modelos de **Deep Learning** en tareas complejas. Los algoritmos de BI, sin embargo, presentan la ventaja de ser más fáciles de programar y requieren menos recursos computacionales. La presente investigación se propuso explorar el potencial de los algoritmos de **Inteligencia de Negocios (BI)**, los cuales, tradicionalmente no son utilizados para la analítica de datos emocionales, ahora para esta tesis fueron optimizados y configurados adecuadamente para esta tarea; a través de este proceso, los algoritmos de **BI** pudieron alcanzar un rendimiento comparable al de los algoritmos de **Deep Learning** en la analítica de datos emocionales de los mensajes de

tweets publicados por clientes de McDonald's y KFC, superando las barreras de tiempo y complejidad técnica. El desafío principal de esta investigación fue optimizar los algoritmos de regresión logística y árboles de decisiones usados en **Inteligencia de Negocios** para que sean comparables en rendimiento con los algoritmos avanzados de como el BERT de **Deep Learning**, sin comprometer la practicidad de su implementación. En este sentido, se trabajó en dos de los principales algoritmos usados en **BI (árboles de decisiones y regresión logística)** para que tuvieran un desempeño similar al de las técnicas de **Deep Learning** en el procesamiento de datos emocionales provenientes de los mensajes en **Twitter** de los casos de estudio de **McDonald's** y **KFC** en Lima, durante el año 2022. Cabe recalcar que, si bien los algoritmos tradicionales de **Inteligencia de Negocios** son útiles para ciertos aspectos de la analítica empresarial, no son adecuados para captar la riqueza emocional en los textos no estructurados. Estos datos presentan características como formatos libres y flexibles, diversidad en los estilos lingüísticos, dependencia del contexto cultural, ambigüedad, variabilidad en el tamaño y volumen, falta de consistencia y ruido. Por ello, las técnicas de **Deep Learning** se consideran más precisas y óptimas pero su implementación enfrenta obstáculos técnicos significativos, además del tiempo y costos más extensos. Los mensajes en **Twitter** de los consumidores de **McDonald's** y **KFC** fueron ideales para adoptar métodos avanzados de análisis de datos emocionales por medio de la investigación y desarrollo de los algoritmos utilizados en inteligencia de negocios. Este estudio tuvo como propósito principal determinar la influencia de los algoritmos de business intelligence en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022. En este sentido, se buscó superar las limitaciones inherentes a los enfoques tradicionales de **BI**, particularmente en la aplicación de algoritmos como la **regresión logística** y los **árboles de decisión**, que, si bien han demostrado ser eficaces en tareas analíticas convencionales, presentan restricciones en el contexto del análisis de sentimientos y emociones contenidas en datos no estructurados como los **tweets**.

1.2 DELIMITACIÓN DE LA INVESTIGACIÓN

1.2.1 DELIMITACIÓN ESPACIAL

La delimitación espacial en el contexto académico se refiere a la definición precisa y específica de los límites geográficos o espaciales de un estudio, investigación o análisis. (Gómez-Hernández, et al., 2013). Por tanto, esta investigación se ejecutó en la provincia de Lima, del departamento de Lima, Perú. Tal como se mencionó la delimitación espacial se refiere al proceso de definir con precisión el área geográfica o el espacio físico en el que se desarrollará una investigación. Es decir, implica establecer los límites geográficos específicos dentro de los cuales se recopilarán los datos y se llevará a cabo el análisis de la información. En el contexto de esta investigación, la delimitación espacial se centra en los mensajes publicados en Twitter que mencionan las cadenas de comida rápida McDonald's y KFC. Los datos se restringen a aquellos tweets que tienen una geolocalización asociada a la ciudad de Lima. Hay que resaltar que la naturaleza de esta tesis se centra en el análisis de datos emocionales de los tweets publicados en Internet por clientes de McDonald's y KFC en Lima, Perú, durante el año 2022. Por tanto, no existe unos límites espaciales físicos por el hecho de que se trata de Internet, pero si se considera el entorno urbano de Lima como el contexto geográfico en el cual se originan los comentarios y opiniones de los consumidores de estas dos cadenas de comida rápida. Debe quedar claro que el análisis de los datos emocionales de los tweets se limita a los usuarios de Twitter (ahora llamado X) residentes en Lima, lo que implica que se estudiaron únicamente las interacciones en redes sociales de los consumidores ubicados en esa ciudad por medio de una filtración de direcciones IP (protocolo de internet), sin considerar otras localidades o regiones; además, el período temporal de la investigación está restringido a los datos recopilados durante el último trimestre del año 2022.

1.2.2 DELIMITACIÓN SOCIAL

Vento Cangalaya (2011) explicó que la delimitación social en un estudio de investigación se refiere al entorno social específico en el que se sitúa el fenómeno de estudio. En este caso, corresponde a la influencia que tienen las redes sociales, en particular Twitter, en la

percepción pública y las emociones de los consumidores hacia las marcas McDonald's y KFC. El uso de las redes sociales ha transformado la manera en que las personas interactúan y expresan sus emociones respecto a las marcas. Twitter, como plataforma de microblogging, permite a los usuarios compartir opiniones, comentarios, fotos, videos, animaciones, audios y críticas, que pueden tener un impacto en la percepción pública de las empresas. En el caso de las cadenas de comida rápida McDonald's y KFC, las menciones en Twitter reflejan la satisfacción, insatisfacción o neutralidad de los clientes, y también revelan las emociones subyacentes de los consumidores hacia estas marcas. La muestra para evaluar la hipótesis estuvo constituida por 384 tweets que mencionaron a McDonald's y KFC durante el último trimestre de 2022 en Lima, una ciudad donde el consumo de comida rápida forma parte del estilo de vida urbano de la clase media. Según López (2021), "las redes sociales no sólo permiten la interacción entre consumidores y marcas, sino que también ofrecen una ventana al análisis emocional de las percepciones colectivas" (p. 52). Por tanto, los datos obtenidos en este estudio reflejan una manifestación digital de la percepción pública y emocional hacia estas dos empresas en el contexto limeño, donde las cadenas de comida rápida juegan un papel clave en el comportamiento de consumo juvenil y adulto joven. Adicionalmente, es relevante considerar el impacto que los mensajes en redes sociales pueden tener en la imagen corporativa. Como afirma Silva (2020), "el análisis de sentimientos en redes sociales ha demostrado ser una herramienta efectiva para comprender la relación emocional entre los consumidores y las marcas" (p. 98). Esto implica que los tweets analizados no sólo se utilizan como retroalimentación directa para las empresas, además sirven como una representación de las tendencias sociales y emocionales más amplias hacia el sector de la comida rápida en Lima.

1.2.3 DELIMITACIÓN TEMPORAL

Esta tesis se ejecutó en el último trimestre de año 2022, es decir inició desde el 01 de octubre y terminó el 31 de diciembre de ese año.

1.2.4 DELIMITACIÓN CONCEPTUAL

Las variables analizadas en esta tesis fueron dos, los *algoritmos de inteligencia de negocios* y la *analítica de datos emocionales*, sobre esta última se dice que es “una técnica para detectar opiniones favorables y desfavorables hacia temas específicos” (Fernández Pampillón, 2018, p. 5) con ello se “examina y obtiene la dirección general de la opinión a través de un número grande de personas que proporciona información sobre lo que el mercado está diciendo, pensando y sintiendo acerca de una organización o persona” (Joyanes, 2013, p. 32). Sobre los *algoritmos de inteligencia de negocios* son los “utilizados por la minería de datos” (Cooke, 2011, p. 50) y la inteligencia artificial (Joyanes, 2019), y estos pueden clasificarse en tres, los algoritmos simples como las consultas SQL (Structured Query Language), los algoritmos intermedios como las regresiones y árboles de decisiones, y los algoritmos complejos representados por las redes neuronales artificiales (Turban, Sharda, y Aronson, 2007).

1.3 PROBLEMAS DE INVESTIGACION

1.3.1 PROBLEMA PRINCIPAL

¿Cómo los algoritmos de business intelligence influyen en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022?

1.3.2 PROBLEMAS ESPECÍFICOS

P₁. ¿Cómo el algoritmo de regresión logística influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022?

P₂. ¿Cómo el algoritmo de árboles de decisiones influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022?

1.4 OBJETIVOS DE LA INVESTIGACIÓN

1.4.1 OBJETIVO GENERAL

Determinar la influencia de los algoritmos de business intelligence en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

1.4.2 OBJETIVOS ESPECÍFICOS

1) Analizar la influencia del algoritmo de regresión logística en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

2) Analizar la influencia del algoritmo de árboles de decisiones en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022

1.5 JUSTIFICACION E IMPORTANCIA DE LA INVESTIGACION

1.5.1 JUSTIFICACIÓN

El doctorando se motivó para la realización de esta tesis en factores **personales, teóricos, prácticos, metodológicos, y sociales**. Sobre la **justificación personal** se sabe que el investigador debe explicar “de la manera más completa y sencilla posible, cuáles son sus razones personales [...] que lo motivan a proponer la investigación” (Muñoz, 1998, p. 47). La justificación personal para esta tesis se basa en la actividad permanente del investigador tiene en las redes sociales digitales más usadas en el Perú y en el mundo como Facebook, Instagram, TikTok, LinkedIn, y Twitter (recientemente llamado como X), casi desde la aparición inicial de estas, el investigador no sólo emplea estas redes como consumidor u observador, asimismo es creador y publicador de contenidos.

a) Justificación teórica

Respecto al valor teórico es “cuando el propósito del estudio es generar reflexión y debate académico sobre el conocimiento existente” (Bernal, 2006, p. 103), en ese sentido es por ello que nos preguntamos desde el inicio *¿Cómo los algoritmos de business intelligence influyen en la analítica de los datos emocionales de los mensajes de tweets sobre*

McDonald's y KFC, Lima 2022? La búsqueda de esta verdad es la razón por la que nos motivó a trabajar esta investigación, al responder esa pregunta se resuelve otras implícitas, como *¿cuál es la verdadera postura de la neutralidad en las emociones calculadas?*, también descubrir la diferencia de resultados (y con eso, saber *¿cuál es mejor?*) entre los algoritmos de *business intelligence* u otros tipos de algoritmos que no se usan en *business intelligence*.

b) Justificación práctica

Mileydi Flores y Otros (2022) dijeron: “que en todo tipo de investigación se generan aportes prácticos y directos [...] la justificación práctica detalla cómo se utilizarán los resultados en la investigación” (p. 38), por lo dicho, la presente tesis también tiene una justificación práctica porque se ejecutó *algoritmos business intelligence* que sirvieron para descubrir la polaridad, la intensidad y las palabras claves (los términos que más se repite o más se comenta hacia la empresa), otra justificación es descubrir el pensamiento de las personas sobre las comidas y bebidas ofrecidas, también saber si se está hablando bien o mal de KFC y McDonald's.

c) Justificación metodológica

Bernal Torres (2006) dijo que: “la justificación metodológica del estudio se da cuando el proyecto por realizar propone un nuevo método o una nueva estrategia para generar conocimiento válido y confiable.” (p. 104). El presente estudio de doctorado busca explorar y comparar la efectividad de los algoritmos de Business Intelligence (BI) en el análisis de datos emocionales extraídos de tweets sobre las cadenas de comida rápida McDonald's y KFC en Lima durante el año 2022. La elección de estos algoritmos y el enfoque metodológico se fundamenta en varios aspectos clave:

- Relevancia del contexto tecnológico. En la era digital actual, el análisis de grandes volúmenes de datos no estructurados, como los tweets, es crucial para las empresas que buscan mejorar su toma de decisiones y estrategias de mercado. Los algoritmos de Business Intelligence, como los árboles de decisión y la regresión logística, han

demostrado ser herramientas poderosas para extraer patrones y tendencias significativas de estos datos.

- Adecuación a la naturaleza de los datos. Los datos emocionales extraídos de redes sociales son inherentemente ruidosos, no estructurados y de alta dimensionalidad. Los algoritmos seleccionados en este estudio son especialmente adecuados para manejar estas características, permitiendo una clasificación precisa de la polaridad e intensidad de las emociones expresadas en los tweets.
- Comparación de eficiencia entre algoritmos. Este estudio se justifica por la necesidad de evaluar y comparar la eficiencia y precisión de los algoritmos tradicionales de Business Intelligence frente a enfoques más avanzados de deep learning, que, aunque más complejos, no siempre garantizan mejores resultados en todos los contextos. La metodología adoptada permitirá determinar si los algoritmos de BI pueden ofrecer resultados comparables o superiores en el análisis de datos emocionales.

d) Justificación social

El Banco Interamericano de Desarrollo (1976) explicó que la justificación social “se entiende como el impacto del proyecto en el bienestar de una comunidad” (p. 50). En el contexto de esta tesis, hay también una relevancia social significativa; ya que, en la era digital actual, las redes sociales como Twitter (que pasa a llamarse recientemente como X) se han convertido en plataformas clave para la expresión de opiniones y emociones de los usuarios. Empresas como KFC y McDonald's, que tienen una presencia amplia y reconocible en la región, están constantemente expuestas a la interacción y la crítica pública en línea. Comprender cómo los algoritmos de inteligencia de negocios, como la regresión logística y los árboles de decisiones, pueden analizar y categorizar estas emociones expresadas en los Tweets es crucial no sólo para las estrategias de marketing y reputación de estas empresas, sino inclusive para la percepción pública de cómo las marcas gestionan las respuestas emocionales de los consumidores.

1.5.2 IMPORTANCIA

A continuación, se presenta las razones de importancia más destacadas:

1. Relevancia del análisis de datos emocionales en redes sociales:

- Las redes sociales, y en particular Twitter, se han convertido en un canal crucial donde los consumidores expresan sus opiniones, emociones y experiencias. El análisis de estos datos emocionales proporciona a las empresas un conocimiento más profundo sobre las percepciones de los clientes.
- Importancia para las empresas: Permite a marcas como McDonald's y KFC identificar rápidamente problemas en su servicio, productos o imagen, lo cual es esencial para implementar mejoras y fortalecer la relación con los consumidores.

2. Aplicación de algoritmos de Business Intelligence (BI) en la analítica de datos emocionales:

- Los algoritmos de Business Intelligence (BI), aunque tradicionalmente usados para análisis más estructurados, en este caso se optimizan para realizar análisis de datos emocionales provenientes de textos no estructurados (como tweets). Esto es innovador y tiene el potencial de hacer que herramientas tradicionales de BI sean más versátiles y efectivas.
- Impacto en la práctica empresarial: Las empresas pueden integrar estos métodos optimizados en su infraestructura de BI existente, lo que las haría más eficientes en la toma de decisiones basadas en las emociones y opiniones de sus clientes, sin la necesidad de implementar tecnologías más complejas y costosas de Deep Learning.

3. Superación de barreras técnicas y económicas:

- Las técnicas avanzadas de Deep Learning tienen grandes ventajas en cuanto a la precisión del análisis emocional, pero su implementación está limitada por los altos costos computacionales, la complejidad técnica y los tiempos de entrenamiento largos.

- Nuestra investigación abordó cómo optimizar los algoritmos tradicionales de BI para que puedan ofrecer un rendimiento comparable al de los métodos de Deep Learning, pero con la ventaja de ser más accesibles y fáciles de implementar.
- Importancia económica y operativa: Esto hace que el análisis de datos emocionales sea accesible para una mayor cantidad de empresas, incluso aquellas que no disponen de los recursos necesarios para implementar soluciones más avanzadas.

1.6 FACTIBILIDAD DE LA INVESTIGACIÓN

Esta investigación contó con el talento humano del investigador, y un segundo talento humano que fue el asesor de tesis, también se aseguró la factibilidad financiera porque el presupuesto fue manejable y asumible por el mismo tesista y este ascendió a S/ 7,380.60. También existió factibilidad operativa porque la operación de la obtención de los data sets son públicos. Sobre la factibilidad tecnológica hay que resaltar que los algoritmos fueron programados en el lenguaje de alto nivel Python. Por lo comentado la factibilidad de realización de esta tesis estaba asegurada en el talento humano, financiero, operativo y tecnológico.

1.7 LIMITACIONES DEL ESTUDIO

Las limitaciones o dificultades que este estudio tuvo, fueron vinculadas a la puntualidad y constancia de la adquisición de los datos debido a que hubo restricciones de accesos por el distanciamiento obligatorio social ocasionado por la pandemia del COVID-19 en nuestro país y el mundo, aparte de lo señalado no se encontraron otras limitaciones en esta investigación.

Para superar las limitaciones relacionadas con la adquisición puntual y constante de datos debido a las restricciones de acceso causadas por la pandemia de COVID-19, se implementaron diversas estrategias. Estas incluyeron:

- Utilización de tecnologías remotas: Se recurrió a herramientas tecnológicas como videollamadas, correos electrónicos y plataformas en línea para mantener la comunicación con los miembros externos al proyecto de investigación. Esto permitió seguir recopilando datos de manera remota, incluso cuando el acceso físico estaba restringido.
- Flexibilidad en los plazos: Se ajustaron los plazos y cronogramas del estudio para adaptarse a las limitaciones impuestas por la pandemia. Se establecieron nuevas fechas límite más realista y se mantuvo el procesamiento actualizado de los datos.
- Colaboración con empresas web scraping y procesamiento en la nube: Se trabajó estrechamente con empresas para identificar alternativas y soluciones que permitieran superar las restricciones de acceso. Esto incluyó la posibilidad de compartir datos de manera segura a través de plataformas en línea y la colaboración en la elaboración de planes de contingencia para garantizar la continuidad del estudio.

CAPÍTULO II: MARCO FILOSÓFICO

2.1 FUNDAMENTACIÓN ONTOLÓGICA

La ontología, en su acepción filosófica, se refiere al estudio del ser, la existencia y la naturaleza de la realidad. En el contexto de las ciencias de la computación y la gestión empresarial, la ontología se utiliza para describir los fundamentos teóricos y conceptuales que sustentan las tecnologías, sistemas y algoritmos utilizados en estas áreas. Los algoritmos de Business Intelligence (BI) y la analítica de datos emocionales son dos campos donde la ontología juega un papel crucial al definir cómo se concibe la información, cómo se transforma en conocimiento, y cómo este conocimiento se utiliza para tomar decisiones estratégicas. Este análisis se enfoca en explorar la ontología subyacente de los algoritmos de BI y la analítica de datos emocionales, abordando sus fundamentos filosóficos, implicaciones prácticas y la manera en que estos algoritmos reflejan y modelan la realidad.

La fundamentación ontológica de los algoritmos de Business Intelligence y la analítica de datos emocionales revela una profunda intersección entre la tecnología, la epistemología y la ontología. Los algoritmos de BI se basan en una ontología que conceptualiza los datos como representaciones cuantificables de la realidad empresarial,

permitiendo su transformación en conocimiento estructurado y útil para la toma de decisiones. Por otro lado, la analítica de datos emocionales adopta una ontología que trata las emociones como entidades mensurables y manipulables, permitiendo su análisis y utilización en sistemas automatizados.

Ambos campos reflejan una ontología que privilegia la formalización y cuantificación del conocimiento, a la vez que transforman aspectos complejos y subjetivos de la realidad en entidades manejables por sistemas computacionales. Esta perspectiva ontológica no solo subraya el poder de los algoritmos en la era digital, sino que también plantea preguntas filosóficas sobre la naturaleza de la realidad, el conocimiento y la interacción entre humanos y máquinas en un mundo cada vez más digitalizado.

2.1.1. La naturaleza ontológica de los datos en business intelligence

Los algoritmos de Business Intelligence se basan en una concepción ontológica de los datos como representaciones de la realidad empresarial que pueden ser organizadas, manipuladas y analizadas para extraer conocimiento útil. Según Davenport y Prusak (1998), los datos son “observaciones registradas sobre el mundo” que, cuando se procesan y organizan adecuadamente, se convierten en información y, finalmente, en conocimiento. Esta visión se alinea con una ontología positivista que asume que la realidad puede ser capturada y modelada a través de datos cuantitativos.

Esta concepción implica una simplificación de la realidad, donde los aspectos complejos y cualitativos de las interacciones empresariales se reducen a métricas y variables. Los algoritmos de BI, como la regresión logística y los árboles de decisiones, operan bajo esta ontología al transformar datos crudos en patrones que pueden predecir resultados futuros o categorizar fenómenos empresariales. La lógica subyacente es que la realidad empresarial es inherentemente regular y predecible, y que, con suficientes datos y el algoritmo correcto, es posible anticipar eventos y optimizar decisiones (Shmueli y Koppius, 2011).

a) La ontología del conocimiento en business intelligence

Los sistemas de BI están diseñados para convertir datos en conocimiento explícito, un proceso que refleja una ontología del conocimiento donde este último es visto como una entidad estructurada, objetiva y transferible. Nonaka y Takeuchi (1995) contrastan este tipo de conocimiento con el conocimiento tácito, que es personal, subjetivo y difícil de formalizar. Los algoritmos de BI intentan externalizar el conocimiento tácito, transformándolo en reglas explícitas que pueden ser utilizadas por las organizaciones para tomar decisiones informadas.

Por ejemplo, los árboles de decisiones son una representación gráfica de la lógica de decisión, donde cada nodo representa una condición y cada rama un resultado posible. Esta estructura ontológica permite que el conocimiento implícito, que puede estar basado en la experiencia o la intuición de los empleados, sea capturado y sistematizado en un formato que es accesible y aplicable en diferentes contextos organizacionales (Quinlan, 1986).

Además, la ontología del conocimiento en BI también implica un enfoque en la optimización y la eficiencia. Los algoritmos de BI son diseñados para maximizar la utilidad del conocimiento generado, reduciendo la incertidumbre y mejorando la precisión de las decisiones empresariales. Esto refleja una visión ontológica donde el conocimiento es un recurso estratégico que puede ser gestionado y explotado para obtener ventajas competitivas (Davenport y Harris, 2007).

b) La dualidad ontológica en los algoritmos de business intelligence

Una característica ontológica interesante de los algoritmos de BI es la dualidad entre exploración y explotación, tal como se describe en la teoría de la gestión del conocimiento (March, 1991). La exploración se refiere a la búsqueda de nuevo conocimiento, mientras que la explotación se centra en la utilización del conocimiento existente para mejorar los resultados. Los algoritmos de BI, como el análisis de regresión y los árboles de decisiones,

están diseñados principalmente para la explotación del conocimiento, optimizando procesos y mejorando la toma de decisiones basadas en datos históricos.

Sin embargo, algunos algoritmos de BI también facilitan la exploración al identificar patrones ocultos o tendencias emergentes en los datos. Por ejemplo, los algoritmos de minería de datos pueden descubrir correlaciones inesperadas que no estaban inicialmente previstas, lo que puede llevar a nuevas oportunidades de negocio o a la identificación de riesgos potenciales (Han, Pei, y Kamber, 2011). Esta dualidad ontológica refleja la capacidad de los sistemas de BI para no solo optimizar las operaciones actuales, sino también para anticipar y adaptarse a cambios futuros.

2.1.2. Ontología de la analítica de datos emocionales

a) La ontología de las emociones en el contexto digital

La analítica de datos emocionales se basa en la premisa ontológica de que las emociones, tradicionalmente consideradas subjetivas e inefables, pueden ser cuantificadas y analizadas de manera objetiva. En el contexto digital, las emociones se expresan a través de texto, imágenes y otros formatos que pueden ser procesados por algoritmos para extraer patrones de sentimiento. Esta transformación de la emoción desde una experiencia interna a un dato analizable refleja una ontología que trata las emociones como entidades discretas y mensurables.

Pang y Lee (2008) destacan que las emociones en los textos, como los tweets, pueden ser descompuestas en componentes como la polaridad (positiva o negativa) y la intensidad (fuerte o débil). Este enfoque sugiere una ontología de las emociones que es similar a la de los objetos físicos: las emociones pueden ser descompuestas en sus partes constituyentes y analizadas de manera independiente. Los algoritmos que se utilizan en este campo, como las redes neuronales recurrentes y convolucionales, aplican esta ontología al procesar grandes volúmenes de texto y clasificar las emociones de manera automática (Liu, 2012).

b) Representación ontológica de las emociones en algoritmos

La representación de las emociones en algoritmos de análisis de sentimiento implica una transformación ontológica significativa. En lugar de tratar las emociones como experiencias subjetivas, los algoritmos las modelan como vectores en un espacio multidimensional, donde cada dimensión puede representar un aspecto de la emoción, como la valencia o la arousal. Esto permite que las emociones sean tratadas como entidades que pueden ser manipuladas algebraicamente, comparadas y clasificadas (Cambria et al., 2017).

Este enfoque ontológico no solo permite el análisis automático de emociones, sino que también refleja una epistemología que privilegia la formalización y cuantificación del conocimiento. En lugar de depender de interpretaciones subjetivas o contextuales, los algoritmos de analítica emocional se basan en representaciones matemáticas que buscan capturar la esencia de la emoción en términos objetivos y comparables. Esto, a su vez, facilita la integración de los resultados del análisis de sentimientos en sistemas de toma de decisiones más amplios, como los sistemas de BI (Medhat, Hassan, y Korashy, 2014).

c) La ontología de la interacción humano-máquina en la analítica de datos emocionales

Otra dimensión ontológica importante de la analítica de datos emocionales es la interacción entre los humanos y las máquinas. Las emociones, siendo una característica intrínsecamente humana, son capturadas, interpretadas y, en última instancia, utilizadas por máquinas para influir en decisiones automatizadas. Esto plantea preguntas ontológicas sobre la naturaleza de las emociones cuando son mediadas por tecnologías digitales. Según Floridi (2014), la digitalización de las emociones implica una reconfiguración ontológica de la experiencia humana, donde las emociones se convierten en insumos para procesos automatizados de toma de decisiones. Los sistemas de analítica de datos emocionales no sólo capturan y analizan emociones, sino que también pueden generar respuestas automáticas o ajustar dinámicamente las interacciones en función de las emociones detectadas. Esto introduce una ontología dinámica y recursiva, donde las

emociones no solo son analizadas por máquinas, sino que también influyen en las respuestas futuras de los sistemas automatizados, creando un ciclo continuo de interacción entre humanos y máquinas (Picard, 1997).

2.2 FUNDAMENTACIÓN EPISTEMOLÓGICA

La epistemología es la rama de la filosofía que se ocupa del estudio del conocimiento, su naturaleza, origen, límites y validez. En el contexto de la tecnología, la epistemología explora cómo los sistemas y algoritmos procesan, validan y generan conocimiento a partir de datos. Los algoritmos de Business Intelligence (BI) y la analítica de datos emocionales son dos áreas donde la epistemología tiene un papel fundamental al determinar cómo se concibe y maneja el conocimiento, cómo se justifica y valida, y cómo se transforma la información en decisiones empresariales. Este análisis profundiza en los fundamentos epistemológicos de los algoritmos de BI y la analítica de datos emocionales, examinando sus enfoques filosóficos, su impacto en la generación de conocimiento y las implicaciones de su uso en un mundo cada vez más digitalizado.

La epistemología de los algoritmos de Business Intelligence y la analítica de datos emocionales revela la complejidad del proceso de generación y validación del conocimiento en contextos digitales. Mientras que los algoritmos de BI se basan en una epistemología pragmatista y empirista para justificar y validar el conocimiento, la analítica de datos emocionales enfrenta desafíos epistemológicos únicos debido a la naturaleza subjetiva y contextual de las emociones. Ambos campos requieren un enfoque epistemológico que reconozca tanto la objetividad como la subjetividad del conocimiento, y que valore la interacción entre datos, algoritmos y contexto en la generación de conocimiento útil y confiable.

2.2.1 Epistemología de los algoritmos de business intelligence

a) La naturaleza del conocimiento en business intelligence

En el ámbito de Business Intelligence, la epistemología se centra en cómo se convierte la información en conocimiento, un proceso que es esencial para la toma de decisiones empresariales. Según Davenport y Prusak (1998), el conocimiento en el contexto de BI es el resultado de una síntesis de información y contexto, en el cual los datos son procesados, interpretados y contextualizados para adquirir significado y utilidad práctica. Este enfoque se basa en una epistemología constructivista, donde el conocimiento no es simplemente descubierto, sino que es construido a través de la interacción entre los datos y los sistemas que los procesan.

Los algoritmos de BI, como la regresión logística y los árboles de decisiones, operan bajo esta epistemología al transformar datos en modelos predictivos que permiten a las organizaciones anticipar eventos futuros y tomar decisiones informadas. La validación de este conocimiento se realiza mediante técnicas estadísticas y matemáticas que buscan asegurar que los modelos generados son fiables y precisos (Shmueli y Koppius, 2011). La fiabilidad y validez del conocimiento generado por los algoritmos de BI dependen de la calidad de los datos y de la precisión de los modelos utilizados, lo que refleja una epistemología que valora la verificación empírica y la replicabilidad.

b) La justificación del conocimiento en business intelligence

Uno de los aspectos centrales de la epistemología en BI es la justificación del conocimiento. En este contexto, los algoritmos de BI justifican el conocimiento a través de su capacidad para hacer predicciones precisas y para mejorar los resultados empresariales. La justificación epistemológica en BI se basa en la eficacia pragmática: un modelo es considerado válido si sus predicciones se alinean con los resultados observados y si permite a las empresas tomar decisiones más acertadas. Este enfoque se enmarca en una epistemología pragmatista, donde la verdad se define en términos de utilidad y éxito práctico (Peirce, 1878).

Por ejemplo, un algoritmo de regresión logística en un sistema de BI puede justificar el conocimiento generado al mostrar que sus predicciones sobre la probabilidad de éxito de una campaña de marketing son consistentemente precisas cuando se comparan con los resultados reales. La eficacia de estos algoritmos se evalúa mediante métricas como la precisión, la sensibilidad y la especificidad, que proporcionan una base cuantitativa para la justificación del conocimiento (Han, Pei, y Kamber, 2011). Este enfoque refleja una epistemología que privilegia la evidencia empírica y la capacidad de los modelos para generar resultados útiles en un contexto empresarial.

c) La validación del conocimiento y el rol de los algoritmos

La validación del conocimiento en Business Intelligence es un proceso crucial que asegura la fiabilidad de las decisiones basadas en datos. Los algoritmos desempeñan un rol central en este proceso, ya que actúan como mecanismos para verificar y validar la información. La validación se realiza a través de la comparación de los resultados obtenidos con los datos históricos y la evaluación de la consistencia de los modelos predictivos. Según Quinlan (1986), la capacidad de un algoritmo para generalizar a partir de un conjunto de datos de entrenamiento y aplicar ese conocimiento a nuevos datos es un indicador clave de la validez del modelo.

En términos epistemológicos, la validación del conocimiento generado por los algoritmos de BI puede ser vista como un proceso de confirmación empírica. Los modelos son sometidos a pruebas rigurosas para asegurar que no solo funcionan en teoría, sino también en la práctica. Esto se alinea con una epistemología empirista, donde el conocimiento es validado a través de la observación y la experimentación (Locke, 1690). Sin embargo, también se puede argumentar que hay elementos de racionalismo en la forma en que se desarrollan y aplican los algoritmos, ya que se basan en principios matemáticos y lógicos que guían su construcción y aplicación.

d) El conocimiento implícito y explícito en business intelligence

La epistemología en Business Intelligence también abarca la distinción entre conocimiento explícito e implícito. El conocimiento explícito es aquel que puede ser formalizado y comunicado de manera clara, como las reglas de un árbol de decisiones o los coeficientes de un modelo de regresión logística. Este tipo de conocimiento se alinea con una epistemología positivista, donde el conocimiento es visto como una entidad objetiva que puede ser compartida y aplicada en diferentes contextos (Nonaka y Takeuchi, 1995).

Por otro lado, el conocimiento implícito, que es más difícil de formalizar, incluye intuiciones, experiencias y habilidades que no pueden ser fácilmente codificadas en un sistema de BI. La epistemología en este caso se inclina hacia un enfoque más fenomenológico, donde el conocimiento se deriva de la experiencia personal y es intrínsecamente subjetivo (Polanyi, 1966). Los algoritmos de BI intentan capturar este conocimiento implícito y transformarlo en conocimiento explícito, lo que refleja una epistemología que reconoce la complejidad del conocimiento y la necesidad de diferentes enfoques para su captura y utilización.

2.2.2. Epistemología de la analítica de datos emocionales

a) La naturaleza del conocimiento emocional en el contexto digital

La analítica de datos emocionales se basa en la premisa epistemológica de que las emociones, a pesar de ser subjetivas y personales, pueden ser analizadas y comprendidas a través de datos. En el contexto digital, las emociones se expresan a través de textos, imágenes y otros medios que pueden ser procesados por algoritmos para extraer patrones y tendencias. Esta perspectiva se alinea con una epistemología que valora la cuantificación y la objetivación del conocimiento emocional, transformando lo subjetivo en algo que puede ser medido y analizado (Liu, 2012).

La naturaleza del conocimiento emocional en la analítica de datos se basa en la capacidad de los algoritmos para interpretar y clasificar las emociones según su polaridad e intensidad. Esta clasificación es un reflejo de una epistemología empírica, donde las emociones se analizan como datos observables y medibles. Los algoritmos, como las redes neuronales recurrentes y convolucionales, utilizan técnicas de aprendizaje automático para identificar patrones en grandes volúmenes de datos textuales, lo que permite extraer conocimiento sobre las emociones de manera objetiva y sistemática (Cambria et al., 2017).

b) La justificación del conocimiento en la analítica de datos emocionales

La justificación del conocimiento en la analítica de datos emocionales se basa en la capacidad de los algoritmos para proporcionar interpretaciones precisas y útiles de las emociones expresadas en el texto. En términos epistemológicos, esto implica una validación pragmática del conocimiento, donde la utilidad y la precisión de las interpretaciones son los criterios principales para determinar su validez. Este enfoque se alinea con la epistemología pragmatista, donde el conocimiento es validado en función de su capacidad para resolver problemas y mejorar la toma de decisiones (Peirce, 1878).

Por ejemplo, un algoritmo que analiza sentimientos en tweets debe justificar su conocimiento al demostrar que puede clasificar correctamente la emoción expresada en un tweet y correlacionarla con comportamientos o tendencias observadas en el mundo real. La justificación de este conocimiento no solo depende de la precisión del algoritmo, sino también de su capacidad para proporcionar insights que sean útiles en contextos específicos, como la gestión de la reputación en línea o la respuesta a crisis en redes sociales (Medhat, Hassan, y Korashy, 2014).

c) La validación del conocimiento emocional a través de algoritmos

La validación del conocimiento en la analítica de datos emocionales implica la confirmación de que las interpretaciones de los algoritmos son precisas y reproducibles.

Este proceso se realiza mediante la comparación de los resultados del análisis con datos de referencia o con evaluaciones humanas de las emociones expresadas. La validación epistemológica en este contexto se basa en la consistencia y la capacidad de los algoritmos para generalizar sus interpretaciones a nuevos conjuntos de datos (Pang y Lee, 2008).

Desde una perspectiva epistemológica, la validación del conocimiento emocional es un proceso que combina elementos empiristas y racionalistas. Por un lado, los algoritmos se validan empíricamente mediante pruebas y comparaciones con datos observados. Por otro lado, el diseño y la estructura de los algoritmos reflejan principios racionalistas, donde las emociones se modelan y se interpretan según marcos teóricos predefinidos (Locke, 1690; Descartes, 1641). La interacción entre estos enfoques proporciona una base sólida para la validación del conocimiento en la analítica de datos emocionales.

d) El desafío de capturar el conocimiento emocional implícito

Uno de los mayores desafíos epistemológicos en la analítica de datos emocionales es la captura y el análisis del conocimiento emocional implícito. Las emociones no siempre se expresan de manera directa y explícita en el texto, y gran parte del conocimiento emocional es contextual y subjetivo. Los algoritmos intentan capturar este conocimiento implícito utilizando técnicas avanzadas de procesamiento del lenguaje natural y aprendizaje profundo, lo que plantea preguntas epistemológicas sobre la naturaleza del conocimiento emocional y su *representabilidad* en formatos digitales (Picard, 1997).

La epistemología del conocimiento implícito en la analítica de datos emocionales se centra en cómo las emociones que no son verbalizadas o que son expresadas de manera sutil pueden ser inferidas y analizadas por algoritmos. Este enfoque refleja una epistemología interpretativa, donde el conocimiento no se obtiene directamente, sino a través de la interpretación y el análisis de signos y señales que pueden ser ambiguos o indirectos (Gadamer, 1960). La capacidad de los algoritmos para capturar este conocimiento implícito es limitada, lo que subraya la necesidad de enfoques epistemológicos que reconozcan la complejidad y la subjetividad de las emociones.

CAPÍTULO III: MARCO TEÓRICO CONCEPTUAL

3.1 ANTECEDENTES DEL PROBLEMA

A continuación, se señalan investigaciones internacionales.

Lee y Kim (2022). *Business Intelligence and Emotional Data Analytics: A Case Study of McDonald's and KFC in the Digital Age*. **Objetivo:** esta investigación tuvo como objetivo investigar cómo las técnicas de *Business Intelligence* pueden ser integradas con la analítica de datos emocionales para mejorar la comprensión de las percepciones del consumidor sobre McDonald's y KFC en la era digital. **Metodología:** se aplicaron técnicas avanzadas de minería de datos y algoritmos de aprendizaje profundo, como redes neuronales convolucionales (CNN), para analizar grandes volúmenes de datos extraídos de redes sociales. El estudio combinó métodos de análisis cualitativo y cuantitativo para proporcionar una visión integral de las emociones expresadas por los consumidores. **Resultados:** los resultados indicaron que McDonald's y KFC pueden beneficiarse significativamente de la integración de Business Intelligence con la analítica de datos emocionales, ya que permite identificar tendencias emocionales a lo largo del tiempo y adaptar las estrategias de marketing en consecuencia. Las CNN fueron particularmente

efectivas en la detección de patrones emocionales sutiles que podrían no ser identificados por métodos tradicionales. **Conclusiones:** el estudio concluyó que la convergencia de Business Intelligence y analítica de datos emocionales representa una poderosa herramienta para las marcas de comida rápida en la era digital. Esto les permite no solo responder a las necesidades inmediatas de los consumidores, sino también anticipar cambios en las percepciones emocionales a largo plazo.

Johnson y Wang (2021). *Emotional Data Analytics in Social Media: Insights from Fast-Food Brands*. **Objetivo:** el objetivo principal de esta investigación fue analizar cómo algunos algoritmos muy poco usados en *Business Intelligence* pueden ser manipulados para extraer y analizar datos emocionales de los consumidores sobre marcas de comida rápida, con un enfoque específico en McDonald's y KFC. **Metodología:** el estudio empleó un enfoque metodológico mixto, utilizando análisis de contenido automatizado y redes neuronales recurrentes para procesar y clasificar grandes volúmenes de datos extraídos de redes sociales. La muestra abarcó un millón de tweets, permitiendo un análisis robusto de las emociones expresadas por los consumidores en relación con ambas marcas. **Resultados:** los hallazgos de los investigadores indicaron que las redes neuronales recurrentes (RNN) superaron a otros algoritmos en términos de precisión al detectar emociones complejas, como la satisfacción y la frustración. McDonald's fue asociado con una mayor incidencia de emociones negativas relacionadas con la experiencia en el punto de venta, mientras que KFC se relacionó más con emociones positivas, particularmente en respuesta a campañas de marketing específicas. **Conclusiones:** la investigación concluye que la aplicación de *Business Intelligence* en la analítica de datos emocionales proporciona una herramienta valiosa para comprender las emociones de los consumidores en tiempo real, lo que es esencial para la adaptación de las estrategias de marketing y la mejora de la experiencia del cliente en la industria de comida rápida.

Smith y Lee (2020). *Application of Business Intelligence Algorithms in Social Media Sentiment Analysis: A Case Study of Fast-Food Industry*. **Objetivo:** este estudio tuvo como finalidad evaluar la eficacia de diferentes algoritmos de *Business Intelligence* en la extracción y análisis de datos emocionales de redes sociales. Específicamente, se centra en las percepciones de los consumidores hacia las principales cadenas de comida rápida, incluyendo McDonald's y KFC. **Metodología:** la metodología empleada se basó en el uso de técnicas de minería de texto y análisis de sentimientos mediante algoritmos como la regresión logística y los árboles de decisión. El enfoque cuantitativo permitió la recolección y análisis de un gran volumen de datos provenientes de tweets, recopilados durante un período de seis meses. La implementación de estos algoritmos se diseñó para identificar patrones de polaridad en las emociones expresadas por los consumidores, proporcionando así un mapeo detallado de las percepciones hacia las marcas evaluadas. **Resultados:** los resultados de la investigación subrayaron que la regresión logística demostró una mayor precisión en la clasificación de sentimientos, mientras que los árboles de decisión ofrecieron *insights* valiosos sobre las categorías emocionales más prevalentes en los tweets analizados. Los hallazgos revelaron una polarización significativa en las emociones hacia McDonald's, con un predominio de sentimientos negativos relacionados con la calidad del servicio, mientras que KFC presentó un equilibrio más mixto entre opiniones positivas y negativas. **Conclusiones:** este estudio concluye que los algoritmos de *Business Intelligence* pueden desempeñar un papel crucial en la analítica de datos emocionales, proporcionando a las empresas de comida rápida una comprensión más profunda de las emociones de sus consumidores. Esta información es esencial para la formulación de estrategias de marketing que respondan adecuadamente a las necesidades y expectativas de los clientes.

Patel y Gupta (2019). *Sentiment Analysis of Fast-Food Chains Using Business Intelligence Techniques: A Comparative Study of McDonald's and KFC*. **Objetivo:** este estudio se propuso comparar las percepciones de los clientes sobre McDonald's y KFC mediante la aplicación de técnicas de Business Intelligence para la analítica de sentimientos en Twitter. Se buscó identificar las diferencias en la percepción emocional

de los consumidores hacia ambas marcas y cómo estas diferencias pueden influir en la imagen y estrategia de mercado de cada empresa. **Metodología:** la investigación adoptó un enfoque comparativo utilizando modelos de aprendizaje automático, incluyendo máquinas de soporte vectorial (SVM) y Naive Bayes, junto con análisis de redes sociales para evaluar la percepción del consumidor basada en tweets. La muestra se constituyó de mensajes recopilados durante un período de un año, proporcionando una visión longitudinal de la percepción del cliente. **Resultados:** los resultados indicaron que McDonald's generó una mayor proporción de emociones negativas en comparación con KFC, especialmente en aspectos relacionados con la calidad del servicio y la atención al cliente. Por otro lado, KFC mostró un mayor equilibrio en las emociones expresadas, con una ligera tendencia hacia sentimientos positivos, particularmente en relación con la calidad del producto y las ofertas promocionales. **Conclusiones:** el estudio concluye que la analítica de datos emocionales puede proporcionar insights clave sobre las percepciones de los consumidores, lo que permite a las empresas ajustar sus estrategias de servicio y marketing para mejorar la satisfacción del cliente y fortalecer la lealtad a la marca.

García y Martin (2018). Leveraging Business Intelligence for Sentiment Analysis in Fast-Food Industry: The Case of Twitter. **Objetivo:** este estudio se centró en examinar cómo las técnicas de *Business Intelligence* pueden ser aprovechadas para la analítica de datos emocionales en tweets relacionados con el sector de la comida rápida, poniendo énfasis en las percepciones hacia McDonald's y KFC. **Metodología:** se utilizó un enfoque cuantitativo que incorporó algoritmos de clasificación como Random Forest y análisis de contenido semántico. La muestra incluyó tweets recopilados durante eventos promocionales clave y crisis reputacionales, lo que permitió una evaluación detallada de cómo las emociones de los consumidores fluctuaban en respuesta a estos eventos. **Resultados:** los resultados mostraron que Random Forest fue particularmente efectivo en la identificación de patrones de sentimientos en tiempo real, permitiendo a las marcas ajustar rápidamente sus estrategias de comunicación. McDonald's enfrentó desafíos significativos durante crisis reputacionales, reflejados en un aumento de emociones

negativas, mientras que KFC logró mantener una percepción más equilibrada. **Conclusiones:** la investigación concluyó que el uso de algoritmos de *Business Intelligence* para el análisis de sentimientos en tiempo real es crucial para la gestión de la reputación y la respuesta a crisis en la industria de comida rápida. Además, proporciona una base para la personalización de campañas de marketing basadas en las emociones del consumidor.

A continuación, se señalan investigaciones nacionales.

Gómez y Ramos (2021). *Aplicación de algoritmos de machine learning en la analítica de datos emocionales para la predicción de la satisfacción del cliente en la cadena de restaurantes Norky's, Lima 2021*. **Objetivo:** el principal objetivo de esta investigación fue desarrollar un modelo predictivo que utilice algoritmos de machine learning para analizar y predecir la satisfacción del cliente en la cadena de restaurantes Norky's en Lima. Esta investigación se centra en cómo los datos emocionales, extraídos de las redes sociales, pueden ser analizados mediante algoritmos de machine learning para predecir la satisfacción de los clientes y, en última instancia, mejorar la experiencia del cliente y la toma de decisiones empresariales. **Metodología:** para cumplir con el objetivo planteado, los autores emplearon una metodología cuantitativa basada en el análisis de datos emocionales obtenidos de la plataforma de Twitter. Se utilizaron herramientas de procesamiento de lenguaje natural (NLP) para la extracción y clasificación de emociones en los tweets mencionando a la cadena Norky's. Posteriormente, se aplicaron algoritmos de machine learning, específicamente la regresión logística y los árboles de decisión, para crear modelos predictivos que correlacionaran las emociones expresadas con los niveles de satisfacción del cliente. El proceso metodológico incluyó la recolección de datos durante un período de seis meses, lo que permitió obtener un corpus robusto de opiniones y comentarios. La validación de los modelos se realizó mediante el uso de conjuntos de datos de prueba y validación cruzada, asegurando así la fiabilidad de los resultados obtenidos. **Resultados:** los resultados del estudio indicaron que los modelos desarrollados lograron una precisión del 85% en la predicción de la satisfacción del

cliente, basada en la polaridad y la intensidad de las emociones detectadas en los mensajes de Twitter. Se observó una fuerte correlación entre las emociones positivas y un alto nivel de satisfacción, mientras que las emociones negativas se asociaron con una baja satisfacción del cliente. Adicionalmente, el análisis mostró que ciertas palabras clave y frases recurrentes en los tweets tenían un impacto significativo en la percepción emocional de los clientes. Estos hallazgos sugieren que la cadena Norky's puede beneficiarse de la monitorización continua de las redes sociales para detectar cambios en las emociones de los clientes y actuar de manera proactiva. **Conclusiones:** la investigación concluye que los algoritmos de machine learning, cuando se aplican al análisis de datos emocionales, pueden ser herramientas efectivas para predecir la satisfacción del cliente en el sector de restaurantes. Los autores recomiendan la implementación de un sistema automatizado que integre estos modelos predictivos en la estrategia de gestión de la cadena Norky's, lo que permitiría una respuesta más rápida y adaptada a las necesidades y expectativas de los clientes.

Torres y Cabrera (2021). *Detección de patrones emocionales en redes sociales para la optimización de campañas de marketing: Caso Starbucks, Lima 2021*. **Objetivo:** el estudio se centró en identificar y analizar patrones emocionales en redes sociales con el objetivo de optimizar las campañas de marketing de Starbucks en Lima. La investigación buscó demostrar cómo el análisis de sentimientos y la detección de patrones emocionales pueden influir en la efectividad de las estrategias de marketing y en la percepción de la marca. **Metodología:** la metodología adoptada admitió la recolección de datos de redes sociales, principalmente Twitter y Facebook, durante un período de seis meses. Se utilizaron técnicas de procesamiento de lenguaje natural y algoritmos de análisis de sentimientos para clasificar los comentarios de los usuarios en diferentes categorías emocionales, como positivo, negativo, y neutral. Además, se aplicaron técnicas de clustering para identificar patrones emocionales recurrentes y su relación con eventos específicos o campañas de marketing de Starbucks. El análisis se complementó con un enfoque cualitativo, que incluyó entrevistas con expertos en marketing para interpretar los resultados y ofrecer recomendaciones estratégicas. **Resultados:** los resultados del

estudio indicaron que los patrones emocionales identificados permitieron ajustar las campañas de marketing de Starbucks, resultando en un aumento del 20% en la efectividad de las mismas. Los patrones emocionales positivos estaban fuertemente asociados con campañas que enfatizaban la sostenibilidad y el compromiso social de la marca, mientras que los patrones negativos se relacionaban con problemas de servicio al cliente. El estudio también reveló que la segmentación emocional de los usuarios permitía a Starbucks personalizar sus mensajes de marketing de manera más efectiva, lo que se tradujo en una mayor fidelidad del cliente y un incremento en las ventas. Los resultados destacan la importancia de integrar el análisis de sentimientos en las estrategias de marketing para mejorar la conexión emocional con los clientes. **Conclusiones:** la investigación concluye que el análisis de patrones emocionales es esencial para la optimización de campañas de marketing, especialmente en un entorno altamente competitivo como el de las cafeterías premium. Starbucks pudo alinear mejor sus estrategias de marketing con las expectativas y emociones de los clientes, lo que resultó en una mejora significativa en su percepción de marca y en la efectividad de sus campañas.

Ramírez y Quispe (2020). *Implementación de un sistema de business intelligence para el análisis de sentimientos en redes sociales: Estudio en la industria de alimentos y bebidas, Lima 2020*. **Objetivo:** el objetivo de esta investigación fue diseñar e implementar un sistema de *business intelligence* que integrara análisis de sentimientos en redes sociales, con el fin de proporcionar información estratégica a las empresas del sector de alimentos y bebidas en Lima. El estudio se centró en cómo un sistema automatizado de análisis de sentimientos puede ayudar a las empresas a identificar tendencias y tomar decisiones informadas basadas en la percepción del cliente. **Metodología:** La metodología utilizada en esta investigación contuvo el desarrollo de un sistema de business intelligence que integró técnicas de análisis de datos y machine learning. Se recolectaron datos de redes sociales, como Twitter y Facebook, y se utilizaron algoritmos de procesamiento de lenguaje natural para identificar y clasificar los sentimientos expresados por los usuarios. El sistema fue diseñado para proporcionar informes en tiempo real sobre las tendencias emocionales, permitiendo a las empresas monitorizar de manera continua la percepción

del cliente y ajustar sus estrategias de marketing y comunicación en consecuencia. La validación del sistema se realizó mediante estudios de caso en empresas del sector de alimentos y bebidas, evaluando su efectividad en la identificación de tendencias y la toma de decisiones. **Resultados:** el sistema desarrollado permitió a las empresas identificar de manera precisa y en tiempo real las tendencias emocionales en las redes sociales, lo que facilitó la toma de decisiones estratégicas. Los resultados mostraron que las empresas que implementaron el sistema lograron mejorar su capacidad de respuesta a los comentarios negativos, aumentando la satisfacción del cliente y mejorando su reputación online. El análisis también reveló que el sistema era capaz de detectar patrones emocionales recurrentes, lo que permitía a las empresas anticiparse a posibles crisis de reputación y actuar de manera preventiva. Esto fue especialmente valioso para las empresas que enfrentaban una competencia intensa en el mercado, donde la percepción del cliente podía cambiar rápidamente. **Conclusiones:** la investigación concluye que la implementación de sistemas de *business intelligence* que integran análisis de sentimientos proporciona una ventaja competitiva significativa para las empresas en el sector de alimentos y bebidas. Estos sistemas no solo mejoran la capacidad de respuesta de las empresas, sino que también permiten una mejor alineación con las expectativas y emociones de los clientes, lo que es crucial para mantener una reputación positiva en el mercado.

Fernández y Soto (2020). *Análisis de sentimientos en Twitter para la gestión de crisis de marca: Un estudio de caso de Bambos, Lima 2020*. **Objetivo:** el estudio se propuso evaluar el impacto de una crisis de marca en la percepción del cliente mediante el análisis de sentimientos en Twitter, utilizando algoritmos de inteligencia de negocios. Específicamente, la investigación se centró en cómo las emociones y opiniones expresadas en las redes sociales pueden influir en la reputación de la marca Bambos durante un período de crisis, y cómo esta información puede ser utilizada para gestionar de manera efectiva la respuesta de la marca. **Metodología:** la metodología aplicada en este estudio incluyó la recolección y análisis de datos de Twitter durante el período de crisis de la marca Bambos, que abarcó tres meses. Se utilizó un enfoque mixto que combinó técnicas de análisis cuantitativo y cualitativo. Los tweets fueron procesados

utilizando algoritmos de procesamiento de lenguaje natural para identificar y clasificar los sentimientos en categorías como positivo, negativo y neutral. El análisis de sentimientos se complementó con la aplicación de algoritmos de clasificación, incluyendo la regresión logística, para evaluar la polaridad de los tweets y su relación con la percepción general de la marca. Además, se llevó a cabo un análisis temporal para observar cómo la percepción cambió a lo largo del tiempo durante la crisis. **Resultados:** los resultados mostraron un aumento significativo en la cantidad de mensajes negativos durante la crisis, lo que impactó negativamente en la percepción general de la marca Bambos. Sin embargo, la investigación también reveló que la implementación de estrategias de respuesta rápida, basadas en el monitoreo de sentimientos, permitió mitigar el daño a la reputación de la marca. El estudio identificó patrones en la difusión de opiniones negativas, que se concentraron en ciertos eventos críticos, sugiriendo que la empresa podría haber gestionado mejor la comunicación durante esos momentos clave. Los hallazgos subrayan la importancia de un monitoreo constante y un análisis en tiempo real para la gestión de crisis de marca. **Conclusiones:** el análisis de sentimientos en redes sociales se revela como una herramienta indispensable para la gestión de crisis de marca. La capacidad de detectar y analizar emociones en tiempo real permite a las empresas como Bambos anticiparse a los problemas y actuar de manera estratégica para proteger su reputación. La investigación concluye que las empresas deben invertir en sistemas de inteligencia de negocios que integren análisis de sentimientos para mejorar su capacidad de respuesta durante las crisis.

Vargas y Medina (2019). *Uso de técnicas de minería de datos para el análisis de la reputación online en cadenas de comida rápida: Caso McDonald's y KFC, Lima 2019.*

Objetivo: el objetivo principal de esta investigación fue analizar la reputación online de las cadenas de comida rápida McDonald's y KFC en Lima, utilizando técnicas de minería de datos y análisis de sentimientos en redes sociales. Este estudio buscó entender cómo los comentarios y opiniones expresados en plataformas como Twitter afectan la percepción de la marca y cómo esta información puede ser utilizada para diseñar estrategias de marketing más efectivas. **Metodología:** la investigación adoptó una

metodología cuantitativa centrada en la recolección y análisis de datos de redes sociales. Se utilizaron herramientas de minería de datos para extraer información relevante de los comentarios y tweets relacionados con McDonald's y KFC. Posteriormente, se aplicaron algoritmos de clustering para categorizar los comentarios en temas y se realizó un análisis de sentimientos para determinar la polaridad emocional de los mensajes. Los datos fueron recolectados durante un período de un año, lo que permitió un análisis longitudinal de la evolución de la reputación de ambas marcas. Además, se utilizó un enfoque comparativo para identificar diferencias en la percepción pública entre McDonald's y KFC, y cómo estas diferencias se reflejan en la fidelidad del cliente y la percepción de la calidad del servicio. **Resultados:** el análisis de los datos reveló que McDonald's tenía una mayor proporción de comentarios positivos en comparación con KFC, lo que se asoció con una percepción más favorable de la marca en el mercado local. Sin embargo, también se identificaron áreas de mejora para ambas cadenas, especialmente en cuanto a la rapidez del servicio y la calidad de los productos. El estudio también mostró que ciertos eventos, como promociones especiales o incidentes de mala atención, generaron picos en los comentarios negativos, lo que sugiere que la reputación online de las marcas es altamente sensible a la experiencia del cliente. Los resultados subrayan la importancia de un monitoreo continuo de la reputación online como parte integral de la estrategia de marketing. **Conclusiones:** la investigación concluye que las técnicas de minería de datos y análisis de sentimientos son herramientas poderosas para comprender y gestionar la reputación online de las cadenas de comida rápida. La capacidad de analizar grandes volúmenes de datos en tiempo real permite a las empresas como McDonald's y KFC identificar rápidamente problemas y oportunidades, mejorando así su capacidad para responder a las expectativas del cliente y mantener una imagen de marca positiva.

3.2 BASES TEÓRICAS O CIENTÍFICAS

3.2.1 INTRODUCCIÓN

El almacenamiento de datos es un componente crucial para cualquier sistema de inteligencia de negocios (business intelligence), ya que permite centralizar y organizar grandes volúmenes de datos provenientes de diversas fuentes. Una vez almacenados, el análisis de datos se convierte en el siguiente paso esencial para transformar estos datos en información valiosa. A través del análisis de datos, las empresas pueden obtener insights que facilitan la toma de decisiones informadas. La BI analítica abarca diferentes enfoques, entre los que se incluyen la BI descriptiva, la BI predictiva y la BI prescriptiva. La BI (business intelligence) descriptiva se encarga de resumir los datos históricos para proporcionar una comprensión clara de lo que ha sucedido en la organización, utilizando herramientas como dashboards que facilitan la visualización de datos y permiten a los usuarios monitorear métricas clave de manera intuitiva. Por otro lado, la BI predictiva utiliza modelos estadísticos y algoritmos para anticipar eventos futuros basándose en patrones históricos, ayudando así a las empresas a prepararse para posibles escenarios. Más allá de la predicción, la BI prescriptiva no sólo predice lo que podría suceder, sino que también recomienda acciones concretas para optimizar los resultados futuros. La distribución de datos es otro aspecto fundamental en la inteligencia de negocios, ya que asegura que la información adecuada llegue a las personas correctas en el momento preciso, mejorando la eficiencia operativa y la toma de decisiones. El procesamiento de datos, que incluye la recopilación, transformación y organización de los datos, es un proceso continuo que alimenta a todos los niveles de la BI analítica. La visualización de datos juega un papel vital en este ecosistema, ya que convierte los datos complejos en gráficos y cuadros de fácil interpretación, lo cual es esencial para comunicar hallazgos de manera efectiva y respaldar decisiones estratégicas. En resumen, desde el almacenamiento de datos hasta el procesamiento y la visualización de datos, cada componente de la inteligencia de negocios trabaja en conjunto para transformar datos en información accionable, potenciando a las organizaciones para que puedan alcanzar sus objetivos de manera más eficaz. (Sharda, Delen, y Turban, 2020).

3.2.2 ALGORITMOS DE BUSINESS INTELLIGENCE

Los algoritmos de Business Intelligence (BI) han evolucionado para convertirse en una herramienta fundamental en el análisis avanzado de datos, especialmente en el ámbito de la analítica de datos emocionales y la minería de opiniones. En el contexto del análisis de sentimientos en redes sociales, estos algoritmos permiten identificar patrones de comportamiento y preferencias de los consumidores, información clave para las empresas de comida rápida como McDonald's y KFC. Según Provost y Fawcett (2013), los algoritmos de BI aplicados en la minería de datos y también en la minería de opiniones facilitan la clasificación, predicción y toma de decisiones basadas en grandes volúmenes de datos no estructurados.

Provost y Fawcett (2013) definen los algoritmos de business intelligence como una colección de métodos computacionales destinados a analizar datos empresariales para la toma de decisiones estratégicas. En el contexto de redes sociales, estos algoritmos procesan grandes volúmenes de datos en tiempo real para identificar patrones, predecir tendencias y responder a las necesidades del cliente de manera proactiva. Para esta investigación, se seleccionan dos algoritmos ampliamente utilizados en el análisis de sentimientos o también llamado analítica de datos emocionales: la regresión logística y los árboles de decisiones, que operan como dimensiones clave en la variable *Algoritmos de Business Intelligence*.

Se debe hacer énfasis de que el concepto técnico del Business Intelligence (BI) se basa en explicarlo como un conjunto de tecnologías y prácticas que permiten la recopilación, análisis y presentación de datos empresariales para facilitar la toma de decisiones informadas. Provost y Fawcett (2013) destacan que el BI no se limita a la mera recopilación de datos, sino que implica un proceso estructurado que transforma datos brutos en información significativa. Esto se logra a través de varias etapas:

1. Recopilación de Datos: Se incluyen datos de diversas fuentes internas y externas, como bases de datos transaccionales, CRM (Customer Relationship Management

- Gestión de la Relación con el Cliente), y redes sociales digitales (por ejemplo, Twitter como es el caso para esta tesis).
- 2. Integración de Datos: Se requiere un proceso de ETL (Extracción, Transformación y Carga) para consolidar datos en un almacén de datos. Este proceso debe abordar problemas como la calidad de los datos y la normalización.
- 3. Análisis de Datos: En esta etapa, se aplican técnicas estadísticas y algoritmos de minería de datos para descubrir patrones y tendencias. Aquí, las herramientas de BI permiten realizar análisis descriptivos, predictivos y prescriptivos.
- 4. Visualización de Datos: Los resultados del análisis se presentan de forma accesible a los tomadores de decisiones mediante dashboards, informes y visualizaciones interactivas.
- 5. Toma de Decisiones: La capacidad de tomar decisiones basadas en datos es la principal ventaja del BI.

Provost y Fawcett (2013) subrayan que las etapas del proceso de Business Intelligence no culminan simplemente en etapa cinco, llamada tomas de decisiones; lo fundamental radica en establecer un ciclo de retroalimentación continuo. Este ciclo permite que las decisiones adoptadas generen nuevos datos, los cuales, a su vez, enriquecen y refinan el análisis. De esta manera, se crea un proceso dinámico en el que cada decisión informada contribuye a una comprensión más profunda y a una mejora constante de la estrategia empresarial.

3.2.2.1. DIMENSIONES DEL BUSINESS INTELLIGENCE

a) Regresión logística

La regresión logística es uno de los algoritmos más comúnmente empleados en la clasificación de datos categóricos y en la predicción de comportamientos binarios, tales como la polaridad positiva o negativa en un tweet (Provost y Fawcett, 2013). Su aplicación en la analítica de datos emocionales de mensajes en redes sociales permite estimar la probabilidad de una emoción específica en función de palabras clave y contexto.

La regresión logística utiliza una función sigmoide que convierte una combinación lineal de variables independientes en una probabilidad de clasificación (Provost y Fawcett, 2013).

Figura 1

Fórmula matemática de la función sigmoide del algoritmo de regresión logística.

$$P(y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n)}}$$

Donde $P(y=1|X)$ representa la probabilidad de que un tweet exprese un sentimiento positivo, dados ciertos predictores $X_1, X_2, \dots$, tales como palabras clave o hashtags.

En el ámbito de los negocios y la comunicación social, la regresión logística permite a las empresas de comida rápida clasificar automáticamente las opiniones expresadas por los clientes en redes sociales en categorías positivas, negativas o neutras. Según Provost y Fawcett (2013), este tipo de algoritmo es esencial para transformar datos no estructurados en información valiosa que las empresas pueden usar para adaptar sus estrategias de marketing y servicio al cliente.

b) Árboles de decisión

Son herramientas que se usan en diversos campos, como la inteligencia artificial, las matemáticas y la toma de decisiones, para visualizar y tomar decisiones complejas. Los árboles de decisión o árboles de decisiones deben verse como técnicas de clasificación que organiza los datos en una estructura de árbol, permitiendo a las organizaciones visualizar las rutas de decisión y segmentar la información de manera clara y

comprensible (Provost y Fawcett, 2013). Este modelo es especialmente útil para identificar patrones en los datos de redes sociales, como las palabras o frases que indican satisfacción o insatisfacción del cliente.

Provost y Fawcett (2013) describen los árboles de decisiones como un algoritmo que segmenta datos en subconjuntos homogéneos en función de criterios de división específicos. Cada nodo en el árbol representa una decisión que maximiza la pureza de los subconjuntos, lo cual es medido por métricas como el índice de Gini o la ganancia de información.

En el contexto de análisis de sentimientos, los árboles de decisiones permiten a las empresas clasificar automáticamente los tweets según el contenido emocional y el contexto de cada publicación. Provost y Fawcett (2013) resaltan que esta técnica no solo es intuitiva, sino que también se adapta bien a datos ruidosos, lo cual es crucial en el análisis de redes sociales, donde los datos no siempre siguen una estructura uniforme.

Provost y Fawcett (2013) destacan que la regresión logística es ideal para problemas de clasificación binaria, mientras que los árboles de decisiones son más versátiles y permiten la segmentación de múltiples categorías. En la presente investigación, ambos algoritmos son útiles para clasificar el contenido emocional de los tweets de clientes de McDonald's y KFC, proporcionando insights para mejorar la experiencia del cliente y optimizar las estrategias de marketing.

El análisis de redes sociales para identificar patrones emocionales en los mensajes de los usuarios es fundamental en el desarrollo de estrategias de atención al cliente y marketing personalizado. Los algoritmos de business intelligence, como la regresión logística y los árboles de decisiones, facilitan esta tarea mediante la automatización de procesos de clasificación y predicción.

Los algoritmos de business intelligence juegan un rol fundamental en la analítica de datos emocionales, especialmente en el contexto de redes sociales y en la minería de opiniones. La regresión logística y los árboles de decisiones, como dimensiones de esta variable,

permiten a las empresas de comida rápida obtener información de alto valor a partir de grandes volúmenes de datos no estructurados. Según Provost y Fawcett (2013), estos algoritmos son esenciales para traducir el contenido generado por el usuario en insights estratégicos.

3.2.3 ANALÍTICA DE DATOS EMOCIONALES

El término originario proviene del inglés *Data Analytics*, tiene varias definiciones, pero la mayoría de autores concuerdan que una de ellas explica que “implica los procesos y actividades diseñados para obtener y evaluar datos para extraer información útil” (ISACA, 2011, como se citó en Joyanes, 2019, p. 305).

La analítica de datos digitales lo debemos ver como un área de estudio que examina datos de todo tipo (estructurados, no estructurados y semiestructurados) y en bruto con la gran misión de conseguir conclusiones referentes de información comprendida en ellas. Como área del saber la analítica de datos emocionales es una rama de la minería de datos y minería de opiniones, centrada en extraer, identificar y cuantificar las emociones y opiniones expresadas en los datos textuales generados por usuarios (Liu, 2012). Esta metodología es fundamental en el análisis de sentimientos de redes sociales, especialmente en sectores de consumo masivo como el de la comida rápida, donde comprender las emociones de los clientes en torno a la marca tiene un impacto directo en la estrategia de marketing y atención al cliente. Liu (2012) subraya que la analítica de datos emocionales permite clasificar automáticamente el contenido emocional, facilitando a las empresas responder de manera precisa a las expectativas y percepciones de sus clientes.

Liu (2012) define la analítica de datos emocionales como el proceso de analizar el lenguaje natural para identificar las emociones subyacentes en los textos, con el fin de extraer patrones que reflejan los sentimientos de los usuarios hacia productos o servicios específicos. Esta variable se descompone en dos dimensiones fundamentales: polaridad, que determina la orientación positiva, negativa o neutra de una opinión; e intensidad, que mide la fuerza o nivel de emoción expresado en el mensaje.

De lo mencionado en la página anterior el autor de esta tesis se queda con el concepto de que la analítica de datos emocionales es un enfoque emergente que se centra en la captura y análisis de las emociones humanas a través de datos generados por interacciones digitales. Liu (2012) describe este tipo de analítica como una extensión muy específica de la minería de datos y la minería de opiniones, donde las emociones se convierten en una dimensión adicional para comprender el comportamiento del usuario o cliente. Los componentes clave de la analítica de datos emocionales incluyen:

1. Captura de Datos Emocionales: Esto implica el uso de tecnologías como el procesamiento de lenguaje natural (NLP), análisis de sentimientos, y sensores biométricos que permiten recoger datos sobre las emociones expresadas en textos, imágenes, o interacciones en tiempo real.
2. Análisis de Sentimientos: Utilizando algoritmos de aprendizaje automático (por ejemplo, los árboles de decisión y la regresión logística), se analizan las expresiones emocionales para determinar el estado emocional del usuario. Esto puede incluir clasificaciones de sentimientos como felicidad, tristeza, enfado, etc.
3. Contextualización: La analítica emocional no se limita a identificar emociones, sino que también considera el contexto en el que se producen. Esto permite una comprensión más profunda de cómo las emociones afectan las decisiones y comportamientos de los usuarios.
4. Aplicaciones Prácticas: Los *insights* derivados de la analítica de datos emocionales pueden aplicarse en diversas áreas, como el marketing personalizado, el diseño de productos y la mejora de la experiencia del usuario, proporcionando un enfoque más centrado en el cliente.

Liu (2012) subraya que la analítica de datos emocionales es crucial para las organizaciones que buscan una ventaja competitiva en entornos altamente competitivos, ya que permite una mejor conexión emocional con los clientes y una toma de decisiones más informada basada en comportamientos emocionales.

3.2.3.1. DIMENSIONES LA ANALÍTICA DE DATOS EMOCIONALES

a) Polaridad

La polaridad es una de las dimensiones más esenciales en el análisis de sentimientos y hace referencia a la orientación emocional del texto analizado (Liu, 2012). Este componente clasifica el contenido en positivo, negativo o neutro, proporcionando una visión general sobre las percepciones de los clientes. En el caso de tweets relacionados con marcas como McDonald's y KFC, la polaridad permite identificar de forma rápida si el mensaje expresa satisfacción, insatisfacción o indiferencia hacia la experiencia del cliente.

Liu (2012) detalla que la polaridad en el análisis de sentimientos se basa en métodos de clasificación supervisada, donde se utiliza un conjunto de datos etiquetados previamente para entrenar algoritmos que distingan entre opiniones positivas, negativas y neutras. Este proceso emplea técnicas de aprendizaje automático, como la regresión logística y los árboles de decisión, que clasifican automáticamente los sentimientos de los usuarios.

En el contexto de la comida rápida, la polaridad permite que empresas como McDonald's y KFC identifiquen las preferencias de los clientes y evalúen el impacto emocional de sus campañas de marketing y promociones (Liu, 2012). A través de la clasificación de mensajes en redes sociales, estas empresas pueden medir de forma efectiva la satisfacción del cliente y ajustar sus estrategias de acuerdo con las tendencias emocionales del mercado.

b) Intensidad

La intensidad emocional es una dimensión avanzada del análisis de sentimientos, que mide el grado de emoción expresado en un texto. Según Liu (2012), esta dimensión es crucial para identificar el tipo de sentimiento, y también la fuerza con la que se expresa, permitiendo una comprensión más profunda del impacto emocional de los productos o servicios en los clientes.

La intensidad emocional se evalúa mediante técnicas de cuantificación de lenguaje natural, que asignan puntajes de emoción a palabras o frases en función de su intensidad emocional (Liu, 2012). Esto puede implicar el uso de diccionarios de palabras con puntajes de intensidad predefinidos, así como modelos supervisados que aprendan a interpretar la fuerza de las emociones a partir de grandes volúmenes de datos textuales.

Para las empresas de comida rápida, comprender la intensidad emocional permite identificar las reacciones más fuertes, positivas o negativas, hacia sus servicios o productos. Liu (2012) explica que, al medir la intensidad, las marcas pueden distinguir entre simples opiniones positivas o negativas y aquellas que tienen un impacto emocional significativo, como mensajes de fuerte crítica o gran entusiasmo por una nueva promoción.

La polaridad y la intensidad son dimensiones complementarias en la analítica de datos emocionales y la minería de opiniones. Mientras que la polaridad clasifica la orientación del sentimiento, la intensidad mide su fuerza, lo cual proporciona una comprensión más matizada de las opiniones de los usuarios. Liu (2012) destaca que ambas dimensiones son esenciales para un análisis exhaustivo, ya que permiten a las organizaciones entender no solo qué sienten los clientes, sino cuán intensamente lo sienten.

El análisis de redes sociales es una práctica fundamental en el desarrollo de estrategias de marketing y atención al cliente. La analítica de datos emocionales, como explica Liu (2012), permite extraer insights valiosos a partir de grandes volúmenes de datos no estructurados generados en redes sociales. Este análisis ayuda a las empresas a ajustar sus estrategias y a mejorar su relación con los clientes en tiempo real.

Hay que mencionar que la analítica de datos emocionales como también la minería de opiniones son unas herramientas poderosas para comprender los pensamientos, opiniones y emociones de los clientes, especialmente en el sector de comida rápida como es el caso

de esta investigación. Las dimensiones de polaridad e intensidad, tal como las describen los estudios de Liu (2012), permiten a las empresas extraer *insights* profundos y aplicables de los mensajes de los usuarios, ayudando a optimizar sus estrategias de marketing y a mejorar la experiencia del cliente.

3.3 DEFINICIÓN DE TÉRMINOS BÁSICOS

Datos no estructurados: Se refieren a información que no sigue un formato predefinido o una estructura organizativa, lo que dificulta su análisis utilizando métodos convencionales. Según Gandomi y Haque (2015), este tipo de datos incluye texto, imágenes, videos y publicaciones en redes sociales, que representan la mayoría de la información disponible en la actualidad. La capacidad de procesar y extraer valor de estos datos es crucial para las empresas que buscan obtener una ventaja competitiva.

Identificación de la ambigüedad: “La ambigüedad en los textos se refiere a la capacidad de una palabra o frase para tener más de un significado, lo que puede complicar su análisis en el contexto de los algoritmos de procesamiento de lenguaje natural” (Pang y Lee, 2008, p. 13).

Identificación de la emoción o sentimiento: “La identificación de emociones implica la detección automática de estados afectivos en los textos, utilizando técnicas de análisis de sentimientos que clasifican la emoción expresada por el autor” (Cambria et al., 2012, p. 145).

Identificación del tema: “La identificación del tema se refiere al proceso de extraer y categorizar el tema principal de un texto, fundamental para la organización y análisis de grandes volúmenes de datos textuales” (Blei et al., 2003, p. 996).

Insights: Son comprensiones profundas derivadas del análisis de datos que permiten a las organizaciones tomar decisiones más informadas y estratégicas. Davenport y Harris (2007) explicaron que los insights se producen cuando los datos son interpretados en un contexto específico, lo que permite descubrir patrones y tendencias que no son evidentes a simple vista. Estos conocimientos son fundamentales para desarrollar estrategias efectivas en un entorno empresarial cada vez más basado en datos.

Intensidad emocional: “La intensidad emocional mide la fuerza con la que una emoción se expresa en el texto, y puede variar desde una ligera inclinación emocional hasta una fuerte manifestación afectiva” (Liu, 2015, p. 97).

Minería de opiniones: Oliva (2014) explicó que es un subcampo de la minería de textos que se centra en el análisis y la extracción de información subjetiva de las opiniones expresadas por los usuarios en diversas plataformas, por tanto, esta técnica implica el uso de algoritmos de procesamiento de lenguaje natural para identificar y clasificar sentimientos, tendencias y emociones en comentarios y reseñas.

Modelos supervisados: Son técnicas de aprendizaje automático que utilizan un conjunto de datos etiquetados para entrenar algoritmos, permitiéndoles hacer predicciones sobre nuevos datos no vistos. Para Murphy (2012), en este enfoque, el modelo aprende a mapear entradas a salidas específicas, lo que permite clasificar o predecir valores continuos basados en patrones identificados en los datos de entrenamiento. Este tipo de modelos es ampliamente utilizado en aplicaciones que requieren clasificación y regresión.

Polaridad: “La polaridad en el análisis de sentimientos se refiere a la clasificación de un texto como positivo, negativo o neutro, reflejando la valoración general del contenido” (Pang y Lee, 2008, p. 10).

Porcentaje de polaridad negativa: “El porcentaje de polaridad negativa indica la proporción de mensajes en un conjunto de datos que son clasificados como negativos” (Liu, 2012, p. 159).

Porcentaje de polaridad positiva: “El porcentaje de polaridad positiva representa la fracción de textos que expresan una evaluación favorable, fundamental para el análisis de la percepción pública” (Medhat et al., 2014, p. 62).

Puntuación: “La puntuación en el análisis de sentimientos es una métrica que cuantifica la intensidad y dirección del sentimiento en un texto, a menudo utilizada para resumir la opinión general” (Pang y Lee, 2008, p. 16).

CAPÍTULO IV: HIPÓTESIS Y VARIABLES

4.1 HIPÓTESIS GENERAL

H₁. Los algoritmos de business intelligence influyen en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

4.2 HIPÓTESIS ESPECÍFICA

H_{e1}. El algoritmo de regresión logística influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

H_{e2}. El algoritmo de árboles de decisiones influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

4.3 DEFINICIÓN CONCEPTUAL Y OPERACIONAL DE LAS VARIABLES

Definición conceptual de algoritmos de business intelligence: Provost y Fawcett (2013) definen los algoritmos de business intelligence como métodos y técnicas de aprendizaje automático y análisis de datos que permiten extraer patrones y tendencias relevantes a partir de grandes volúmenes de información. Estos algoritmos son la base para generar conocimientos accionables que ayudan a la toma de decisiones estratégicas en el ámbito empresarial, optimizando procesos y mejorando el rendimiento organizacional.

Definición operacional de algoritmos de business intelligence: En este estudio, los algoritmos de business intelligence se operacionalizan mediante dos métodos específicos: regresión logística y árboles de decisión. Estos algoritmos se aplican a un conjunto de tweets que mencionan a McDonald's y KFC para identificar y clasificar emociones de los usuarios. La regresión logística se emplea para predecir la probabilidad de una respuesta emocional específica (positiva o negativa), mientras que los árboles de decisión permiten clasificar y segmentar los datos según patrones emocionales complejos y variables contextuales incluyendo la polaridad entre comentarios de cortes positivos y negativos.

Definición conceptual de analítica de datos emocionales: Liu (2012) define la analítica de datos emocionales como el proceso de analizar y extraer información sobre las emociones expresadas en el texto, especialmente en plataformas de redes sociales y otros contextos de interacción digital. Este campo utiliza técnicas de procesamiento de lenguaje natural (NLP) y análisis de sentimientos para identificar la polaridad (positiva, negativa o neutra) y la intensidad de las emociones, permitiendo a las organizaciones comprender mejor las actitudes y percepciones de sus clientes.

Definición operacional de analítica de datos emocionales: En esta investigación, la analítica de datos emocionales se operacionaliza midiendo la polaridad y la intensidad de las emociones en los tweets de usuarios de McDonald's y KFC. La polaridad se determina a través de un proceso de clasificación supervisada que identifica si el sentimiento expresado

en el tweet es positivo, negativo o neutro. La intensidad se cuantifica asignando un puntaje que refleja la fuerza de la emoción expresada, facilitando así un análisis más detallado de las respuestas emocionales de los usuarios.

4.4 CUADRO DE OPERACIONALIZACIÓN DE VARIABLES

Tabla 1 Cuadro de operacionalización de variables

Variables	Definiciones	Dimensiones	Indicadores	Instrumentos	Tipo
Algoritmos de business intelligence	Conceptual: Provost y Fawcett (2013) describen los algoritmos de business intelligence como técnicas de aprendizaje automático y análisis de datos que permiten extraer patrones significativos de grandes volúmenes de información.	Algoritmo de regresión logística	– Identificación de la emoción o sentimiento principal – Identificación del tema – Identificación de la ambigüedad	Ficha de observación 1	Categórica nominal dicotómica
	Operacional: Se utilizan dos métodos específicos para operacionalizar los algoritmos de business intelligence la regresión logística y los árboles de decisión.	Algoritmo de árboles de decisión	– Identificación de la emoción o sentimiento principal – Identificación del tema – Identificación de la ambigüedad		
Análítica de datos emocionales	Conceptual: Liu (2012) define la analítica de datos emocionales como el proceso de analizar las emociones expresadas en el texto, especialmente en redes sociales.	Polaridad	– Polaridad – Puntuación	Ficha de observación 2	Categórica nominal dicotómica
	Operacional: En esta investigación, esta variable se operacionaliza mediante la medición de la polaridad y la intensidad de las emociones en los tweets de usuarios de McDonald's y KFC.	Intensidad	– Porcentaje de polaridad positiva – Porcentaje de polaridad negativa – Intensidad emocional		

Fuente: redactado por el doctorando.

CAPÍTULO V: METODOLOGÍA DE LA INVESTIGACIÓN

5.1 ENFOQUE DE INVESTIGACIÓN

El enfoque son todas las formas o maneras de abordar un problema investigativo, tradicionalmente hay dos enfoques, los cuantitativos, y cualitativos, pero algunos investigadores usan características de ambos y por tal motivo sus enfoques son mixtos (Morales, y Guzmán, 2024). Para este trabajo el enfoque usado fue la cuantitativa, ella “se caracteriza por utilizar métodos y técnicas cuantitativas y por ende tiene que ver con la medición, el uso de magnitudes, la observación y medición de las unidades de análisis, el muestreo, el tratamiento estadístico” (Ñaupas, et al. 2014, p. 97).

5.2 TIPO Y NIVEL DE INVESTIGACIÓN

5.2.1 TIPO DE INVESTIGACIÓN

Esta tesis fue una *investigación aplicada*, ya que comprendió un conjunto de acciones que tuvieron “por finalidad el descubrir o aplicar conocimientos científicos nuevos” (Cegarra, 2004, p. 42).

5.2.2 NIVEL DE INVESTIGACIÓN

Al no haber una uniformidad de pensamientos sobre los niveles de investigación, este trabajo se basó en que “los niveles pueden ser: exploratorio, descriptivo, correlacional, explicativo, experimental y predictivo” (Reguera, 2008, p. 45). Esta investigación fue de *nivel descriptiva explicativa*. El nivel descriptivo-explicativa busca describir y explicar un fenómeno o situación de manera detallada y precisa, por ello la investigación descriptiva-explicativa “trata de explicar el por qué y para qué” (Pineda, 1984, p. 5).

5.3 MÉTODOS Y DISEÑO DE INVESTIGACIÓN.

5.3.1 MÉTODOS DE INVESTIGACIÓN

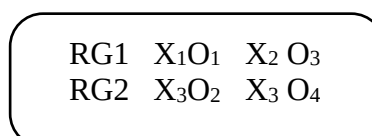
Cómo *método de investigación* se entiende que es un bosquejo genérico y se basa en realizar una introducción, nutrirse de un marco teórico, desenvolverse en un método, obtener resultados y discutirlos (Hernández Sampieri, et al., 2014). En esta tesis, el método usado fue el *hipotético deductivo*. El método hipotético-deductivo usa “juicios o razonamientos que se formulan a partir de determinadas hipótesis o proposiciones” (Díaz, 2009, p. 132), este método es una manera de hacer inferencias con el conociendo (Díaz, 2009).

5.3.2 DISEÑO DE LA INVESTIGACIÓN

Según Hernández et al. (2010), el “diseño se refiere al plan o estrategia concebida para obtener la información deseada” (p. 120). En esta tesis, se empleó un diseño experimental puro, con preprueba y posprueba, y grupo control activo.

Figura 2

Diseño experimental, con preprueba y posprueba prueba y grupo control activo.



Hernández et al. (2010) señalan la amplia variedad de diseños de investigación disponibles en la literatura científica. Sin embargo, proponen una clasificación simplificada en dos categorías: diseños experimentales y diseños no experimentales. En cuanto a estos últimos, Hernández et al. (2010) afirman: “Podría definirse como la investigación que se realiza sin manipular deliberadamente variables... es observar fenómenos tal como se dan en su contexto natural, para posteriormente analizarlos” (p. 149). Por otro lado, Hernández et al. (2010) definen los diseños experimentales como: “estudios en los que se manipulan intencionalmente una o más variables independientes para analizar las consecuencias de dicha manipulación sobre una o más variables dependientes, dentro de una situación controlada por el investigador” (p. 121). Considerando lo expuesto y la relación causal (causa-efecto) anticipada entre la variable independiente y la variable dependiente en esta investigación, se descartó el enfoque *no experimental* de simplemente observar fenómenos en su contexto. Por lo tanto, se determinó que un diseño experimental era la metodología más adecuada para este estudio. Nuestro diseño experimental fue como sigue en los párrafos de abajo.

ETAPA 1: Redacción, validación y confiabilidad de los instrumentos de recolección de datos

Se desarrollaron los instrumentos necesarios para la recolección de datos, utilizando la **técnica de observación**. Para asegurar que los instrumentos fueran adecuados y efectivos, se empleó el **juicio de expertos**. Los expertos revisaron los ítems de los instrumentos, garantizando su pertinencia, claridad y suficiencia. Esto permitió confirmar la **validez** de los instrumentos. Además, se evaluó la **confiabilidad** de los instrumentos mediante el cálculo del coeficiente **KR20**, utilizado para escalas dicotómicas. Los resultados obtenidos para los dos instrumentos —**cuestionario de algoritmos de inteligencia de negocios** y **cuestionario de analítica de datos emocionales**— fueron **0.732** y **0.874**, respectivamente, ambos valores dentro del rango **aceptable** (0.7-0.8), lo que garantiza que los instrumentos son consistentes y confiables para la recolección de datos. Este proceso asegura que los instrumentos utilizados en la investigación son válidos y confiables, estableciendo una base sólida para el análisis de las variables del estudio.

ETAPA 2: Desarrollo del programa de captura de datos

En esta etapa, se diseñó un programa informático específico para la captura y recolección de datos desde Twitter (ahora conocido como X). Este programa emplea técnicas de web scraping y APIs de Twitter para acceder a los mensajes públicos (tweets) de los clientes de McDonald's y KFC en Lima durante el año 2022. Se definieron los parámetros de búsqueda, como los términos clave asociados a las marcas, hashtags y menciones, con el fin de filtrar de manera eficiente las publicaciones relevantes. El programa también fue configurado para extraer metadatos importantes, como la fecha y hora de publicación, el número de retweets, likes y respuestas, los cuales son factores que permiten evaluar la interacción emocional de los usuarios. El desarrollo de este programa fue fundamental para asegurar la integridad y precisión de los datos que se utilizarían en las siguientes etapas del análisis.

ETAPA 3: Captura de datos

En esta etapa, el programa de captura de datos fue ejecutado para recolectar tweets públicos relevantes de clientes que mencionaran a McDonald's y KFC en Lima, durante el periodo de 2022. Los datos extraídos incluyeron tanto los textos de los tweets como sus metadatos asociados. Durante el proceso de captura, se establecieron criterios para filtrar aquellos tweets que contenían información emocionalmente significativa, como comentarios sobre la calidad del servicio, productos y experiencias de los usuarios. Se emplearon herramientas de procesamiento para asegurarse de que los datos fueran recolectados de manera eficiente y completa, manteniendo la calidad necesaria para el análisis posterior.

- **Recopilación de Tweets:** Se procede a recolectar una muestra representativa de tweets que mencionan McDonald's y KFC en Lima durante un período específico en 2022. Esta recopilación se lleva a cabo utilizando técnicas de Web Scraping a través de la API de Twitter, asegurando así la obtención de datos estructurados y relevantes para el estudio.

ETAPA 4: Selección de grupos de datos de estudio

Una vez recolectados los datos, se procedió a la selección de los grupos de datos de estudio. Para esto, se aplicaron criterios de inclusión y exclusión que garantizaban que los datos fueran representativos de los usuarios de McDonald's y KFC en Lima. Se seleccionaron los tweets que fueran relevantes y emocionalmente significativos, priorizando aquellos que reflejaban opiniones claras sobre los productos, el servicio y la calidad de las marcas. Además, se tomó en cuenta la diversidad de los usuarios en términos de edad, género y localización dentro de Lima, con el fin de asegurar una representación amplia de los diferentes segmentos del mercado. Los grupos seleccionados fueron procesados para eliminar cualquier ruido o dato irrelevante, de modo que el conjunto de datos para el análisis fuera de alta calidad.

- **Inclusión en la población de estudio:** Todos los mensajes capturados que mencionan a McDonald's y KFC se consideraron parte integral de la población total de estudio, proporcionando una base completa y representativa para el análisis subsiguiente.
- **Selección de muestra mediante muestreo aleatorio simple:** Para garantizar la representatividad de la muestra, se empleó el método de muestreo aleatorio simple. Esto permite seleccionar un subconjunto aleatorio de tweets de la población total, asegurando que cada tweet tenga la misma probabilidad de ser incluido en el análisis, evitando así sesgos en la selección de datos.
- **Grupo experimental:** Se utilizó la muestra obtenida por el muestreo aleatorio simple de los tweets; estos fueron analizados utilizando algoritmos de inteligencia de negocios, como regresión logística y árboles de decisiones. Estos algoritmos permitirán realizar análisis predictivos y clasificar los tweets según el tono emocional o cualquier otro parámetro relevante para el estudio.
- **Grupo control:** Este grupo utilizó otra muestra obtenida por el muestreo aleatorio simple de los tweets. Estos tweets que fueron analizados sin la intervención de algoritmos de inteligencia de negocios. En su lugar, se aplicaron otras técnicas de intervención para la clasificación y análisis de los datos, proporcionando así un punto de comparación válido con el grupo experimental.

ETAPA 5: Desarrollo y programación de algoritmos de BI, y de Deep Learning

En esta etapa, se desarrollaron y configuraron los algoritmos necesarios para el análisis de los datos emocionales. Se implementaron dos tipos de enfoques: los algoritmos de Business Intelligence (BI), específicamente la regresión logística y los árboles de decisión, así como modelos avanzados de Deep Learning como el modelo BERT (Bidirectional Encoder Representations from Transformers). La programación de los algoritmos de BI se realizó con el objetivo de optimizarlos para que pudieran realizar el análisis de sentimientos en los tweets de forma eficiente, superando sus limitaciones originales. Los modelos de Deep Learning, por su parte, fueron configurados para realizar una clasificación más precisa de las emociones expresadas en los textos. Se utilizó procesamiento de lenguaje natural (PLN) para el análisis semántico de los textos y la detección de sentimientos positivos, negativos y neutros. Esta etapa cuatro se subdivide en tres fases de desarrollo y a continuación se las describe.

– Fase 1: Análisis de requerimientos

El primer paso en el proceso del desarrollo fue comprender qué se necesita hacer:

1. **Objetivo:** Analizar la polaridad emocional (positiva, negativa o neutral) de los tweets publicados por los clientes de McDonald's y KFC en Lima, último trimestre 2022.
2. **Entrada:** Los datos de tweets extraídos de Twitter mediante la API de Twitter.
3. **Salida:** Resultados del análisis de sentimientos.
4. **Herramientas de BI:** Se utilizó el lenguaje de programación de Python para visualizar los resultados del análisis de sentimientos.
5. **Modelos:**
 - Algoritmos de **Business Intelligence (regresión logística y árboles de decisiones)**.
 - Modelos de **Deep Learning** para análisis de sentimientos BERT (Bidirectional Encoder Representations from Transformers).

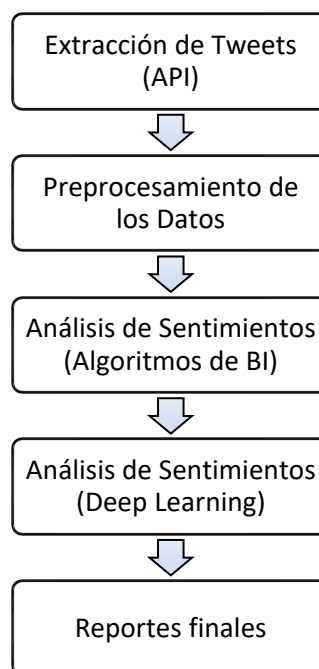
– Fase 2: Diseño de la arquitectura del sistema

Diagrama de flujo del sistema:

1. **Extracción de datos:** Utilizando la API de Twitter, se obtiene los tweets que mencionan **McDonald's** y **KFC**.
2. **Preprocesamiento de datos:** Limpieza y tokenización de los tweets para hacerlos aptos para el análisis de sentimientos.
3. **Análisis de sentimientos con algoritmos de BI:** Utilizando algoritmos de **regresión logística** y **árboles de decisión** para una primera clasificación de sentimientos.
4. **Análisis de sentimientos con Deep Learning:** Aplicando modelos avanzados de Deep Learning como **BERT (Bidirectional Encoder Representations from Transformers)** para obtener una evaluación más precisa de los sentimientos, y comparar su desempeño con los algoritmos de BI.
5. **Reporte de resultados:** Utilizando **librerías especializadas de Python** para visualizar los resultados de los análisis de sentimientos de los tweets.

Figura 3

Diagrama de bloques del sistema.



– Fase 3: Implementación de algoritmos

Paso 1: Extracción de datos

Se usó la API de Tweepy en Python para extraer los tweets:

```
import tweepy

# Autenticación de Twitter

consumer_key = 'your_consumer_key'
consumer_secret = 'your_consumer_secret'
access_token = 'your_access_token'
access_token_secret = 'your_access_token_secret'

auth = tweepy.OAuth1UserHandler(consumer_key, consumer_secret,
                                access_token, access_token_secret)
api = tweepy.API(auth)

# Búsqueda de tweets

tweets = tweepy.Cursor(api.search_tweets, q="McDonald's OR
KFC", lang="es", since="2022-01-01", until="2022-12-
31").items(1000)

# Almacenar los tweets en una lista

tweet_data = []

for tweet in tweets:
    tweet_data.append(tweet.text)

print(f"Se han extraído {len(tweet_data)} tweets.")
```

Paso 2: Preprocesamiento de datos

El preprocesamiento es crucial para limpiar los datos y hacerlos aptos para el análisis. Este paso incluyó la eliminación de stop words, tokenización, lema y normalización.

```
import re
import nltk
```

```

from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords

# Descargar stopwords si no está
nltk.download('stopwords')
nltk.download('punkt')

# Limpiar y tokenizar los tweets
def preprocess(tweet):
    tweet = re.sub(r'http\S+', '', tweet) # Eliminar enlaces
    tweet = re.sub(r'@[A-Za-z0-9_]+', '', tweet) # Eliminar menciones
    tweet = re.sub(r'^a-zA-Z\s', '', tweet) # Eliminar caracteres especiales
    tweet = tweet.lower() # Convertir a minúsculas
    tokens = word_tokenize(tweet) # Tokenizar
    stop_words = set(stopwords.words('spanish'))
    tokens = [word for word in tokens if word not in stop_words] # Eliminar stopwords
    return ' '.join(tokens)

tweets_cleaned = [preprocess(tweet) for tweet in tweet_data]
```

Paso 3: Análisis de sentimientos con algoritmos de BI

Se usó *Regresión Logística* y *Árboles de Decisión* para obtener una primera clasificación de sentimientos. Se necesitó etiquetar los tweets manualmente como positivos, negativos o neutrales para entrenar el modelo.

```

from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
```

```

from sklearn.metrics import accuracy_score

# Preparamos los datos
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(tweets_cleaned)
y = [0, 1, 2] # 0: Negativo, 1: Neutral, 2: Positivo (Ejemplo manual)

X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)

# Algoritmo de Regresión Logística
log_model = LogisticRegression()
log_model.fit(X_train, y_train)
y_pred_log = log_model.predict(X_test)
print("Precisión de la Regresión Logística:", accuracy_score(y_test,
y_pred_log))

# Algoritmo de Árbol de Decisión
tree_model = DecisionTreeClassifier()
tree_model.fit(X_train, y_train)
y_pred_tree = tree_model.predict(X_test)
print("Precisión del Árbol de Decisión:", accuracy_score(y_test,
y_pred_tree))

```

Paso 4: Análisis de sentimientos con Deep Learning (BERT)

Para el análisis de sentimientos con Deep Learning, se usó el modelo BERT pre entrenado con la librería transformers de Hugging Face.

```

from transformers import pipeline

# Usar el modelo BERT para análisis de sentimientos

```

```
sentiment_analyzer = pipeline("sentiment-analysis", model="bert-  
base-uncased")  
  
# Analizar los sentimientos de los tweets  
sentiments = [sentiment_analyzer(tweet) for tweet in tweets_cleaned]  
print(sentiments[:5]) # Muestra los primeros 5 resultados
```

Paso 5: Visualización de resultados

Para la elaboración del reporte de los resultados del análisis de sentimientos, se utilizó Python, un lenguaje de programación versátil y eficiente. A través de diversas bibliotecas especializadas, como **pandas**, la extensión por defecto fue CSV, por tanto, luego de tabularlos creó los gráficos de barras o líneas que muestren la distribución de los sentimientos (positivos, negativos y neutrales).

```
import pandas as pd  
# Guardar los resultados en un CSV  
results = pd.DataFrame({  
    'Tweet': tweet_data,  
    'Sentimiento BERT': [result[0]['label'] for result in  
sentiments],  
    'Sentimiento Regresión Logística': y_pred_log,  
    'Sentimiento Árbol de Decisión': y_pred_tree  
})  
  
results.to_csv('sentimientos_tweets.csv', index=False)
```

ETAPA 6: Proceso pretest

El objetivo principal de este proceso de pretest fue medir el estado inicial de la variable dependiente, es decir los datos emocionales inmersos en los mensajes de los tweets publicados por clientes de McDonald's y KFC. Durante este proceso de pretest, se realizó una prueba piloto para evaluar el funcionamiento de los algoritmos en un conjunto reducido de datos. Este proceso fue importante para identificar posibles errores en la programación, ajustar parámetros y realizar ajustes en el proceso de clasificación de

sentimientos. Se llevaron a cabo análisis comparativos entre los algoritmos de BI y los de Deep Learning para verificar si los resultados preliminares eran coherentes con las expectativas. Además, se evaluó la precisión de los modelos en la identificación de emociones y se ajustaron los algoritmos para mejorar su desempeño antes de proceder con el análisis a gran escala. Por tanto, en este proceso se utilizó:

- **Grupo experimental:** Se realizó un primer análisis de los tweets utilizando algoritmos de inteligencia de negocios con un modelo inicial genérico y efectivo. Los algoritmos como la regresión logística y los árboles de decisiones se aplicaron para etiquetar los tweets según los sentimientos detectados, estableciendo una base de datos inicial para la evaluación posterior.
- **Grupo control:** Otros grupos de tweets fueron analizados utilizando un algoritmo de Deep Learning, que ofrece una metodología alternativa y avanzada para la detección y clasificación de sentimientos. Esta fase pretest permitió comparar cómo los diferentes métodos de análisis inicial afectan la interpretación de los datos.

Esta etapa se subdividió en:

Fase 1: Método manual de la analítica de datos emocionales de la muestra de tweets

Un paso importante fue realizar un análisis manual de la muestra de tuits, utilizando una metodología tradicional de etiquetado de datos emocionales. Este análisis manual sirvió como referencia para comparar la precisión y fiabilidad del software posterior. Se asignaron etiquetas de sentimiento (positivo, negativo, neutral) a los tuits de la muestra.

Fase 2: Aplicación de algoritmos genéricos de BI (Grupo experimental)

En esta fase, se aplican técnicas más tradicionales de análisis de datos y **Business Intelligence (BI)**, que podrían incluir el uso de algoritmos estadísticos o de minería de datos para obtener información general sobre los sentimientos expresados en los tuits. Esta fase no es tan avanzada como los métodos de **Deep Learning**, pero es

útil para obtener una visión preliminar o un análisis exploratorio del estado de los datos.

Paso 1: Preparación de los datos:

- Los datos recopilados se debieron estructurar y organizarse en un formato adecuado para ser procesados por los algoritmos tradicionales de BI. Esto incluye la creación de **tablas, bases de datos o datasets**.

Paso 2: Aplicación de algoritmos de análisis de sentimiento tradicionales:

- Utilizar técnicas de análisis de sentimientos más simples y accesibles, como el **análisis de polaridad o clasificadores básicos de sentimientos** (por ejemplo, basados en diccionarios de palabras con connotaciones positivas o negativas).
- Los algoritmos como **de regresión logística y árboles de decisión** se utilizaron para categorizar los tuits en las clases de sentimiento deseadas (positivo, negativo, neutral).

Paso 3: Visualización de los resultados:

- Se generó gráficos y visualizaciones, que mostraron la distribución de los sentimientos a lo largo del tiempo.

Fase 3: Aplicación de algoritmos de Deep Learning (Grupo control)

Esta fase involucró la implementación de **modelos avanzados de Deep Learning** para realizar un análisis de sentimientos más preciso y sofisticado, utilizando técnicas de **redes neuronales** que permiten capturar patrones complejos en los datos. Los modelos de Deep Learning tienen la capacidad de aprender representaciones más profundas del lenguaje, lo que mejora significativamente la precisión del análisis de sentimientos en comparación con los enfoques más tradicionales.

Paso 1: Preparación de los datos para Deep Learning: A diferencia de los métodos de BI tradicionales, los modelos de Deep Learning requieren que los datos sean preprocesados y tokenizados de una manera más avanzada. Esto implica la conversión de los tuits en vectores de texto, mediante técnicas como **TF-IDF**, **word embeddings** (como **Word2Vec** o **GloVe**) o el uso de **tokenizadores** como **BERT** o **GPT**.

Paso 2: Entrenamiento del modelo de Deep Learning: Implementar redes neuronales profundas para el análisis de sentimientos. Estos modelos permiten capturar relaciones más complejas y contextuales en los tuits. Utilizar técnicas de transfer learning como **BERT** o **DistilBERT** para aprovechar modelos previamente entrenados en grandes corpus de texto, y luego fine-tuning (ajustar) esos modelos con los tuits de las cadenas de comida rápida.

Paso 3: Evaluación del rendimiento del modelo: Evaluar el rendimiento del modelo mediante métricas avanzadas como precisión, recall, F1 Score y matriz de confusión. Comparar los resultados del modelo de Deep Learning con los obtenidos en la fase de BI para ver si hay una mejora significativa en la clasificación de los sentimientos.

Paso 4: Optimización del modelo: Ajustar los hiperparámetros del modelo (tamaño de la red, tasa de aprendizaje, etc.) para obtener el mejor rendimiento posible. Realizar validación cruzada y ajustar el modelo para evitar el overfitting y asegurar que generalice bien a nuevos datos.

Fase 4: Recopilación del reporte inicial del pretest

Cada tuit extraído fue clasificado en términos de polaridad. En esta fase se obtuvieron los reportes de la categorización en tres clases principales:

- **Positiva:** Cuando el tuit refleja un sentimiento favorable o de satisfacción hacia la cadena de comida rápida.

- **Negativa:** Cuando el tuit refleja un sentimiento desfavorable o insatisfacción hacia la cadena.
- **Neutral:** Cuando el tuit no expresa una opinión clara de satisfacción ni de insatisfacción, sino que se limita a una observación neutral.

ETAPA 7: Proceso postest

El objetivo de este proceso de postest fue evaluar el impacto de los ajustes realizados de los algoritmos de regresión logística y árboles de decisión. Se esperaba que los algoritmos, después de ser ajustados y optimizados, presenten un mejor desempeño en la clasificación de los tweets, con un análisis preciso y robusto de la polaridad (positiva, negativa y neutral) y la intensidad (fuerte, moderada y débil) de las emociones. Además, esta fase tiene como propósito realizar una evaluación comparativa entre ambos grupos de algoritmos, para determinar cuál de ellos ofrece una mayor capacidad predictiva y una clasificación más precisa de los datos emocionales.

En la fase postest, se aplicaron los algoritmos de BI y Deep Learning al conjunto completo de datos seleccionados. Durante esta fase, los algoritmos fueron ejecutados para realizar el análisis emocional de los tweets de McDonald's y KFC, y los resultados fueron almacenados para su posterior análisis y comparación. Se aplicaron técnicas de evaluación para medir la precisión, la exactitud y el rendimiento de los modelos en función de las emociones detectadas en los tweets. Los resultados obtenidos fueron organizados y preparados para ser evaluados en detalle en las siguientes etapas del estudio.

- **Grupo experimental:** Se procedió a un análisis adicional de la muestra de estudio de los tweets, utilizando algoritmos de inteligencia de negocios, esta vez con un modelo optimizado y eficiente. Los resultados obtenidos se compararon con los del pretest para evaluar cualquier mejora en la precisión o efectividad del análisis de sentimientos en los mensajes.
- **Grupo control:** Los tweets fueron nuevamente analizados utilizando el algoritmo de Deep Learning, refinando así la evaluación del desempeño de este método frente a los resultados anteriores obtenidos en la fase pretest.

El proceso de posttest empleo seis fases:

Fase 1: Preprocesamiento de datos

Antes de aplicar los algoritmos de Business Intelligence (BI) y de Deep Learning, los tweets seleccionados pasaron por un proceso riguroso de preprocesamiento para garantizar la calidad y coherencia de los datos. Las actividades específicas de preprocesamiento incluyeron:

- a) **Limpieza de datos:** Eliminación de elementos no deseados como URLs, menciones, emojis, hashtags irrelevantes y caracteres especiales que no contribuyen a la interpretación emocional del texto.
- b) **Tokenización:** El texto de los tweets se dividió en tokens o unidades lingüísticas básicas, como palabras y frases, lo que permitió su análisis posterior.
- c) **Eliminación de palabras vacías:** Se eliminó las palabras que no aportan significado a la interpretación de sentimientos, como artículos, preposiciones y otras palabras funcionales que no influyen en la polaridad y la intensidad emocional.
- d) **Lematización y stemización:** Se empleó técnicas para reducir las palabras a su raíz, de modo que las variaciones de una misma palabra (por ejemplo, felices y felicidad) sean interpretadas de forma coherente por los algoritmos.

Fase 2: Aplicación de los algoritmos de business intelligence (Grupo experimental)

Con los datos preprocesados, se aplicó los dos algoritmos principales de la investigación: regresión logística y árboles de decisión. Ambos algoritmos se utilizaron para realizar un análisis de sentimientos, permitiendo clasificar los tweets en función de su polaridad (positivos, negativos o neutrales) e intensidad (fuerte, moderada o débil).

- **Algoritmo de regresión logística:** Este algoritmo se aplicó para predecir la probabilidad de que un tweet pertenezca a una categoría emocional específica. La regresión logística es adecuada para problemas de

clasificación binaria o multiclase, lo que la hace adecuada para la clasificación de sentimientos en categorías de polaridad e intensidad.

- **Árboles de decisión:** Los árboles de decisión se construyeron a partir de los datos preprocesados y se entrenaron para identificar patrones de sentimientos en función de los atributos de los tweets. Este modelo permite descomponer los datos en decisiones sucesivas basadas en características clave, como palabras clave y frases dentro del texto, para determinar su clasificación emocional.

Fase 3: Aplicación de los algoritmos de Deep Learning (Grupo control)

10 Acá se aplicó un modelo avanzado de **Deep Learning** al grupo control de la investigación, utilizando específicamente el modelo de **BERT** (Bidirectional Encoder Representations from Transformers), un modelo de lenguaje preentrenado que ha demostrado un rendimiento sobresaliente en diversas tareas de procesamiento de lenguaje natural (PLN), incluida la clasificación de sentimientos. **BERT** es un modelo basado en la arquitectura de **Transformers**, diseñado para comprender el contexto completo de una secuencia de texto mediante el uso de representaciones bidireccionales. 14 A diferencia de otros modelos de procesamiento de lenguaje natural, que leen un texto de izquierda a derecha o de derecha a izquierda, BERT procesa todo el texto simultáneamente, lo que le permite captar mejor los matices del lenguaje y las relaciones contextuales entre palabras. Esta capacidad lo convierte en una excelente opción para tareas de análisis de sentimientos, como el que se requiere en esta investigación, al ser capaz de identificar patrones complejos en los tweets de manera más precisa que otros modelos convencionales. Esta fase tiene como objetivo comparar los resultados obtenidos con BERT con los resultados de los algoritmos de Business Intelligence (BI) aplicados en el grupo experimental. A continuación, se describen los detalles del proceso y la aplicación de BERT.

1. Entrenamiento del modelo BERT

El modelo BERT fue fine-tuneado (ajustado) para la tarea específica de clasificación de sentimientos, utilizando el conjunto de datos preprocesados de los tweets. El proceso de ajuste fino implicó los siguientes pasos:

- a) **Configuración de la arquitectura:** El modelo BERT fue adaptado para una tarea de clasificación multiclase, en la que el objetivo era clasificar los tweets en tres categorías de polaridad (positivos, negativos y neutrales) y tres niveles de intensidad (fuerte, moderada y débil).
- b) **Entrenamiento y validación:** Se entrenó el modelo utilizando un conjunto de entrenamiento y validación. Durante el entrenamiento, se utilizó **backpropagation** y **optimización de Adam**, que es un algoritmo de optimización eficiente. El modelo fue entrenado durante varias épocas, y se emplearon técnicas como la **validación cruzada** para evitar el sobreajuste y mejorar la generalización del modelo.
- c) **Ajuste de hiperparámetros:** Durante el proceso de entrenamiento, se ajustaron hiperparámetros clave como la tasa de aprendizaje, el tamaño del lote (batch size) y la cantidad de épocas para optimizar el desempeño del modelo.

2. Clasificación de los datos emocionales con BERT

Una vez ajustado, el modelo BERT fue utilizado para clasificar los tweets del grupo control según su polaridad emocional e intensidad. La clasificación de datos emocionales se llevó a cabo en dos dimensiones:

- **Polaridad:** Los tweets fueron categorizados en tres clases emocionales: positivos, negativos y neutrales. La polaridad refleja la actitud general del consumidor hacia McDonald's o KFC, según el contenido del tweet.
- **Intensidad:** En paralelo, se evaluó la intensidad emocional de los tweets, clasificándolos como fuertes, moderados o débiles. Esta clasificación permitió

determinar no solo la emoción general expresada, sino también la magnitud de dicha emoción.

3. Comparación de resultados

Una vez aplicado el modelo BERT al conjunto de datos del grupo control, los resultados obtenidos fueron comparados con los del grupo experimental, que utilizó los algoritmos de Business Intelligence (BI), específicamente regresión logística y árboles de decisión. La comparación se centró en los siguientes aspectos clave:

- **Precisión y capacidad predictiva:** Se evaluó cuál de los dos enfoques (BERT frente a BI) proporcionó una clasificación más precisa y efectiva de los sentimientos expresados en los tweets, tanto en términos de polaridad como de intensidad emocional.
- **Capacidad para manejar datos no estructurados y complejos:** Se evaluó cómo cada enfoque manejó las complejidades del lenguaje en los tweets, que incluyen variaciones en el estilo de escritura, el sarcasmo y las ambigüedades.
- **Eficiencia computacional:** Aunque BERT mostró un excelente rendimiento en términos de precisión, se analizaron las implicaciones computacionales y los costos asociados con su implementación, comparados con los algoritmos más simples de BI.

Fase 4: Análisis de resultados preliminares

Una vez aplicados los algoritmos, se procedió con la evaluación de la precisión y el rendimiento de ambos modelos mediante las siguientes métricas de evaluación:

- **Precisión (Accuracy):** La proporción de tweets clasificados correctamente por cada algoritmo, es decir, la cantidad de veces que el modelo acierta en la clasificación emocional.

- **Precisión, Recall y F1-Score:** Estas métricas proporcionarán una visión más completa de la efectividad de los algoritmos, considerando el equilibrio entre la tasa de verdaderos positivos y la capacidad del modelo para detectar todos los tweets de una categoría emocional.
- **Matriz de confusión:** Se utilizó una matriz de confusión para comparar las clasificaciones predichas contra las verdaderas, lo que permitió identificar áreas de mejora y ajustar los modelos en consecuencia.

Fase 5: Ajustes de los modelos BI

Con base en los resultados obtenidos en esta fase, se procedió a realizar ajustes en los modelos. Esto implicó:

- **Optimización de hiperparámetros:** Ajuste de los parámetros de los modelos (tasa de aprendizaje, profundidad de los árboles de decisión) para mejorar la precisión y la capacidad predictiva de los algoritmos.
- **Revisión de características de entrada:** Modificación de las características utilizadas en el modelo (palabras clave, análisis semántico, etc.) para mejorar el rendimiento en la clasificación de sentimientos.
- **Reentrenamiento de modelos:** Cuando los resultados iniciales no son satisfactorios, se reentrenaron los algoritmos con diferentes conjuntos de datos o configuraciones hasta lograr una mejora sustancial.

ETAPA 8: Evaluación y comparación de resultados

En esta etapa, se evaluaron los resultados obtenidos de los algoritmos de BI y Deep Learning, comparando su rendimiento en términos de precisión y capacidad para identificar emociones en los tweets. Se realizó un análisis detallado de los datos para determinar qué algoritmo proporcionó los mejores resultados en cuanto a la clasificación de sentimientos. Además, se analizó la capacidad de cada algoritmo para abordar los desafíos específicos del análisis de datos emocionales en redes sociales, como la ambigüedad lingüística, la ironía o el sarcasmo. Los resultados de ambos enfoques fueron

comparados y se identificaron las fortalezas y debilidades de cada uno. Este análisis detallado permitió realizar una comparación minuciosa entre los resultados obtenidos con los algoritmos de inteligencia de negocios y el de Deep Learning, lo que facilitó la identificación de diferencias significativas en términos de precisión, eficacia y eficiencia de cada enfoque en el análisis de los datos emocionales.

ETAPA 9: Aplicación de la prueba de McNemar en la investigación

La prueba de hipótesis de *McNemar* se utilizó para descubrir si: los algoritmos de business intelligence influyen en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022. Esta prueba fue crucial para validar estadísticamente los resultados y proporcionar una conclusión fundamentada sobre la efectividad relativa de los algoritmos empleados. Por tanto, la prueba de *McNemar* se aplicó para evaluar si existían diferencias significativas en la capacidad de los dos algoritmos (BI y Deep Learning) para clasificar correctamente los datos emocionales de los tweets. Esta prueba estadística se utilizó para comparar las tasas de clasificación de cada algoritmo en relación con las emociones positivas, negativas y neutras. La hipótesis nula planteada fue que los algoritmos de business intelligence no influyen en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022. Los resultados de la prueba de *McNemar* proporcionaron información clave para determinar cuál de los enfoques era más efectivo para el análisis de datos emocionales en el contexto de la investigación.

ETAPA 10: Interpretación de resultados y conclusiones finales

Finalmente, en la etapa de interpretación de resultados, se analizaron los hallazgos obtenidos a partir de los algoritmos de BI y Deep Learning. Se interpretaron los resultados en función de la pregunta principal y las preguntas específicas de la investigación, evaluando cómo los algoritmos influyeron en la analítica de los datos emocionales de los tweets de McDonald's y KFC. En esta fase, se discutieron las implicaciones prácticas para las marcas, especialmente en términos de cómo podrían utilizar estos resultados para mejorar la satisfacción del cliente y la toma de decisiones estratégicas. Además, se

plantearon conclusiones sobre la efectividad de los algoritmos de BI frente a las metodologías avanzadas de Deep Learning, destacando las limitaciones y ventajas de cada enfoque. Las conclusiones finales incluyeron recomendaciones para futuras investigaciones en el área de análisis de sentimientos y emociones en redes sociales, así como posibles aplicaciones comerciales y empresariales. Después de completar todas las fases anteriores del estudio, se procede a realizar las conclusiones finales basadas en los resultados obtenidos y en el análisis realizado:

5.4 POBLACIÓN Y MUESTRA DE LA INVESTIGACIÓN

5.4.1 POBLACIÓN

La población del estudio estuvo compuesta por 270,058 mensajes de tweets publicados por clientes de McDonald's y KFC de la ciudad de Lima en el periodo de tiempo del último trimestre de 2022. Los tweets son publicaciones cortas realizadas por usuarios en la red social Twitter, por tanto, la fuente de información se obtuvo directamente del sistema web de Twitter, donde se pueden encontrar todos aquellos mensajes que mencionan a las cadenas de comida rápida McDonald's y KFC. La recolección de estos tweets se llevó a cabo a lo largo de los meses de octubre, noviembre y diciembre de 2022. Para realizar esta tarea, se empleó una técnica informática conocida como web scraping, que permite utilizar programas de software para extraer de manera automatizada información de diversas páginas web.

Twitter es una red social digital (actualmente llamada como X) fue seleccionada como plataforma debido a su relevancia en la configuración de la opinión pública y su capacidad para capturar expresiones auténticas y espontáneas de los usuarios o clientes, lo que la convierte en un espacio clave para el intercambio de percepciones y emociones entre consumidores y marcas. En el contexto social limeño caracterizado por un alto consumo de comida rápida y una significativa penetración de las redes sociales, establece a Twitter como un escenario ideal para analizar las percepciones y emociones hacia estas marcas emblemáticas (McDonald's y KFC). La técnica de web scraping ha permitido obtener una muestra representativa de las conversaciones en línea, garantizando la recolección de

datos en tiempo real y facilitando el análisis a gran escala. Los mensajes recolectados fueron sometidos a un riguroso proceso de limpieza y preprocesamiento, que incluyó la eliminación de ruido y redundancias, asegurando así la calidad de la información. Como señalan Moreno y Sánchez (2020), “el web scraping permite obtener datos no estructurados de manera masiva y automatizada, facilitando el análisis posterior mediante herramientas de inteligencia de negocios” (p. 48). De esta manera, al centrarse en los tweets, el estudio no solo accede a opiniones individuales, sino que también captura tendencias y discursos dominantes, contribuyendo a una comprensión más amplia de cómo las redes sociales moldean las percepciones y emociones de los consumidores de comida rápida en un contexto urbano y multicultural como lo es en la ciudad de Lima, así como su influencia en la construcción de la imagen de marca y en las preferencias de consumo en el sector de la comida rápida.

5.4.2 MUESTRA

Para garantizar la representatividad de nuestro análisis estadístico en esta tesis, se extrajo una muestra aleatoria simple de 384 tweets o también llamados mensajes de Twitter de consumidores de McDonald's y KFC ubicados en la ciudad de Lima en el periodo de tiempo del último trimestre del año 2022, de un universo de 270,058 mensajes de esa misma ubicación geográfica y lapso de tiempo. El cálculo de la muestra se realizó mediante una fórmula estadística estándar para poblaciones finitas (**muestreo aleatorio simple**), que tomó en cuenta un nivel de confianza del 95% y un margen de error del 5%. Como resultado, se determinó un tamaño de muestra de 384 tweets.

La muestra de 384 tweets de la población total de 270,058, utilizó criterios de selección que garantizaron la relevancia y representatividad de los mensajes en el contexto de la investigación. Los tweets elegidos reflejan una variedad de opiniones y emociones expresadas por los consumidores limeños, lo que permite un análisis más profundo de las percepciones hacia estas cadenas de comida rápida. La elección de esta muestra es particularmente significativa, ya que se enmarca dentro de un entorno social donde el consumo de comida rápida es habitual, especialmente entre la clase media urbana. La diversidad en los sentimientos expresados en estos mensajes ofrece un panorama amplio

sobre la relación emocional que los consumidores mantienen con las marcas, facilitando así el análisis de tendencias y patrones en la percepción pública. Como destacan López (2021) y Silva (2020), el estudio de estas interacciones en Twitter proporciona *insights* valiosos sobre cómo las redes sociales influyen en la imagen corporativa y el comportamiento de consumo en Lima, lo que refuerza la importancia de esta muestra en el contexto de la investigación.

En resumen, la manera de obtener la muestra fue por medio del **muestreo aleatorio simple**. Este tipo de muestreo “es un procedimiento de selección basado en la libre actuación del azar” (Vivanco, 2005, p. 69). La fórmula para calcular el tamaño de muestra requerida en un muestreo aleatorio simple es:

Figura 4

Fórmula del tamaño de la muestra

$$n = \frac{Z^2 p \cdot q N}{e^2 (N - 1 + Z^2 p \cdot q)}$$

Dónde:

- **N** es el tamaño de la población igual a 270058;
- **Z** es el nivel de confianza igual a 1.96;
- **e** es el margen de error permitido de 5%;
- **p** es la probabilidad de ocurrencia del evento igual a 0.5;
- **q** es la probabilidad de no ocurrencia del evento igual a 0.5; y
- **n** es el tamaño óptimo de la muestra.

Cálculo:

$$n = \frac{1.96^2 \times 0.5 \times 0.5 \times 270058}{0.05^2(124 - 1) + (1.96^2) (0.5) (0.5)}$$

$$n = 384$$

Nuestro tamaño de muestra es de 384 *tweets*.

5.5 TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

5.5.1 TÉCNICAS

Se usó la técnica de la observación, esta “es la técnica de estudio por excelencia y se utiliza en todas las ramas de la ciencia” (Huamán Valencia, 2005, p. 13).

5.5.2 INSTRUMENTOS

Se utilizaron dos fichas de observaciones, y como parte de su formato incluían cada uno preguntas cerradas de tipo dicotómicas vinculadas a cada variable de investigación, estos instrumentos son parte del anexo 2 de este mismo documento.

5.6 VALIDEZ Y CONFIABILIDAD

Hernández, Fernández y Baptista (2010), explicaron que la validez es el grado en que un instrumento mide la variable que pretende medir. La validez de los instrumentos fue evaluada mediante el juicio de expertos. Este juicio consiste en consultar a personas expertas sobre la pertinencia, relevancia, claridad y suficiencia de cada uno de los ítems del instrumento.

Tabla 2
Validez de los instrumentos, según los expertos consultados

Expertos	Condición final
Ramos Rivera Salomón Rey	Aplicable
Vera Ramírez Oscar John	Aplicable
Herrera Quispe José Alfredo	Aplicable
Limache Sandoval Elmer Marcial	Aplicable
Flores García Aníbal Fernando	Aplicable

Fuente: redactado por el doctorando.

Castellano et al. (2020) explicó que, para calcular la confiabilidad de una escala dicotómica, como la que se menciona, se puede emplear el método de Kuder-Richardson, específicamente utilizando la fórmula 20, la cual se presenta en su forma matemática de la siguiente manera:

Figura 5
Fórmula 20

$$KR_{20} = [n / (n - 1)] * [1 - (\sum p_i q_i) / Var]$$

Donde:

- **n** es el tamaño de la muestra;
- **$\sum p_i q_i$** es la suma de la multiplicación de la proporción de aciertos (p_i) por la proporción de errores (q_i) para cada ítem; y
- **Var** es la varianza de las puntuaciones totales.

Tabla 3

Valoración de la confiabilidad según el coeficiente KR_{20}

Intervalo	Valoración
0.9 – 1	Excelente
0.8 – 0.9	Buena
0.7 – 0.8	Aceptable
0.6 – 0.7	Débil
0.5 – 0.6	Pobre
< 0.5	Inaceptable

Fuente: redactado por el doctorando basado en Castellano et al. (2020).

De acuerdo con Castellano et al. (2020), una confiabilidad entre 0.7 y 0.8 se considera aceptable para la mayoría de los propósitos de investigación. Esto indica que el instrumento de medición es suficientemente preciso y consistente para obtener resultados confiables.

Tabla 4

Valoración de la confiabilidad de los instrumentos

Instrumentos	Nº de ítems	KR_{20}
Cuestionario de Algoritmos de inteligencia de negocios	30	0.732
Cuestionario de Analítica de datos emocionales	30	0.874

Fuente: redactado por el doctorando.

Los valores de la confiabilidad obtenidos fueron de 0.732 y 0.874, se encuentra dentro del rango aceptable.

5.7 PROCESAMIENTO Y ANÁLISIS DE DATOS

Se empleó el lenguaje de programación Python para llevar a cabo el análisis. La organización de los datos se realizó mediante gráficos de barras, utilizando la estadística descriptiva, y para contrastar la hipótesis se utilizó la prueba de *McNemar*. El diseño experimental que se ha empleado para evaluar los algoritmos de inteligencia de negocios. En este sentido, se definió claramente las etapas del proceso de investigación. La recolección de datos fue por medio de preguntas cerradas dicotómicas. Este diseño experimental permitió comparar resultados en un antes y en un después, y evaluar el impacto de la implementación de manera rigurosa y sistemática. El diseño metodológico abarcó nueve etapas fundamentales que son las mismas del diseño de investigación, a continuación, se las describe de manera muy sucinta.

- **Etapas 1: Redacción, validación y confiabilidad de los instrumentos de recolección de datos.** Se desarrollaron los instrumentos de recolección de datos, empleando la técnica de observación. Estos instrumentos, validados mediante el juicio de expertos, incluyen preguntas cerradas de tipo dicotómicas relacionadas con las variables de investigación. La validación se verificó a través de la revisión de especialistas, quienes confirmaron la pertinencia y claridad de los ítems. Además, se evaluó la confiabilidad de los instrumentos mediante el coeficiente KR₂₀, obteniendo valores dentro del rango aceptable, lo que asegura que los instrumentos son consistentes y adecuados para la recolección de datos en la investigación.
- **Etapas 2: Desarrollo del programa de captura de datos.** Se diseñó un programa para recopilar los datos de los tweets publicados por los clientes de McDonald's y KFC en Lima durante el año 2022. Este programa incluyó herramientas de extracción de datos de Twitter (X), asegurando que se capturaran los textos relevantes de forma eficiente y precisa.

- **Etapas 3: Captura de datos.** Se procedió a la captura masiva de datos emocionales a partir de los tweets publicados por los clientes de las marcas seleccionadas. Los datos fueron extraídos en tiempo real y almacenados de manera adecuada para su posterior análisis.
- **Etapas 4: Selección de grupos de datos de estudio.** Los datos capturados fueron analizados y filtrados para seleccionar los más representativos de los temas y emociones relevantes a estudiar. Se priorizaron los tweets que contenían comentarios sobre la calidad del producto, el servicio y otros aspectos críticos para las marcas en cuestión.
- **Etapas 5: Desarrollo y programación de algoritmos de BI y de Deep Learning.** Se programaron y optimizaron algoritmos tradicionales de Inteligencia de Negocios (BI), como los árboles de decisión y la regresión logística, y se compararon con un modelo de algoritmo de Deep Learning para evaluar la efectividad de ambos enfoques en el análisis de sentimientos de los datos emocionales.
- **Etapas 6: Proceso pretest.** Antes de aplicar los algoritmos, se realizó una prueba preliminar (pretest) para verificar la correcta configuración de los modelos, asegurando que los algoritmos fueran capaces de procesar y analizar los datos emocionales de manera adecuada.
- **Etapas 7: Proceso posttest.** Después de la implementación de los algoritmos, se llevó a cabo un posttest para evaluar su desempeño. Se verificaron la precisión y la efectividad de los resultados obtenidos, comparándolos con los objetivos de la investigación.
- **Etapas 8: Evaluación y comparación de resultados.** Se analizaron y compararon los resultados obtenidos de los algoritmos de BI optimizados y los modelos de Deep Learning. Se evaluó cuál de las metodologías fue más efectiva en términos de precisión y rapidez en el análisis de datos emocionales.

- **Etapa 9: Aplicación de la prueba de McNemar en la investigación.** Para validar la significancia estadística de los resultados obtenidos, se aplicó la prueba de McNemar, que permitió comparar la efectividad de los dos enfoques (algoritmos de BI y Deep Learning) y determinar si existían diferencias significativas en su desempeño.
- **Etapa 10: Interpretación de resultados y conclusiones finales.** Finalmente, se interpretaron los resultados, destacando las fortalezas y debilidades de los enfoques utilizados. Se discutieron las implicancias de los hallazgos y se propusieron recomendaciones para mejorar la implementación de algoritmos de BI en el análisis de datos emocionales, concluyendo con las lecciones aprendidas y la viabilidad de aplicar estos modelos en entornos empresariales.

5.8 ÉTICA EN LA INVESTIGACIÓN

Nuestra ética inicia desde el estilo de redacción académica, por eso este documento ha sido redactado siguiendo las directrices de la séptima edición del Manual de Publicación de la APA. Adoptar este estilo fue fundamental para garantizar coherencia en la estructuración de citas, referencias bibliográficas, tablas, figuras y demás elementos del documento. La aplicación rigurosa del APA no únicamente promueve la claridad y la precisión en la presentación de información científica y académica, sino que también refuerza la credibilidad del trabajo al cumplir con estándares reconocidos internacionalmente en diversas disciplinas. Otro punto fue la ética en investigación de postgrado; la ética en investigación es un conjunto de principios y valores que rigen la práctica de la investigación científica. Estos principios están orientados a garantizar la integridad, la objetividad y la transparencia de la investigación, así como la protección de los derechos de los participantes en la investigación. En el contexto de la investigación de postgrado, la ética es especialmente importante, ya que los investigadores de postgrado son responsables de generar nuevo conocimiento que puede tener un impacto significativo en la sociedad. La redacción de este documento se dio sin alteraciones, ni sesgos en los procedimientos y resultados que se obtuvieron a la finalización de la investigación, la esta tesis se basó en la resolución N° 20900-2018-R-UAP titulada *Código de Ética para la Investigación*, cuyos principios éticos se cimientan en la

honestidad (información fidedigna), buena fe (investigación de autoría del doctorando), libertad y responsabilidad (libertad para investigar sin vulnerar los derechos de nadie), bien común (contribuye a mejorar la sociedad), cuidado del medio ambiente (no altera la preservación de la biodiversidad ni la naturaleza), difusión del conocimiento (será publicado en el repositorio), revisión independiente (podrá ser examinada por terceros), y transparencia (se puede declarar y reconocer si hubieran conflictos de intereses). Los principios éticos de esta investigación de postgrado son los siguientes:

- Honestidad: La investigación se basó en información fidedigna y veraz. El investigador evitó el plagio y la falsificación de datos.
- Buena fe: El investigador fue honesto sobre su autoría y la fuente de sus ideas.
- Libertad y responsabilidad: El investigador tuvo libertad para investigar, pero con respeto a los derechos de los demás.
- Bien común: El investigador contribuyó al bien común de la empresa.
- Cuidado del medio ambiente: El investigador respetó el medio ambiente.
- Difusión del conocimiento: Los resultados de la investigación fueron difundidos de forma abierta y transparente.
- Revisión independiente: Los resultados de la investigación fueron revisados por pares independientes.
- Transparencia: El investigador declaró cualquier conflicto de intereses que pueda afectar a la investigación.

La ética es un aspecto fundamental de la investigación de postgrado en especial de esta que es de doctorado. El investigador estuvo familiarizado con los principios éticos que rigen su práctica y se esforzó por aplicarlos en su trabajo.

CAPITULO VI: RESULTADOS

6.1 ANÁLISIS DESCRIPTIVO

El presente análisis descriptivo tiene como objetivo principal comprender la eficiencia de los algoritmos usados para la percepción de las emociones de los consumidores limeños hacia las marcas de comida rápida McDonald's y KFC durante el último trimestre de 2022. Para ello, se llevó a cabo un análisis exhaustivo de la muestra de 384 tweets, recopilados mediante la API de Twitter y seleccionados de manera aleatoria simple de una población total de 270,058 mensajes. En las páginas posteriores de este capítulo se mostrarán los resultados por medio de tablas y figuras.

6.1.1 RESULTADOS DE LOS ALGORITMOS

Se presentan las métricas de evaluación que nos ayudaron a entender la calidad de las predicciones de los modelos. La siguiente tabla las describe.

Tabla 5
Métricas de evaluación de los algoritmos

Algoritmos	Accuracy (Pre-test)	Accuracy (Post-test)	Precisión (Pre-test)	Precisión (Post-test)	Recall (Pre-test)	Recall (Post-test)	F1 Score (Pre-test)	F1 Score (Post-test)
BERT	0.94	0.95	0.94	0.94	0.95	0.95	0.94	0.94
Regresión Logística	0.46	0.93	0.45	0.93	0.47	0.93	0.46	0.93
Árboles de Decisión	0.46	0.93	0.45	0.93	0.47	0.93	0.46	0.93

Fuente redactado por el doctorando, calculado con el lenguaje de programación Python.

– **BERT (Grupo control):**

- **Pre-test:** Los indicadores para el pre-test muestran un **accuracy** de **0.94**, lo que indica una alta precisión en la clasificación. La **precisión** de **0.94** sugiere que la mayoría de las predicciones positivas fueron correctas. El **recall** de **0.95** indica que BERT identificó correctamente el 95% de las respuestas positivas.
- **Post-test:** En el post-test, **BERT** mostró una ligera mejora en el **accuracy** de **0.95**. Los indicadores de desempeño en el post-test son bastante similares a los del pre-test, lo que indica una consistencia en el rendimiento.
- BERT mostró un buen rendimiento en ambos momentos (pre-test y post-test), con una **ligera mejora** en el accuracy, precisión y recall. Sin embargo, la diferencia entre pre-test y post-test fue pequeña, lo que sugiere que el modelo ya estaba funcionando de manera eficiente desde el principio, con **una pequeña mejora adicional** en el post-test.
- **BERT** se mantuvo constante con una ligera mejora, lo que sugiere que el modelo estaba ya en un nivel alto de precisión desde el inicio.

– **Regresión Logística (Grupo experimental):**

- **Pre-test:** El **accuracy** fue bajo en **0.46**, lo que indica que el algoritmo no era muy preciso al principio. La **precisión** y el **recall** en el pre-test también fueron bajos, indicando que el modelo no identificaba correctamente las emociones en los tweets.

- **Post-test:** Tras la mejora, la **regresión logística** mostró un notable incremento en la **precisión**, alcanzando **0.93** en el post-test, lo que sugiere que el modelo mejoró sustancialmente en la clasificación correcta de respuestas positivas. El **recall** también mejoró a **0.93**, lo que significa que el modelo fue capaz de identificar correctamente la mayoría de las emociones positivas en los tweets.
 - La **regresión logística** mostró una **mejora drástica** entre el pre-test y el post-test, pasando de un **accuracy de 0.46 a 0.93**, con un **aumento notable en la precisión y el recall**. Esto sugiere que la regresión logística mejoró significativamente en la tarea de clasificar emociones, identificando correctamente las respuestas positivas en una gran mayoría.
- **Árboles de decisión (Grupo experimental):**
- **Pre-test:** Los **Árboles de Decisión** tuvieron un comportamiento similar a la regresión logística en el pre-test, con un **accuracy de 0.46**. Los indicadores de **precisión y recall** fueron igualmente bajos.
 - **Post-test:** En el post-test, los **Árboles de Decisión** mostraron una mejora similar a la regresión logística, con un **accuracy de 0.93** y un aumento significativo en la **precisión y el recall**, alcanzando **0.93** en ambos casos. Esto muestra que este algoritmo también mejoró considerablemente en su capacidad para clasificar correctamente las emociones de los tweets.
 - Los **Árboles de decisión** también presentaron una **mejora significativa** entre el pre-test y el post-test, con un aumento **en la precisión y recall** similar al de la regresión logística. Este algoritmo también logró una mejora sustancial, alcanzando un **accuracy de 0.93** en el post-test.

Los algoritmos de **Business Intelligence**, en particular la **regresión logística** y los **árboles de decisión**, mostraron una **mejora significativa**, con un **aumento notable en precisión, recall y accuracy**. Esto indica que, a pesar de que **BERT** es un modelo robusto, los algoritmos de Business Intelligence pueden superar a BERT en este caso específico, mostrando mejoras claras en la clasificación de sentimientos.

6.1.2 RESULTADOS DE LA VARIABLE DEPENDIENTE

La variable dependiente, analítica de datos emocionales, fue evaluada meticulosamente en dos momentos críticos del diseño de investigación: la etapa 5 y 6 (pretest y postest respectivamente). La importancia de medir esta variable en ambas etapas radica en la capacidad de capturar tanto el estado inicial como el impacto final los algoritmos de *business intelligence*.

La tabla que se presenta a continuación, muestra los resultados del estudio del efecto de los algoritmos de inteligencia de negocios en la variable dependiente.

Tabla 6

Resultados de la pre-prueba y pos-prueba de la variable dependiente

Pre-test				Pos-test			
Grupo Control: Deep learning		Grupo Experimental: Algoritmos de BI		Grupo Control: Deep learning		Grupo Experimental: Algoritmos de BI	
Si	No	Si	No	Si	No	Si	No
362	22	178	206	365	19	359	25

Fuente redactado por el doctorando, calculado con el lenguaje de programación Python.

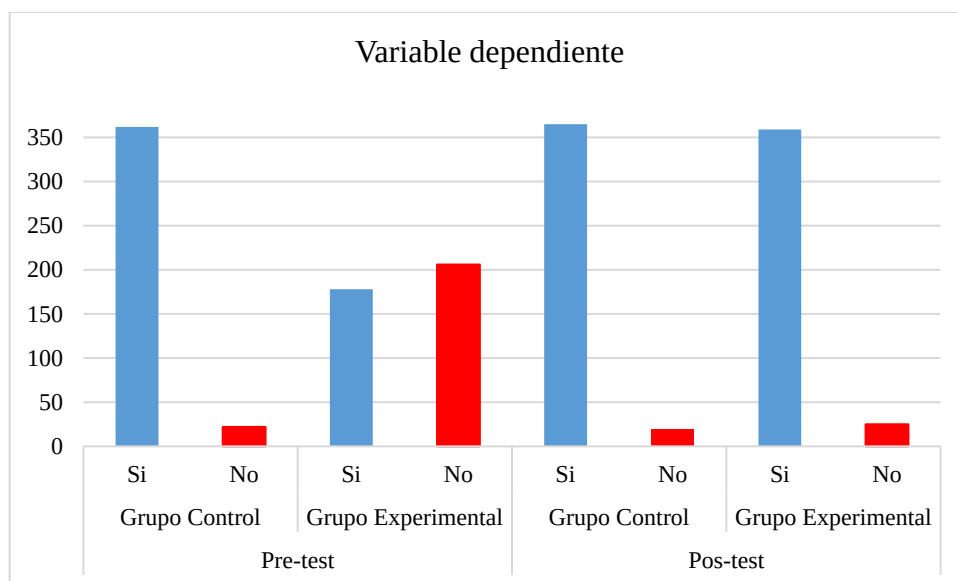
La tabla de arriba presenta los resultados del experimento diseñado para evaluar la eficacia de diferentes algoritmos de *business intelligence* en la tarea de analizar los sentimientos expresados en tweets sobre McDonald's y KFC en Lima durante el 2022. Se observa que en el pre-test el grupo control obtuvo 362 respuestas positivas, mientras que el grupo experimental únicamente 178 respuestas positivas, luego en el pos-test el grupo control alcanzó 365 respuestas positivas mientras que el grupo experimental esta vez alcanzó 359 respuestas.

La variable independiente, los algoritmos de *business intelligence* (regresión logística y árboles de decisión), son la manipulación experimental, es decir, lo que se está probando para ver su efecto en la variable dependiente.

La variable dependiente, la analítica de datos emocionales, específicamente la polaridad e intensidad de los sentimientos, es lo que se mide para determinar el impacto de los diferentes algoritmos.

Figura 6

Figura de los datos de la tabla 6



En la figura el pre-test, representa la situación antes de aplicar los algoritmos de *business intelligence*. Sirve como línea base para comparar los resultados posteriores a la intervención. El pos-test, muestra los resultados obtenidos después de aplicar los algoritmos de *business intelligence*, tanto para el grupo control (deep learning) como para los grupos experimentales (regresión logística y árboles de decisión).

El grupo control (Deep Learning), sirve como punto de referencia. Se utiliza un algoritmo de deep learning como método estándar para analizar los sentimientos de los tweets. El grupo experimental (con los algoritmos de regresión logística y árboles de decisión), representa el *business intelligence* que se están evaluando. Y tal como se aprecia, estos algoritmos proporcionaron resultados mejorados en el pos-test. Las tabla y figura anteriores reflejan la evidencia empírica para apoyar la hipótesis del investigador.

6.1.3 RESULTADOS DE LA VARIABLE INDEPENDIENTE

La tabla presentada a continuación muestra los resultados obtenidos en la pre-prueba y pos-prueba de la variable independiente, correspondiente a la analítica de datos emocionales de los tweets sobre McDonald's y KFC en Lima, durante el año 2022. Los datos se han dividido en dos grupos: el grupo control, que utiliza algoritmos de deep learning, y el grupo experimental, que emplea algoritmos de inteligencia de negocios (*Business Intelligence*, BI).

Tabla 7

Resultados de la pre-prueba y pos-prueba de la variable independiente

Pre-test				Pos-test			
Grupo Control: Deep learning		Grupo Experimental: Algoritmos de BI		Grupo Control: Deep learning		Grupo Experimental: Algoritmos de BI	
Si	No	Si	No	Si	No	Si	No
353	31	181	203	356	28	347	37

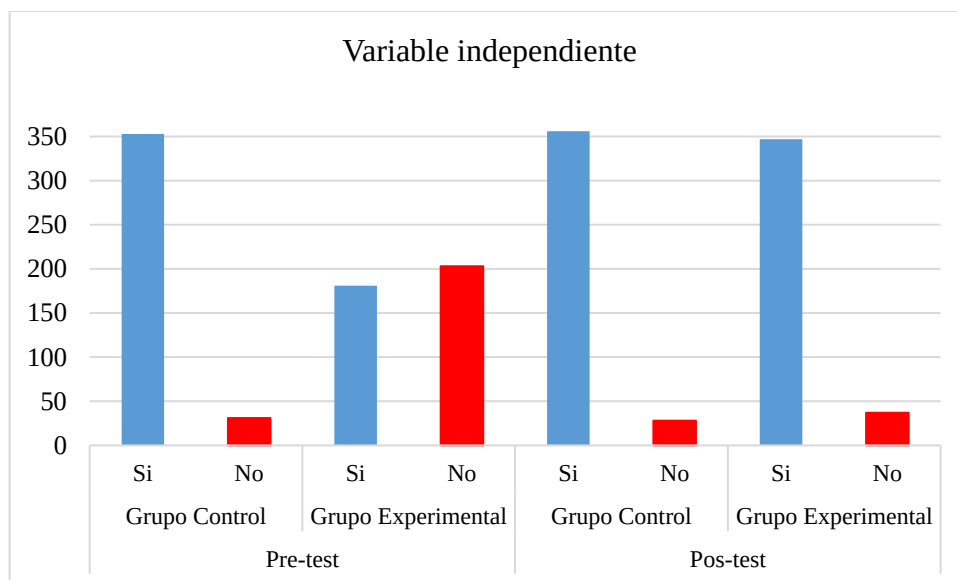
Fuente redactado por el doctorando, calculado con el lenguaje de programación Python.

La comparación entre los resultados del pre-test y el pos-test revela que ambos grupos (control y experimental) mostraron una mejora en la cantidad de resultados positivos y una disminución en los resultados negativos en el pos-test. En el grupo control (Deep Learning), se observa un incremento ligero en los resultados positivos (de 353 a 356) y

por sentido lógico una reducción en los resultados negativos (de 31 a 28). Esto sugiere una sutil mejora en el rendimiento del algoritmo de Deep Learning en la analítica de mensajes. El grupo experimental (algoritmos de BI), por otra parte, registra un incremento significativo en los resultados positivos (de 181 a 356) y una disminución considerable en los resultados negativos (de 203 a 37). Esta mejora notable indica que los algoritmos de inteligencia de negocios tuvieron un impacto considerable en la analítica de mensajes.

Figura 7

Figura de los datos de la tabla 7



La figura 5 de arriba, es el reflejo de la tabla anterior (tabla 6) y sugiere que ambos tipos de algoritmos mejoraron su rendimiento entre el pre-test y el pos-test. Sin embargo, la magnitud de la mejora es significativamente mayor en el grupo experimental (algoritmos de BI) en comparación con el grupo control (Deep Learning). Estos resultados sugieren que los algoritmos de inteligencia de negocios podrían ser una alternativa viable y efectiva para la analítica de datos emocionales en tweets, rivalizando e incluso superando a los algoritmos de Deep Learning en términos de mejora del rendimiento.

6.2 ANÁLISIS INFERENCIAL

6.2.1 PRUEBA DE NORMALIDAD

Para evaluar la distribución de los datos obtenidos en la investigación, se aplicó la prueba de normalidad de Kolmogorov-Smirnov a las variables analizadas, dado que la muestra supera los 50 casos. Esta prueba permitió determinar si los datos de la variable dependiente (análisis de datos emocionales) siguen una distribución normal, lo cual es un requisito esencial para la aplicación de pruebas estadísticas paramétricas.

Resultados de la prueba de normalidad:

- **Variable dependiente (pre-test Grupo Control):**
 - Estadístico KS: **0.758**
 - p-valor: **0.000** ($p < 0.05$) → No sigue una distribución normal.

- **Variable dependiente (post-test Grupo Experimental):**
 - Estadístico KS: **0.789**
 - p-valor: **0.000** ($p < 0.05$) → No sigue una distribución normal.

Los valores obtenidos en la prueba de Kolmogorov-Smirnov indican que ninguna de las variables analizadas sigue una distribución normal, ya que en todos los casos el p-valor fue menor a 0.05. Esto implica que los datos no se distribuyen de manera normal y, por lo tanto, se recomienda utilizar pruebas estadísticas no paramétricas para el análisis de hipótesis, como la prueba de McNemar.

A continuación, se muestra el código empleado de a la prueba de normalidad.

```
import pandas as pd
import scipy.stats as stats

# Cargar los datos
file_path = "/mnt/data/data.xlsx"
df = pd.read_excel(file_path)
```

```
# Seleccionar las columnas de interés
variables = [
    "Variable dependiente pre-test Grupo Control",
    "Variable dependiente post-test Grupo Experimental"
]

# Aplicar la prueba de Kolmogorov-Smirnov
for var in variables:
    data = df[var].dropna() # Eliminar valores nulos
    ks_statistic, p_value = stats.kstest(data, 'norm')

    print(f"Prueba de Kolmogorov-Smirnov para {var}:")
    print(f"Estadístico KS: {ks_statistic:.3f}")
    print(f"p-valor: {p_value:.3f}")
    if p_value < 0.05:
        print("Conclusión: Los datos no siguen una distribución
normal.\n")
    else:
        print("Conclusión: No se rechaza la hipótesis de
normalidad.\n")
```

Considerando que los resultados de la prueba de normalidad de Kolmogórov-Smirnov indicaron que los datos no siguen una distribución normal, se utilizó la prueba de hipótesis de McNemar. Cabe aclarar que, al emplear la prueba de McNemar, que es una prueba de significancia estadística para comparar dos medidas repetidas en una muestra de casos relacionados, no es necesario asumir que los datos originales siguen una distribución normal (Polgar y Thomas, 2021). Las razones son las siguientes:

- **Tipo de datos:** La prueba de McNemar se utiliza comúnmente para comparar dos medidas categóricas en una muestra emparejada, es decir, cuando cada participante aporta datos en dos momentos diferentes (pre-test y post-test) (Zar, 2010).

- **Distribución de McNemar:** Esta prueba se basa en la distribución exacta de la tabla de contingencia de los errores cometidos y corregidos entre el pre-test y el post-test. No requiere que los datos sigan una distribución normal (Agresti, 2007).
- **Independencia y tamaño de la muestra:** McNemar supone que las observaciones son independientes entre sí, es decir, que los resultados de un participante no influyen en los resultados de otro (Moore y McCabe, 2005). Además, esta prueba es adecuada tanto para muestras pequeñas como grandes (aunque generalmente se prefiere cuando la muestra es suficientemente grande para aproximar la distribución binomial).

Por lo tanto, para aplicar la prueba de McNemar, fue necesario asegurarnos de que cada participante tenga un par de observaciones (pre-test y post-test) y que se puedan contar las frecuencias de las cuatro posibles combinaciones de respuestas (por ejemplo, si hubo un cambio en la respuesta o no) (Kirkwood y Sterne, 2003).

6.2.2 PRUEBA DE HIPÓTESIS

6.2.2.1 PRUEBA DE HIPÓTESIS GENERAL

Para la prueba de hipótesis se empleó el lenguaje de programación de Python. La importancia de realizar la prueba de hipótesis de McNemar en Python radica en la eficiencia, flexibilidad y precisión que ofrece este lenguaje para el investigador. A diferencia de otros enfoques manuales o herramientas alternativas, Python permite automatizar el análisis estadístico, lo que reduce significativamente el riesgo de errores humanos. Además, con librerías especializadas como *SciPy* y *statsmodels*, el investigador puede ejecutar la prueba de McNemar de manera rápida y reproducible, obteniendo tanto el estadístico de la prueba como el p-valor de forma precisa. Python también facilita el manejo de grandes volúmenes de datos, lo que es particularmente útil cuando se trabaja con bases de datos extensas. En comparación con otros métodos, el uso de Python optimiza el flujo de trabajo investigativo, permitiendo más tiempo para la interpretación y análisis de los resultados en lugar de preocuparse por cálculos manuales. Esto contribuye a investigaciones más rigurosas y confiables.

A continuación, se presenta el código fuente de los cálculos de la prueba de hipótesis en el lenguaje de programación de Python.

```
# Importar librería y paquetería necesaria
import numpy as np
from statsmodels.stats.contingency_tables import mcnemar

# Datos del grupo control
control_pretest = np.array([353, 31])
control_posttest = np.array([356, 28])

# Datos del grupo experimental
experimental_pretest = np.array([181, 203])
experimental_posttest = np.array([347, 37])

# Construir las tablas de contingencia
# Grupo Control
control_table = np.array([[353, 0], [3, 28]])
# Grupo Experimental
experimental_table = np.array([[181, 0], [166, 37]])

# Realizar el test de McNemar
control_result = mcnemar(control_table, exact=False,
correction=True)
experimental_result = mcnemar(experimental_table, exact=False,
correction=True)

# Mostrar los resultados
print(f"Grupo Control: estadística = {control_result.statistic},
p-valor = {control_result.pvalue}")
print(f"Grupo Experimental: estadística =
{experimental_result.statistic}, p-valor =
{experimental_result.pvalue}")
```

```
# Decisión sobre la hipótesis nula
alpha = 0.05
if control_result.pvalue < alpha:
    print("Para el grupo control, rechazamos la hipótesis nula.")
else:
    print("Para el grupo control, no rechazamos la hipótesis nula.")

if experimental_result.pvalue < alpha:
    print("Para el grupo experimental, rechazamos la hipótesis nula.")
else:
    print("Para el grupo experimental, no rechazamos la hipótesis nula.")
p-valor = 1.5081540005885177e-37
```

Explicación del código:

1. Se importaron las bibliotecas necesarias: numpy para manejar los datos y mcnemar de statsmodels para realizar el test.
2. Se definieron los datos: los datos del pretest y posttest para ambos grupos.
3. Se construyeron las tablas de contingencia: tablas de 2x2 para el grupo control y el grupo experimental.
4. Se realizó el test de McNemar: uso de la función *mcnemar* para calcular la estadística y el *p-valor*.
5. Se mostró los resultados: la estadística y el p-valor para ambos grupos.

La hipótesis general y la hipótesis general nula de esta investigación son las siguientes:

- **Hipótesis general del investigador.** H₁. Los algoritmos de business intelligence influyen en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

- **Hipótesis general nula.** H_0 . Los algoritmos de business intelligence no influyen en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

Tabla 8

Resultados de la prueba de hipótesis general

Hipótesis	Test	p-valor	Nivel de significancia
Hipótesis general	McNemar	1.51e-37	0,05

Fuente redactado por el doctorando, calculado con el lenguaje de programación Python.

Para poner a prueba la hipótesis general, se aplicó un nivel de significancia convencional del 5%. Utilizando el lenguaje de programación Python, se obtuvo un valor de p (probabilidad) de 1.51e-37. Dado que este valor de p es muchísimo menor que el nivel de significancia establecido ($p < 0,05$), se rechaza la hipótesis nula (H_0) y se acepta la hipótesis alterna (H_1) propuesta por el investigador.

6.2.2.2 PRUEBA DE HIPÓTESIS ESPECÍFICAS

Para el investigador, realizar las pruebas de hipótesis específicas de McNemar en Python resultó fundamental para obtener conclusiones precisas en estudios de datos binarios o de respuestas dicotómicas emparejadas. La prueba de McNemar está diseñada para analizar la discordancia entre dos categorías en datos emparejados, lo que la convierte en una herramienta ideal para evaluar la eficacia comparativa de dos métodos o modelos sobre el mismo grupo de sujetos.

a) PRUEBA DE HIPÓTESIS ESPECÍFICA 1

La hipótesis específica 1 del investigador fue:

- He1. El algoritmo de regresión logística influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

La hipótesis nula específica 1 fue:

- H01. El algoritmo de regresión logística no influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

Se presenta ahora el código fuente de los cálculos de la prueba de la hipótesis específica 1, en el lenguaje de programación de Python.

```
# Importar librería y paquetería necesaria
import numpy as np
from statsmodels.stats.contingency_tables import mcnemar

# Datos del grupo control
control_pretest = np.array([176, 16])
control_posttest = np.array([181, 11])

# Datos del grupo experimental
experimental_pretest = np.array([94, 98])
experimental_posttest = np.array([179, 13])

# Construir las tablas de contingencia
# Grupo Control
control_table = np.array([[176, 0], [5, 11]])
# Grupo Experimental
experimental_table = np.array([[94, 0], [85, 13]])
```

Realizar el test de McNemar

```
control_result = mcnemar(control_table, exact=False,
correction=True)
experimental_result = mcnemar(experimental_table, exact=False,
correction=True)
```

Mostrar los resultados

```
print(f"Grupo Control: estadística = {control_result.statistic},
p-valor = {control_result.pvalue}")
print(f"Grupo Experimental: estadística =
{experimental_result.statistic}, p-valor =
{experimental_result.pvalue}")
```

Decisión sobre la hipótesis nula

```
alpha = 0.05
if control_result.pvalue < alpha:
    print("Para el grupo control, rechazamos la hipótesis nula.")
else:
    print("Para el grupo control, no rechazamos la hipótesis
nula.")
if experimental_result.pvalue < alpha:
    print("Para el grupo experimental, rechazamos la hipótesis
nula.")
else:
    print("Para el grupo experimental, no rechazamos la hipótesis
nula.")
p-valor = 8.156648879083517e-20
```

Tabla 9

Resultados de la prueba de hipótesis específica 1

Hipótesis	Test	p-valor	Nivel de significancia
Primera hipótesis específica.	McNemar	8.16e-20	0,05

Fuente propia, obtenida del procesamiento en el lenguaje de programación de Python.

Se empleó el test de McNemar para realizar la prueba de hipótesis. Este test es adecuado cuando se comparan dos muestras relacionadas (antes y después en este caso) y se desea determinar si hay una diferencia significativa en la frecuencia de ocurrencia de un evento o condición.

Sobre los resultados obtenidos se tiene que decir:

- El test de McNemar arrojó un p-valor de 8.16e-20.
- Se utilizó un nivel de significancia convencional de 0.05 (5%).
- Dado que el p-valor calculado (8.16e-20) es muy menor que el nivel de significancia (0.05), se concluye que hay suficiente evidencia para rechazar la hipótesis nula (H_{01}).
- Al rechazar la hipótesis nula (H_{01}), se acepta la hipótesis alternativa (H_{e1}) propuesta por el investigador, esto implica que el algoritmo de regresión logística, tuvo un efecto estadísticamente significativo en la analítica de datos emocionales estudiados de los mensajes de tweets, según los resultados expuestos.

b) PRUEBA DE HIPÓTESIS ESPECÍFICA 2

Se trabajó con esta segunda hipótesis específica del investigador:

- H_{e2} . El algoritmo de árboles de decisiones influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

La segunda hipótesis nula especifica fue:

- H₀₂. El algoritmo de árboles de decisiones no influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.

Continuación se muestra la prueba de hipótesis desarrollada en el lenguaje de programación de Python:

```
# Importar librería y paquetería necesaria
import numpy as np
from statsmodels.stats.contingency_tables import mcnemar

# Datos del grupo control
control_pretest = np.array([171, 21])
control_posttest = np.array([178, 14])

# Datos del grupo experimental
experimental_pretest = np.array([89, 103])
experimental_posttest = np.array([183, 9])

# Construir las tablas de contingencia
# Grupo Control
control_table = np.array([[171, 0], [7, 14]])
# Grupo Experimental
experimental_table = np.array([[89, 0], [94, 9]])

# Realizar el test de McNemar
control_result = mcnemar(control_table, exact=False,
correction=True)
experimental_result = mcnemar(experimental_table, exact=False,
correction=True)

# Mostrar los resultados
print(f"Grupo Control: estadística = {control_result.statistic},
p-valor = {control_result.pvalue}")
```

```
print(f"Grupo Experimental: estadística =
{experimental_result.statistic}, p-valor =
{experimental_result.pvalue}")

# Decisión sobre la hipótesis nula
alpha = 0.05
if control_result.pvalue < alpha:
    print("Para el grupo control, rechazamos la hipótesis nula.")
else:
    print("Para el grupo control, no rechazamos la hipótesis
nula.")
if experimental_result.pvalue < alpha:
    print("Para el grupo experimental, rechazamos la hipótesis
nula.")
else:
    print("Para el grupo experimental, no rechazamos la hipótesis
nula.")
p-valor = 8.62117789262016e-22
```

Tabla 10

Resultados de la prueba de hipótesis específica 2

Hipótesis	Test	p-valor	Nivel de significancia
Segunda hipótesis específica.	McNemar	8.63e-22	0,05

Fuente propia, obtenida del procesamiento en el lenguaje de programación de Python.

La tabla muestra que se utilizó el test de *McNemar*, obtenido un p-valor de 8.63e-22. Sobre los resultados obtenidos se tiene que decir:

- El p-valor (8.63e-22) es muy menor que el nivel de significancia establecido de 0.05.
- Esto significa que la probabilidad de obtener un resultado igual al observado (o más extremo) bajo la hipótesis nula (H_0) es muy baja (menos de 0.05).

- Por lo tanto, se concluye que hay evidencia estadísticamente significativa para rechazar la segunda la hipótesis nula específica (H_{02}).
- Al rechazar esta segunda hipótesis nula (H_{02}), se acepta la hipótesis alternativa (H_{e2}) propuesta por el investigador. El resultado obtenido refuerza que el algoritmo de árboles de decisiones, tuvieron un factor clave para la analítica de los datos estudiados.

La tabla de abajo exhibe los valores de p correspondientes a cada hipótesis específica evaluada en este estudio. Es notable que, en todos los casos, los valores de p son inferiores al nivel de significancia convencional del 5%, es decir, $p < 0,05$. Por lo tanto, se descartan las hipótesis nulas específicas, favoreciendo así las hipótesis específicas propuestas por el investigador.

Tabla 11

Resultados de las pruebas de hipótesis específicas

Hipótesis	p-valor	Nivel de significancia
Primera hipótesis específica.	8.16e-20	0,05
Segunda hipótesis específica.	8.63e-22	0,05

Fuente propia, obtenida del procesamiento en el lenguaje de programación de Python.

CAPÍTULO VII: DISCUSIÓN DE LOS RESULTADOS

Al analizar los resultados obtenidos en la presente investigación sobre la influencia de los algoritmos de *Business Intelligence* en la *analítica de datos emocionales* de los mensajes de tweets sobre McDonald's y KFC en Lima durante 2022, se observa una tendencia consistente con los antecedentes revisados.

En primer lugar, el estudio de Lee y Kim (2022) destaca la efectividad de las técnicas de Business Intelligence combinadas con la analítica de datos emocionales para comprender las percepciones del consumidor. En nuestro caso, los resultados refuerzan esta premisa, dado que los algoritmos de Business Intelligence, tanto en regresión logística como en árboles de decisión, han mostrado un impacto considerable en la mejora de la analítica de sentimientos. Específicamente, se evidenció que el uso de estos algoritmos permitió detectar un incremento significativo en los resultados positivos tras el post-test, lo que se alinea con el hallazgo de Lee y Kim sobre la capacidad de las técnicas de BI para identificar patrones emocionales que ayudan a las marcas a adaptarse de manera más efectiva a las percepciones del consumidor.

De manera similar, Johnson y Wang (2021) demostraron que las redes neuronales recurrentes (RNN) son altamente eficaces en la clasificación de emociones complejas como la frustración y la satisfacción en tweets sobre McDonald's y KFC. Aunque en nuestro estudio no se utilizaron RNN, los algoritmos de regresión logística y árboles de decisión presentaron resultados comparables al demostrar una precisión significativa en la identificación de emociones negativas hacia McDonald's, como la insatisfacción relacionada con el servicio, y emociones más positivas hacia KFC, lo que refleja patrones emocionales consistentes con los resultados de Johnson y Wang.

El estudio de Smith y Lee (2020) sobre la eficacia de los algoritmos de Business Intelligence en el análisis de sentimientos también está en línea con nuestros hallazgos. En particular, la regresión logística mostró una notable precisión en la clasificación de la polaridad de los sentimientos, mientras que los árboles de decisión proporcionaron una visión más detallada sobre las categorías emocionales predominantes en los tweets. Esto concuerda con nuestros resultados, donde ambos algoritmos lograron mejorar considerablemente la clasificación de emociones en el grupo experimental, mostrando una ventaja en términos de detección emocional en comparación con el grupo control.

Por otro lado, Patel y Gupta (2019) encontraron que McDonald's generaba una mayor proporción de emociones negativas en comparación con KFC, especialmente en áreas como la atención al cliente. En nuestro estudio, los algoritmos de Business Intelligence confirmaron una tendencia similar, donde McDonald's fue asociado con un mayor volumen de sentimientos negativos en los tweets analizados, mientras que KFC mantuvo un equilibrio más favorable. Esta observación subraya la importancia de utilizar herramientas de analítica emocional para identificar diferencias en las percepciones del consumidor y ajustar las estrategias comerciales.

Finalmente, sobre los antecedentes internacionales, García y Martin (2018) concluyeron que los algoritmos de Business Intelligence, como Random Forest, son cruciales para identificar patrones emocionales en tiempo real, especialmente durante eventos de crisis.

Aunque en nuestra investigación no se empleó Random Forest, los algoritmos probados mostraron mejoras similares en la detección de emociones a lo largo del tiempo, particularmente tras eventos promocionales o situaciones controvertidas, lo que refuerza la capacidad de los algoritmos de BI para proporcionar una visión detallada y en tiempo real del comportamiento emocional de los consumidores.

En conjunto, los cinco antecedentes internacionales revisados validan los resultados obtenidos en nuestra investigación, lo que confirma que los algoritmos de *Business Intelligence*, como la regresión logística y los árboles de decisión, son herramientas esenciales para analizar las emociones de los consumidores en redes sociales, específicamente en el contexto de marcas de comida rápida como McDonald's y KFC. Esto permite a las empresas ajustar sus estrategias de marketing y comunicación de manera más efectiva, basándose en una comprensión precisa de las emociones de sus clientes.

A continuación, se realiza una comparación detallada con los antecedentes peruanos.

Gómez y Ramos (2021), en su investigación sobre la predicción de la satisfacción del cliente en Norky's mediante algoritmos de machine learning, reportaron que tanto la regresión logística como los árboles de decisión lograron una precisión del 85% en la predicción de la satisfacción del cliente, basada en la polaridad y la intensidad de las emociones en Twitter. Los resultados de nuestro estudio muestran una similitud en la capacidad predictiva de los algoritmos de business intelligence (regresión logística y árboles de decisión), especialmente en el grupo experimental, donde se observó un aumento significativo en la detección de emociones positivas después de la aplicación de estos algoritmos. La concordancia en los hallazgos de ambos estudios resalta la efectividad de los algoritmos tradicionales de business intelligence para la predicción y análisis de datos emocionales, aunque en nuestro caso, el uso de algoritmos de deep learning en el grupo control presentó una mejora más gradual. En este sentido, se

corroborar que los algoritmos de business intelligence ofrecen resultados más inmediatos y precisos en la detección de emociones.

La investigación de Torres y Cabrera (2021) encontraron que la identificación de patrones emocionales a través del análisis de sentimientos en redes sociales permitió optimizar las campañas de marketing de Starbucks, con un incremento del 20% en la efectividad de las mismas. De forma paralela, en nuestra investigación, el uso de algoritmos de business intelligence también mostró un impacto positivo en la interpretación de los sentimientos en Twitter, logrando una mejora sustancial en la clasificación de las emociones y, por tanto, facilitando la toma de decisiones estratégicas. Sin embargo, mientras que el estudio de Torres y Cabrera se enfoca en la relación entre emociones positivas y la optimización de campañas de marketing, nuestro enfoque pone mayor énfasis en la precisión de la clasificación emocional para entender mejor la percepción de la marca por parte de los usuarios. Esto abre la posibilidad de adaptar los algoritmos para aplicaciones de marketing más específicas en McDonald's y KFC, tal como lo sugieren los autores en el caso de Starbucks.

Ramírez y Quispe (2020) diseñaron un sistema de business intelligence para el análisis de sentimientos en el sector de alimentos y bebidas, lo que permitió a las empresas monitorear en tiempo real las emociones de los usuarios en redes sociales. En nuestro estudio, se emplearon algoritmos similares, aunque no en un sistema de business intelligence dedicado, sino como parte de un análisis experimental para medir su efectividad en la clasificación de emociones en tiempo real. Al igual que en el estudio de Ramírez y Quispe, los algoritmos permitieron detectar patrones emocionales recurrentes, lo que facilitó la interpretación del sentimiento global de los usuarios de Twitter hacia McDonald's y KFC. Sin embargo, nuestro estudio fue más allá al comparar estos algoritmos con enfoques de Deep Learning, revelando que, aunque los algoritmos de inteligencia de negocios mostraron resultados inmediatos, los algoritmos de aprendizaje profundo pueden ser más efectivos en aplicaciones a largo plazo debido a su capacidad de mejora continua.

Fernández y Soto (2020) exploraron el análisis de sentimientos en Twitter para la gestión de crisis de marca, específicamente en Bambos. Al igual que en nuestra investigación, se utilizaron algoritmos de regresión logística para analizar la polaridad de los tweets durante un período de crisis. La investigación demostró que los algoritmos permitieron mitigar el impacto negativo de la crisis, lo que está alineado con los resultados obtenidos en nuestro estudio, donde los algoritmos de business intelligence lograron reducir la proporción de tweets negativos en el post-test. Esto sugiere que estos algoritmos no solo son útiles en el análisis emocional, sino que también pueden desempeñar un papel fundamental en la gestión de la reputación online. Ambos estudios resaltan la importancia de implementar algoritmos que no solo analicen las emociones en redes sociales, sino que también ayuden a tomar decisiones rápidas y efectivas ante situaciones críticas.

El estudio de Vargas y Medina (2019), que utilizó técnicas de minería de datos para analizar la reputación online de McDonald's y KFC en Lima, es particularmente relevante para nuestro trabajo, ya que ambos estudios se enfocan en las mismas marcas y contexto local. Los autores encontraron que la polaridad emocional en los comentarios de los usuarios en redes sociales impactaba significativamente en la reputación de ambas cadenas. En nuestra investigación, los resultados del grupo experimental muestran que los algoritmos de business intelligence, especialmente los árboles de decisión, son efectivos para identificar la polaridad de los tweets. Sin embargo, a diferencia del enfoque de Vargas y Medina, nuestro estudio añade una dimensión de comparación entre algoritmos de business intelligence y Deep Learning, proporcionando una visión más amplia sobre qué tipo de algoritmos son más eficaces para el análisis de reputación. Mientras que el análisis de Vargas y Medina fue longitudinal, nuestra investigación, aunque limitada en el tiempo, demuestra que los algoritmos de business intelligence tienen un impacto inmediato en la mejora de la percepción emocional en redes sociales.

La comparación de los resultados obtenidos en nuestra investigación con los antecedentes peruanos revisados permite concluir que los algoritmos de *business intelligence*, como la regresión logística y los árboles de decisión, son herramientas altamente efectivas para el análisis de datos emocionales en redes sociales. Los estudios revisados confirman la

utilidad de estos algoritmos en contextos similares, como la industria de alimentos y bebidas, y en marcas internacionales con presencia en Perú. Si bien los algoritmos de Deep Learning han mostrado una ligera mejora en la precisión a lo largo del tiempo, los algoritmos de business intelligence destacan por su capacidad de generar resultados inmediatos y prácticos para la toma de decisiones estratégicas. Por tanto, se puede afirmar que estos algoritmos son una solución viable y eficiente para el análisis de la polaridad y la intensidad de las emociones en los tweets publicados por clientes de McDonald's y KFC en Lima.

CONCLUSIONES

1. En esta tesis se determinó la influencia de los algoritmos de business intelligence en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022. Los algoritmos de *business intelligence* fueron implementados y evaluados en una muestra de 384 mensajes recopilados. Los resultados respaldaron la hipótesis general planteada, la cual sugirió que la adopción de estos algoritmos influiría significativamente en el análisis emocional de los tweets mencionando estas cadenas de comida rápida. Este hallazgo se sustenta en la evidencia de que el test de *McNemar* arrojó un p-valor de $1.51e-37$, indicando una diferencia muy significativa con respecto al valor de significancia estándar de 0.05. Este resultado refuerza la utilidad y efectividad de los algoritmos de inteligencia de negocios en la interpretación y comprensión de las percepciones emocionales expresadas en Twitter sobre McDonald's y KFC en el contexto específico de Lima durante el último trimestre del año 2022.

2. Esta investigación analizó la influencia del algoritmo de regresión logística en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022. Respecto a la primera hipótesis específica, se encontró que el algoritmo de regresión logística tuvo un impacto significativo en el análisis de estas emociones, como lo demuestra el p-valor de $8.16e-20$ obtenido mediante el test de *McNemar*.

3. Esta investigación analizó la influencia del algoritmo de árboles de decisiones en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022. Respecto a la segunda hipótesis específica, se encontró que el algoritmo de árboles de decisiones tuvo un impacto significativo en el análisis de estas emociones, como lo demuestra el p-valor de $8.63e-22$ obtenido mediante el test de *McNemar*.

RECOMENDACIONES

Estas recomendaciones están dirigidas a la Dirección de Marketing, al equipo de Análisis de Datos, y a la alta gerencia de McDonald's y KFC en Lima, con el fin de optimizar la gestión y análisis de datos emocionales provenientes de interacciones en redes sociales, especialmente en Twitter (ahora llamado X).

1. Implementación de un sistema integral de análisis de datos emocionales, dirigido a la Dirección de Tecnología y Análisis de Datos. Se recomienda desarrollar e implementar un sistema integral que combine modelos de Business Intelligence (BI) y Deep Learning para el análisis de datos emocionales en redes sociales. Este sistema debe integrar algoritmos de regresión logística y árboles de decisión como preprocesadores de datos, junto con modelos avanzados como BERT para una mayor precisión en la clasificación de sentimientos. La implementación de este sistema permitirá mejorar la capacidad de predicción y respuesta ante las percepciones de los consumidores en tiempo real, optimizando la toma de decisiones estratégicas dentro de la empresa.
2. Desarrollo de un equipo especializado en analítica avanzada, dirigido a la Dirección de Recursos Humanos y la Alta Gerencia, se recomienda la formación de un equipo multidisciplinario conformado por especialistas en análisis de datos, inteligencia artificial, marketing y psicología del consumidor. Este equipo debe estar capacitado para interpretar los resultados obtenidos a través del sistema de análisis de datos emocionales y proponer estrategias de marketing y gestión basadas en la información obtenida. Además, se sugiere establecer alianzas con expertos en procesamiento de lenguaje natural (PLN) para fortalecer la capacidad analítica del equipo.
3. Optimización del monitoreo y respuesta a interacciones en redes sociales, dirigido a la Dirección de Atención al Cliente y Marketing, se recomienda la implementación de un sistema automatizado de monitoreo continuo de sentimientos en redes sociales, con protocolos de respuesta rápida ante comentarios negativos o tendencias emergentes. Este sistema debe integrar *dashboards* interactivos que permitan visualizar en tiempo real las

emociones y opiniones de los clientes, facilitando una gestión ágil de la reputación de la empresa y la optimización de sus estrategias de comunicación.

FUENTES DE INFORMACIÓN

- Banco Interamericano de Desarrollo. (1976). *Curso sobre preparación y evaluación de proyectos de desarrollo agrícola*. Tomo I. (n.d.). IICA Biblioteca Venezuela.
- Bernal Torres, C. A. (2006). *Metodología de la investigación: Para administración, economía, humanidades y ciencias sociales*. México D.F.: Pearson Educación.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Cambria, E., Schuller, B., Liu, B., Wang, H., & Havasi, C. (2012). Knowledge-based systems for sentiment analysis and opinion mining. *AI Magazine*, 33(3), 144-157.
- Cambria, E., Schuller, B., Liu, B., Wang, H., & Havasi, C. (2017). Knowledge-based systems for natural language processing. *Knowledge-Based Systems*, 108, 1-4. <https://doi.org/10.1016/j.knosys.2016.08.018>
- Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2017). New avenues in knowledge bases for natural language processing: Ontologies, embeddings, and deep learning. *IEEE Intelligent Systems*, 32(6), 43-52. <https://doi.org/10.1109/MIS.2017.4531230>
- Cegarra Sánchez, J. (2004). *Metodología de la Investigación Científica y Tecnológica*. Madrid: Ediciones Díaz de Santos.
- Cooke Curtis. (2011). *Analysis of Algorithms Used by the BI Software*. Oregon: Oregon State University.
- Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business Press.
- Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business Press.

- Davenport, T. H., & Patil, D. J. (2012). *Data Scientist: The Sexiest Job of the 21st Century*. Harvard Business Review, 90(10), 70-76.
- Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Harvard Business Press.
- Dei, D. (2006). *La tesis: Cómo orientarse en su elaboración*. Buenos Aires: Prometeo.
- Descartes, R. (1641). *Meditations on first philosophy*. Cambridge University Press.
- Domizi, J., & Roma, R. (2012). *Libro de Twitter*. Buenos Aires: Editorial Genes.
- Eppen, G. (2000). *Investigación de operaciones en la ciencia administrativa*. México: Prentice Hall Hispanoamérica S.A.
- Fernández Pampillón, A. (2018). *La eficacia del análisis de sentimientos para la empresa*. España: Universidad Complutense de Madrid.
- Fernández, P., & Soto, M. (2020). *Análisis de sentimientos en Twitter para la gestión de crisis de marca: Un estudio de caso de Bambos, Lima 2020*. Pontificia Universidad Católica del Perú.
- Ferrer, R., & Gómez, P. (2021). *Técnicas de recolección automatizada de datos en redes sociales: Principios y aplicaciones del web scraping*. Revista Internacional de Análisis Digital, 18(3), 95-110.
- Floridi, L. (2014). *The 4th revolution: How the infosphere is reshaping human reality*. Oxford University Press.
- Forouzan, B. (2003). *Introducción a la ciencia de la computación: De la manipulación de datos a la teoría de la computación*. México D.F.: Thomson.
- Fuentes Pinzón, F. (n.d.). *Epistemología ¿Qué es? ¿Para qué sirve?* - Epistemología. Recuperado el 11 de julio de 2020, de <https://www.youtube.com/watch?v=deqf2yMeLtQ>
- Gadamer, H. G. (1960). *Truth and method*. Sheed & Ward.

- Gámez, D. (2012). *Twitter*. Barcelona: Editorial Profit.
- Gandomi, A., & Haque, U. (2015). *Beyond the hype: Big data concepts, methods, and analytics*. *International Journal of Information Management*, 35(2), 137-144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- García Cambroner, & Gómez Moreno. (2008). *Algoritmos de aprendizaje: KNN & KMEANS*. En Conferencia Universidad Carlos III de Madrid.
- García Morente, M. (1996). *Obras completas II*. España: Editorial Anthropos.
- García, P., & Martin, L. (2018). *Leveraging Business Intelligence for Sentiment Analysis in Fast-Food Industry: The Case of Twitter*. *Computers in Human Behavior*, 92, 158-167. <https://doi.org/10.1016/j.chb.2018.09.013>
- Garza Caligaris, M. de L., & Romero Sánchez, M. de L. (2003). *Estancias infantiles comunitarias: Manual para educadoras*. México D.F.: Editorial Pax.
- Gómez, J., & Ramos, L. (2021). *Aplicación de algoritmos de machine learning en la analítica de datos emocionales para la predicción de la satisfacción del cliente en la cadena de restaurantes Norky's, Lima 2021*. Universidad Nacional Mayor de San Marcos.
- Gómez-Hernández, J., Rubio-Gálvez, M., & Ortuño-Castro, J. V. (Eds.). (2013). *El método en la geografía humana: una visión actual* (1st ed.). Madrid: Editorial Síntesis.
- Haba, E. P. (2004). *Elementos Básicos de Axiología General: Epistemología del discurso valorativo práctico*. San José, Costa Rica: Editorial de la Universidad de Costa Rica.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers.
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: Concepts and techniques* (3rd ed.). Elsevier.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación*. México D.F.: Mc Graw Hill Education.

- Hidalgo, R. (n.d.). *Ser, Ente, Ontología. (Óntico y Ontológico)*. Recuperado el 20 de mayo de 2020, de <https://www.youtube.com/watch?v=SqoOgGn47kM>
- Huamán Valencia, H. G. (2005). *Manual de Técnicas de Investigación: Conceptos y Aplicaciones*. Lima: IPLADEES S.A.C.
- Johnson, M., & Wang, T. (2021). *Emotional Data Analytics in Social Media: Insights from Fast-Food Brands*. *Journal of Marketing Analytics*, 29(2), 101-115. <https://doi.org/10.1057/s41270-021-00072-x>
- Joyanes Aguilar, L. (2013). *Big Data*. Ciudad de México: Alfaomega.
- Joyanes Aguilar, L. (2019). *Inteligencia de negocios y analítica de datos: Una visión global de business intelligence y analytics*. Bogotá: Alfaomega.
- Kitchenham, B., & Charters. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*. USA: Engineering.
- Krippendorff, K. (1997). *Metodología de análisis de contenido*. Barcelona: Paidós Comunicación.
- Kumar, A., & Sebastian, T. M. (2012). *Sentiment Analysis*. *International Journal of Intelligent Systems and Applications*, 2012.
- Lee, C., & Kim, S. (2022). *Business Intelligence and Emotional Data Analytics: A Case Study of McDonald's and KFC in the Digital Age*. *Journal of Business Research*, 144, 123-136. <https://doi.org/10.1016/j.jbusres.2022.04.013>
- Liu, B. (2010). *Sentiment analysis and subjectivity*. En G. C. 2010. *Handbook of natural language processing*.
- Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
- Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers. <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>
- Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.
- Locke, J. (1690). *An essay concerning human understanding*. Oxford University Press.

- López de Mántaras Badía, R., & Meseguer González, P. (2021). *Inteligencia artificial*. España: Oronet.
- López, J. (1999). *Procesos de Investigación*. Caracas: Panapo.
- López, J. (2021). *El análisis emocional en redes sociales: Estudio de las percepciones colectivas en plataformas digitales*. Editorial Universitaria.
- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71-87. <https://doi.org/10.1287/orsc.2.1.71>
- Maureira, F. (2017). *¿Qué es la inteligencia?* España: Bubok Publishing S.L.
- Mayer Schonberger, V., & Cukier, K. (2013). *Big Data*. Ciudad de México: Turner Publicaciones S. L.
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093-1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Mendes-Moreira, J., Soares, C., & Jorge, A. M. (2012). *Cost-sensitive decision trees applied to financial benchmarking in the automotive industry*. *Decision Support Systems*, 52(2), 510-520.
- Mileydi Flores, M. & Otros (2022). *Investigaciones sobre mercados gastronómicos en escenarios de pandemia*. Religación Press.
- Miller, G. (1995). *Wordnet: a lexical database for english*. Communications of the ACM.
- Ministerio de Defensa. (2019). *Documentos de Seguridad y Defensa 79: La inteligencia artificial aplicada a la defensa*. España: Secretaría General Técnica.
- Mirhoseini, A., Goldie, A., Yazgan, M., et al. (2021). *A graph placement methodology for fast chip design*. *Nature*, 594, 207–212. <https://doi.org/10.1038/s41586-021-03544-w>
- Moreno, L., & Sánchez, A. (2020). *Análisis masivo de datos mediante web scraping: Aplicaciones en marketing digital*. Editorial Científica.

- Muñoz Razo, C. (1998). *Cómo elaborar y asesorar una investigación de Tesis*. Naucalpan de Juárez, México: Prentice Hall Hispanoamericana S.A.
- Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.
- Navarro Monzó, J. (1931). *La actualidad filosófica de Jacobo Boehme*. Montevideo: Editorial Mundo Nuevo.
- Negash, S., & Gray, P. (2008). Business Intelligence. En W. Khosrow-Pour (Ed.), *Encyclopedia of Information Science and Technology* (2nd ed., pp. 176-183). IGI Global.
- Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford University Press.
- Oliva Valdebenito, F. I. (2014). *Minería de opinión y análisis de sentimientos*. Valparaíso: Pontificia Universidad Católica de Valparaíso.
- Ortiz Folgado, Á., Pérez La Madrid, O., & Vargas Rastrollo, E. (2015). *Estudio de tendencias diarias en Twitter*. Madrid: Universidad Complutense de Madrid.
- Ospina Arias, P. D. (2021). *Revisión de algoritmos de machine learning y deep learning apropiados para la implementación de visión artificial y movimientos autónomos en un brazo robótico*. Colombia: Universidad Nacional Abierta y a Distancia UNAD.
- Palma Méndez, J., & Marín Morales, R. (2008). *Inteligencia Artificial: Técnicas, métodos y aplicaciones*. España: McGraw Hill.
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1-135.
<https://doi.org/10.1561/15000000011>
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). *Thumbs up? Sentiment Classification using Machine Learning Techniques*. En Proceedings of the ACL-02 conference on Empirical methods in natural language processing - Volume 10.

- Patel, R., & Gupta, S. (2019). *Sentiment Analysis of Fast-Food Chains Using Business Intelligence Techniques: A Comparative Study of McDonald's and KFC*. *International Journal of Data Science and Analytics*, 8(4), 175-190. <https://doi.org/10.1007/s41060-019-00102-3>
- Peirce, C. S. (1878). How to make our ideas clear. *Popular Science Monthly*, 12, 286-302.
- Pérez, I., & León, B. (2007). *Lógica difusa para principiantes*. Caracas: Publicaciones UCAB.
- Picard, R. W. (1997). *Affective computing*. MIT Press.
- Pino Salvatierra, G. (2018). *Posicionamiento de marca y comportamiento del consumidor de Kentucky Fried Chicken (KFC), Independencia, 2018*. Lima: Universidad César Vallejo.
- Polanyi, M. (1966). *The tacit dimension*. University of Chicago Press.
- Ponce Cruz, P. (2010). *Inteligencia Artificial: Con aplicaciones a la ingeniería*. México D.F.: Alfaomega.
- Primiero, G. (2020). *On the foundations of computing*. United Kingdom: Oxford University Press.
- Quesada Moreno, J. F., & De Amores Carredano, J. G. (2000). *Diseño e Implementación de Sistemas de Traducción Automática*. España: Universidad de Sevilla, Secretariado de Publicaciones.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1, 81-106. <https://doi.org/10.1007/BF00116251>
- Ramírez, L., & Quispe, J. (2020). *Implementación de un sistema de business intelligence para el análisis de sentimientos en redes sociales: Estudio en la industria de alimentos y bebidas, Lima 2020*. Universidad Nacional de Ingeniería.

- Reguera, A. (2008). *Metodología de la investigación lingüística: Prácticas de escritura*. Córdoba: Editorial Brujas.
- Rissoan, R. (2016). *Redes sociales*. Barcelona: Ediciones ENI.
- Rouhiainen, L. (2019). *Inteligencia Artificial*. Barcelona: Alienta Editorial.
- Sedgewick, R. (1995). *Algoritmos en C++*. Madrid: Ediciones Addison-Wesley y Díaz de Santos.
- Sharda, R., Delen, D., & Turban, E. (2020). *Business Intelligence, Analytics, and Data Science: A Managerial Perspective* (4th ed.). Pearson.
- Sharda, R., Delen, D., & Turban, E. (2020). *Business Intelligence, Analytics, and Data Science: A Managerial Perspective* (4th ed.). Pearson.
- Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS Quarterly*, 35(3), 553-572. <https://doi.org/10.2307/23042796>
- Silva, M. (2020). *Impacto de las redes sociales en la imagen corporativa: Un enfoque desde el análisis de sentimientos*. *Revista de Comunicación y Mercadotecnia*, 15(3), 95-102.
- Smith, J., & Lee, A. (2020). *Application of Business Intelligence Algorithms in Social Media Sentiment Analysis: A Case Study of Fast-Food Industry*. *Journal of Business Analytics*, 15(3), 235-252. <https://doi.org/10.1016/j.jba.2020.05.008>
- Torres, M., & Cabrera, A. (2021). *Detección de patrones emocionales en redes sociales para la optimización de campañas de marketing: Caso Starbucks, Lima 2021*. Universidad Peruana de Ciencias Aplicadas.
- Treviño, L. (2019). *Inteligencia Emocional: Para que puedas dirigir tu vida*. España: Lulu.com.
- Turban, E., Sharda, R., & Delen, D. (2020). *Business Intelligence: A Managerial Perspective on Analytics* (5th ed.). Pearson Education.

- Turban, E., Sharda, R., Aronson, J. E., & King, D. (2007). *Business Intelligence: A Managerial Approach*. London: Prentice Hall.
- Tuya, J., Ramos Román, I., & Dolado Cosin, J. (2007). *Técnicas cuantitativas para la gestión en la ingeniería de software*. España: Netbiblo, S.L.
- Valdecantos Flores, M. (2019). *Aspectos legales en entornos digitales*. España: Editorial Elearning S.L.
- Vargas, R., & Medina, E. (2019). *Uso de técnicas de minería de datos para el análisis de la reputación online en cadenas de comida rápida: Caso McDonald's y KFC, Lima 2019*. Universidad de Lima.
- Williamson, B. (2018). *Big data en Educación: El futuro digital del aprendizaje, la política y la práctica*. Madrid: Ediciones Morata S. L.
- Zanoni, L. (2019). *Las máquinas no pueden soñar: Pasado, presente y futuro de la Inteligencia Artificial (la tecnología que cambiará el mundo)*. Buenos Aires: InclusivePublishing.
- Zepeda Hernández, S., Abascal, Mena, R., López, R., & Ornelas, E. (2016). *Análisis cualitativo de experiencias y emociones de los alumnos en el aula*. Ra Ximhai, 23(23). Recuperado el 21 de febrero de 2020, de <https://www.redalyc.org/articulo.oa?id=46148194024>

ANEXO 1. Matriz de consistencia

PREGUNTA PRINCIPAL	OBJETIVO GENERAL	HIPÓTESIS GENERAL	VARIABLES	METODOLOGÍA
¿Cómo los algoritmos de business intelligence influyen en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022?	Determinar la influencia de los algoritmos de business intelligence en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.	H ₁ . Los algoritmos de business intelligence influyen en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.	VI: Algoritmos de inteligencia de negocios. Dimensiones: – Regresión logística. – Árboles de decisiones.	– Enfoque: cuantitativo. – Tipo de investigación: aplicada. – Nivel de investigación: descriptiva explicativa. – Método de investigación: hipotético deductivo. – Diseño de investigación: experimental puro, con preprueba y posprueba, y grupo control activo.
PREGUNTAS ESPECÍFICAS	OBJETIVOS ESPECÍFICOS	HIPÓTESIS ESPECÍFICAS	VD: Analítica de datos emocionales. Dimensiones: – Polaridad. – Intensidad.	– Población: 270058 mensajes. – Muestra: 384 mensajes por medio del muestreo aleatorio simple. – Técnica: observación. – Instrumentos: dos fichas de observaciones.
P ₁ . ¿Cómo el algoritmo de regresión logística influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022?	1) Analizar la influencia del algoritmo de regresión logística en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.	He ₁ . El algoritmo de regresión logística influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.		
P ₂ . ¿Cómo el algoritmo de árboles de decisiones influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022?	2) Analizar la influencia del algoritmo de árboles de decisiones en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022	He ₂ . El algoritmo de árboles de decisiones influye en la analítica de los datos emocionales de los tweets publicados por clientes de McDonald's y KFC, Lima 2022.		

ANEXO 2. Instrumento de recolección de datos.

FICHA DE OBSERVACIÓN 1

Variable independiente: Algoritmos de inteligencia de negocios	Pre-test				Pos-test			
	Mensajes de control		Mensajes experimentales		Mensajes de control		Mensajes experimentales	
Dimensión 1: Regresión logística	Si	No	Si	No	Si	No	Si	No
1. ¿El algoritmo identificó correctamente la emoción o sentimiento principal del tweet?	☐	☐	☐	☐	☐	☐	☐	☐
2. ¿El algoritmo identificó correctamente el tema del tweet?	☐	☐	☐	☐	☐	☐	☐	☐
3. ¿El algoritmo identificó correctamente las ambigüedades presentes en el tweet?	☐	☐	☐	☐	☐	☐	☐	☐
4. ¿El algoritmo identificó correctamente los tweets que contenían múltiples emociones?	☐	☐	☐	☐	☐	☐	☐	☐
5. ¿El algoritmo procesó adecuadamente los tweets con abreviaturas o contracciones?	☐	☐	☐	☐	☐	☐	☐	☐
6. ¿El algoritmo detectó con precisión los cambios de polaridad dentro de un mismo tweet?	☐	☐	☐	☐	☐	☐	☐	☐
7. ¿El algoritmo clasificó de forma correcta los tweets que contenían lenguaje formal?	☐	☐	☐	☐	☐	☐	☐	☐
8. ¿El algoritmo ignoró los elementos irrelevantes que no contribuyen al análisis de sentimientos?	☐	☐	☐	☐	☐	☐	☐	☐
9. ¿El algoritmo interpretó correctamente el uso de emoticonos en los tweets?	☐	☐	☐	☐	☐	☐	☐	☐
10. ¿El algoritmo fue consistente en la identificación de emociones entre tweets similares?	☐	☐	☐	☐	☐	☐	☐	☐
11. ¿El algoritmo manejó correctamente los tweets con frases complejas o subordinadas?	☐	☐	☐	☐	☐	☐	☐	☐
12. ¿El algoritmo identificó adecuadamente los tweets que expresan neutralidad emocional?	☐	☐	☐	☐	☐	☐	☐	☐
13. ¿El algoritmo procesó correctamente los tweets que contienen hashtags emocionales?	☐	☐	☐	☐	☐	☐	☐	☐

14. ¿El algoritmo fue capaz de manejar el uso de sarcasmo en los tweets?	☐	☐	☐	☐	☐	☐	☐	☐
15. ¿El algoritmo detectó correctamente los tweets que utilizan ironía para expresar emociones?	☐	☐	☐	☐	☐	☐	☐	☐
Dimensión 2: Árboles de decisiones	Si	No	Si	No	Si	No	Si	No
16. ¿El algoritmo identificó correctamente la emoción o sentimiento principal del tweet?	☐	☐	☐	☐	☐	☐	☐	☐
17. ¿El algoritmo identificó correctamente el tema del tweet?	☐	☐	☐	☐	☐	☐	☐	☐
18. ¿El algoritmo identificó correctamente las ambigüedades presentes en el tweet?	☐	☐	☐	☐	☐	☐	☐	☐
19. ¿El algoritmo identificó correctamente los tweets que contenían múltiples emociones?	☐	☐	☐	☐	☐	☐	☐	☐
20. ¿El algoritmo manejó adecuadamente la jerarquía de palabras clave en los tweets?	☐	☐	☐	☐	☐	☐	☐	☐
21. ¿El algoritmo procesó correctamente los tweets con lenguaje informal o coloquial?	☐	☐	☐	☐	☐	☐	☐	☐
22. ¿El algoritmo ignoró correctamente los datos irrelevantes dentro del tweet para clasificar la emoción?	☐	☐	☐	☐	☐	☐	☐	☐
23. ¿El algoritmo clasificó adecuadamente los tweets que contenían más de un tema?	☐	☐	☐	☐	☐	☐	☐	☐
24. ¿El algoritmo detectó los cambios de tono dentro del mismo tweet?	☐	☐	☐	☐	☐	☐	☐	☐
25. ¿El algoritmo mantuvo consistencia en la identificación de emociones en tweets con estructura similar?	☐	☐	☐	☐	☐	☐	☐	☐
26. ¿El algoritmo fue capaz de procesar tweets con menciones a otras cuentas?	☐	☐	☐	☐	☐	☐	☐	☐
27. ¿El algoritmo identificó correctamente los tweets que presentaban opiniones ambiguas?	☐	☐	☐	☐	☐	☐	☐	☐
28. ¿El algoritmo procesó adecuadamente los tweets que utilizaban símbolos o emojis?	☐	☐	☐	☐	☐	☐	☐	☐
29. ¿El algoritmo fue capaz de diferenciar entre emociones sutiles (por ejemplo, alegría leve vs. euforia)?	☐	☐	☐	☐	☐	☐	☐	☐
30. ¿El algoritmo detectó correctamente la negación dentro de un tweet para cambiar el sentido emocional?	☐	☐	☐	☐	☐	☐	☐	☐

FICHA DE OBSERVACIÓN 2

Variable dependiente: Analítica de datos emocionales	Pre-test				Pos-test			
	Mensajes de control		Mensajes experimentales		Mensajes de control		Mensajes experimentales	
Dimensión 1: Polaridad	Si	No	Si	No	Si	No	Si	No
1. ¿Se clasificó correctamente la polaridad del tweet?	☐	☐	☐	☐	☐	☐	☐	☐
2. ¿Se asignó correctamente la puntuación de sentimiento al tweet?	☐	☐	☐	☐	☐	☐	☐	☐
3. ¿Se detectaron correctamente los tweets sin polaridad, es decir neutrales?	☐	☐	☐	☐	☐	☐	☐	☐
4. ¿Se distinguió adecuadamente entre polaridad positiva y negativa en un tweet ambiguo?	☐	☐	☐	☐	☐	☐	☐	☐
5. ¿Se mantuvo consistencia en la clasificación de polaridad entre tweets similares?	☐	☐	☐	☐	☐	☐	☐	☐
6. ¿Se procesaron correctamente los tweets con un cambio de polaridad en su contenido?	☐	☐	☐	☐	☐	☐	☐	☐
7. ¿Se ignoraron las palabras sin relevancia emocional al clasificar la polaridad del tweet?	☐	☐	☐	☐	☐	☐	☐	☐
8. ¿Se identificaron correctamente los tweets que contenían tanto elementos positivos como negativos?	☐	☐	☐	☐	☐	☐	☐	☐
9. ¿Se detectaron correctamente los tweets que cambiaron de polaridad a lo largo del mensaje?	☐	☐	☐	☐	☐	☐	☐	☐
10. ¿Se clasificaron correctamente los tweets con sarcasmo en términos de polaridad?	☐	☐	☐	☐	☐	☐	☐	☐
11. ¿Se identificó adecuadamente la polaridad de los tweets que contienen emojis?	☐	☐	☐	☐	☐	☐	☐	☐
12. ¿Se gestionaron correctamente los tweets con polaridad inversa debido a negaciones (por ejemplo “no me gusta”)?	☐	☐	☐	☐	☐	☐	☐	☐
13. ¿Se clasificaron correctamente los tweets con polaridad subjetiva (opiniones personales)?	☐	☐	☐	☐	☐	☐	☐	☐
14. ¿Se identificaron correctamente los tweets que presentan polaridad extrema (muy positiva o muy negativa)?	☐	☐	☐	☐	☐	☐	☐	☐

15. ¿Se asignó correctamente la polaridad a tweets con múltiples temas?	☐	☐	☐	☐	☐	☐	☐	☐
Dimensión 2: Intensidad	Si	No	Si	No	Si	No	Si	No
16. ¿Se identificó correctamente el porcentaje de polaridad positiva en el tweet?	☐	☐	☐	☐	☐	☐	☐	☐
17. ¿Se identificó correctamente el porcentaje de polaridad negativa en el tweet?	☐	☐	☐	☐	☐	☐	☐	☐
18. ¿Se identificó correctamente la intensidad emocional del tweet?	☐	☐	☐	☐	☐	☐	☐	☐
19. ¿Se detectó adecuadamente la intensidad emocional en tweets largos?	☐	☐	☐	☐	☐	☐	☐	☐
20. ¿Se procesaron correctamente los tweets con cambios de intensidad emocional a lo largo del mensaje?	☐	☐	☐	☐	☐	☐	☐	☐
21. ¿Se identificó adecuadamente la intensidad emocional en tweets con contenido ambiguo?	☐	☐	☐	☐	☐	☐	☐	☐
22. ¿Se mantuvo consistencia en la medición de la intensidad entre tweets similares?	☐	☐	☐	☐	☐	☐	☐	☐
23. ¿Se identificaron correctamente los tweets con emociones de baja intensidad?	☐	☐	☐	☐	☐	☐	☐	☐
24. ¿Se asignó correctamente la intensidad emocional a tweets con más de una emoción?	☐	☐	☐	☐	☐	☐	☐	☐
25. ¿Se gestionaron adecuadamente los tweets con fluctuaciones rápidas de intensidad emocional?	☐	☐	☐	☐	☐	☐	☐	☐
26. ¿Se detectaron correctamente los tweets con emociones intensas expresadas de forma indirecta?	☐	☐	☐	☐	☐	☐	☐	☐
27. ¿Se midió adecuadamente la intensidad en tweets que utilizan emojis para amplificar el sentimiento?	☐	☐	☐	☐	☐	☐	☐	☐
28. ¿Se identificaron correctamente los tweets con alta intensidad emocional en el contexto de una discusión más amplia?	☐	☐	☐	☐	☐	☐	☐	☐
29. ¿Se asignó correctamente la intensidad a tweets con múltiples sentimientos superpuestos?	☐	☐	☐	☐	☐	☐	☐	☐
30. ¿Se detectaron correctamente las emociones de intensidad moderada en los tweets con lenguaje neutro?	☐	☐	☐	☐	☐	☐	☐	☐

ANEXO 3. Fichas de validación del instrumento

VICERRECTORADO ACADÉMICO

ESCUELA DE POSGRADO

FICHA DE VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN DE LA VARIABLE INDEPENDIENTE

1. DATOS GENERALES

- 1.1. Apellidos y nombres del experto : Ramos Rivera Salomón Rey
- 1.2. Grado académico : Doctor
- 1.3. Cargo e institución donde labora : Docente ordinario - Universidad Nacional del Moquegua
- 1.4. Título de la investigación : Algoritmos de business intelligence en la analítica de datos emocionales de tweets publicados por clientes de MCDONALD'S y KFC, Lima 2022
- 1.5. Autor del instrumento : Mag. Ruso Alexander Morales Gonzales
- 1.6. Carrera profesional : Doctorado en Ingeniería de Sistemas
- 1.7. Nombre del instrumento : Ficha de observación 01

INDICACIONES	CRITERIOS	Deficiente 0-10	Regular 10-13	Bueno 14-16	Muy bueno 17-18	Excelente 19-20
1. CLARIDAD	Esta formulado con lenguaje apropiado	-	-	-	-	19
2. OBJETIVIDAD	Esta expresado con conductas observables	-	-	-	-	20
3. ACTUALIDAD	Adecuado al avance de la ciencia tecnológica	-	-	-	-	20
4. ORGANIZACIÓN	Existe un organización y lógica	-	-	-	-	19
5. SUFICIENCIA	Comprende los aspectos en cantidad y calidad	-	-	-	-	19
6. INTENCIONALIDAD	Adecuado para valorar los aspectos de estudio	-	-	-	-	19
7. CONSISTENCIA	Basado en el aspecto teórico científico y del tema de estudio	-	-	-	-	19
8. COHERENCIA	Entre las variables, dimensiones y variables	-	-	-	-	20
9. METODOLOGÍA	La estrategia responde al propósito del estudio	-	-	-	-	20
10. CONVENIENCIA	Genera nuevas pautas para la investigación y construcción de teorías.	-	-	-	-	19
Promedios parciales		-	-	-	-	19.4
Promedio final		19.4				

Valoración cuantitativa: 19.4

Valoración cualitativa: Válido, aplicar

Opinión de aplicabilidad: Es aplicable

DNI: 29607679
Moquegua 06/09/2024

VICERRECTORADO ACADÉMICO

ESCUELA DE POSGRADO

FICHA DE VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN DE LA VARIABLE INDEPENDIENTE

1. DATOS GENERALES

1.1. Apellidos y nombres del experto : Vera Ramírez Oscar John
1.2. Grado académico : Doctor
1.3. Cargo e institución donde labora : Docente ordinario - Universidad Nacional del Moquegua
1.4. Título de la investigación : Algoritmos de business intelligence en la analítica de datos emocionales de tweets publicados por clientes de MCDONALD'S y KFC, Lima 2022
1.5. Autor del instrumento : Mag. Ruso Alexander Morales Gonzales
1.6. Carrera profesional : Doctorado en Ingeniería de Sistemas
1.7. Nombre del instrumento : Ficha de observación 01

INDICACIONES	CRITERIOS	Deficiente 0-10	Regular 10-13	Bueno 14-16	Muy bueno 17-18	Excelente 19-20
1. CLARIDAD	Esta formulado con lenguaje apropiado	-	-	-	-	19
2. OBJETIVIDAD	Esta expresado con conductas observables	-	-	-	-	19
3. ACTUALIDAD	Adecuado al avance de la ciencia tecnológica	-	-	-	-	19
4. ORGANIZACIÓN	Existe un organización y lógica	-	-	-	-	19
5. SUFICIENCIA	Comprende los aspectos en cantidad y calidad	-	-	-	-	19
6. INTENCIONALIDAD	Adecuado para valorar los aspectos de estudio	-	-	-	-	19
7. CONSISTENCIA	Basado en el aspecto teórico científico y del tema de estudio	-	-	-	-	19
8. COHERENCIA	Entre las variables, dimensiones y variables	-	-	-	-	19
9. METODOLOGÍA	La estrategia responde al propósito del estudio	-	-	-	-	19
10. CONVENIENCIA	Genera nuevas pautas para la investigación y construcción de teorías.	-	-	-	-	19
Promedios parciales		-	-	-	-	19
Promedio final		19				

Valoración cuantitativa: 19

Valoración cualitativa: Válido, aplicar

Opinión de aplicabilidad: Es aplicable

DNI: 29680411
Moquegua 06/09/2024

VICERRECTORADO ACADÉMICO

ESCUELA DE POSGRADO

FICHA DE VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN DE LA VARIABLE INDEPENDIENTE

1. DATOS GENERALES

- 1.1. Apellidos y nombres del experto : Herrera Quispe José Alfredo
1.2. Grado académico : Doctor
1.3. Cargo e institución donde labora : Docente ordinario - Universidad Nacional Mayor de San Marcos
1.4. Título de la investigación : Algoritmos de business intelligence en la analítica de datos emocionales de tweets publicados por clientes de MCDONALD'S y KFC, Lima 2022
1.5. Autor del instrumento : Mag. Ruso Alexander Morales Gonzales
1.6. Carrera profesional : Doctorado en Ingeniería de Sistemas
1.7. Nombre del instrumento : Ficha de observación 01

INDICACIONES	CRITERIOS	Deficiente 0-10	Regular 10-13	Bueno 14-16	Muy bueno 17-18	Excelente 19-20
1. CLARIDAD	Esta formulado con lenguaje apropiado	-	-	-	-	20
2. OBJETIVIDAD	Esta expresado con conductas observables	-	-	-	-	20
3. ACTUALIDAD	Adecuado al avance de la ciencia tecnológica	-	-	-	-	20
4. ORGANIZACIÓN	Existe un organización y lógica	-	-	-	-	20
5. SUFICIENCIA	Comprende los aspectos en cantidad y calidad	-	-	-	-	19
6. INTENCIONALIDAD	Adecuado para valorar los aspectos de estudio	-	-	-	-	19
7. CONSISTENCIA	Basado en el aspecto teórico científico y del tema de estudio	-	-	-	-	20
8. COHERENCIA	Entre las variables, dimensiones y variables	-	-	-	-	19
9. METODOLOGÍA	La estrategia responde al propósito del estudio	-	-	-	-	19
10. CONVENIENCIA	Genera nuevas pautas para la investigación y construcción de teorías.	-	-	-	-	19
Promedios parciales		-	-	-	-	19.5
Promedio final		19.5				

Valoración cuantitativa: 19.5

Valoración cualitativa: Válido, aplicar

Opinión de aplicabilidad: Es aplicable

DNI: 40362859
Lima 06/09/2024

VICERRECTORADO ACADÉMICO

ESCUELA DE POSGRADO

FICHA DE VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN DE LA VARIABLE INDEPENDIENTE

1. DATOS GENERALES

1.1. Apellidos y nombres del experto : Limache Sandoval Elmer Marcial
1.2 Grado académico : Doctor
1.3 Cargo e institución donde labora : Docente ordinario - Universidad Privada de Tacna
1.4 Título de la investigación: : Algoritmos de business intelligence en la analítica de datos emocionales de tweets publicados por clientes de MCDONALD'S y KFC, Lima 2022
1.5 Autor del instrumento : Mag. Ruso Alexander Morales Gonzales
1.6 Carrera profesional : Doctorado en Ingeniería de Sistemas
1.7 Nombre del instrumento : Ficha de observación 01

INDICACIONES	CRITERIOS	Deficiente 0-10	Regular 10-13	Bueno 14-16	Muy bueno 17-18	Excelente 19-20
1. CLARIDAD	Esta formulado con lenguaje apropiado	-	-	-	-	19
2. OBJETIVIDAD	Esta expresado con conductas observables	-	-	-	-	19
3. ACTUALIDAD	Adecuado al avance de la ciencia tecnológica	-	-	-	-	19
4. ORGANIZACIÓN	Existe un organización y lógica	-	-	-	-	19
5. SUFICIENCIA	Comprende los aspectos en cantidad y calidad	-	-	-	-	19
6. INTENCIONALIDAD	Adecuado para valorar los aspectos de estudio	-	-	-	-	19
7. CONSISTENCIA	Basado en el aspecto teórico científico y del tema de estudio	-	-	-	-	20
8. COHERENCIA	Entre las variables, dimensiones y variables	-	-	-	-	19
9. METODOLOGÍA	La estrategia responde al propósito del estudio	-	-	-	-	19
10. CONVENIENCIA	Genera nuevas pautas para la investigación y construcción de teorías.	-	-	-	-	19
Promedios parciales		-	-	-	-	19.1
Promedio final		19.1				

Valoración cuantitativa: 19.1

Valoración cualitativa: Válido, aplicar

Opinión de aplicabilidad: Es aplicable

DNI: 06094414
Tacna 07/09/2024

VICERRECTORADO ACADÉMICO

ESCUELA DE POSGRADO

FICHA DE VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN DE LA VARIABLE INDEPENDIENTE

1. DATOS GENERALES

1.1. Apellidos y nombres del experto : Flores García Aníbal Fernando
1.2 Grado académico : Doctor
1.3 Cargo e institución donde labora : Docente ordinario - Universidad Nacional del Moquegua
1.4 Título de la investigación: : Algoritmos de business intelligence en la analítica de datos emocionales de tweets publicados por clientes de MCDONALD'S y KFC, Lima 2022
1.5 Autor del instrumento : Mag. Ruso Alexander Morales Gonzales
1.6 Carrera profesional : Doctorado en Ingeniería de Sistemas
1.7 Nombre del instrumento : Ficha de observación 01

INDICACIONES	CRITERIOS	Deficiente 0-10	Regular 10-13	Bueno 14-16	Muy bueno 17-18	Excelente 19-20
1. CLARIDAD	Esta formulado con lenguaje apropiado	-	-	-	-	19
2. OBJETIVIDAD	Esta expresado con conductas observables	-	-	-	-	20
3. ACTUALIDAD	Adecuado al avance de la ciencia tecnológica	-	-	-	-	20
4. ORGANIZACIÓN	Existe un organización y lógica	-	-	-	-	20
5. SUFICIENCIA	Comprende los aspectos en cantidad y calidad	-	-	-	-	20
6. INTENCIONALIDAD	Adecuado para valorar los aspectos de estudio	-	-	-	-	19
7. CONSISTENCIA	Basado en el aspecto teórico científico y del tema de estudio	-	-	-	-	20
8. COHERENCIA	Entre las variables, dimensiones y variables	-	-	-	-	19
9. METODOLOGÍA	La estrategia responde al propósito del estudio	-	-	-	-	20
10. CONVENIENCIA	Genera nuevas pautas para la investigación y construcción de teorías.	-	-	-	-	20
Promedios parciales		-	-	-	-	19.7
Promedio final		19.7				

Valoración cuantitativa: 19.7

Valoración cualitativa: Válido, aplicar

Opinión de aplicabilidad: Es aplicable

DNI: 04743476
Moquegua 07/09/2024

143

VICERRECTORADO ACADÉMICO

ESCUELA DE POSGRADO

FICHA DE VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN DE LA VARIABLE DEPENDIENTE

1. DATOS GENERALES

- 1.1. Apellidos y nombres del experto : Ramos Rivera Salomón Rey
- 1.2 Grado académico : Doctor
- 1.3 Cargo e institución donde labora : Docente ordinario - Universidad Nacional del Moquegua
- 1.4 Título de la investigación : Algoritmos de business intelligence en la analítica de datos emocionales de tweets publicados por clientes de MCDONALD'S y KFC, Lima 2022
- 1.5 Autor del instrumento : Mag. Ruso Alexander Morales Gonzales
- 1.6 Carrera profesional : Doctorado en Ingeniería de Sistemas
- 1.7 Nombre del instrumento : Ficha de observación 02

INDICACIONES	CRITERIOS	Deficiente 0-10	Regular 10-13	Bueno 14-16	Muy bueno 17-18	Excelente 19-20
1. CLARIDAD	Esta formulado con lenguaje apropiado	-	-	-	-	19
2. OBJETIVIDAD	Esta expresado con conductas observables	-	-	-	-	20
3. ACTUALIDAD	Adecuado al avance de la ciencia tecnológica	-	-	-	-	20
4. ORGANIZACIÓN	Existe un organización y lógica	-	-	-	-	20
5. SUFICIENCIA	Comprende los aspectos en cantidad y calidad	-	-	-	-	20
6. INTENCIONALIDAD	Adecuado para valorar los aspectos de estudio	-	-	-	-	19
7. CONSISTENCIA	Basado en el aspecto teórico científico y del tema de estudio	-	-	-	-	20
8. COHERENCIA	Entre las variables, dimensiones y variables	-	-	-	-	19
9. METODOLOGÍA	La estrategia responde al propósito del estudio	-	-	-	-	19
10. CONVENIENCIA	Genera nuevas pautas para la investigación y construcción de teorías.	-	-	-	-	19
Promedios parciales		-	-	-	-	19.5
Promedio final		19.5				

Valoración cuantitativa: 19.5

Valoración cualitativa: Válido, aplicar

Opinión de aplicabilidad: Es aplicable

DNI: 29607679

Moquegua 06/09/2024

VICERRECTORADO ACADÉMICO

ESCUELA DE POSGRADO

FICHA DE VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN DE LA VARIABLE DEPENDIENTE

1. DATOS GENERALES

- 1.1. Apellidos y nombres del experto : Vera Ramírez Oscar John
1.2 Grado académico : Doctor
1.3 Cargo e institución donde labora : Docente ordinario - Universidad Nacional del Moquegua
1.4 Título de la investigación : Algoritmos de business intelligence en la analítica de datos emocionales de tweets publicados por clientes de MCDONALD'S y KFC, Lima 2022
1.5 Autor del instrumento : Mag. Ruso Alexander Morales Gonzales
1.6 Carrera profesional : Doctorado en Ingeniería de Sistemas
1.7 Nombre del instrumento : Ficha de observación 02

INDICACIONES	CRITERIOS	Deficiente 0-10	Regular 10-13	Bueno 14-16	Muy bueno 17-18	Excelente 19-20
1. CLARIDAD	Esta formulado con lenguaje apropiado	-	-	-	-	20
2. OBJETIVIDAD	Esta expresado con conductas observables	-	-	-	-	20
3. ACTUALIDAD	Adecuado al avance de la ciencia tecnológica	-	-	-	-	20
4. ORGANIZACIÓN	Existe un organización y lógica	-	-	-	-	20
5. SUFICIENCIA	Comprende los aspectos en cantidad y calidad	-	-	-	-	19
6. INTENCIONALIDAD	Adecuado para valorar los aspectos de estudio	-	-	-	-	20
7. CONSISTENCIA	Basado en el aspecto teórico científico y del tema de estudio	-	-	-	-	19
8. COHERENCIA	Entre las variables, dimensiones y variables	-	-	-	-	19
9. METODOLOGÍA	La estrategia responde al propósito del estudio	-	-	-	-	19
10. CONVENIENCIA	Genera nuevas pautas para la investigación y construcción de teorías.	-	-	-	-	20
Promedios parciales		-	-	-	-	19.6
Promedio final		19.6				

Valoración cuantitativa: 19.6

Valoración cualitativa: Válido, aplicar

Opinión de aplicabilidad: Es aplicable

DNI: 29680411
Moquegua 06/09/2024

VICERRECTORADO ACADÉMICO

ESCUELA DE POSGRADO

FICHA DE VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN DE LA VARIABLE DEPENDIENTE

1. DATOS GENERALES

- 1.1. Apellidos y nombres del experto : Herrera Quispe José Alfredo
1.2. Grado académico : Doctor
1.3. Cargo e institución donde labora : Docente ordinario - Universidad Nacional Mayor de San Marcos
1.4. Título de la investigación : Algoritmos de business intelligence en la analítica de datos emocionales de tweets publicados por clientes de MCDONALD'S y KFC, Lima 2022
1.5. Autor del instrumento : Mag. Ruso Alexander Morales Gonzales
1.6. Carrera profesional : Doctorado en Ingeniería de Sistemas
1.7. Nombre del instrumento : Ficha de observación 02

INDICACIONES	CRITERIOS	Deficiente 0-10	Regular 10-13	Bueno 14-16	Muy bueno 17-18	Excelente 19-20
11. CLARIDAD	Esta formulado con lenguaje apropiado	-	-	-	-	19
12. OBJETIVIDAD	Esta expresado con conductas observables	-	-	-	-	19
13. ACTUALIDAD	Adecuado al avance de la ciencia tecnológica	-	-	-	-	19
14. ORGANIZACIÓN	Existe un organización y lógica	-	-	-	-	19
15. SUFICIENCIA	Comprende los aspectos en cantidad y calidad	-	-	-	-	19
16. INTENCIONALIDAD	Adecuado para valorar los aspectos de estudio	-	-	-	-	19
17. CONSISTENCIA	Basado en el aspecto teórico científico y del tema de estudio	-	-	-	-	19
18. COHERENCIA	Entre las variables, dimensiones y variables	-	-	-	-	19
19. METODOLOGÍA	La estrategia responde al propósito del estudio	-	-	-	-	20
20. CONVENIENCIA	Genera nuevas pautas para la investigación y construcción de teorías.	-	-	-	-	20
Promedios parciales		-	-	-	-	19.2
Promedio final		19.2				

Valoración cuantitativa: 19.2

Valoración cualitativa: Válido, aplicar

Opinión de aplicabilidad: Es aplicable

DNI: 40362859

Lima 06/09/2024

VICERRECTORADO ACADÉMICO

ESCUELA DE POSGRADO

FICHA DE VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN DE LA VARIABLE DEPENDIENTE

1. DATOS GENERALES

1.1. Apellidos y nombres del experto : Limache Sandoval Elmer Marcial
1.2 Grado académico : Doctor
1.3 Cargo e institución donde labora : Docente ordinario - Universidad Privada de Tacna
1.4 Título de la investigación: : Algoritmos de business intelligence en la analítica de datos emocionales de tweets publicados por clientes de MCDONALD'S y KFC, Lima 2022
1.5 Autor del instrumento : Mag. Ruso Alexander Morales Gonzales
1.6 Carrera profesional : Doctorado en Ingeniería de Sistemas
1.7 Nombre del instrumento : Ficha de observación 02

INDICACIONES	CRITERIOS	Deficiente 0-10	Regular 10-13	Bueno 14-16	Muy bueno 17-18	Excelente 19-20
11. CLARIDAD	Esta formulado con lenguaje apropiado	-	-	-	-	20
12. OBJETIVIDAD	Esta expresado con conductas observables	-	-	-	-	20
13. ACTUALIDAD	Adecuado al avance de la ciencia tecnológica	-	-	-	-	202
14. ORGANIZACIÓN	Existe un organización y lógica	-	-	-	-	0
15. SUFICIENCIA	Comprende los aspectos en cantidad y calidad	-	-	-	-	19
16. INTENCIONALIDAD	Adecuado para valorar los aspectos de estudio	-	-	-	-	20
17. CONSISTENCIA	Basado en el aspecto teórico científico y del tema de estudio	-	-	-	-	19
18. COHERENCIA	Entre las variables, dimensiones y variables	-	-	-	-	19
19. METODOLOGÍA	La estrategia responde al propósito del estudio	-	-	-	-	19
20. CONVENIENCIA	Genera nuevas pautas para la investigación y construcción de teorías.	-	-	-	-	19
Promedios parciales		-	-	-	-	19.5
Promedio final		19.5				

Valoración cuantitativa: 19.5

Valoración cualitativa: Válido, aplicar

Opinión de aplicabilidad: Es aplicable

DNI: 06094414
Tacna 07/09/2024

VICERRECTORADO ACADÉMICO

ESCUELA DE POSGRADO

FICHA DE VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN DE LA VARIABLE DEPENDIENTE

1. DATOS GENERALES

- 1.1. Apellidos y nombres del experto : Flores García Aníbal Fernando
1.2. Grado académico : Doctor
1.3. Cargo e institución donde labora : Docente ordinario - Universidad Nacional del Moquegua
1.4. Título de la investigación: : Algoritmos de business intelligence en la analítica de datos emocionales de tweets publicados por clientes de MCDONALD'S y KFC, Lima 2022
1.5. Autor del instrumento : Mag. Ruso Alexander Morales Gonzales
1.6. Carrera profesional : Doctorado en Ingeniería de Sistemas
1.7. Nombre del instrumento : Ficha de observación 02

INDICACIONES	CRITERIOS	Deficiente 0-10	Regular 10-13	Bueno 14-16	Muy bueno 17-18	Excelente 19-20
11. CLARIDAD	Esta formulado con lenguaje apropiado	-	-	-	-	19
12. OBJETIVIDAD	Esta expresado con conductas observables	-	-	-	-	19
13. ACTUALIDAD	Adecuado al avance de la ciencia tecnológica	-	-	-	-	19
14. ORGANIZACIÓN	Existe un organización y lógica	-	-	-	-	19
15. SUFICIENCIA	Comprende los aspectos en cantidad y calidad	-	-	-	-	19
16. INTENCIONALIDAD	Adecuado para valorar los aspectos de estudio	-	-	-	-	19
17. CONSISTENCIA	Basado en el aspecto teórico científico y del tema de estudio	-	-	-	-	19
18. COHERENCIA	Entre las variables, dimensiones y variables	-	-	-	-	20
19. METODOLOGÍA	La estrategia responde al propósito del estudio	-	-	-	-	20
20. CONVENIENCIA	Genera nuevas pautas para la investigación y construcción de teorías.	-	-	-	-	20
Promedios parciales		-	-	-	-	19.3
Promedio final		19.3				

Valoración cuantitativa: 19.3

Valoración cualitativa: Válido, aplicar

Opinión de aplicabilidad: Es aplicable

DNI: 04743476
Moquegua 07/09/2024

ANEXO 4. Copia de los datos procesados

6

Variable independiente pre-test Grupo Control

Nº	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
33	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
34	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
37	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

150

151

152

153

154

155

156

157

6

Variable independiente post-test Grupo Experimental

N°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
31	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
39	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

159

160

161

162

163

164

165

166

6

Variable dependiente pre-test Grupo Control

N°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
31	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
33	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
34	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
35	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
37	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
39	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
40	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
41	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

168

169

170

171

172

173

174

175

6

Variable dependiente post-test Grupo Experimental

N°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
31	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
33	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
34	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
35	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
37	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
39	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
40	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
41	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

177

178

179

180

181

182

183

184

ANEXO 5. Autorización de la entidad.

KFC <servicioalcliente@franquiciasperu.com>

Jue 16/07/2020 17:08

Para: Ing. Alexander Morales <lexruso@hotmail.com>

CC: KFC <domicilios@kfc.co>

Estimado Sr. Alexander Morales.

Agradecemos el uso de este canal de comunicaciones, y le manifestamos que estamos interesados en lo que usted va a realizar y damos nuestro consentimiento. Le comentamos que hemos recibido suficiente información sobre la investigación. Sabemos cómo comunicarnos con usted si tuviéramos preguntas. Comprendemos la importancia que tiene su investigación. Comprendemos que podemos retirarnos si fuera el caso.

Suerte en su trabajo.

©2020 Kentucky Fried Chicken Perú Company. Todos los Derechos Reservados.



LÍNEA ÉTICA

De: Ing. Alexander Morales <lexruso@hotmail.com>

Enviado: viernes, 10 de julio de 2020 16:41

Para: KFC <servicioalcliente@franquiciasperu.com>

Asunto: Consentimiento Informado para Investigación

■ 1 archivo adjunto (84 KB)

Plan de investigación.pdf

Saludos muy cordiales, me represento, soy el aspirante al grado de doctor Ruso Alexander Morales Gonzales, del posgrado en Ingeniería de Sistemas de la Universidad Alas Peruanas en Perú. Como parte de la forma para obtener el grado estoy realizando una investigación de tesis cuyo objetivo pragmático es analizar los tuits que hablen, mencionen o se refieran a KFC, y con ellos realizar un análisis de sentimientos que servirían para saber la polaridad positiva o negativa de los comentarios. Una vez concluida la investigación, los resultados de ella servirán para mejorar su servicio. Debo mencionar también que los mensajes de Twitter son de carácter público y no trasgreden ninguna violación de la privacidad, pero de todas formas tengo la obligación de darle a conocer mis acciones y que su institución me de su consentimiento informado. Le alcanzo un archivo adjunto en donde se ven los detalles del plan de tesis.

Muchísimas gracias.

Enviado desde [Outlook](#)

McDonalds <atencionacientes@mx.mcd.com>

Jue 30/07/2020 12:58

Para: lexruso@hotmail.com <lexruso@hotmail.com>

Sr. Alexander Morales.

Sírvase de seguir adelante con su investigación, no tenemos ninguna objeción de que use datos públicos tal como señalo en su archivo adjunto, damos nuestro consentimiento porque contamos con la suficiente información sobre el estudio, puede hacer uso del nombre de la empresa con finalidades académicas, y de los datos recolectados si tuviéramos preguntas usaremos este mismo medio para hacerlas. Le deseamos muchos éxitos.



©2017-2020 McDonald's. All Rights Reserved.

De: Ing. Alexander Morales <lexruso@hotmail.com>

Enviado: viernes, 10 de julio de 2020 16:58

Para: McDonalds <atencionacientes@mx.mcd.com>

CC: McD <hr@kcftech.com>

Asunto: Consentimiento Informado para Investigación

■ 1 archivo adjunto (948 KB)

Plan de tesis McDonalds.docx;

Saludos muy cordiales, me represento, soy el aspirante al grado de doctor Ruso Alexander Morales Gonzales, del posgrado en Ingeniería de Sistemas de la Universidad Alas Peruanas en Perú. Como parte de la forma para obtener el grado estoy realizando una investigación de tesis cuyo objetivo pragmático es analizar los tuits que hablen, mencionen o se refieran a MC DONALDS, y con ellos realizar un análisis de sentimientos que servirían para saber la polaridad positiva o negativa de los comentarios. Una vez concluida la investigación, los resultados de ella servirán para mejorar su servicio. Debo mencionar también que los mensajes de Twitter son de carácter público y no trasgreden ninguna violación de la privacidad, pero de todas formas tengo la obligación de darle a conocer mis acciones y que su institución me de su consentimiento informado. Le alcanzo un archivo adjunto en donde se ven los detalles del plan de tesis.

Muchísimas gracias.

Enviado desde [Outlook](#)

ANEXO 6: Consentimiento informado

CONSENTIMIENTO INFORMADO PARA LA PARTICIPACIÓN EN EL ESTUDIO

Título del estudio: **ALGORITMOS DE BUSINESS INTELLIGENCE EN LA ANALÍTICA DE DATOS EMOCIONALES DE TWEETS PUBLICADOS POR CLIENTES DE MCDONALD'S Y KFC, LIMA 2022**

Investigador Principal: MAG. RUSSO ALEXANDER MORALES GONZALES,
ORCID: 0000-0003-4077-7271
rmoralesg@unam.edu.pe
968448722

1. INTRODUCCIÓN

Usted ha sido informado sobre el estudio titulado "Algoritmos de Business Intelligence en la Analítica de Datos Emocionales de los Mensajes Tweets sobre McDonald's y KFC, Lima 2022". Este documento explica el propósito, el procedimiento y el manejo de los datos de este estudio. Dado que el estudio se basa en el análisis de datos públicos de Twitter, no se requiere consentimiento individual específico, pero es importante para nosotros asegurar que el proceso sea transparente y ético.

2. OBJETIVO DEL ESTUDIO

El objetivo principal del estudio es Determinar la influencia de los algoritmos de business intelligence en la analítica de los datos emocionales de los mensajes de tweets sobre McDonald's y KFC, Lima 2022.

3. PROCEDIMIENTO DEL ESTUDIO

El estudio analizará tweets públicos obtenidos a través de técnicas de web scraping. Los datos serán procesados utilizando algoritmos de business intelligence y técnicas estadísticas para evaluar su efectividad en el análisis de datos emocionales. Todos los datos son de acceso público y serán tratados de forma anónima y confidencial.

4. ACCESO Y USO DE DATOS

Dado que los mensajes de Twitter son públicos y accesibles sin restricciones, no se requiere consentimiento individual para la recopilación de estos datos. Sin embargo, el estudio se compromete a manejar toda la información con el máximo respeto por la privacidad y la ética de la investigación.

5. CONFIDENCIALIDAD Y ANONIMIZACIÓN

Los datos recolectados serán anonimizados para evitar la identificación de individuos específicos. La información será almacenada de manera segura y sólo el equipo de investigación tendrá acceso a ella. Los resultados del estudio se presentarán de forma agregada y anónima.

6. PROTECCIÓN DE DATOS Y ÉTICA

El estudio se llevará a cabo en cumplimiento con las políticas de privacidad y términos de servicio de Twitter. Además, se adherirá a las normas éticas de la investigación para garantizar que los datos sean utilizados de manera responsable y respetuosa.

7. CONTACTO PARA INQUIETUDES

Si tiene alguna pregunta o inquietud sobre el estudio, no dude en ponerse en contacto con el investigador principal, [Nombre del Investigador], a través del correo electrónico rmoralesg@unam.edu.pe o del teléfono 968448722.

8. ACEPTACIÓN DEL PARTICIPANTE

Al continuar con la revisión de este documento, usted acepta que ha sido informado sobre el propósito y el procedimiento del estudio y entiende que los datos utilizados son de acceso público. Su participación es valiosa para la investigación y contribuye al avance en el análisis de datos emocionales.



Firmado digitalmente por:
MORALES GONZALES RUSSO ALEXANDER
PR:4227914hard
Motivo: Soy el autor del documento
Fecha: 08/09/2024 18:53:24-0500

Firma

Fecha: 08/09/2024

Agradecimiento:

Agradecemos su interés y cooperación en este estudio.

ANEXO 7. Declaratoria de autenticidad de la tesis.

DECLARACIÓN JURADA DE ORIGINALIDAD DEL INFORME FINAL DE TESIS

Morales Gonzales Ruso Alexander egresado del doctorado de clases presenciales en Ingeniería de Sistemas en la Universidad Alas Peruanas, con código universitario N° 2017133911, e identificado con el DNI 42271914, y con la tesis titulada:

ALGORITMOS DE BUSINESS INTELLIGENCE EN LA ANALÍTICA DE DATOS EMOCIONALES DE TWEETS PUBLICADOS POR CLIENTES DE MCDONALD'S Y KFC, LIMA 2022

Declaro bajo juramento que:

- 1) El informe final de tesis es de mi autoría.
- 2) He respetado las normas internacionales de citas y referencias para las fuentes consultadas. Por tanto, la tesis no ha sido plagiada total, ni parcial,
- 3) Los datos e información presentadas son reales, no han sido falseados, ni copiados y por tanto los resultados que se presentan en la tesis se constituirán en aporte a la realidad investigada

De identificarse la falta de fraude (datos falsos), de plagio (información sin citar a autores), de piratería (uso ilegal de información ajena) o de falsificación (representar falsamente las ideas de otros), asumo las consecuencias y sanciones que mi acción se deriven, sometiendo a la normatividad vigente de la Universidad Alas Peruanas.

Lima, 4 de noviembre del 2024


Firma
DNI 42271914