

Análisis de Sentimientos en Redes Sociales para Mattelsa

Laura Riveros Quintero
ltriversq@eafit.edu.co

Mariel Viviana Sánchez Gamba
mvsanchezg@eafit.edu.co

10 de junio de 2025



Github: Repositorio del Proyecto

Índice

| | |
|--|-----------|
| 1. Introducción | 3 |
| 1.1. Problema del negocio | 3 |
| 1.2. Impacto de la solución | 4 |
| 1.3. Objetivos | 4 |
| 2. Marco teórico y Referencias | 5 |
| 2.1. Modelos de Clasificación de imágenes | 8 |
| 2.1.1. MobileNet | 9 |
| 2.1.2. ResNet | 11 |
| 2.1.3. Xception | 12 |
| 2.2. Análisis de sentimientos | 13 |
| 2.2.1. Pseudo-labeling | 14 |
| 2.2.2. RoBERTuito: | 14 |
| 3. Desarrollo metodológico | 15 |
| 3.1. Entendimiento del problema, pregunta de negocio o hipótesis | 15 |
| 3.2. Análisis exploratorio de los datos | 18 |
| 3.2.1. Entendimiento de los datos | 18 |
| 3.2.2. Análisis Exploratorio de Datos | 19 |
| 3.2.3. Preparación de los datos | 20 |
| 3.3. Selección de modelos | 21 |

| | |
|--|-----------|
| 3.3.1. Modelos | 21 |
| 3.4. Evaluación | 25 |
| 3.4.1. Clasificación de Imágenes | 25 |
| 3.4.2. Análisis de Sentimientos | 27 |
| 3.5. Análisis y conclusiones | 28 |
| 4. Tecnologías | 31 |
| 4.1. Descripción de características | 31 |
| 4.1.1. Ingesta de los datos | 31 |
| 4.1.2. Procesamiento - validación | 31 |
| 4.1.3. Ingeniería de Características | 32 |
| 4.1.4. Ingeniería de Características | 35 |
| 4.2. Desarrollo del Proyecto | 36 |
| 4.3. Despliegue del Proyecto | 37 |
| 4.4. Detalles Técnicos | 39 |
| 5. Conclusiones generales del Proyecto | 40 |
| 5.1. Clasificación de imágenes: | 40 |
| 5.2. Análisis de sentimientos: | 40 |

1. Introducción

Mattelsa es una marca de ropa que ha logrado diferenciarse en el mercado no solo por sus productos, sino por los mensajes de transformación social y emocional que comunica constantemente. Desde sus inicios como comercializadora de ropa nacional e internacional, ha evolucionado hacia la creación y distribución de prendas propias, acompañadas de un discurso activista centrado en el cuidado del medio ambiente, el bienestar emocional y la vida en comunidad (El Colombiano, 2023).

Más allá de su propuesta comercial, Mattelsa promueve una filosofía de vida que se refleja en sus publicaciones y acciones: alimentación consciente, actividad física, descanso, meditación, ocio y comunidad. Estos pilares se integran en su narrativa digital y forman parte de la identidad que proyectan en sus canales (Mattelsa, s.f.).

En este trabajo, nos proponemos investigar cómo los usuarios perciben a Mattelsa a través de las redes sociales, con base en los mensajes que la marca publica. Analizaremos las reacciones, comentarios y sentimientos expresados por su audiencia para comprender si estos contenidos generan identificación, afinidad o incluso rechazo. Esta investigación busca aportar al entendimiento del papel que juega la comunicación digital en la construcción de marca con propósito y en la conexión emocional con sus seguidores.

1.1. Problema del negocio

Mattelsa ha construido una marca que va más allá de la ropa promueve una forma de vivir basada en el bienestar, la libertad, el cuidado del entorno y la conexión entre personas. Esta identidad se refleja en los mensajes que comparte en redes sociales, donde constantemente comunica su visión del mundo y los valores que la representan.

Sin embargo, no se tiene claridad sobre cómo las personas realmente perciben y sienten esa identidad. Aunque los mensajes están presentes, no hay una forma concreta de saber qué emociones generan, qué tan auténtico resulta ese vínculo para la comunidad, ni si los contenidos están conectando de verdad con quienes siguen la marca.

En este contexto, surgen tres preguntas clave que guían esta investigación:

-
- ¿Sabemos cómo se sienten realmente las personas con nuestra filosofía?
 - ¿Medimos el impacto emocional de nuestras campañas?
 - ¿Detectamos a tiempo las señales de alerta?

El reto no está solo en comunicar, sino en escuchar y comprender las emociones que esos mensajes despiertan.

1.2. Impacto de la solución

El análisis de sentimientos permitirá comprender con mayor profundidad cómo reaccionan emocionalmente las personas ante los mensajes que Mattelsa publica sobre su filosofía, cultura organizacional y propósito de marca. Esta herramienta facilitará la identificación de las emociones más frecuentes, los temas que generan mayor conexión y posibles señales de alerta en la percepción del público.

Con base en estos hallazgos, se podrán formular recomendaciones que ayuden a reforzar los aspectos más valorados por la comunidad y ajustar aquellos mensajes que no estén generando el impacto esperado. Así, la marca podrá alinear mejor su comunicación con las emociones reales de su audiencia y fortalecer el vínculo emocional que ha construido en torno a su identidad.

1.3. Objetivos

- **Objetivo General:** Realizar un análisis de sentimientos sobre los contenidos relacionados con la identidad y cultura de marca publicados por Mattelsa en redes sociales, para comprender la percepción emocional de sus consumidores y apoyar la mejora de la estrategia de comunicación digital.
- **Objetivos Específicos:**
 - Identificar automáticamente cuáles imágenes publicadas por Mattelsa corresponden a contenidos de identidad y cultura de marca.
 - Analizar los comentarios de los usuarios relacionados con estos contenidos, clasificándolos según el sentimiento expresado y asignándoles un puntaje que refleje el grado de emoción manifestada.

-
- Evaluar el impacto emocional de las campañas y contenidos de identidad de marca mediante un tablero de control interactivo que facilite el monitoreo y la interpretación de la percepción de la comunidad.

2. Marco teórico y Referencias

Para fundamentar el análisis realizado sobre la marca Mattelsa, es importante considerar dos procesos principales que sustentan este estudio: la clasificación de imágenes y el análisis de sentimientos, ambos aplicados a los contenidos publicados en redes sociales. La marca presenta una variedad de publicaciones que pueden dividirse en dos categorías: por un lado, contenidos relacionados con la identidad y cultura de marca, que incluyen imágenes de activismo social, animaciones, memes y mensajes que refuerzan los valores y filosofía de la marca, como el cuidado ambiental, el bienestar emocional y la vida en comunidad. Por otro lado, están los contenidos de producto, que se enfocan en la promoción y comercialización directa de las prendas y accesorios del catálogo de Mattelsa, con el objetivo de captar la atención y el interés del consumidor en el mercado de la moda urbana.

Estos dos tipos de contenido tienen características y objetivos distintos, lo que implica la necesidad de aplicar enfoques analíticos diferenciados. Mientras que las publicaciones orientadas al producto buscan promover la compra a través de la presentación de diseños, precios o calidad, los contenidos de identidad y cultura de marca apuntan a generar conexiones emocionales mediante mensajes que reflejan valores, estilo de vida y propósito. En este sentido, aunque inicialmente se clasificaron ambos tipos de publicaciones, el análisis de sentimientos se enfoca exclusivamente en los contenidos que expresan la filosofía de Mattelsa, como el bienestar, el activismo social, la sostenibilidad y la vida en comunidad.

Este enfoque permite observar cómo reaccionan los usuarios ante las publicaciones que representan el núcleo ideológico y emocional de la marca. Para ello, los comentarios se analizaron según el sentimiento expresado, clasificándolos como positivos, negativos o neutros, y asignándoles un puntaje que refleja el grado de emoción manifestada. Este proceso facilita la evaluación del nivel de identificación, afinidad o rechazo que generan dichos mensajes, así como la detección de temas o valores que resuenan con mayor fuerza en la comunidad digital de Mattelsa. En consecuencia, el análisis se convierte en una herramienta clave para comprender el impacto emocio-

nal de la narrativa de marca y su capacidad de construir vínculos auténticos con su audiencia.

- Imágenes de producto

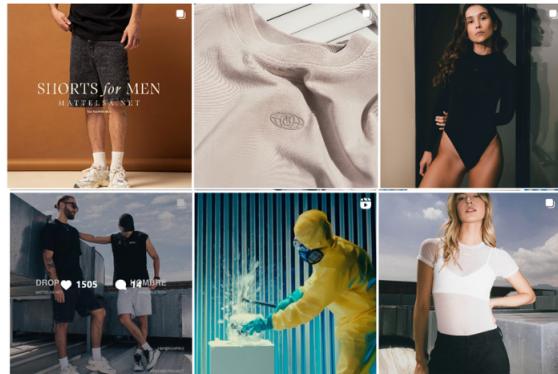


Figura 1: Imágenes de producto

- Imágenes de activismo y cultura popular



Figura 2: Imágenes de activismo y cultura popular

Para respaldar este enfoque, se realizó una revisión exhaustiva de la literatura científica a través de la base de datos Scopus, reconocida por su amplitud y rigor académico. Esta revisión permitió identificar estudios actuales y relevantes en las áreas de clasificación automática de imágenes, análisis de sentimientos en redes sociales y marketing

digital, garantizando una base sólida para el desarrollo del análisis. Entre los conceptos y teorías clave destacan los modelos de clasificación automática de imágenes que emplean técnicas de aprendizaje automático para categorizar contenido visual, así como el análisis de sentimientos como herramienta fundamental para detectar emociones y actitudes en textos generados por usuarios, lo cual es esencial para comprender el impacto emocional de las estrategias de comunicación digital. Además, se consideraron las teorías de marketing con propósito y construcción de marca emocional, que resaltan la importancia de conectar con la audiencia a través de valores compartidos y mensajes significativos.

| Campo | Contenido |
|---------------|---|
| Base de Datos | Scopus |
| Query | TITLE-ABS-KEY (Sentiment analysis) AND TITLE-ABS-KEY (Image classification) AND TITLE-ABS-KEY (Social media) OR TITLE-ABS-KEY (Deep learning) OR TITLE-ABS-KEY (Image-based sentiment analysis) OR TITLE-ABS-KEY (Multimodal sentiment analysis) OR TITLE-ABS-KEY (Nlp) OR TITLE-ABS-KEY (Feature engineering) OR TITLE-ABS-KEY (Social media analysis) OR TITLE-ABS-KEY (Natural language processing) OR TITLE-ABS-KEY (Visual data mining) OR TITLE-ABS-KEY (Knn) OR TITLE-ABS-KEY (Dbscan) OR TITLE-ABS-KEY (Latent dirichlet allocation (TextBold)) OR TITLE-ABS-KEY (TextBold) OR TITLE-ABS-KEY (Bert) OR TITLE-ABS-KEY (Cnn) OR TITLE-ABS-KEY (Roberta) |
| Subject Area | Computer Science, Engineering, Mathematics, Decision Sciences |
| Años | 2016-2024 |

Cuadro 1: Query con especificación de la base de datos, áreas temáticas y rango de años.

Como complemento, se presenta información bibliométrica clave derivada de la búsqueda documental. El rango de análisis abarca publicaciones entre 2016 y 2024, con un total de 164 documentos recopilados desde 97 fuentes, incluyendo revistas académicas, libros y conferencias. Se identificó un crecimiento anual del 38.66 %, lo que refleja un interés creciente en las temáticas abordadas. La antigüedad promedio de los documentos es de 2.28 años, indicando la relevancia y actualidad de las fuentes consideradas. Además, el promedio de citas por documento es de 52.41, acumulando un total de 4,126 referencias, lo que demuestra un impacto significativo en la comunidad académica. Se identificaron también 746 palabras clave adicionales (Keywords Plus) y 332 palabras clave proporcionadas directamente por los autores, lo que evidencia una diversidad temática importante.

| Descripción | Resultados |
|------------------------------------|------------|
| MAIN INFORMATION ABOUT DATA | |
| Timespan | 2016:2024 |
| Sources (Journals, Books, etc) | 97 |
| Documents | 164 |
| Annual Growth Rate (%) | 38.66 |
| Document Average Age | 2.28 |
| Average citations per doc | 52.41 |
| References | 4126 |
| DOCUMENT CONTENTS | |
| Keywords Plus (ID) | 746 |
| Author's Keywords (DE) | 332 |
| AUTHORS | |
| Authors | 382 |
| Authors of single-authored docs | 2 |
| AUTHORS COLLABORATION | |
| Single-authored docs | 2 |
| Co-Authors per Doc | 2.56 |
| International co-authorships (%) | 12.2 |
| DOCUMENT TYPES | |
| article | 52 |
| book | 1 |
| book chapter | 1 |
| conference paper | 56 |
| conference review | 52 |
| erratum | 2 |

Cuadro 2: Información general sobre los datos bibliométricos.

2.1. Modelos de Clasificación de imágenes

La clasificación de imágenes es una de las tareas fundamentales en el campo de la visión por computador, y ha experimentado un notable avance gracias al desarrollo de las redes neuronales convolucionales (CNN, por sus siglas en inglés). Estas redes están diseñadas para aprender representaciones jerárquicas de los datos visuales mediante la aplicación sucesiva de filtros convolucionales, funciones de activación no lineales y técnicas de agrupamiento espacial, lo que les permite identificar patrones

complejos y estructurados en las imágenes (LeCun, Bengio & Hinton, 2015). Su éxito ha sido especialmente notable en tareas como la detección de objetos, la segmentación semántica y la clasificación de escenas, superando ampliamente los métodos tradicionales basados en ingeniería manual de características.

Para abordar el análisis de las imágenes compartidas por Mattelsa en redes sociales, se emplearon modelos de CNN preentrenados, adaptados mediante técnicas transfer learning. Este enfoque consiste en reutilizar conocimientos adquiridos por redes entrenadas en grandes conjuntos de datos, como ImageNet (Deng et al., 2009), y transferirlos a tareas más específicas con conjuntos de datos más pequeños y especializados. De esta manera, se mejora el rendimiento predictivo, se acelera el proceso de entrenamiento y se reduce la necesidad de grandes recursos computacionales, aspectos particularmente relevantes cuando se trabaja con colecciones de imágenes limitadas o altamente especializadas (Pan & Yang, 2010).

En este trabajo, se han utilizado algunas de las arquitecturas CNN preentrenadas más reconocidas y eficientes para la clasificación de imágenes en entornos reales: MobileNet, ResNet y Xception. Estas arquitecturas se seleccionaron por su balance entre precisión, eficiencia computacional y facilidad de adaptación a nuevos dominios mediante técnicas de fine-tuning. A continuación, se describen sus principales características estructurales y ventajas en el contexto del análisis de comunicación visual de marca.

2.1.1. MobileNet

MobileNet es una familia de arquitecturas de redes neuronales convolucionales desarrollada por Google, optimizada para tareas de visión por computador en dispositivos con recursos limitados, como teléfonos móviles y sistemas embebidos. Su innovación principal es el uso de *convoluciones separables en profundidad* (*depthwise separable convolutions*), que dividen una convolución estándar en dos operaciones más simples: una convolución *depthwise*, que aplica un filtro por canal de entrada, y una convolución *pointwise* (de 1×1), que combina la salida de los canales anteriores. Esta descomposición reduce significativamente el costo computacional y la cantidad de parámetros del modelo, con una pérdida mínima de precisión [?].

Además de su estructura liviana, MobileNet introduce dos hiperparámetros clave que permiten ajustar la eficiencia del modelo según las restricciones del entorno: el

multiplicador de anchura (α), que controla la cantidad de canales en cada capa, y el *multiplicador de resolución* (ρ), que ajusta el tamaño de entrada de las imágenes. Por ejemplo, si $\alpha = 0,5$, el número de filtros se reduce a la mitad, disminuyendo considerablemente el tamaño del modelo y su tiempo de inferencia. De manera similar, una imagen de entrada de 224×224 píxeles puede reducirse a 160×160 mediante ρ , reduciendo así la carga computacional sin modificar la arquitectura general del modelo.

Gracias a esta flexibilidad, MobileNet es una arquitectura ideal para aplicaciones en tiempo real, como reconocimiento de imágenes en dispositivos móviles, sin sacrificar en exceso la precisión del modelo. En el contexto del presente estudio, se empleó MobileNetV2, que incluye bloques lineales de expansión y atajos residuales para mejorar aún más la eficiencia y el aprendizaje de representaciones útiles con menor sobreajuste, especialmente en conjuntos de datos reducidos y heterogéneos, como los provenientes de redes sociales.

- El multiplicador de anchura (α) controla la cantidad de canales en cada capa, afectando directamente el número de parámetros y operaciones. Por ejemplo, si $\alpha = 0,5$, el número de filtros se reduce a la mitad, lo que disminuye drásticamente el tamaño del modelo y su tiempo de inferencia.
- El multiplicador de resolución (ρ) reduce el tamaño de entrada de las imágenes, disminuyendo así la carga computacional. Por ejemplo, una imagen de entrada de 224×224 píxeles puede reducirse a 160×160 sin modificar la arquitectura general del modelo [?].

La arquitectura de MobileNet v1 comienza con una capa de convolución estándar seguida de una serie de bloques de convolución depthwise y pointwise, cada uno acompañado por normalización por lotes y activación ReLU. Al final, se aplica una capa de agrupamiento promedio global y una capa densa con activación softmax para la clasificación. Esta estructura modular permite lograr una alta eficiencia y velocidad, lo que convierte a MobileNet en una opción ideal para aplicaciones en tiempo real y entornos con capacidad de cómputo limitada.

En términos de parámetros, MobileNetV1 contiene aproximadamente 4,2 millones de parámetros cuando se entrena sobre imágenes de tamaño 224×224 con $\alpha = 1$ y $\rho = 1$. Esta cifra es significativamente menor en comparación con arquitecturas más complejas como VGG-16, que cuenta con aproximadamente 138 millones de

parámetros, manteniendo una precisión competitiva en tareas de clasificación. Las versiones más avanzadas, como MobileNetV2 y MobileNetV3, optimizan aún más el rendimiento mediante el uso de bloques residuales invertidos, capas lineales sin activación (*linear bottlenecks*) y técnicas de cuantización, lo que permite reducir el uso de memoria y acelerar la inferencia sin sacrificar significativamente la precisión [?, ?].

2.1.2. ResNet

ResNet (Residual Network) es una arquitectura de red neuronal convolucional propuesta por He et al. (2016) que marcó un hito en el desarrollo de modelos profundos de aprendizaje profundo al resolver de manera efectiva el problema de la degradación del rendimiento en redes neuronales muy profundas. Este problema se manifiesta cuando, al aumentar la profundidad de la red, la precisión en entrenamiento comienza a disminuir en lugar de mejorar, incluso cuando no hay síntomas claros de sobreajuste.

La innovación central de ResNet radica en la introducción de los bloques residuales (residual blocks), los cuales incorporan conexiones de atajo (skip connections) que permiten que la entrada de un bloque se sume directamente a su salida, sin pasar por las capas intermedias. Esta estrategia facilita el flujo del gradiente durante el entrenamiento, permitiendo que la red aprenda las funciones de mapeo residuales en lugar de funciones directas, lo cual mejora significativamente la convergencia y el rendimiento en redes profundas (He et al., 2016).

La estructura de un bloque residual básico se compone de dos o más capas convolucionales, seguidas de operaciones de normalización por lotes y activaciones ReLU, con una conexión directa que suma la entrada del bloque con su salida transformada. Esta arquitectura permite construir redes extremadamente profundas, como ResNet-50, ResNet-101 y ResNet-152, sin caer en problemas de sobreajuste o pérdida de señal.

En términos de parámetros, ResNet mantiene una relación escalable entre profundidad y complejidad. Por ejemplo:

- ResNet-50 cuenta con aproximadamente 25,6 millones de parámetros,
- ResNet-101 con 44,5 millones,

-
- ResNet-152 con más de 60 millones (He et al., 2016).

Estos modelos utilizan bloques convolucionales con filtros de tamaño 1x1, 3x3 y 1x1, diseñados para reducir y expandir dimensionalidades, optimizando así el uso de parámetros sin comprometer la capacidad de aprendizaje. A pesar de su gran tamaño, ResNet sigue siendo más eficiente y precisa que arquitecturas anteriores como VGG, que requerían más parámetros para lograr un rendimiento inferior.

ResNet fue evaluada por primera vez en el conjunto de datos ImageNet, logrando una precisión de clasificación superior a la de modelos anteriores, como VGG o Inception, al mismo tiempo que facilitaba la implementación de redes con más de 100 capas sin deterioro del desempeño (He et al., 2016; Szegedy et al., 2016). Desde entonces, ha sido adoptada como una de las arquitecturas base en numerosas tareas de visión por computador, como clasificación de imágenes, detección de objetos, segmentación semántica, y más.

2.1.3. Xception

Xception (Extreme Inception), propuesta por Chollet (2017), representa una evolución conceptual y arquitectónica del modelo Inception, al llevar al extremo el principio de factorizar convoluciones para mejorar la eficiencia. Mientras que Inception combina filtros de diferentes tamaños en paralelo para capturar patrones espaciales diversos, Xception simplifica este enfoque al separar completamente el procesamiento espacial del procesamiento de canales, basándose en la idea de que estos dos tipos de correlación pueden aprenderse de manera más eficaz por separado.

El elemento central de Xception es la convolución separable en profundidad (depthwise separable convolution), que divide el proceso en dos pasos consecutivos: primero, una convolución depthwise que aplica un único filtro por canal de entrada (extrayendo características espaciales por canal de forma independiente); y luego, una convolución pointwise (1×1) que combina linealmente las salidas a lo largo de los canales. Esta descomposición permite reducir considerablemente la complejidad computacional sin sacrificar la capacidad de aprendizaje del modelo.

La arquitectura de Xception consta de 36 capas convolucionales, organizadas en tres secciones principales: un bloque de entrada, donde se reduce la dimensionalidad inicial de la imagen; bloques intermedios residuales repetitivos, donde ocurre la mayor

parte del procesamiento jerárquico; y un bloque de salida, que condensa las características extraídas para la etapa de clasificación. Cada uno de estos bloques emplea capas de normalización por lotes, activación ReLU, y conexiones residuales, lo que facilita la convergencia durante el entrenamiento y evita el problema de la desaparición del gradiente en redes profundas.

A pesar de su simplicidad estructural en comparación con Inception-v3, Xception logra mayor precisión en tareas de clasificación en conjuntos como ImageNet, al mismo tiempo que mantiene una eficiencia computacional competitiva. El modelo tiene aproximadamente 22.9 millones de parámetros, lo cual lo posiciona como una opción intermedia entre modelos ligeros como MobileNet y arquitecturas más pesadas como ResNet-152. Estas características hacen de Xception un modelo ideal para aplicaciones de aprendizaje por transferencia, especialmente en contextos donde se cuenta con datos limitados pero se requiere alto rendimiento predictivo (Chollet, 2017).

2.2. Análisis de sentimientos

El análisis de sentimientos es una disciplina del procesamiento del lenguaje natural que tiene como objetivo identificar y clasificar automáticamente las opiniones, emociones y actitudes expresadas en textos. Esta técnica es especialmente relevante en el contexto de la comunicación digital, donde los usuarios generan gran cantidad de contenido en plataformas como redes sociales, blogs y foros, constituyendo una fuente esencial para comprender la percepción del público respecto a marcas, productos o campañas (Liu, 2012; Pang Lee, 2008). El análisis de sentimientos permite a las organizaciones evaluar la respuesta emocional de su audiencia, optimizar estrategias comunicacionales y anticipar posibles crisis reputacionales.

Sin embargo, en muchos escenarios prácticos, la cantidad de datos etiquetados disponibles es limitada, dificultando el entrenamiento óptimo de los modelos. Para superar esta limitación, se emplean técnicas de aprendizaje semi-supervisado, entre las que destaca el pseudo-labeling. Esta técnica permite utilizar grandes volúmenes de datos no etiquetados generando etiquetas artificiales que se incorporan al entrenamiento, mejorando así la generalización y robustez del modelo (Lee, 2013).

2.2.1. Pseudo-labeling

El pseudo-labeling es un método semi-supervisado que comienza entrenando un modelo con un conjunto de datos etiquetados. Luego, este modelo predice etiquetas para un conjunto adicional de datos no etiquetados; estas etiquetas predichas se denominan pseudo-etiquetas. A continuación, las pseudo-etiquetas se utilizan como si fueran etiquetas reales para ampliar el conjunto de entrenamiento, permitiendo al modelo aprender de una mayor cantidad de datos sin necesidad de anotación manual adicional (Lee, 2013).

Esta técnica es especialmente útil en el análisis de sentimientos, donde la obtención de datos etiquetados puede ser costosa o lenta, y donde los textos pueden variar mucho en vocabulario y estilo según el contexto digital. El pseudo-labeling ayuda a adaptar el modelo a estas variaciones, mejorando su desempeño y capacidad para captar emociones en diversos tipos de contenido.

2.2.2. RoBERTuito:

En el ámbito del procesamiento del lenguaje natural en español, RoBERTuito es un modelo basado en la arquitectura de transformadores BERT, desarrollado y preentrenado específicamente para textos en español, con adaptaciones orientadas a tareas como el análisis de sentimientos y la clasificación emocional (Pérez et al., 2021). Gracias a su entrenamiento en grandes corpus representativos del idioma, RoBERTuito es capaz de comprender el contexto bidireccional de las palabras y captar las sutilezas del lenguaje, características fundamentales para un análisis de sentimientos efectivo.

Para el análisis de comunicación digital de Mattelsa, se implementa una estrategia que combina RoBERTuito con técnicas de fine-tuning y pseudo-labeling. El fine-tuning consiste en ajustar los parámetros del modelo preentrenado para que se adapte específicamente al conjunto de datos particular de la empresa, mejorando la precisión en la clasificación de emociones y opiniones expresadas en los textos generados por los usuarios en redes sociales y otras plataformas digitales.

Al incorporar pseudo-labeling, el modelo amplía su capacidad de aprendizaje utilizando datos no etiquetados de la comunicación digital, lo que contribuye a una mejor generalización y a la detección más precisa de las emociones y tendencias en los mensajes. Este enfoque integrado es crucial para monitorear la reputación online,

evaluar el impacto de campañas y anticipar posibles crisis comunicacionales a partir del análisis sistemático del sentimiento expresado en el contenido digital (Devlin et al., 2019; Lee, 2013; Pérez et al., 2021).

3. Desarrollo metodológico

3.1. Entendimiento del problema, pregunta de negocio o hipótesis

Mattelsa es una marca con una alta presencia en redes sociales, consideradas como un canal estratégico para su reconocimiento de marca. Contamos con la hipótesis de que este reconocimiento aporta a las ventas obtenidas en los distintos canales comerciales. Además, Mattelsa cuenta con una filosofía enfocada en generar impacto social, la cual es un pilar clave en su estrategia de comunicación. La dirección de marca se enfoca en la creación de un alto volumen de contenido, principalmente en la plataforma Instagram. Dicho contenido puede agruparse en dos categorías principales:

1. Publicaciones de producto: Estas publicaciones están orientadas a dar a conocer las prendas disponibles en las tiendas físicas o en el ecommerce, así como nuevas colecciones, aperturas de tiendas y cualquier comunicación relacionada con los productos. Este tipo de publicaciones incluye:

- Fotografías de modelos vistiendo las prendas en diversos escenarios.
- Imágenes exclusivamente del producto.
- Fotografías de la visualización de las tiendas.

2. Publicaciones de marca: Este tipo de contenido busca transmitir los valores y la filosofía de Mattelsa, promoviendo una conexión emocional con las personas que consumen dicho contenido. Incluye:

- Imágenes basadas en ilustraciones.
- Textos inspiradores o creativos.

-
- Memes.

El entendimiento del problema se basa en el hecho de que, semanalmente, se generan en cierto número de publicaciones, lo que implica la necesidad de gestionar, clasificar y analizar un volumen considerable de contenido.

Problemas identificados Problema 1: Clasificación manual de publicaciones Actualmente, el equipo de marca elabora un informe mensual que evalúa el rendimiento de las publicaciones mediante indicadores obtenidos a través de la red social. Estas métricas permiten conocer el impacto de las publicaciones las cuales se dividen en dos categorías principales: fotografía y marca. El objetivo de estos indicadores es generar retroalimentación específica para los diferentes equipos, fotografía y diseñadores de moda que se enfocan en todo lo relacionado a producto y diseñadores gráficos, que son los encargados de las ilustraciones. Sin embargo, este proceso de clasificación se realiza de forma manual, lo que resulta complejo y consume una cantidad significativa de tiempo. Los pasos actuales son los siguientes:

- 1. Consultar carpeta on the Mattelsa service with all the imágenes publicadas.
- 2. Verificar cuáles de estas imágenes corresponden a las publicaciones de la semana en análisis.
- 3. Visualizar cada imagen para clasificarla manualmente como fotografía o marca.

Dado el volumen de publicaciones y la necesidad de agilizar el proceso, se propone entrenar un modelo de clasificación de imágenes capaz de automatizar este flujo de trabajo, mejorando la eficiencia y reduciendo la carga manual del equipo.

Problema 2: Percepción de la marca y análisis de sentimientos Aunque el equipo de marca extrae indicadores sobre el rendimiento de las publicaciones y sus categorías, actualmente no se cuenta con una evaluación clara de la percepción que los clientes tienen sobre la filosofía de marca. Es crucial comprender cómo los usuarios perciben la filosofía de la empresa. Para abordar este desafío, se plantea realizar un análisis de sentimientos aplicado a los comentarios generados en las publicaciones. Este análisis permitirá:

- Identificar la percepción de las personas sobre la filosofía de marca.

-
- Detectar oportunidades de mejora para generar mayor engagement y conexión con las personas.

Dado que las publicaciones se segmentan en dos grupos (fotografía y marca), se decide enfocar el análisis de sentimientos exclusivamente en los comentarios de las publicaciones clasificadas como marca, ya que estas son las que generan mayor conexión con las personas, sea positiva o negativa, por el contrario las publicación enfocadas al producto suelen general comentarios mas neutrales.

Enfoque metodológico

1. Clasificación automática de imágenes:

- Desarrollo de un modelo de clasificación de imágenes que asigne automáticamente las publicaciones a las categorías fotografía o marca.
- Integración de este modelo en el flujo de trabajo actual para optimizar tiempos y reducir el esfuerzo manual del equipo.

2. Análisis de sentimientos:

- Implementación de un modelo de procesamiento de lenguaje natural (NLP) para analizar los comentarios asociados a las publicaciones clasificadas como fotografía.
- Generación de insights valiosos sobre la percepción de las personas sobre el tipo de mensaje que se esta transmitiendo.

Es importante señalar que el análisis de sentimientos esta relacionado directamente con la precisión del modelo de clasificación de imágenes. Este planteamiento busca aportar valor estratégico al proceso de toma de decisiones en Mattelsa.

3.2. Análisis exploratorio de los datos

3.2.1. Entendimiento de los datos

La información inicial se obtiene directamente desde las bases de datos de Mattelsa. Este proceso requiere permisos de propietario.

Extracción de comentarios: Esta función recopila información sobre los comentarios realizados en cada publicación. Los datos se guardan se extrae directamente de las bases de datos de Mattelsa, que contiene las siguientes variables:

- rowidPublicacionMeta: Identificador único de la publicación a la que pertenece el comentario.
- Comentario: Texto del comentario.
- Fecha: Fecha y hora en la que se realizó el comentario.
- Año: Año donde se realizó la publicación.
- ComentarioLimpio: Comentario sin caracteres especiales, ni emojis.
- urlImagen: URL de la imagen de la publicación.

Esta información obtenida será utilizada tanto para entrenar y evaluar el modelo de clasificación como para realizar el análisis de sentimientos.

Fuentes adicionales de datos Además de los datos obtenidos de los comentarios, se cuenta con información complementaria igualmente proporcionada directamente por Mattelsa. Esta información consiste en un conjunto de datos previamente estructurados y clasificados, disponibles en un archivo .csv. Este archivo incluye las siguientes variables relevantes para los análisis:

Bd.imagen: DataFrame con información previa de publicaciones:

- NOMBRE: Identificador único de cada imagen.
- ÁREA: Clasificación asignada según la categoría de la publicación (Fotografía o Marca).

Cabe destacar que este DataFrame no contiene las imágenes asociadas, pero sí se tiene acceso al repositorio de imágenes de Mattelsa en formato .jpg. Los nombres de los archivos de imagen coinciden con los identificadores únicos especificados en el archivo Bd.imagen.csv, lo que permite su correcta asociación durante el procesamiento.

Esta estructura de datos garantiza que se cuente con los insumos necesarios para los procesos de clasificación de imágenes y análisis de sentimientos que se desarrollarán posteriormente.

3.2.2. Análisis Exploratorio de Datos

Como se mencionó anteriormente, se dispone de diferentes fuentes de información para llevar a cabo la clasificación de imágenes. Inicialmente, se cuenta con un total de 5,194 imágenes previamente clasificadas en cinco categorías. Sin embargo, para este análisis se realizará una reducción de dichas categorías, dado que en la actualidad el negocio trabaja únicamente con dos clasificaciones principales: fotografía y marca. Estas categorías representan las más relevantes para el modelo, siendo fotografía el 70 % de las imágenes clasificadas, es decir, un conjunto de 3,636 imágenes y marca el 30 % que corresponde a un total de 1,558 imágenes.

En el DataFrame bd.imagen, todas las imágenes cuentan con una clasificación previa, por lo que no se identificaron datos faltantes en las etiquetas. No obstante, se encontraron únicamente 5,056 imágenes en el repositorio de Mattelsa. Debido a esto, este será el total de registros que se empleará para la creación y entrenamiento del modelo.

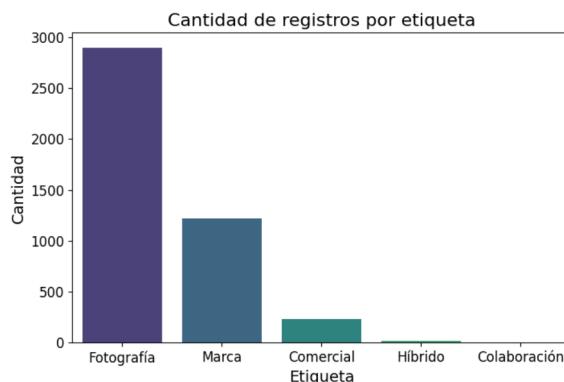


Figura 3: Etiquetas de cada uno de los registros obtenidos

3.2.3. Preparación de los datos

Clasificación de imágenes Para utilizar las imágenes en los modelos que se implementarán, es necesario someterlas a un proceso de preprocesamiento. Este proceso incluye:

Decodificación de la imagen: Se dejan las imágenes a color, esto con el fin de poder usar modelos preentrenados basado en imágenes a color.

Redimensionamiento: Se ajusta el tamaño de cada imagen a 100x100 píxeles para garantizar uniformidad.

Limpieza y Normalización: Se normalizan los valores de los píxeles para llevarlos a una escala de [0, 1]. La normalización se realiza dividiendo cada valor de píxel entre 255, ya que los valores originales de los píxeles suelen estar en el rango de [0, 255],

Conversión a un arreglo: Cada imagen se convierte en un arreglo para poder ser procesadas por el modelo.

Este procedimiento tiene como objetivo optimizar el rendimiento durante el procesamiento de las imágenes y asegurar una entrada homogénea a los modelos.

Cada imagen, tras procesarse como se describió previamente, se visualiza de la siguiente manera:

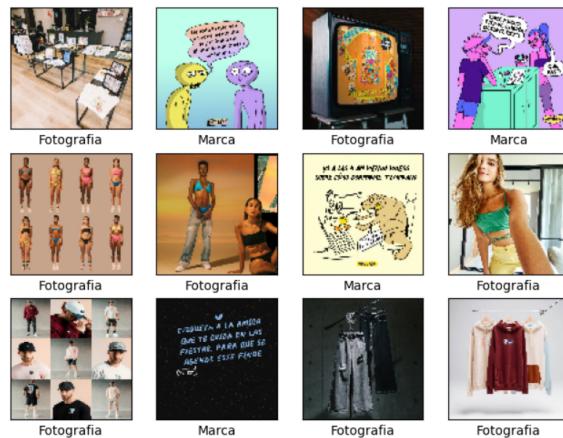


Figura 4: Imágenes procesadas

Análisis de Sentimientos Antes del análisis, los datos se sometieron a un proceso de preprocesamiento para normalizar y estructurar los textos. Este paso incluyó:

Tokenización: División de cada comentario en palabras individuales, eliminando puntuaciones y palabras irrelevantes (stopwords).

Bag of Words (BoW): Representación de los textos en vectores numéricos basados en la frecuencia de palabras dentro de un vocabulario definido.

Limpieza y Normalización: Conversión a minúsculas y eliminación de caracteres especiales para garantizar la uniformidad de los datos.

3.3. Selección de modelos

3.3.1. Modelos

Clasificación de imágenes

Redes Neuronales Convolucionales (CNNs) Las Redes Neuronales Convolucionales fueron seleccionadas como modelo principal para la clasificación de imágenes debido a su alta precisión y capacidad para manejar datos visuales complejos. A continuación, se detallan las razones clave para esta elección:

- **Clasificación de imágenes complejas:** Las CNNs son capaces de identificar diferencias sutiles entre imágenes de productos y de activismo, gracias a su habilidad para capturar patrones visuales jerárquicos.
- **Impacto comercial:** La alta precisión de las CNNs garantiza una clasificación confiable, minimizando errores que podrían afectar decisiones estratégicas relacionadas con campañas publicitarias y análisis de productos.
- **Flexibilidad y adaptabilidad:** Mediante técnicas como transferencia de aprendizaje, las CNNs pueden aprovechar redes preentrenadas (por ejemplo, ResNet o VGG), lo que las hace adecuadas para trabajar con volúmenes moderados de datos y mejorar el rendimiento del modelo.

-
- **Detección de características relevantes:** Las CNNs trabajan directamente con píxeles de la imagen y extraen automáticamente características visuales relevantes, eliminando la necesidad de preprocesamiento intensivo o diseño manual de características.
 - **Robustez frente a variaciones:** Este modelo es resiliente frente a variaciones en las imágenes, como rotaciones, cambios de escala o diferentes condiciones de iluminación, lo que asegura una mayor generalización al clasificar nuevos datos.
 - **Reconocimiento visual avanzado:** Dado que las imágenes de activismo pueden incluir elementos abstractos (como memes o animaciones) y las de productos presentan detalles específicos, las CNNs destacan en la identificación y clasificación basándose en relaciones visuales jerárquicas.

Se emplearán modelos preentrenados con imágenes del conjunto ImageNet, lo que proporciona una base sólida de características visuales generales. En particular, se utilizarán arquitecturas como MobileNet y ResNet50, sobre las cuales se aplicará fine-tuning para ajustar y reentrenar los modelos, permitiendo que se adapten mejor a los datos específicos de Mattelsa y mejoren su capacidad de predicción.

Modelo Xception -Modelo 1:

- Se emplea como base el modelo **Xception**, una arquitectura profunda desarrollada por Google que mejora el rendimiento mediante el uso de convoluciones separables en profundidad (depthwise separable convolutions). Este modelo ha sido preentrenado con el conjunto de datos **ImageNet**, lo cual permite reutilizar características previamente aprendidas. Se establece `include_top=False` para excluir la capa de clasificación original y adaptarlo a una nueva tarea.
- `input_shape=(100, 100, 3)`: Define el tamaño de entrada de las imágenes como 100x100 píxeles y 3 canales de color (RGB), lo cual permite procesar imágenes a color de resolución reducida.
- `GlobalAveragePooling2D()`: Esta capa resume cada mapa de activación generado por Xception en un único valor promedio, reduciendo así la dimensionalidad y previniendo el sobreajuste.
- `Dense(1, activation='sigmoid')`: Capa densa de salida con una sola neurona y función de activación sigmoide, adecuada para problemas de

clase. Clasifica la imagen en una de las dos clases.

- Las capas del modelo base se configuran como no entrenables (`trainable=False`), de forma que solo se ajustan los pesos de las capas añadidas durante el proceso de fine-tuning, reduciendo el riesgo de sobreentrenamiento y acelerando la convergencia.

Modelo MobileNet -Modelo 2:

- Se utiliza como base el modelo **MobileNetV2**, una arquitectura eficiente diseñada para dispositivos con recursos limitados. Entre sus parámetros se encuentra `input_shape=(100, 100, 3)`, que indica el tamaño de las imágenes en píxeles (100x100) y los tres canales de color (RGB). El parámetro `weights='imagenet'` especifica que se emplearán los pesos preentrenados con el conjunto de datos ImageNet, lo cual permite aprovechar características visuales previamente aprendidas.
- `GlobalAveragePooling2D()`: Esta capa reduce la dimensionalidad del mapa de características generado por el modelo base, promediando los valores espaciales de cada filtro. Esto disminuye el número de parámetros y ayuda a prevenir el sobreajuste.
- `Dropout(0.5)`: Se aplica regularización mediante la técnica de dropout, desactivando aleatoriamente el 50 % de las neuronas durante el entrenamiento. Esta estrategia mejora la generalización del modelo y evita el sobreentrenamiento.
- `Dense(1, activation='sigmoid')`: Capa de salida densa con una sola neurona y función de activación sigmoide, adecuada para tareas de clasificación binaria. El resultado será una probabilidad entre 0 y 1, indicando la pertenencia a una de las dos clases.

ResNet50 - Modelo 3:

- Se utiliza como base el modelo **ResNet50**, una arquitectura de redes neuronales profundas con 50 capas, reconocida por su uso de conexiones residuales que facilitan el entrenamiento de redes muy profundas. El modelo se inicializa con pesos preentrenados sobre el conjunto de datos **ImageNet**, lo cual proporciona una base robusta de características visuales.

-
- `input_shape=(100, 100, 3)`: Especifica el tamaño de entrada de las imágenes (100x100 píxeles con 3 canales RGB), adecuado para imágenes a color de resolución reducida.
 - `include_top=False`: Se omite la capa de clasificación original de ResNet50 para permitir una personalización del modelo hacia una tarea específica.
 - `GlobalAveragePooling2D()`: Esta capa transforma los mapas de activación en un vector de características compacto, reduciendo la dimensionalidad del modelo y mitigando el riesgo de sobreajuste.
 - `Dropout(0.3)`: Se aplica una regularización que desactiva aleatoriamente el 30 % de las neuronas durante el entrenamiento, promoviendo la generalización del modelo.
 - `Dense(1, activation='sigmoid')`: Capa de salida densa con una sola neurona y activación sigmoide, diseñada para tareas de clasificación binaria.
 - Las capas del modelo base son `no entrenables (trainable=False)`, lo cual preserva los pesos preentrenados y permite ajustar únicamente las capas superiores mediante *fine-tuning*.

Análisis de Sentimientos

Pseudo-Labeling con Robertuito + Fine-Tuning Se implementó un enfoque de **pseudo-etiquetado** (pseudo-labeling) utilizando el modelo preentrenado *pysentimiento/robertuito-sentiment-analysis*, diseñado específicamente para análisis de sentimientos en español. Este modelo se aplicó inicialmente a un conjunto de comentarios no etiquetados para asignar automáticamente una de las tres categorías: "Positivo", "Negativo." o "Neutral".

- **Limitaciones:** El etiquetado automático puede estar sujeto a sesgos del modelo preentrenado. Además, los comentarios con baja confianza en la predicción podrían introducir ruido si no se filtran adecuadamente.
- **Razón de Selección:** Robertuito fue seleccionado por estar optimizado para lenguaje informal y corto como el que se encuentra en redes sociales. El uso de pseudo-etiquetado permitió generar un conjunto de datos confiables a partir de textos originalmente no etiquetados, sin necesidad de intervención humana.

■ Flujo Implementado:

- Se aplicó el modelo preentrenado a todos los comentarios, extrayendo las etiquetas predichas junto con las probabilidades asociadas.
- Se filtraron las predicciones cuya probabilidad superaba un umbral de confianza (por ejemplo, 0.90), obteniendo un subconjunto de datos pseudoetiquetados con alta precisión.
- Este subconjunto fue dividido estratificadamente en conjuntos de entrenamiento y validación.
- Luego, se realizó un *fine-tuning* sobre el mismo modelo Robertuito, utilizando únicamente estos datos filtrados, para adaptar el modelo a las características particulares del dominio analizado.
- Se empleó la métrica **f1-score** como criterio principal de evaluación durante el entrenamiento y se habilitó el guardado automático del mejor modelo.

■ Parámetros:

- **Modelo Utilizado:** *pysentimiento/robertuito-sentiment-analysis*, basado en RoBERTa preentrenado para sentimiento en español.
- **Tokenización:** Máximo de 128 tokens con truncamiento y padding.
- **Batch Size:** 8 ejemplos por lote, para balancear precisión y uso de memoria.
- **Épocas de Entrenamiento:** 3, dado el tamaño moderado del conjunto de datos filtrado.
- **Métrica de Selección:** F1-score ponderado, utilizada para guardar el mejor modelo.

3.4. Evaluación

3.4.1. Clasificación de Imágenes

En esta sección se presentan las evaluaciones realizadas a los diferentes modelos utilizados para la clasificación de imágenes. A continuación, se muestran las métricas de desempeño junto con los resultados gráficos para cada modelo evaluado:

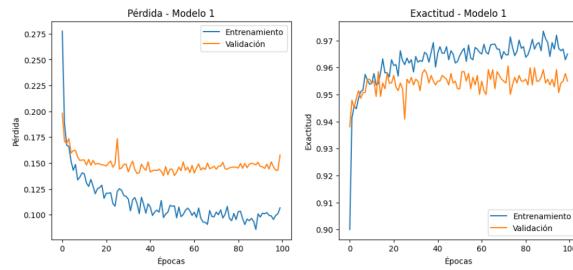


Figura 5: Evaluación del Modelo Xception

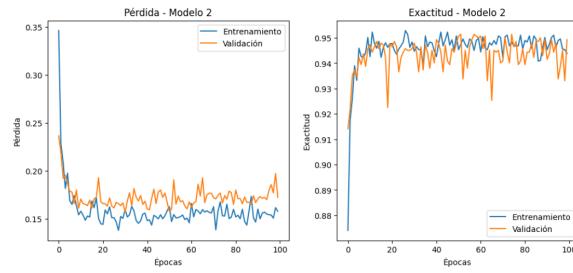


Figura 6: Evaluación del Modelo MobileNet

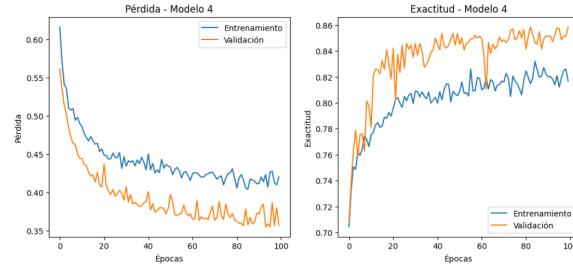


Figura 7: Evaluación del Modelo RedNet50

Los modelos fueron evaluados utilizando métricas clave como el **F1-Score** y el **AUC** (Área Bajo la Curva ROC), las cuales son esenciales para comprender el rendimiento en tareas de clasificación:

- El **F1-Score** combina la precisión (la proporción de predicciones correctas entre las positivas) y el recall (la proporción de verdaderos positivos identificados).
- El **AUC** mide la capacidad del modelo para distinguir entre las clases positivas y negativas; valores más cercanos a 1 indican un mejor desempeño.

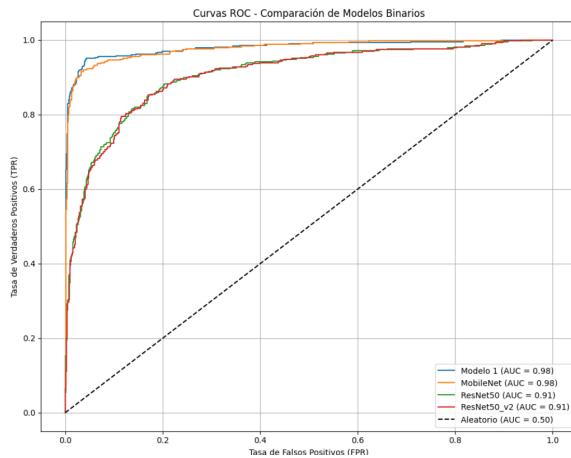


Figura 8: Curva ROC

Por ejemplo:

- El modelo **Xception** alcanzó un F1-Score de 0.9559 y una perdida de 0.1553, un AUC 0.98, reflejando una buena precisión en el ajuste.
- El modelo **MobilNet** presentó un F1-Score de 0.9527, con una perdida de 0.1487, un AUC de 0.98 lo que indica un desempeño similar al Xception pero destacándose con un área bajo la curva superior.
- Las **RedNet50** lograron un F1-Score de 0.8653 y una perdida de 0.3574, un AUC de 0.91, siendo el modelo que representa un menor ajuste.

Estos resultados destaca al modelo preentrenado Xception y MobilNet con finetuning como los modelos más efectivos para la tarea de clasificación de imágenes, gracias a su alta precisión y capacidad para identificar patrones visuales complejos.

3.4.2. Análisis de Sentimientos

RoBERTuito es un modelo de procesamiento de lenguaje natural especialmente entrenado para el idioma español. Su arquitectura está basada en RoBERTa, una variante optimizada del modelo BERT, y ha sido ajustado para comprender mejor el lenguaje informal y cotidiano que suele encontrarse en redes sociales.

A diferencia de BERT, que fue entrenado principalmente en inglés y sobre textos más formales, RoBERTuito fue alimentado con grandes volúmenes de datos en español, propias del lenguaje digital. Gracias a esto, ofrece una mayor precisión en tareas como el análisis de sentimientos, ya que logra captar mejor el contexto emocional en los textos.

3.5. Análisis y conclusiones

Con el propósito de analizar y visualizar el comportamiento emocional de los comentarios dirigidos hacia la marca Mattelsa, se desarrolló una aplicación interactiva utilizando el framework Dash de la biblioteca Plotly. Esta herramienta permite integrar gráficos dinámicos en una interfaz web amigable, lo cual facilita la exploración y el análisis profundo de datos sentimentales extraídos de redes sociales u otras plataformas digitales.

La aplicación está estructurada para ofrecer un panorama detallado del impacto emocional de las publicaciones de la marca, a partir de los comentarios previamente clasificados. En la interfaz, el usuario puede seleccionar una fecha específica mediante un calendario desplegable. A partir de esa selección, se actualizan automáticamente los gráficos que muestran el análisis emocional correspondiente a ese día.



Figura 9: Selección de Fechas en Interfaz

Uno de los componentes principales es un gráfico de dispersión, en el que cada punto representa un comentario publicado en la fecha seleccionada. En este gráfico, el eje horizontal indica la hora de publicación, mientras que el eje vertical representa el score de sentimiento. El color de cada punto indica la emoción (verde para positivo y rojo para negativo), y el tamaño varía de acuerdo con la intensidad del sentimiento. Este recurso visual permite identificar momentos del día con mayor carga emocional, así como detectar comentarios especialmente intensos o extremos.



Figura 10: Gráfico de dispersión de comentarios

Otro elemento clave es la generación de nubes de palabras, que se dividen en dos: una para los comentarios positivos y otra para los negativos. Estas visualizaciones muestran las palabras más frecuentes empleadas por los usuarios, excluyendo palabras vacías del español. Gracias a esto, es posible identificar de forma rápida los temas predominantes en cada tipo de emoción, lo cual aporta información cualitativa sobre la percepción de la audiencia.



Figura 11: Nube de palabras por día

A nivel temporal, la aplicación presenta dos gráficos de barras separados: uno que muestra la frecuencia diaria de comentarios negativos, y otro para los positivos. Esto permite monitorear la evolución del sentimiento en el tiempo, identificar picos emocionales, y vincularlos con campañas, eventos o publicaciones específicas. Complementariamente, se incluye un gráfico tipo boxplot que muestra la distribución mensual de los scores de sentimiento, diferenciando entre emociones. Esta visualización permite observar variaciones en la intensidad emocional por mes, identificar valores atípicos, y analizar cambios en la percepción general de la audiencia con el paso del tiempo.

Finalmente, se incorpora un gráfico de barras apiladas que muestra las 10 publicaciones más comentadas, desglosadas por emoción. Esta representación permite evaluar qué contenidos generan mayor interacción emocional, así como distinguir si dicha

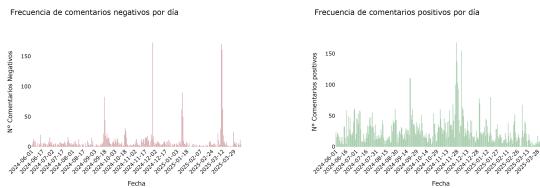


Figura 12: Frecuencia de Comentarios negativos y Positivos

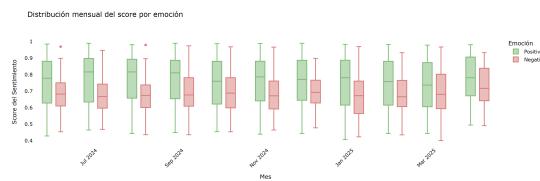


Figura 13: Distribución mensual del Score por emoción

interacción fue predominantemente positiva o negativa. Este último análisis resulta crucial para afinar el tono de la comunicación de marca, ajustar estrategias y responder a posibles crisis de reputación digital.

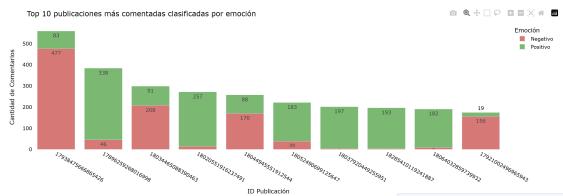


Figura 14: Top Publicaciones más comentadas

El análisis de los comentarios en redes sociales hacia la marca Mattelsa, clasificados por emoción y puntuación de sentimiento, evidencia una percepción mayoritariamente positiva por parte de los usuarios. A lo largo del periodo analizado, se observa una frecuencia sostenida de comentarios positivos con puntuaciones altas y baja dispersión, lo que sugiere una comunidad comprometida y generalmente satisfecha con la marca.

No obstante, también se identifican momentos puntuales con incrementos significativos en comentarios negativos, tanto en frecuencia diaria como en publicaciones específicas. Estos picos pueden corresponder a eventos o contenidos que generaron

descontento o controversia, lo que representa una oportunidad para revisar el contexto y el tipo de respuesta emitida por la marca.

4. Tecnologías

4.1. Descripción de características

4.1.1. Ingesta de los datos

Se utiliza la base de datos formato .xlsx de los históricos de las publicaciones realizadas en Mattelsa que cuenta con el identificador único de la publicación y la clasificación de esa publicación (Marca, Fotografía) y a su vez una carpeta que cuenta con todas las imágenes de esas publicaciones .jpg, esto sirve como insumo para entrenar el modelo de clasificación de imagen.

4.1.2. Procesamiento - validación

Clasificación de Imágenes

La clasificación de imágenes es una parte esencial del proyecto, permitiendo categorizar las imágenes según sus características visuales. A continuación, se detalla el proceso de conexión, preprocesamiento, modelado y evaluación aplicado.

Características de los Datos

- **Origen y Almacenamiento de los Datos:**

- **Fuente de datos:** Los datos consisten en imágenes almacenadas en una carpeta del FileServer de Mattelsa.
- **Formato de los datos:** Las imágenes están principalmente en formato JPEG (.jpg) y algunas en otros formatos como .JPG y .jpeg.

-
- **Metadatos asociados:** Se cuenta con un archivo CSV (`BD_Instagram.csv`) que contiene metadatos de las imágenes, incluyendo el nombre (`NOMBRE`) y la clasificación (`ÁREA`).
 - **Tamaño del conjunto de datos:** El DataFrame original contiene 4.590 registros con 19 variables diferentes.
- **Características de las Imágenes:**
- **Dimensiones y resolución:** Las imágenes tienen diferentes dimensiones y resoluciones, lo que requiere un procesamiento previo para estandarizar su tamaño.
 - **Canales de color:** Las imágenes pueden tener múltiples canales de color.
 - **Etiquetas de clasificación:** Las imágenes están clasificadas en categorías como `Marca` y `Producto`, lo que las hace aptas para un modelo de clasificación supervisada.
 - **Distribución de clases:** La distribución de las etiquetas se analiza para identificar posibles desbalanceos que puedan afectar el rendimiento del modelo.

4.1.3. Ingeniería de Características

Preprocesamiento de los Datos

- **Listar y cargar archivos:** Se filtran las imágenes con extensiones `.jpg` para su posterior procesamiento.
- **Limpieza de metadatos:** Las columnas del dataframe `bd_imagen` se limpian eliminando espacios en blanco y reemplazando caracteres acentuados por sus equivalentes sin tilde.

Creación y Unión de DataFrames

- **Construcción del DataFrame de imágenes:** Se crea un nuevo dataframe `df_imagen` que contiene dos columnas: `NOMBRE` (nombre de la imagen) e `IMAGEN` (ruta de la imagen).

-
- **Unión de DataFrames:** Se realiza una unión (`merge`) entre `df_imagen` y `bd_imagen` utilizando la columna `NOMBRE`, incorporando así las clasificaciones (`ÁREA`) en el dataframe de imágenes.

Conversión y Preprocesamiento de Imágenes

- **Conversión de formatos:** Las imágenes en formato `.jpg` se convierten a `.jpeg` para estandarizar el formato y facilitar su procesamiento.
- **Carga y preprocesamiento de imágenes:** Las imágenes se cargan desde sus rutas, se decodifican a escala RGB, se redimensionan a 100x100 píxeles, se convierten a arrays y se normalizan para que los valores de píxel estén en el rango [0, 1].

Análisis Exploratorio de Datos (EDA)

- **Visualización de imágenes:** Se visualizan un subconjunto de imágenes preprocesadas junto con sus etiquetas para verificar la correcta carga y clasificación.
- **Distribución de etiquetas:** Se analiza la frecuencia de cada etiqueta (`ÁREA`) para identificar posibles desbalanceos que podrían requerir técnicas de balanceo en etapas posteriores del modelado.

Limpieza y Recodificación de Etiquetas

- **Verificación de datos faltantes:** Se verifica la presencia de valores nulos en las columnas de rutas de imágenes y etiquetas para asegurar la integridad de los datos.

Analisis de sentimientos

El análisis de sentimientos es parte crucial de este proyecto, ya que permite comprender las emociones expresadas en los textos asociados a las imágenes. A continuación, se detalla el proceso de preprocesamiento y modelado aplicado.

Tipo de datos e idioma del texto:

- **Tipo de datos:** El conjunto de datos (`df['comentario']`) contiene textos en formato libre, comentarios de Instagram, en el cual se incluye elementos no estructurados como emojis, menciones, hashtags y URLs.
- **Idioma:** Los textos están en español, los cuales usan recursos lingüísticos específicos, como stopwords en español y modelos de procesamiento de lenguaje natural adaptados a este idioma.

Elementos Especiales:

- **Emojis:** Los textos contienen emojis que expresan emociones adicionales.
- **Menciones y hashtags:** Presencia de menciones (`@usuario`) y hashtags (`#tema`), que pueden influir en el análisis de sentimientos.
- **URLs y números:** Inclusión de enlaces web y números que no aportan valor semántico al análisis de sentimientos.

Longitud del Texto:

- **Variedad en longitud:** Los textos varían desde frases cortas hasta párrafos más extensos, lo que puede afectar la tokenización y el procesamiento subsiguiente.

Distribución de Sentimientos:

- **Balance de clases:** Distribución equilibrada o desbalanceada entre clases de sentimiento (positivo, negativo, neutro), lo que puede influir en las técnicas de balanceo aplicadas.

4.1.4. Ingeniería de Características

Preprocesamiento de Texto

- Conversión a tipo de dato string:

- `df['comentario'] = df['comentario'].astype(str)`
- **Descripción:** Se asegura que todos los datos en la columna de texto sean tratados como cadenas de caracteres, evitando errores durante el procesamiento.

- Manejo de emojis:

- **Conservación de emojis:** Se Limpia el texto manteniendo los emojis, lo que permite su uso en análisis emocionales más detallados.

Tokenización

- Uso de Autokenizador de `texttpysentimiento/robertuito-sentiment-analysis`: Divide los textos en tokens (palabras individuales) de manera eficiente, considerando la naturaleza informal y abreviada de los comentarios de Instagram.

Análisis de frecuencia y visualización

Clustering y modelado de tópicos

- **Pseudo label con Robertuito para análisis de sentimientos:** Realiza un etiquetado automático para clasificar los comentarios según su polaridad

(positivo, negativo, neutral) y subjetividad. Esto ayuda a entender el tono emocional detrás de los comentarios, proporcionando un contexto adicional para los tópicos identificados.

- **Fine-tuning con Robertuito para análisis de sentimientos:** Realiza un entrenamiento basado en el modelo pre entrenado aplicando fine-tuning utilizando las etiquetas generadas anteriormente por la etiqueta automática.

Evaluación y validación

- **Métricas de evaluación:** Proporciona métricas como precisión, recall y F1-score para evaluar el desempeño del modelo en la clasificación de sentimientos.
- **Validación cruzada:** Evalúa la consistencia del modelo a través de diferentes subconjuntos de datos, asegurando una mejor generalización.
- **Análisis de errores:** Inspecciona ejemplos mal clasificados para identificar posibles mejoras en el preprocesamiento o en el modelo.

4.2. Desarrollo del Proyecto

El desarrollo del proyecto se centró en aprovechar tecnologías modernas de procesamiento de datos, aprendizaje automático y visualización para garantizar un flujo continuo y eficiente desde la extracción de datos hasta la generación de insights accionables.

- **Extracción de datos:**

- Bases de datos de Mattelsa para extraer comentarios de Instagram.

-
- Complemento de datos estructurados proporcionados por Mattelsa en formato CSV, garantizando integridad en la asociación de imágenes y metadatos.

- **Preprocesamiento de datos:**

- Limpieza y estandarización de datos textuales mediante **Python** (NLTK, Dataset de Hugginface, Autokenizador de Robertuito).
- Normalización y conversión de imágenes a un formato compatible para los modelos de machine learning (escala de RGB y redimensionamiento).

- **Entrenamiento de modelos:**

- Implementación de un enfoque modular para utilizar modelos preentrenados de clasificación (**Xception**, **MobilNet** y **RedNet50**) y análisis de sentimientos (**Robertuito para pseudo label** y **Finet -tuning**).
- Validación y evaluación del rendimiento mediante métricas como **F1-score**, **AUC**, **Loss** y **Accuracy**.

- **Naturaleza:**

- Batch con datos estructurados y no estructurados.

4.3. Despliegue del Proyecto

Un escenario ideal para el despliegue de esta solución implica la integración en una arquitectura escalable y robusta, que permita automatizar los procesos de clasificación de imágenes y análisis de sentimientos, mientras se garantizan resultados en tiempo real para la toma de decisiones.

- **Propuesta de despliegue:**

- A partir de la API de Meta se correrá cada cierto tiempo bajo una automatización que alojara los datos en el DataWarehouse de Mattelsa que contiene el comentario y la URL de la imagen.
- Se deberá correr pipeline para capturar la URL de la imagen.
- Esta URL de la información se irá al modelo de clasificación de imágenes.
- Se guarda la clasificación de la imagen de nuevo en la tabla de donde se extrajo la información.
- Posterior a esto, se procesan los comentarios en el modelo de análisis de sentimientos. Para obtener la clasificación del comentario (positivos, negativos o neutros).
- Se alojara nuevamente en la tabla la clasificación obtenida por el análisis de sentimientos.
- Se genera el ciclo en donde se extraen los comentarios nuevos que están llegando de la API de instagram y que aun no contiene ninguna clasificación ni de imagen, ni emoción y se realiza de nuevo el proceso.
- Estos datos alimentarán un dashboard de percepción general al área de marketing de la compañía.

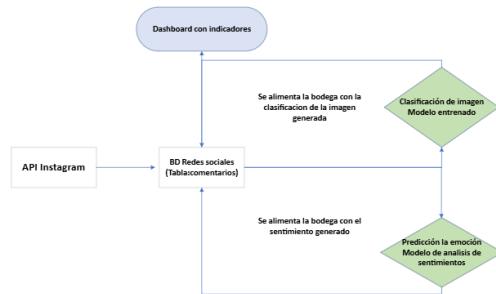


Figura 15: Flujo propuesta de despliegue

4.4. Detalles Técnicos

- **Fuentes de Datos:**

- **Batch:** Datos estructurados y no estructurados de publicaciones y comentarios de Instagram extraídos periódicamente en formato URL y string.
- **Streaming:** Posible integración futura con flujos en actualización diaria mediante automatización de flujo creada por python para procesar nuevos datos en el momento.

- **Ingesta de Datos:**

- Recolección a través de la API de Meta (batch).
- Almacenamiento directo en Datawarehouse de Mattelsa.

- **Almacenamiento:**

- Lago de datos en **Datawarehouse** con capas diferenciadas:
 - **Trusted:** Datos limpios y listos para modelado.
 - **Processed:** Resultados del análisis de sentimientos y clasificación de imágenes.

- **Ambiente de Procesamiento:**

- **Python:** Librerías principales incluyen TensorFlow, NLTK.
- **Apache Spark:** Procesamiento distribuido para análisis exploratorio y manejo de grandes volúmenes de datos.

- **Propuesta de aplicaciones:**

- **APIs:** Desarrollo de una API con Flask para exponer resultados del análisis a otras aplicaciones.

-
- **Dashboards:** Implementación en Tableau o Power BI para la visualización interactiva de resultados, incluyendo análisis por periodo, tipo de publicación y sentimiento predominante.

5. Conclusiones generales del Proyecto

5.1. Clasificación de imágenes:

- El conjunto de datos total consistió en 5,056 imágenes, las cuales se dividieron de manera estratégica en dos partes: un 70 % para el entrenamiento, y un 30 % para la validación. Esta división permitió entrenar al modelo con una cantidad considerable de datos, mientras que el conjunto de validación proporcionó una evaluación objetiva de su capacidad para generalizar y realizar predicciones precisas sobre datos no vistos.
- De todos los modelos evaluados, el **modelo Xception y MobilNet** fueron los modelos que tuvieron mejor desempeño, con un accuracy de 0.95 y un loss de 0.15. Por recursos se decide que el modelo a utilizar es MobileNet debido a que tiene un menor tiempo de ejecución.

5.2. Análisis de sentimientos:

Se utilizaron dos metodologías: una basada en pseudoetiquetado y otra en fine-tuning directo utilizando el modelo RoBERTuito. Al aplicar ambos procedimientos, se obtuvo un F1-score de 1, lo cual evidenció un posible sobreajuste. A partir de estos

resultados, se concluye que no es necesario realizar un etiquetado automático previo al entrenamiento debido a que se obtienen los mismos resultados, sino que es más conveniente utilizar directamente el modelo preentrenado RoBERTuito para el ajuste fino con los datos disponibles.

Es decir, como respuesta a la pregunta de negocio original, se puede concluir que la marca Mattelsa, en general, mantiene una imagen positiva frente a sus clientes.

Referencias

Cambria, Erik, Li, Yang, Xing, Frank Z., Poria, Soujanya, and Kwok, Kenneth. (2020). *Ensemble Application of Symbolic and Subsymbolic AI for Sentiment Analysis*. Proceedings of the 29th ACM International Conference on Information and Knowledge Management. Liao, Shiyang, Wang, Junbo, Yu, Ruiyun, Sato, Koichi, and Cheng, Zixue. (2017). *CNN for Situations Understanding Based on Sentiment Analysis of Twitter Data*. Procedia Computer Science. Ray, Paramita, and Chakrabarti, Amlan. (2019). *A Mixed Approach of Deep Learning Method and Rule-Based Method to Improve Aspect Level Sentiment Analysis*. Applied Computing and Informatics. Wei, S., & Song, S. (2022). Sentiment classification of tourism reviews based on visual and textual multifeature fusion. *Wireless Communications and Mobile Computing*, 2022, Article ID 9940817. <https://doi.org/10.1155/2022/9940817>