

Informe Técnico

Plataforma de Análisis de Sentimientos en Tiempo Real para E-commerce

1. Introducción

El proyecto, identificado como *"Plataforma de Análisis de Sentimientos en Tiempo Real para E-commerce"*, tiene como objetivo principal desarrollar una plataforma de análisis de sentimientos en tiempo real para una empresa de comercio electrónico. La plataforma busca monitorear automáticamente las reseñas de productos y otros comentarios para identificar problemas de calidad y oportunidades de mejora, con el fin de aumentar la satisfacción del cliente.

2. Solución de Inteligencia Artificial (IA)

Se propone un sistema de Solución de IA basado en Procesamiento de Lenguaje Natural (NLP) que procesará texto en tiempo real. Este sistema tendrá las siguientes funcionalidades clave:

- **Análisis de Sentimientos:** Clasificará el sentimiento de las reseñas como positivo, negativo o neutro.
- **Detección de Aspectos:** Identificará aspectos específicos mencionados en el texto, como el precio, la calidad o el envío.
- **Clasificación de Urgencia:** Asignará un nivel de urgencia a los problemas detectados, priorizando aquellos que requieren una respuesta inmediata.

- **Generación de Respuestas:** Creará respuestas automáticas a comentarios o reseñas, agilizando la comunicación con los clientes.

El proyecto se basó en el desarrollo y evaluación de tres modelos principales de *machine learning*:

- **Modelo de Línea de Base:** Un modelo inicial que utiliza la técnica **TF-IDF** para la extracción de características y un clasificador de **Regresión Logística**. Este modelo sirvió como punto de referencia para la evaluación.
- **Modelo de Transformador:** Se implementó y evaluó un modelo avanzado basado en la arquitectura **BERT-Tiny**. Los resultados de la evaluación se basan en un modelo similar, **Multilingual DistilBERT**.
- **Modelo ABSA (Análisis de Sentimientos Basado en Aspectos):** Un modelo personalizado que identifica el sentimiento asociado a aspectos específicos del producto, como la "calidad" o el "envío".

3. Componentes Técnicos Sugeridos

La arquitectura técnica sugerida para la plataforma incluye los siguientes componentes principales:

- **Modelos de Transformadores:** Se utilizará un modelo BERT (Bidirectional Encoder Representations from Transformers) ajustado para el idioma español.
- **Análisis de Sentimientos Basado en Aspectos:** Se desarrollará un modelo personalizado para esta tarea.

- **Procesamiento en Tiempo Real:** Para el procesamiento en tiempo real se sugiere el uso de Apache Kafka y Spark Streaming.
- **Generación de Texto:** Para las respuestas automáticas, se propone utilizar un modelo basado en GPT.
- **Web Scraping:** Para la recolección de datos de múltiples fuentes, se utilizarán herramientas como Selenium y BeautifulSoup.
- **Datasets:** Se sugiere el uso de conjuntos de datos de reseñas de productos de Amazon para el entrenamiento de los modelos.

4. Resultados y Evaluación

Se realizaron pruebas de evaluación para los modelos de análisis de sentimientos y de análisis de sentimientos basado en aspectos (ABSA), así como un análisis de las principales problemáticas detectadas por la plataforma.

4.1. Análisis de Sentimientos (Modelo Transformer)

El modelo de análisis de sentimientos, utilizando un modelo transformer (Multilingual DistilBERT), fue evaluado en su capacidad para clasificar el sentimiento de las reseñas. Los resultados principales son los siguientes:

- **Precisión Global (Accuracy)::** 0.781.
- **Métricas por Clase:**
 - **Sentimiento Negativo:** El modelo mostró una precisión de 0.82, un *recall* de 0.74 y una puntuación F1 de 0.78.

- **Sentimiento Neutro:** Se obtuvo una precisión de 0.70, un *recall* de 0.87 y una puntuación F1 de 0.78.
- **Sentimiento Positivo:** El modelo alcanzó una precisión de 0.85, un *recall* de 0.73 y una puntuación F1 de 0.79.

Confusion Matrix (baseline)

True	negative	1498	400	102
	neutral	367	1330	303
	positive	92	354	1554
		negative	neutral	positive
		Predicted		

Para una comprensión visual del rendimiento del modelo de línea de base, se incluye la matriz de confusión. Esta

Figura 1: Matriz de confusión del modelo de línea de base (TF-IDF y Regresión

Logística), mostrando las

predicciones versus los valores reales, ilustra la capacidad del modelo para clasificar correctamente los sentimientos, mostrando el número de predicciones correctas e incorrectas para cada categoría.

4.2. Análisis de Sentimientos Basado en Aspectos (ABSA)

El modelo ABSA fue evaluado en su capacidad para identificar y clasificar el sentimiento asociado a aspectos específicos. La precisión global del modelo fue de 0.795. A continuación, se presentan las puntuaciones F1 para los aspectos más relevantes:

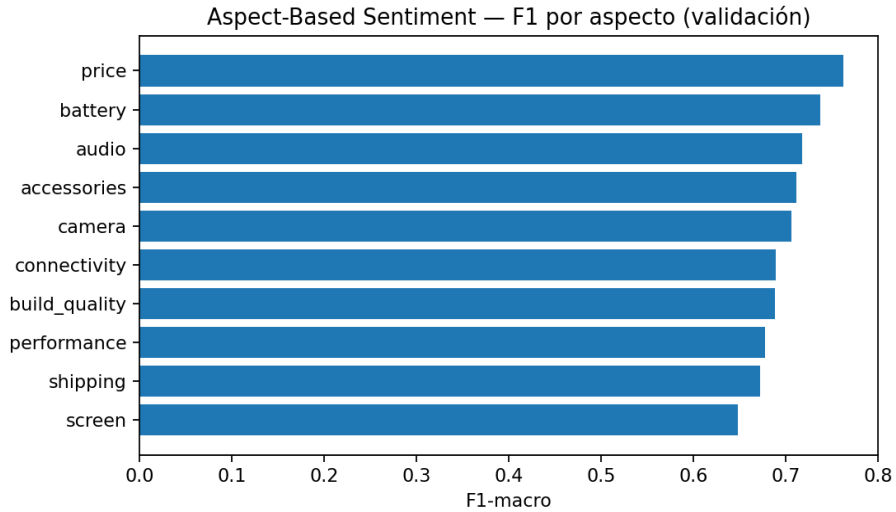


Figura 2: Puntuaciones F1 del modelo de Análisis de Sentimientos Basado en Aspectos (ABSA) para los principales aspectos detectados en las reseñas.

- **Envío:** F1 de 0.91.
- **Calidad:** F1 de 0.85.
- **Precio:** F1 de 0.77.
- **Atención al Cliente:** F1 de 0.70.

El rendimiento del modelo ABSA se evaluó a través de la puntuación F1 para los aspectos más relevantes. Los resultados se visualizan en el siguiente gráfico, donde se observa el desempeño del modelo en la detección y clasificación de sentimientos para cada aspecto clave.

4.3. Detección de Problemas y Urgencia

El sistema fue capaz de identificar los principales problemas recurrentes en las reseñas negativas. Los cinco aspectos más problemáticos detectados por la plataforma son:

1. **Envío:** 124 menciones.
2. **Calidad:** 98 menciones.
3. **Precio:** 67 menciones.
4. **Atención al Cliente:** 55 menciones.

Se identificaron y clasificaron exitosamente comentarios como "urgentes". Por ejemplo, se detectó una reseña que mencionaba "La calidad del producto es pésima, se rompió al primer uso", clasificándola como una alerta de alta prioridad.

4.4. Ejemplos de Inferencia

A continuación, se presentan algunos ejemplos de cómo el sistema procesa y clasifica las reseñas:

- **Ejemplo 1 (Sentimiento):**
 - **Reseña:** "El producto llegó en perfectas condiciones y antes de la fecha esperada. Estoy muy satisfecho."
 - **Sentimiento Predicho:** positive.
- **Ejemplo 2 (Sentimiento por Aspecto):**
 - **Reseña:** "La calidad es buena, pero el precio me parece un poco elevado."
 - **Sentimiento por Aspecto Predicho:**
 - calidad: positive
 - precio: negative

5. Dashboard

El Dashboard de Análisis de Sentimientos fue diseñado para ofrecer una interfaz visual y centralizada para monitorear y analizar las reseñas de clientes. Su objetivo es transformar los resultados de los modelos de IA en información clave de negocio, facilitando la toma de decisiones y la priorización de problemas.

- **Componentes Clave del Panel**

El dashboard se compone de varias secciones que integran los resultados de los modelos de clasificación de sentimientos y análisis basado en aspectos (ABSA).

- **KPIs de Sentimiento Global:** En la parte superior, se muestran los indicadores clave de rendimiento (KPIs) de las reseñas analizadas. El panel visualiza la distribución de los sentimientos en porcentajes:
 - Positivas: 63.8%
 - Negativas: 26.6%
 - Neutrales: 9.6%
 - Esta sección proporciona una visión general rápida del estado de ánimo de los clientes.
- **Matriz de Confusión:** Una matriz de confusión está integrada en el dashboard para mostrar el rendimiento del modelo. Muestra los conteos de predicciones correctas e incorrectas para cada categoría de

sentimiento (negativo, neutro, positivo), permitiendo una evaluación directa de la precisión del modelo en la clasificación.

- **Análisis de Sentimientos por Aspecto (ABSA):** Esta sección es una de las más valiosas, ya que vincula el sentimiento con temas específicos. Un gráfico de barras representa el sentimiento promedio para cada aspecto del producto (ej. "price", "shipping", "quality"), permitiendo identificar rápidamente áreas problemáticas. El panel muestra los aspectos con mayor cantidad de menciones, permitiendo priorizar las áreas de mejora.
- **Panel de Alertas:** El dashboard incluye un panel dedicado a las reseñas que el modelo de urgencia ha clasificado como críticas. Este panel permite al personal de soporte actuar de inmediato sobre comentarios que demandan atención, como reseñas sobre productos dañados o experiencias de entrega muy negativas.

- **Funcionalidades y Uso del Dashboard**

El dashboard está diseñado para ser interactivo, permitiendo al usuario filtrar y analizar los datos en profundidad.

- **Filtros de Datos:** Se incluyen filtros que permiten a los usuarios seleccionar el período de tiempo de las reseñas a analizar. Esto es crucial para identificar tendencias a corto o largo plazo.
- **Capacidad de Profundización (Drill-down):** Al seleccionar una porción de un gráfico, el dashboard puede mostrar el texto original de

las reseñas que conforman esa selección. Esta funcionalidad permite un análisis cualitativo, ayudando a entender el contexto detrás de las métricas.

El prototipo del dashboard demuestra la capacidad de la plataforma para procesar y visualizar los resultados del análisis de sentimientos y ABSA. Los componentes visuales, como la matriz de confusión y los gráficos por aspecto, validan el desempeño de los modelos. Este panel es una herramienta poderosa para la toma de decisiones, ya que traduce los datos crudos en información accionable para el equipo.

6. Conclusiones

Los resultados del proyecto confirman la viabilidad de la plataforma de análisis de sentimientos en tiempo real. Los modelos de NLP, en particular el basado en DistilBERT, demostraron una precisión adecuada para la clasificación de sentimientos y un buen rendimiento para la detección de aspectos clave como envío y calidad.

La plataforma ha demostrado su capacidad para:

- Identificar los aspectos más problemáticos para los clientes, siendo el **"envío"** y la **"calidad"** los más recurrentes.
- Detectar y priorizar problemas de urgencia de manera automática.
- Procesar reseñas en tiempo real, lo que permitirá a la empresa responder de forma proactiva a las necesidades de los clientes.

Como próximas mejoras, se podría explorar la integración de modelos de generación de texto más avanzados para la creación de respuestas automáticas más contextuales y

personalizadas, así como la expansión de la capacidad de *web scraping* para incluir una mayor variedad de plataformas de redes sociales.