



TECNOLOGIA SUPERIOR

BIG DATA E INTELIGENCIA DE NEGOCIO

Prácticas de Análisis de Sentimientos – Ciencia de Datos

WILLIAM ESTUARDO JIMÉNEZ MIGUEZ

william.jimenez@cenestur.edu.ec

Profesor(a): YADIRA FRANCO

yadira.franco@cenestur.edu.ec

Quito, Ecuador

2025

a. Introducción

Este informe presenta el análisis de sentimientos de 2.000 comentarios escritos en español extraídos de reseñas de productos. El objetivo principal es clasificar cada comentario según su carga emocional (positivo, negativo o neutro), identificar patrones textuales relevantes y extraer conclusiones útiles para la toma de decisiones estratégicas o comerciales.

b. Objetivos

Objetivo Principal

Analizar los comentarios escritos en español mediante técnicas de procesamiento de lenguaje natural y modelos de análisis de sentimientos, con el fin de clasificar la percepción de los usuarios respecto a productos o servicios y extraer conclusiones útiles para la toma de decisiones.

Objetivos Específicos

1. Explorar y comprender la estructura, tipo y contenido del dataset sentimiento_2.csv, identificando las columnas relevantes para el análisis textual.
2. Aplicar técnicas de preprocesamiento de lenguaje natural (limpieza, tokenización, normalización y lematización) a los comentarios con el fin de preparar los datos para su análisis semántico.
3. Implementar modelos automáticos de análisis de sentimientos, como BERT multilingüe, para clasificar los comentarios en categorías emocionales: positivo, negativo y neutro.
4. Visualizar los resultados obtenidos a través de gráficos estadísticos e indicadores clave, para facilitar su interpretación y apoyar el análisis exploratorio.
5. Identificar frases clave y patrones lingüísticos recurrentes, que permitan caracterizar el tipo de comentarios asociados a cada categoría de sentimiento.
6. Elaborar conclusiones y recomendaciones basadas en los hallazgos del análisis, orientadas a mejorar la experiencia del cliente y la gestión de productos o servicios.

d. Desarrollo

Metodología

Se aplicó el siguiente flujo de trabajo:

1. **Carga y exploración del dataset**

Se identificaron 10 columnas relevantes, entre ellas Score, Summary y Text. La columna Text fue seleccionada para el análisis de sentimiento.

2. *Preprocesamiento del texto*

- Conversión a minúsculas
- Eliminación de signos de puntuación y símbolos
- Eliminación de stopwords
- Tokenización
- Lematización

3. *Análisis de sentimientos*

- Se utilizó el modelo preentrenado nlptown/bert-base-multilingual-uncased-sentiment, que clasifica textos en 5 niveles (de “1 estrella” a “5 estrellas”).
- Posteriormente, los resultados fueron simplificados a 3 categorías: Negativo, Neutro y Positivo.

Análisis Exploratorio

- *Total de registros:* 2000
- *Longitud promedio del texto:* 257 caracteres
- *Valores nulos en columna Text:* 0
- *Distribución de Score original:*

Score	
5	945
4	650
3	220
2	105
1	80

Resultados del Análisis de Sentimientos

Distribución de Sentimientos (Simplificados)

SENTIMIENTO	CANTIDAD	PORCENTAJE
Positivo	1120	56%
Negativo	520	26%
Neutro	360	18%

- ✓ Más de la mitad de los comentarios reflejan experiencias positivas.
- ✓ El 26% refleja insatisfacción o quejas, lo cual debe investigarse más a fondo.

Comparación con el Score original

Se encontró una alta concordancia entre los puntajes de usuario (Score) y el sentimiento detectado por BERT. Por ejemplo, la mayoría de los comentarios con Score = 5 fueron clasificados como Positivos.

Frases Destacadas por Sentimiento

Comentarios Positivos:

- ✓ “Este producto superó mis expectativas, definitivamente lo volvería a comprar.”
- ✓ “La atención al cliente fue excelente y el envío rápido.”

Comentarios Neutros:

- ✓ “El producto cumple su función, pero nada extraordinario.”
- ✓ “Recibí lo que esperaba, sin más ni menos.”

Comentarios Negativos:

- ✓ “Llegó dañado y no cumplieron con lo prometido.”
- ✓ “Muy mala calidad, me decepcionó completamente.”

e. Conclusiones

- ✓ El análisis de sentimientos permite evaluar con precisión la satisfacción del cliente a partir de texto libre.
- ✓ Existe una clara predominancia de comentarios positivos, lo cual habla bien del producto o servicio ofrecido.
- ✓ Sin embargo, el 26% de opiniones negativas constituye un porcentaje crítico que requiere atención.
- ✓ Las palabras más comunes en reseñas negativas incluyen: “malo”, “dañado”, “tarde”, “decepcionado”.
- ✓ Las palabras más frecuentes en comentarios positivos fueron: “excelente”, “rápido”, “recomiendo”, “calidad”.

f. Recomendaciones

- ✓ Revisar los productos asociados a comentarios negativos, ya que podrían estar dañando la reputación general.
- ✓ Optimizar procesos logísticos, ya que muchas críticas apuntan a retrasos o empaques defectuosos.
- ✓ Usar los comentarios positivos como testimonios en campañas de marketing.

- ✓ Incorporar dashboards interactivos en Power BI o Looker Studio para monitoreo continuo de sentimientos.

g. Entregables

- ✓ dashboard_sentimiento.csv: contiene los comentarios con los sentimientos clasificados (Sentimiento_Label, Sentimiento_Simple).
- ✓ Gráficos de barra y nube de palabras para cada grupo de sentimiento (opcional a exportar).
- ✓ Se desarrollo en Google Colab (Análisis_Sentimientos.ipynb)
- ✓ Link GitHub: <https://github.com/Willyejm/SentimientosTexto.git>