

+

Laboratorio 3

Laboratorio 3

Título del Laboratorio: Laboratorio análisis de sentimiento y predicción del comportamiento del cliente

Duración: 2 horas

Objetivos del Laboratorio:

- *Familiarizarse con el análisis de datos y visualización usando Pandas y Seaborn.*
- *Construir y evaluar un modelo básico de clasificación de sentimientos usando CountVectorizer y Naive Bayes.*

Materiales Necesarios:

1. *Archivo del dataset: Reviews.csv*
2. *Python 3.x*
3. *Pandas, seaborn, matplotlib, scikit-learn*
4. *Google Colab o Jupyter*

Documento de Ayuda:

1. *Reviews.csv*

Estructura del Laboratorio:

1. Carga el dataset.

- Abre un archivo nuevo en Jupyter Notebook, Google Colab o tu editor preferido.
- Cargar el dataset con la librería pandas
- Ver la cabecera del csv
- Seleccionar las columnas importantes ("Text", "Score")
- Crear columna de sentimiento
- Mostrar distribución de sentimientos

2. Preguntas rápidas para reflexión:

- ¿Qué tipo de datos encontramos en las columnas?
- ¿Cuál es la proporción de reseñas positivas vs. negativas?

3. Preprocesamiento y vectorización

- División de datos en entrenamiento y prueba
- Dividir los datos
- Vectorización del texto

4. Reflexión guiada

- ¿Qué hace el CountVectorizer?
- ¿Por qué dividimos los datos en conjuntos de entrenamiento y prueba?

5. Entrenamiento y evaluación del modelo

- Importar librerías para entrenamiento
- Entrenar el modelo
- Evaluar el modelo
- Prueba el modelo con ejemplos reales