

INSTRUCCIONES PARA ENTRENAR Y PROBAR EL MODELO DE SENTIMIENTO

Guía de uso del sistema de entrenamiento y predicción

1. Estructura de carpetas esperada (desde el root del proyecto):

```
/machine_learning
  data/
    balanced_reviews.csv
  models/
    best_model.pkl
    vectorizer.pkl
  scripts/
    data_prep.py
    evaluate_model.py
    inference.py
    test_sentiment.py
    train_models.py
```

2. Entrenamiento del modelo (solo si es necesario):

- Asegúrate de tener el dataset en `machine_learning/data/balanced_reviews.csv`.
- Desde el root del proyecto, ejecuta:

```
$ python -m machine_learning.scripts.train_models
```

Esto realiza los siguientes pasos:

- Carga y limpia los datos.
- Prepara el corpus.
- Vectoriza usando TF-IDF.
- Entrena y evalúa modelos (Logistic Regression, Linear SVM).
- Imprime métricas de rendimiento y matriz de confusión.
- Instrucciones para Entrenar y Probar el Modelo de Sentimiento
- Guarda el mejor modelo y vectorizador en `machine_learning/models/`.

3. Reentrenamiento: ¿Es necesario volver a ejecutar `train_models.py`?

Solo si:

- Cambió el corpus (nuevo preprocesamiento o dataset).
- Se ajustaron hiperparámetros.
- Se agregó un nuevo modelo o métrica.

4. **Prueba de inferencia** (modelo ya entrenado):

- Desde el root del proyecto, ejecuta:

```
$ python -m machine_learning.scripts.test_sentiment
```

Esto utiliza `inference.py` para cargar el modelo y vectorizador desde `machine_learning/models/`, limpia y vectoriza texto de ejemplo, y predice su sentimiento.