

Práctica 2

Inteligencia Artificial
Facultad de Ciencias, UNAM

Elabora un **reconocedor de entidades nombradas**, con base en los siguientes puntos:

1. Toma el dataset proporcionado, donde la primera columna corresponde a los datos de entrada y la segunda a los datos de salida (emisiones). Las etiquetas a considerar son de la forma:

Etiqueta	Uso
O	Son palabras que sin interés
B-tag	Indican el inicio de una entidad nombrada del tipo “tag”
I-tag	Indican el interior de una entidad nombrada del tipo “tag”

2. Pasa a minúsculas los datos de entrada únicamente y separa los datos por espacios para obtener las palabras y las etiquetas.
3. Usa la función `train_test_split` de la paquetería `sklearn` para separar los datos en un 70% de entrenamiento y un 30% de evaluación.
4. Con los datos de entrenamiento, obtene las estimaciones de las probabilidades de los modelos ocultos de Márkov.
5. Con los datos de evaluación, obtén las etiquetas a partir del algoritmo de Viterbi (si no se reconoce una palabra usa probabilidad 1).
6. Evalúa el resultado comparando las predicciones del modelo con las etiquetas reales. Puede usarse `sklearn.metrics.classification_report`.