

TEMA 5

MÓDULO:
TÉCNICAS AVANZADAS DE
MACHINE LEARNING

MODELOS DE SENTIMIENTOS

ALBERTO JULIÁN RIGAU

Ingeniero Superior de Telecomunicaciones y MBA.

MARC FRANCO SALVADOR

Doctor of Philosophy (PhD), programa de doctorado en informática UPV. Ingeniero Informático por UPV.



Institut de Formació Contínua-IL3
UNIVERSITAT DE BARCELONA

ÍNDICE

Objetivos Específicos

Modelos de sentimientos

Ideas clave



OBJETIVOS ESPECÍFICOS

- Adquirir un conocimiento básico del Procesado de Lenguaje Natural o NLP, previo al Deep Learning.
- Conocer algunos modelos de Deep Learning aplicados a NLP.
- Aprender dónde se pueden encontrar corpus de texto y practicar con ellos.

MODELOS DE SENTIMIENTOS

En este último tema, veremos qué es el Procesado de Lenguaje Natural o NLP (Natural Language Processing), tanto en su versión “tradicional” como algunos modelos de Deep Learning aplicado a NLP.

Como libro de referencia podemos citar Jurafsky y Martin, 2021 (previsión).

En el enfoque de NLP con modelos tradicionales, realizaremos las actividades siguientes:

- Extraer y clasificar entidades o NER (Named Entity Recognition).
- Etiquetar partes de un texto o PoS (Part of Speech).
- Realizar un árbol sintáctico.
- Clasificar texto.

Por otro lado, veremos distintos modelos de Deep Learning que han producido una disrupción en NLP.

En primer lugar, podemos citar el modelo word2vec (Mikolov et al., 2013), en el que, a cada palabra, se le asigna un vector (por eso reciben el nombre de word vectors), de forma que las dimensiones del vector dependen de las relaciones (semánticas, sintácticas, etc.) que tenga esa palabra con otras de un corpus lingüístico de referencia. Debido a esas relaciones, también se conoce a estos vectores como **embeddings**.

Por otro lado, se recuperó y evolucionó el modelo de **LSTM** (tutorial en Olah, 2015) por su capacidad para recordar relación entre palabras de una secuencia relativamente alejadas entre sí. En combinación con modelos de word embeddings se aplicó a casos de uso como traducción, análisis de sentimiento, generación de texto, etc.

Desde 2017, están generando mucha expectación los modelos de atención y los basados en arquitecturas de transformers (Bloem, 2019).

Con algunos de estos modelos, realizaremos las actividades siguientes:

- Análisis de sentimiento.
- Generación de texto.

A partir de este punto, vamos a continuar en Colab, herramienta con la que ya sabes trabajar.



IDEAS CLAVE

- El Procesado de Lenguaje Natural sigue superando expectativas año tras año.
- Hay muchos datasets de texto de acceso libre.
- La aplicación de modelos de Deep Learning al NLP ha producido un salto cualitativo en muchos aspectos.