

Actividad. Implementación de un clasificador de sentimiento con Transformers

Unidad	2
Entrega	Documento en formato PDF: La actividad deberá ser entregada en la fecha indicada por el profesor y siguiendo las indicaciones recogidas en el apartado Detalles de la entrega de este mismo documento.

1. Enunciado

En esta actividad, utilizaremos la biblioteca Transformers para implementar un clasificador de sentimiento utilizando un modelo preentrenado de Transformers. Los estudiantes trabajarán con conjuntos de datos de reseñas de películas para entrenar y evaluar el rendimiento del clasificador de sentimiento.

Objetivos de aprendizaje de esta actividad:

- Familiarizarse con la biblioteca Transformers y comprender su uso en el procesamiento del lenguaje natural.
- Aprender a cargar y procesar conjuntos de datos de reseñas de películas para entrenar un clasificador de sentimiento.
- Implementar un clasificador de sentimiento utilizando un modelo preentrenado de Transformers, como BERT o DistilBERT.
- Evaluar el rendimiento del clasificador de sentimiento utilizando métricas como precisión, recall y F1-score.

Entregables:

1. Código fuente del clasificador de sentimiento implementado en Python utilizando la biblioteca Transformers.
2. Informe de análisis de resultados que incluya métricas de evaluación del rendimiento del clasificador y conclusiones sobre la efectividad del modelo implementado.

Recursos necesarios:

1. Entorno de desarrollo Python con acceso a bibliotecas como Transformers, pandas y scikit-learn.
2. Conjuntos de datos de reseñas de películas etiquetadas con sentimientos positivos y negativos.
3. Material de referencia, tutoriales y documentación sobre el uso de Transformers para clasificación de sentimientos.

La idea inicial es recoger un conjunto de textos/opiniones de usuarios y realizar las fases descritas en esta unidad:

Actividad. Implementación de un clasificador de sentimiento con Transformers

- Recopilación de datos.
- Preprocesado.
- Entrenamiento.
- Comparación de resultados.
- Análisis de métricas.
- Arranque de un potencial modelo.

2. Detalles de la entrega

- Elabora un documento que incluya la información detallada en el apartado anterior.
- Es recomendable que, aunque el documento se realice en formato pptx o bien en formato Word, se cumplimente la entrega en pdf para evitar posibles alteraciones de formato que afecten a su correcta lectura.
- Subir de forma individual el documento a la actividad del campus virtual.

3. Anexo

Al ser un ejercicio puntuable, se evaluarán varios puntos. Algunos de ellos los ha formulado el profesor antes del comienzo de la semana de la actividad, pudiendo presentar una rúbrica u otro criterio de corrección que el profesor estime. Pero, además de ellos y de forma implícita, se puntuarán muy firmemente los siguientes:

- **La originalidad:** es decir, las copias de Internet, de un compañero o de otra bibliografía no serán válidas.
- **La simplicidad:** enunciados claros y soluciones fáciles de seguir.
- **La solución:** no es suficiente con unos buenos ejercicios si estos están mal resueltos.
- **La planificación:** entregas a tiempo y consultar con el profesor la evolución de la tarea.
- **La documentación:** se realizará en formato pptx o bien en formato Word, siempre se hablará de un contexto de estudio, la metodología empleada, herramientas utilizadas, el análisis de los resultados obtenidos y las conclusiones. Será muy importante aportar referencias relativas al estado del arte acerca de la metodología empleada.

Es importante dedicarle el tiempo necesario a cada uno de los ejercicios, ya que esta tarea, en su totalidad, es puntuable para la nota final de la materia. (Peso: 10 %).