

RECONOCIMIENTO DE TEXTO Y ANÁLISIS DE SENTIMIENTOS

Proyecto final computación blanda

Juan Camilo Olarte Betancourt - 1004734205
Carlos Daniel Morales Isaza - 1004680247

Librerías utilizadas

Textblob

Transformers

Textblob

Esta librería está enfocada en el análisis de sentimientos a textos, por medio de redes neuronales.

Textblob

El análisis de sentimientos puede ayudarnos a descifrar el estado de ánimo y las emociones del público en general y recopilar información detallada sobre el contexto. El análisis de sentimiento es un proceso de análisis de datos y clasificación en función de la necesidad de la investigación.

Textblob

El enfoque que aplica el paquete TextBlob al análisis de opiniones difiere en que se basa en reglas y, por lo tanto, requiere un conjunto predefinido de palabras categorizadas. Estas palabras pueden, por ejemplo, cargarse desde la base de datos NLTK. Además, los sentimientos se definen en función de las relaciones semánticas y la frecuencia de cada palabra en una oración de entrada que permite obtener como resultado una salida más precisa.

Textblob

Una vez que se logra el primer paso y un modelo de Python se alimenta con los datos de entrada necesarios, un usuario puede obtener los puntajes de sentimiento en forma de polaridad y subjetividad

Textblob

Salida de TextBlob para un **polaridad** la tarea es un flotador dentro del rango $[-1.0, 1.0]$ donde -1.0 es una polaridad negativa y 1.0 es positivo. Esta puntuación también puede ser igual a 0 , que representa una evaluación neutral de una declaración, ya que no contiene ninguna palabra del conjunto de entrenamiento.

Transformers

Transformers (anteriormente conocido como pytorch-transformers y pytorch-pretrained-bert) proporciona miles de modelos previamente entrenados para realizar tareas en diferentes modalidades como texto, visión y audio.

El transformador se refiere a un modelo de aprendizaje profundo, que utiliza el mecanismo de atención utilizado en el procesamiento del lenguaje natural (PNL).

Transformers



Respuesta a preguntas extractivas

La Respuesta Extractiva a Preguntas es la tarea de extraer una respuesta de un texto dado una pregunta. Un conjunto de datos que se usan en el proyecto de respuesta a preguntas es el conjunto de datos SQuAD, que se basa completamente en esa tarea.

Clasificación de texto

Reconocimiento de entidades

El reconocimiento de entidad con nombre (NER) es la tarea de clasificar tokens según una clase, por ejemplo, identificar un token como una persona, una organización o una ubicación. Un ejemplo de un conjunto de datos de reconocimiento de entidad con nombre es el conjunto de datos CoNLL-2003

Clasificación de texto

Reconocimiento de entidades

Se trata de identificar los tokens como pertenecientes a una de las 9 clases:

- O, fuera de una entidad nombrada
- B-MIS, comienzo de una entidad miscelánea inmediatamente después de otra entidad miscelánea
- I-MIS, entidad miscelánea
- B-PER, comienzo del nombre de una persona justo después del nombre de otra persona
- I-PER, nombre de la persona
- B-ORG, comienzo de una organización inmediatamente después de otra organización
- I-ORG, Organización
- B-LOC, comienzo de una ubicación inmediatamente después de otra ubicación
- I-LOC, ubicación

Muchas Gracias