

Proyecto Integrador 1

En esta actividad se explora los principios del procesamiento de lenguaje natural desarrollando un analizador de sentimientos.

La herramienta que utilizarán para su proyecto es *Natural Language Toolkit*, la cual es una librería de Python que incorpora muchas funcionalidades estadísticas para procesar texto.

Pasos a seguir:

1. Familiarizarse con la librería NLTK, mi recomendación es revisar los primeros capítulos del libro *Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit* escrito por Steven Bird:

<https://www.nltk.org/book/>

2. Revisar la documentación de NLTK para complementar su aprendizaje de la librería:

<https://www.nltk.org/>

3. Seguir el siguiente tutorial para hacer un analizador de sentimientos de Twitter:

<https://www.youtube.com/playlist?list=PLQVvva0QuDf2JswnfGkIbInZnIC4HL>

4. Hacer una recolección de datos para generar el set de entrenamiento. (e.g. encontrar una base de datos con tweets ya catalogados como positivos o negativos).

5. Opcionalmente está el curso de NLP de la universidad de Stanford para quienes tengan interés de seguir aprendiendo del tema:

https://www.youtube.com/playlist?list=PLLssT5z_DsK8HbD2sPcUIDfQ7zmBarMYv

Entregable:

El entregable base es el código generado por el programa de Sentdex, en el cual se clasifican los tweets como positivos y negativos. La entrega base amerita un 80 de calificación, los equipos requieren innovar para entregar un proyecto final con funcionalidades adicionales.

Posibles formas de mejorar el analizador de sentimientos:

- Incluir más redes sociales además de Twitter, e.g. clasificar comentarios de Youtube, Instagram o Facebook.
- Incorporar modelos más avanzados para la clasificación de texto, e.g. diseñar una red neuronal recurrente con memoria a corto y largo plazo como el modelo clasificador.
- Incorporar la funcionalidad de transcribir audio a texto y que se haga el análisis de sentimiento.
- Realizar el analizador de sentimiento para diferentes idiomas.

Estas son sólo algunas ideas para mejorar el analizador de sentimientos, el equipo puede proponer ideas nuevas para incrementar su calificación.

Fecha de entrega: 02 de Abril