



UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS

CARRERA DE INGENIERÍA EN SISTEMAS

Asignatura Inteligencia Artificial

Proyecto N°2: Métodos para el Mapeo Palabras a Vectores (Word Embedding)

Descripción

El word embedding es una técnica de Procesamiento de Lenguaje Natural que permite el mapeo de las palabras de un texto a vectores con el fin de poder usarlas para entrenar modelos de IA como las redes neuronales. En este proyecto se desea que el estudiante implemente tres de los métodos más utilizados para realizar el Word Embedding, que son: Bag of Words, TD-IDF y Word2Vec. La implementación de estos métodos se sugiere se base en el siguiente artículo:

<https://stackabuse.com/implementing-word2vec-with-gensim-library-in-python/>

dicha implementación deberá hacerse escribiendo los métodos de la clase definida en el archivo word_embeddings.py proporcionado junto a este documento.

Requisitos de Presentación

- El ejercicio se desarrollará de forma grupal.
- La entrega se realizará por uno sólo de los miembros del grupo, utilizando el enlace habilitado en el campus virtual, la fecha de entrega está definida ahí mismo.
- Para cada uno de los entregables deberá enviar únicamente un directorio comprimido que llevará por nombre:

Grupo n

Donde n es reemplazado por el número de grupo al que pertenece, por ejemplo Grupo 3. Asegúrese de poner este nombre tanto al directorio (antes de comprimirlo) como al archivo comprimido, de modo que al descomprimirlo siga manteniendo el nombre Grupo n.

Incluya dentro del directorio un fichero llamado "integrantes.txt", donde se listen sólo los miembros del grupo que colaboraron, incluyendo sus números de cuenta.

- No está permitido cambiar el nombre o el número de parámetros a las clases y métodos proporcionados. En caso de hacerlo el software evaluador no podrá asignarle los puntos correspondientes.
- Sí es permitido añadir atributos y métodos privados a las clases proporcionadas, así como importar módulos estándar del lenguaje (los mismos que se usan en el laboratorio).
- Para poder comprobar la correctitud de su trabajo, podrá ejecutar el siguiente comando, estando ubicado en el directorio de su proyecto:

```
python3 autograder.py
```

Dicho comando intentará importar el módulo principal de su trabajo y validará la correctitud del mismo. Es importante recalcar que aunque este programa le proporcione una nota, la misma estará sujeta a revisión por parte del profesor.

- El plagio se penalizará con una nota de 0.
- No cumplir con alguno de los puntos anteriores implicará una disminución en su nota, favor si alguno no está claro, consultar en clase.

Elaborado por: Raúl José Palma Mendoza.