

Laboratorio 4



Autores:

Kristopher Javier Alvarado López | Carné 21188
Mario Antonio Guerra Morales | Carné 21008

Curso:

Deep Learning y Sistemas Inteligentes

Catedrático:

Luis Alberto Suriano Saravia

Sección:

10

Universidad del Valle de Guatemala
11 calle 15-79 Zona 15 Vista Hermosa III
Guatemala, C. A.
Facultad de Ingeniería

Laboratorio 4 - Deep Learning

Enlace a repositorio: https://github.com/MarioGuerra21008/Deep_Learning-Lab04

Parte 1

- **¿Cuál es la implicación del overfitting en modelos como Word2Vec?**

La implicación del overfitting en Word2Vec indica que, el modelo aprendió muy bien los datos de entrenamiento, por lo que, en datos de prueba o no vistos, genera un poco de pérdida de generalización.

- **¿Qué tan bien encontró palabras cercanas a su modelo Word2Vec? ¿Podría mejorar? ¿Cómo podría mejorar?**

En base a las palabras cercanas, se podría decir que las encontró relativamente bien, pero con ciertos inconvenientes en palabras como por ejemplo, "América" en lo malo. Podría mejorar utilizándose más datos o por medio de distintos hiperparámetros a considerar.

- **A grandes rasgos, ¿cuál es la diferencia entre Word2Vec y BERT?**

Word2Vec se basa en el contexto local de las palabras, al considerarse como un vector en un espacio de dimensión fija. Mientras que BERT, generalmente considera a las palabras en base al contexto dependiendo de las palabras que están antes o después de ella, con tal de obtener el contexto completo.