

**Trabajo Practico Final de Modelos Grandes de Lenguaje**

*“Comparación de modelo LLM contra Basado en Bag of Words en análisis de sentimiento”*

*Carlos Montiel*

*CEIA – 2024*

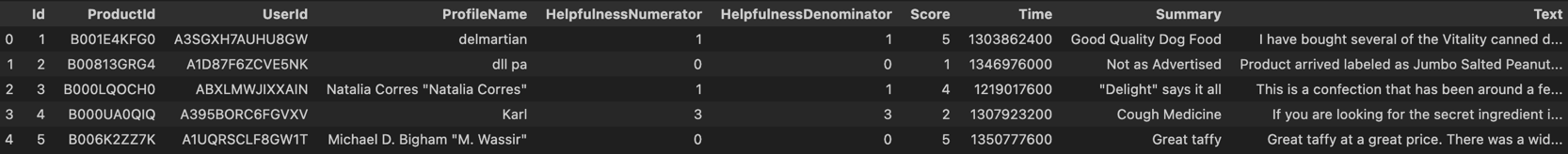
*Introducción.*

Estaremos verificando dos implementaciones de modelos basados en técnicas de inteligencia artificial para la inferencia de análisis de sentimiento sobre los reviews de compradores de amazon.

Por un lado constamos con un modelo basado en *“Bag of words”* llamado VADER (Valence Aware Dictionary and sEntiment Reasoner) y por el otro lado roBERTa (A Robustly Optimized BERT Pretraining Approach) basado en el modelo de Google BERT del 2018.

*Desarrollo.*

Analisamos primeramente el dataset el cual pesa unos 300mb y cuenta con 568.455 reviews de compradores de amazon.



Analizamos la distribución del dataset en función de las estrellas registradas para cada review.



Luego de este pequeño análisis del dataset implementamos el *`SentimentIntensityAnalyzer`* de VADER y también roBERTa.

Utilizamos una frase de ejemplo para la comparación de las inferencias de los modelos

***“This oatmeal is not good. Its mushy, soft, I don't like it. Quaker Oats is the way to go.”***

Con el cual registramos es siguiente resultado con VADER:

***{'neg': 0.22, 'neu': 0.78, 'pos': 0.0, 'compound': -0.5448}***

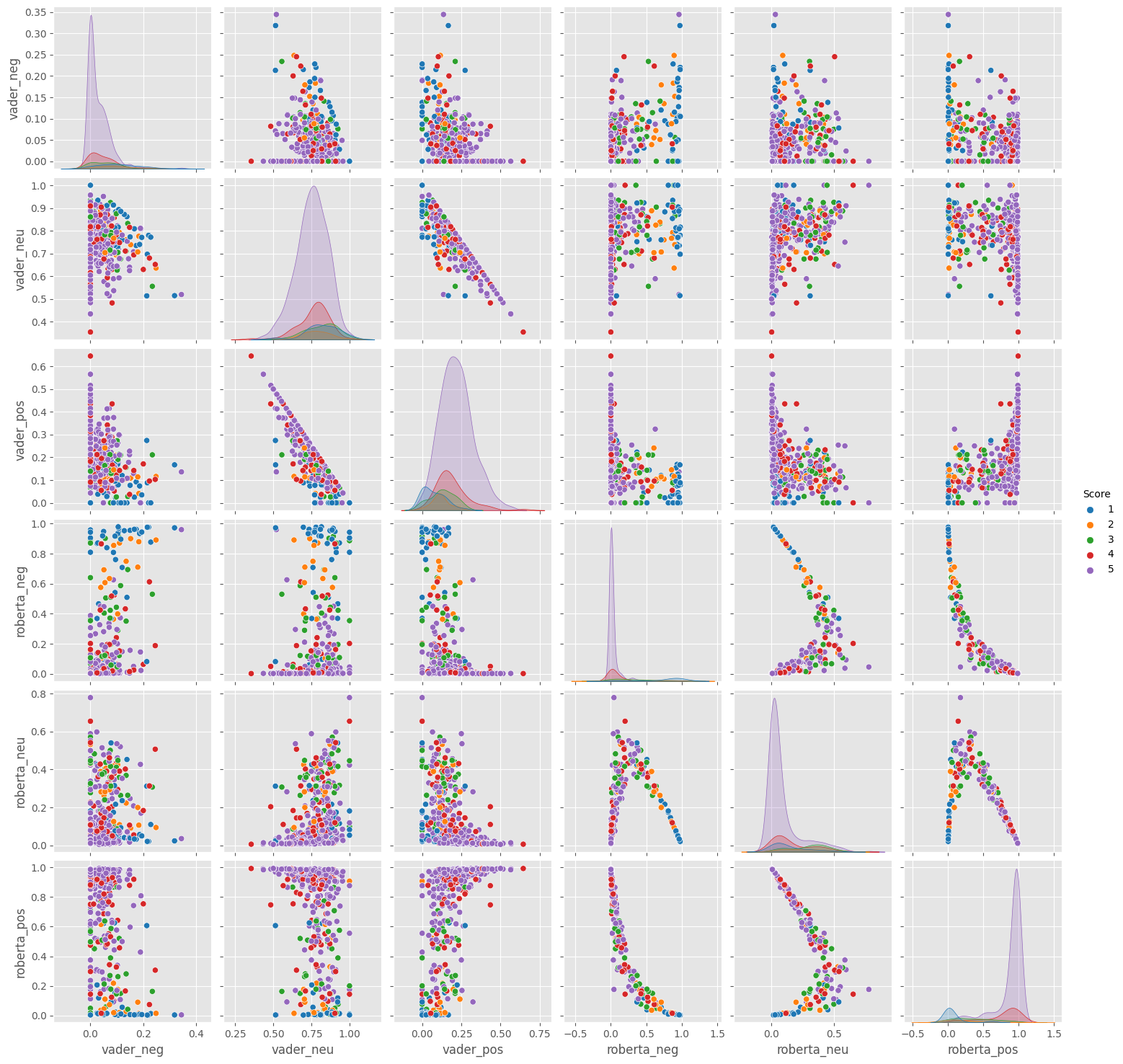
Y el siguiente resultado con roBERTa

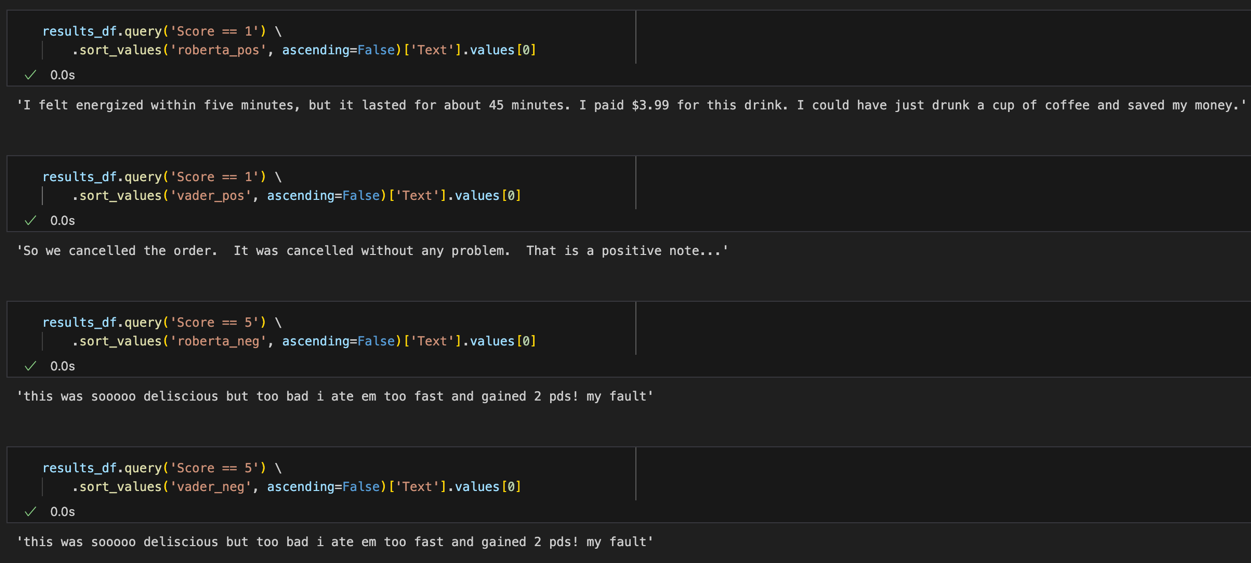
***{'roberta\_neg': 0.97635514, 'roberta\_neu': 0.020687465, 'roberta\_pos': 0.0029573692}***

Aquí notamos diferencias en las inferencias de los diferentes modelos dando a roBERTa un scoring muy negativo a la frase de prueba y en contra partida VADER da un scoring neutro.

*Resultados.*

Realizamos un pairplot de los resultados obtenidos de VADER contra roBERTa, obteniendo los siguientes graficos.



También encontramos en los dos modelos casos de falsos positivos y falsos negativos

*Conclusiones.*

Comparando los modelos propuestos llegamos a la conclusión con este dataset que los dos infieren relativamente bien los reviews de los productos salvo ciertos casos que obtenemos falsos positivos y falsos negativos con los dos modelos.

Sería interesante comparar más modelos de análisis de sentimiento como la librería Transformers la cual es relativamente fácil de utilizar y no requiere de demasiadas configuraciones.