**ASSESSMENT FOR DATA SCIENCE**

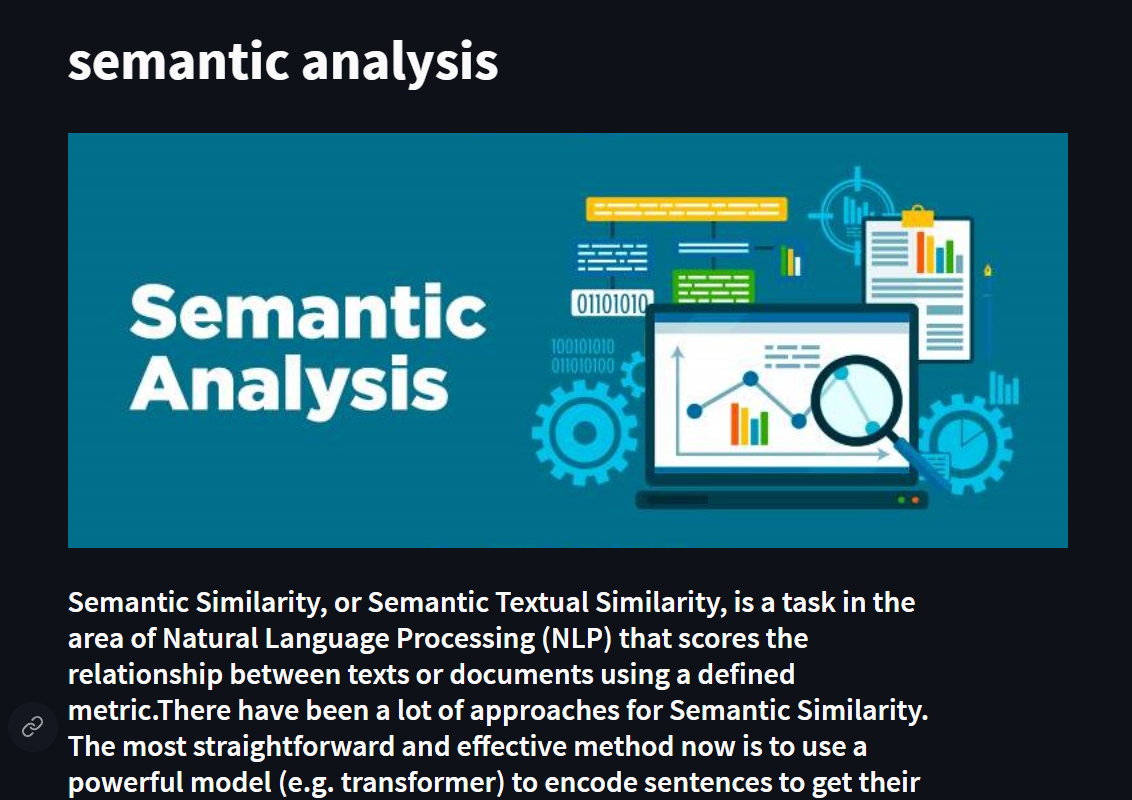
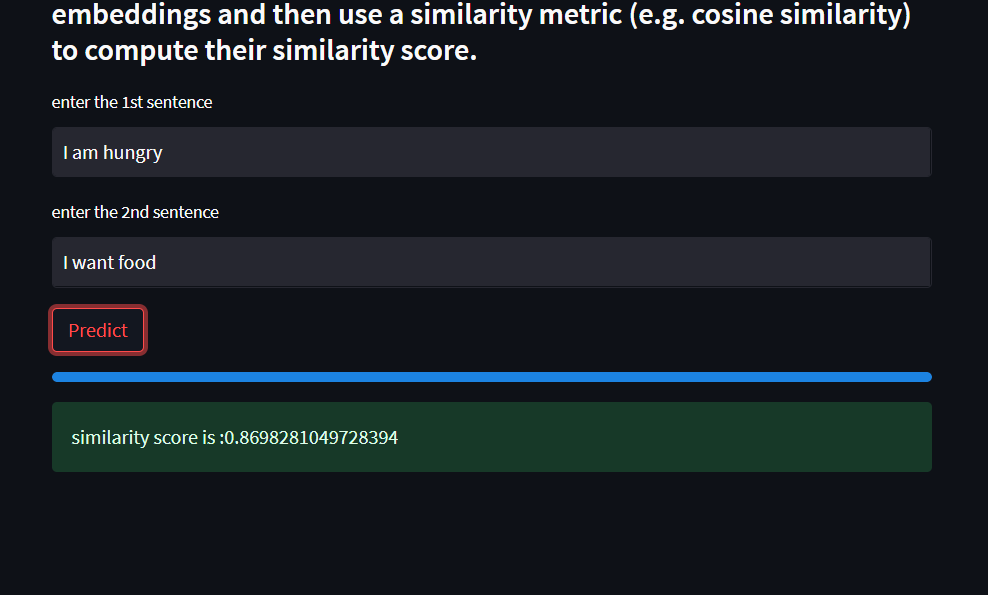
First of all, thank you so much for giving me this opportunity.

Part A:

* So, in this part I load the given model in data frame
* Then imported necessary libraries like transformer for downloading the pre-trained model.
* Here I downloaded ‘stsb-mpnet-base-v2’ model.
* The size of the model is 500mb
* This is a sentence-transformers model. It maps sentences & paragraphs to a 768-dimensional dense vector space and can be used for tasks like clustering or semantic search.
* Then I encode the sentences to get their embeddings
* After that I compute similarity scores of given two embeddings
* Similarity score lies between 0 and 1
* Closer to 0 means less similarity and closer to 1 means more similarity

Part B:

* I made a web application using streamlit.
* As we know our model size is large so its takes time to load
* So, I made app.py file
* And my web application running fine in my local system
* But I am not able to deploy my web application on cloud platform.
* The reason is my model size is more than 500 mb so I can’t deploy on Heroku as it only deploys file size up to 500.
* I also tried to deploy on streamlit.io but unfortunately, I am not able to deploy because of some unknown error
* you can see my web application on local host



If there any improvement needed then please guide me. I am open to learn new things.

Thank you,

Sahil shaikh