NAME : THIRSHA SREE H

REG NO: 20BCE2518

DEGREE: B.TECH COMPUTER SCIENCE AND ENGINEERING

GITHUB : https://github.com/THIRSHASREE5803/20BCE2518_THIRSHA-SREE-H_NLP_TASK-2_FARMWISEAI

PLATFORM USED GOOGLE COLLAB

NLP TASK 2

Task Description: Building a Topic-based Post Recommendation System using Generative AI

LIBRARIES:

- pandas,
- langchain,
- openai,
- openaiembeddings,
- numpy,
- faiss

1. The contextual embeddings generated by the model differ from traditional word embeddings in that they take into account the surrounding context of the word in a given sentence.
2. This dynamic nature of the embeddings allows the model to capture nuances and context-specific meanings.
3. Unlike fixed vectors assigned to each word in traditional word embeddings, the embeddings generated by the model change based on the position of the word in the context.
4. An embedding is represented as a vector, which is essentially a list of floating point numbers.
5. The relatedness between two vectors can be measured by calculating the distance between them.
6. Smaller distances indicate higher relatedness, while larger distances suggest lower relatedness.
7. FAISS, which stands for Facebook AI Similarity Search, is a library that enables developers to efficiently search for embeddings of multimedia documents that are similar to each other.
8. It overcomes the limitations of traditional query search engines that are optimized for hash-based searches and provides more scalable similarity search functions.
9. To facilitate fast search, Faiss creates an index structure that organizes the vectors. It utilizes various distance metrics, such as Euclidean distance, cosine similarity, and inner product, to calculate the similarity or distance between vectors.
10. When given a query vector, Faiss efficiently searches for the nearest neighbors based on the defined distance metric.