

Your grade: 100%

Your latest: 100% • Your highest: 100% • To pass you need at least 66%. We keep your highest score.

Next item →

1. Facebook AI Similarity Search, also known as Faiss, is essentially a tool to:

1 / 1 point

- ☐ Design training and inference for machine learning models
- ☒ Calculate the distance between the question embedding and the vector database of context vector embeddings
- ☐ Take care of tasks related to data management and preprocessing
- ☐ Address high-level data retrieval and operations like SQL querying

✔ Correct

Faiss is a library developed by Facebook AI Research that offers efficient algorithms for searching through large collections of high-dimensional vectors by calculating the distance between the question embedding and the vector database of context vector embeddings.

2. What is the primary function of a contextual tokenizer with respect to tokenizers in natural language processing (NLP)?

1 / 1 point

- ☐ To translate text from one language to another
- ☐ To split text into individual characters
- ☒ To break text into numeric representations or contextual embeddings
- ☐ To summarize long paragraphs into concise sentences

✔ Correct

A context tokenizer not only splits text into tokens but also takes into account the surrounding context. It generates meaningful and accurate text representations as the context provides additional information.

3. How do the inserted prompts typically get encoded before being used to retrieve relevant documents using the retrieval-augmented generation (RAG) process?

1 / 1 point

- ☐ You can convert the entire prompt into a single unique token.
- ☐ You can use one-hot encoding to represent each word as a unique vector.
- ☒ You can apply token embedding and vector averaging to convert the prompts into a vector representation.
- ☐ You can use bag-of-words to create a frequency matrix of the words in the prompts.

✔ Correct

In the RAG process, the inserted prompts are encoded using token embedding and vector averaging. This helps capture the semantic meaning of the prompts and convert them into a vector representation that can be effectively used for retrieving relevant documents.