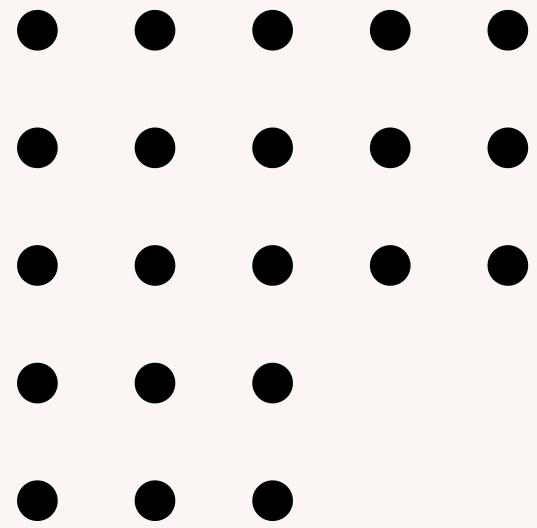
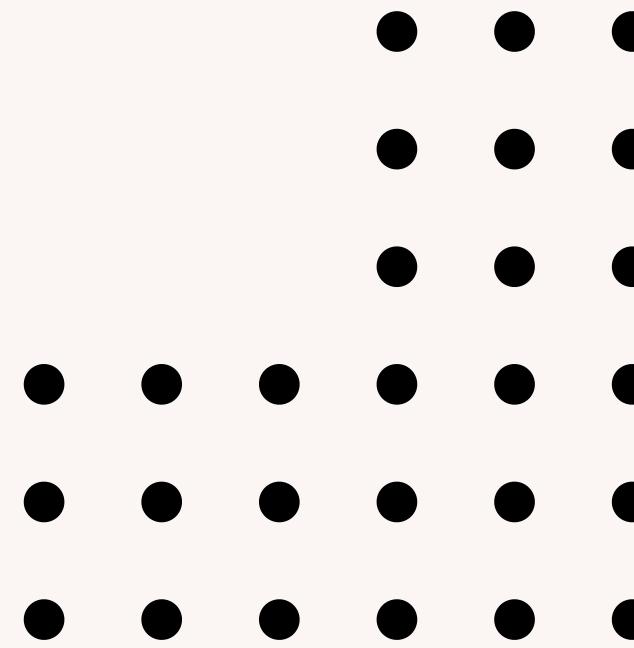


ATIVIDADE AF11

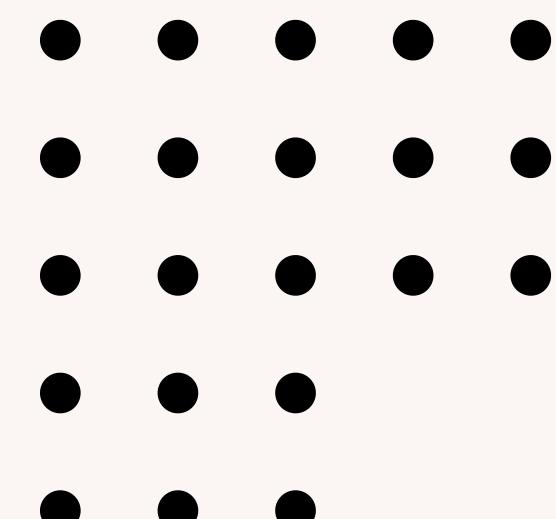
Higor rocha, Julio cesar, Matheus Paiva, Mateo Fabio, Gabriel Cabreira,
Nicolas Medina





Problema

Dificuldade na localização de arquivos,
dentro de uma grande base de dados



Solução

HuggingFace (para gerar embeddings)

- Transformamos nomes de arquivos em vetores numéricos que representam o significado do texto.

FAISS (Vector Database)

- Criamos um índice vetorial para fazer busca rápida por similaridade.

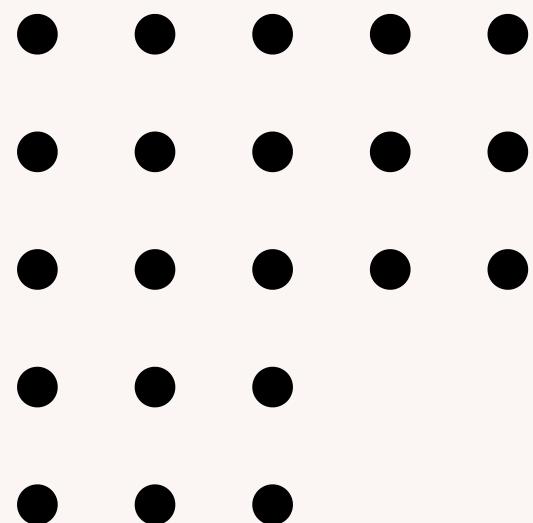
O usuário pode:

digitar o que procura

o modelo gera embeddings

o FAISS encontra o arquivo mais semântico

o sistema pergunta se ele quer continuar



```
atividadeIA.py > 
1 # Instalar:
2 # pip install transformers sentence-transformers faiss-cpu
3
4 from transformers import AutoTokenizer, AutoModel
5 import torch
6 import faiss
7 import numpy as np
8
9 # 1. Carregar modelo HuggingFace
10 model_name = "sentence-transformers/all-MiniLM-L6-v2"
11
12 tokenizer = AutoTokenizer.from_pretrained(model_name)
13 model = AutoModel.from_pretrained(model_name)
14
15 # Função para gerar embeddings
16 def encode(texts):
17     tokens = tokenizer(texts, padding=True, truncation=True, return_tensors="pt")
18     with torch.no_grad():
19         output = model(**tokens)
20     embeddings = output.last_hidden_state.mean(dim=1)
21     return embeddings.numpy()
22
23 # 2. Base de documentos
24 docs = [
25     "trabalho_vector_database.py",
26     "projeto_busca_semantica_vector_db.py",
27     "trabalho_huggingface_faiss.py",
28     "atividade_pratica_vector_db.py",
29     "implementacao_vector_database.py",
]
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

PS C:\Users\mathe\OneDrive\Documentos\ATIVIDADE-IA> & C:/Users/mathe/AppData/Local/Programs/Python/Python311/python.exe c:/Users/mathe/OneDrive/Documentos/ATIVIDADE-IA/atividadeIA.py

Digite sua pergunta: trabalho_IA

Pergunta: trabalho_IA
Documento mais parecido: trabalho_huggingface_faiss.py
Distância: 25.490952

Deseja fazer outra pergunta? (s/n): 
```