fitz
PyMuPDF
openai
sklearn
sentence_transformers
numpy