# 6-Week Learning Roadmap for RAG Pipelines

## Week 1 – Fundamentals

* Review Python basics (APIs, file I/O, JSON handling).
* Learn how LLMs work at a high level (prompts, context length, tokens).
* Read about embeddings and vector similarity.
* Resource: OpenAI Embeddings Guide (https://platform.openai.com/docs/guides/embeddings).
* Mini-Task: Call an embedding API for a few sentences and compute cosine similarity in Python.

## Week 2 – Text Processing & Chunking

* Learn text preprocessing (cleaning, splitting, chunking).
* Study why chunk size matters in retrieval.
* Practice with nltk, spaCy, or langchain.text\_splitter.
* Mini-Task: Load a PDF/Word doc → chunk into passages → store in a simple list.

## Week 3 – Vector Databases

* Learn about FAISS (local) and Pinecone/Weaviate/Chroma (cloud).
* Practice inserting embeddings into a vector DB.
* Learn how to perform similarity search (top-k retrieval).
* Mini-Task: Build a mini 'search engine' → query returns top 3 relevant chunks.

## Week 4 – Retrieval + Generation

* Learn how to pass retrieved chunks + query to an LLM (OpenAI, Anthropic, or HuggingFace).
* Study LangChain’s RetrievalQA pipeline.
* Explore LlamaIndex for connecting multiple data sources.
* Mini-Project: Q&A over PDFs – Upload a PDF → store chunks in FAISS → retrieve → feed into GPT → get contextual answers.

## Week 5 – Improving RAG

* Learn prompt engineering for RAG (e.g., 'You are a helpful assistant. Use ONLY the provided context to answer').
* Add filters: metadata search, hybrid search (keyword + vector).
* Learn evaluation techniques (RAGAS, precision/recall, hallucination tests).
* Mini-Project: Chat with Website Docs – Scrape website text, build a chatbot with retrieval.

## Week 6 – Scaling & Deployment

* Learn how to store data efficiently for large corpora.
* Add caching (Redis, LlamaIndex cache).
* Deploy a simple RAG pipeline as a Streamlit or FastAPI app.
* Explore advanced optimizations (rerankers, fine-tuning retrievers).
* Capstone Project: Internal Knowledge Base Chatbot – Load docs from multiple sources (PDF + CSV + website), deploy chatbot with RAG.

## After 6 Weeks

- A strong understanding of RAG internals.  
- Multiple working mini-projects.  
- A deployable chatbot or knowledge assistant you can showcase in interviews.