

# ■ Generative AI Engineer – 2-Month Roadmap & Weekly Plan (with Resources + Capstone Guide)

## ■ Roadmap Overview

- Week 1: Python for AI + Data Processing → Mini Project: Parse PDFs/Word/Excel and store in DB.
- Week 2: Prompt Engineering + Generative AI Basics → Mini Project: Chatbot using OpenAI API.
- Week 3–4: RAG + LangChain + HuggingFace → Mini Project: Q&A; chatbot over PDFs.
- Week 5: Azure AI & Databricks → Mini Project: Deploy RAG chatbot on Azure.
- Week 6: Integration + Deployment → Mini Project: Dockerize chatbot + CI/CD pipeline.
- Week 7–8: Capstone Projects → Showcase-ready apps (Enterprise Document Assistant, Generative Insights API).

## ■ Weekly Learning Plan with Resources

### Week 1

- Python refresh: list/dict comprehensions, OOP, NumPy, Pandas.
- Learn parsing: PyPDF2, python-docx, openpyxl.
- Mini Project: Extract text from PDFs and load into SQLite/DB.

#### ■ Resources:

- Python Crash Course: <https://docs.python.org/3/tutorial/>
- NumPy: <https://numpy.org/doc/stable/>
- Pandas: <https://pandas.pydata.org/docs/>
- PyPDF2: <https://pypi.org/project/PyPDF2/>
- python-docx: <https://python-docx.readthedocs.io/>
- openpyxl: <https://openpyxl.readthedocs.io/>

### Week 2

- Prompt engineering basics: zero-shot, few-shot, chain-of-thought.
- Explore OpenAI API + simple chatbot.
- Mini Project: Prompt-based chatbot with templates.

#### ■ Resources:

- OpenAI Prompt Engineering Guide: <https://platform.openai.com/docs/guides/prompt-engineering>
- OpenAI API Docs: <https://platform.openai.com/docs/>
- Example Chatbot Repo: <https://github.com/openai/openai-quickstart-python>

### Week 3

- LangChain: chains, agents, memory.
- Vector databases (FAISS/Chroma).
- Mini Project: Basic RAG pipeline.

#### ■ Resources:

- LangChain Docs: <https://python.langchain.com/docs/>
- FAISS: <https://github.com/facebookresearch/faiss>
- ChromaDB: <https://docs.trychroma.com/>

### Week 4

- HuggingFace: embeddings, LLMs (LLaMA, Falcon, Mistral).
- Combine LangChain + HuggingFace for RAG.
- Mini Project: Q&A; chatbot over PDFs.

#### ■ Resources:

- HuggingFace Transformers: <https://huggingface.co/docs/transformers/index>
- HuggingFace Datasets: <https://huggingface.co/docs/datasets/>
- LangChain + HuggingFace Integration: [https://python.langchain.com/docs/integrations/llms/huggingface\\_hub](https://python.langchain.com/docs/integrations/llms/huggingface_hub)

### Week 5

- Azure OpenAI & Cognitive Search.

- Databricks basics: ETL pipelines, ML training.
- Mini Project: Deploy RAG chatbot on Azure.

#### ■ Resources:

- Azure OpenAI: <https://learn.microsoft.com/en-us/azure/cognitive-services/openai/>
- Azure Cognitive Search: <https://learn.microsoft.com/en-us/azure/search/>
- Databricks Getting Started: <https://learn.microsoft.com/en-us/azure/databricks/getting-started/>

## Week 6

- Docker basics + containerizing your chatbot.
- CI/CD: GitHub Actions or Azure DevOps.
- Mini Project: End-to-end chatbot deployment with web UI (Streamlit/Flask).

#### ■ Resources:

- Docker Docs: <https://docs.docker.com/get-started/>
- GitHub Actions: <https://docs.github.com/en/actions>
- Azure DevOps Pipelines: <https://learn.microsoft.com/en-us/azure/devops/pipelines/>
- Streamlit: <https://docs.streamlit.io/>
- Flask: <https://flask.palletsprojects.com/>

## Week 7–8

- Capstone 1: Enterprise Document Assistant (multi-file knowledge base).
- Capstone 2: Generative Insights API (summarization + insights).
- Polish GitHub repo + write LinkedIn posts about projects.
- Prepare for Capstone Projects (detailed below).

#### ■ Resources:

- LangChain End-to-End Examples: [https://python.langchain.com/docs/use\\_cases/](https://python.langchain.com/docs/use_cases/)
- Awesome Generative AI Projects: <https://github.com/steven2358/awesome-generative-ai>

## ■ Capstone Project Guide

### Capstone 1: Enterprise Document Assistant

#### ■ Features:

- Upload multiple documents (PDF, Word, Excel).
- Perform semantic search and Q&A; over knowledge base.
- Support context-aware conversations with memory.
- Provide citations/sources for answers.

#### ■ Tech Stack:

- LangChain for orchestration.
- Vector DB (FAISS/Chroma/Azure Cognitive Search).
- HuggingFace or Azure OpenAI for embeddings + LLM.
- Flask/Streamlit for web UI.
- Docker + Azure for deployment.

#### ■ Expected Outcomes:

- Functional enterprise-ready assistant for internal knowledge bases.
- Deployed demo on Azure with simple UI.
- GitHub repo with clean documentation + LinkedIn showcase post.

### Capstone 2: Generative Insights API

#### ■ Features:

- Accept structured + unstructured input (CSV, PDF, SQL DB).
- Summarize, generate insights, and answer business questions.
- Provide both text and visual output (charts/graphs).
- REST API endpoints for integration.

#### ■ Tech Stack:

- FastAPI/Flask for REST API.
- LangChain + HuggingFace/Azure OpenAI for LLM integration.
- Pandas/Databricks for data wrangling.
- Matplotlib/Plotly for visualization.
- Docker + CI/CD pipeline for deployment.

■ Expected Outcomes:

- Deployed API service capable of handling enterprise queries.
- Integration-ready system for downstream applications.
- Showcase GitHub repo + LinkedIn post describing architecture.