Roadmap for Learning and Developing AI Projects

Month 1: Building the Foundation (Weeks 1-4)

Week 1- Session 1 & 2: Python Basics

- Introduction to Python syntax, data types, and variables.
- Control structures (if-else, loops) and basic functions.
- Practice Exercises

Week 1- Session 3 & 4: Intermediate Python

- Working with file I/O, error handling, and exceptions.
- Functions, arguments, and return values.
- Modules and Packages
- Practice Exercises

Week 2 - Session 5: Object-Oriented Programming (OOP)

- Understanding classes, objects, and methods.
- Inheritance, polymorphism, and encapsulation in Python.
- Building reusable and modular code using OOP principles.
- Project: A CMD-based Project

Week 2- Session 6: Introduction to Numpy

- Working with arrays, vectors, and matrices in Numpy.
- Vectorized operations and broadcasting.
- Practice Exercise

Week 3- Session 7: Introduction to Pandas

- Data manipulation using DataFrames and Series.
- Importing/exporting data, handling missing values, and data slicing.
- Aggregating, filtering, and transforming data efficiently.
- Practice Exercise

Week 4- Session 8 & 9: Databases (SQL & NoSQL)

- Fundamentals of SQL queries: SELECT, JOIN, WHERE, and GROUP BY.
- Introduction to NoSQL databases (MongoDB) and their use cases.
- Understanding data modeling for structured vs unstructured data.
- Project: Hospital Management System with Python

Week 4- Session 10 & 11: Feature Engineering I

- Detecting and handling outliers and missing values.
- Scaling numerical data (normalization, standardization).

- Identifying important features for machine learning models.

Week 4- Session 12: Feature Engineering II

- Categorical encoding techniques (one-hot, label encoding).
- Feature combination and extraction methods.
- Using domain knowledge to improve feature selection.

Month 2: Machine Learning & First Project (Weeks 5-8)

Week 5- Session 13: Introduction to Machine Learning

- Overview of machine learning types (supervised vs unsupervised).
- Common algorithms: linear regression, decision trees, and SVM.
- Model evaluation metrics (accuracy, precision, recall, F1-score).

Week 5- Session 14 & 15: Basic ML Project I

- Data collection, preprocessing, and exploratory data analysis.
- Training and testing an ML model using Scikit-learn.
- Building a simple web interface for the project using Streamlit.

Week 6- Session 16: Basic ML Project II

- Model explainability using SHAP and LIME.
- Improving the model: hyperparameter tuning and cross-validation.
- Project deployment with Streamlit

Week 7- Session 17: Introduction to Deep Learning

- Basics of neural networks: perceptrons, activation functions, and layers.
- Key differences between deep learning and traditional ML.
- Forward Propagation
- How deep learning models learn through backpropagation.

Week 7- Session 18: Deep Learning Architectures

- Introduction to CNNs, RNNs, autoencoders, and transformers.
- Theoretical understanding of how these architectures work.
- A bit of mathematics behind neural network learning.

Week 8- Session 19: PyTorch for Neural Networks

- Introduction to PyTorch
- Encoding a Feature Vector for PyTorch Deep Learning
- Early Stopping and Network Persistence
- Sequences vs Classes in PyTorch

Week 8- Session 20: Training on Tabular Data with PyTorch

- Using K-Fold Cross-validation with PyTorch
- Training Schedules for PyTorch
- Dropout Regularization
- Batch Normalization
- RAPIDS for Tabular Data

Month 3: Advanced AI Concepts & Second Project (Weeks 9-12)

Week 9- Session 21: CNNs & Computer Vision

- Image Processing in Python
- Using Convolutional Neural Networks
- Using Pretrained Neural Networks
- Looking at Generators and Image Augmentation
- Recognizing Multiple Images with YOLO

Week 9- Session 22: ChatGPT & Large Language Models

- Introduction to Transformers
- Accessing the ChatGPT API
- LLM Memory
- Introduction to Embeddings
- Prompt Engineering

Week 10- Session 23: Image Generative Models

- Introduction to Generative AI
- Generating Faces with StyleGAN3
- GANS to Enhance Old Photographs Deoldify
- Text to Images with StableDiffusion
- Finetuning with Dreambooth

Week 10- Session 24: Time Series in PyTorch

- Time Series Data Encoding for Deep Learning, PyTorch
- Seasonality and Trend
- LSTM-Based Time Series with PyTorch
- CNN-Based Time Series with PyTorch
- Predicting with Meta Prophet

Week 11- Session 25: Natural Language Processing (NLP)

- Introduction to Natural Language Processing

- Fundamentals of text processing: tokenization, stopword removal, and stemming.
- Hugging Face Introduction
- Hugging Face Tokenizers
- Hugging Face Data Sets
- Training a Model in Hugging Face

Week 12- Session 26: Deployment & Monitoring

- Deploying models to production environments using platforms like AWS and Heroku.
- Using Denoising AutoEncoders
- Anomaly Detection
- Model Drift and Retraining

Projects

Project 1: Simplifying PDF Information Retrieval using RAG

Skills and Technologies You'll Learn:

- Natural Language Processing (NLP)
- Python Programming
- Web Development with Streamlit
- OpenAI API Integration
- Database Management
- Cloud Services Integration
- Embeddings and Vector Databases
- Application Deployment

Project Structure:

```
├── notebooks/
│   ├── 01_preprocessing.ipynb
│   ├── 02_intent_detection.ipynb
│   ├── 03_information_retrieval.ipynb
│   ├── 04_response_generation.ipynb
│   ├── main.ipynb
│
├── README.md
├── requirements.txt
├── streamlit_app.py
```

By working on **this project**, you'll gain hands-on experience with:

Libraries and Frameworks:

- Streamlit
- OpenAI API
- SQLAlchemy

Tools:

- PostgreSQL
- Supabase
- pgvector
- Google Drive API

Concepts:

- Retrieval Augmented Generation (RAG)
- Embeddings
- Natural Language Understanding
- Content Moderation

—----------

Project 1: ResumeMatcher

Skills and Technologies You'll Learn:

- Web Scraping and Data Collection
- Database Management with MongoDB
- Optical Character Recognition (OCR)
- Natural Language Processing (NLP)
- Web Development with Streamlit
- Machine Learning
- Version Control with Git
- Deployment

Project Structure:

```
├── Dockerfile
├── main.py
├── app/
│   ├── __init__.py
│   ├── config.py
│   ├── courses.py
│   ├── exceptions.py
```

```
│   ├── text_preprocessing.py
│   ├── database/
│   │   ├── __init__.py
├── README.md
├── requirements.txt
├── setup.py
├── data/
│   ├── processed/
│   └── raw/
├── notebooks/
│   ├── data_collection.ipynb
│   ├── main_notebook.ipynb
│   ├── Working_with_OCR.ipynb
│   └── text_preprocessing.ipynb
├── pages/
│   ├── candidate_page.py
│   ├── resume_analyzer.py
│   └── recruiter_page.py
```

By working on this project, you'll gain experience with:

Tools and Libraries:

- Apify, MongoDB, PyTesseract, NLTK, spaCy, Streamlit, Scikit-learn, Pandas, NumPy

Concepts:

- Recommendation systems, data preprocessing, OCR, NLP, web development, deployment, and version control.

==========

These two projects are optional, but they will have a significant impact on your resume, as both are related to real-world problems.

1. **Unlimited Chat Assistant**
    - An AI assistant with unlimited free plugins.

2. **Easy File Manager**
    - Simplify uploading and processing client files.