Data Science, Machine Learning & Generative AI Mastery

From Data Foundations to Fine-Tuned AI Systems
Instructor: Rohan Patil

⏳ Duration & Format

- **Total Duration**: 3 months
- **Mode**: Weekend Batch (Saturday & Sunday)
- **Session Length**: 2 – 2.5 hrs/session
- **Total Hours**: ~60–65 hrs

🎯 Course Objectives

- Build strong foundations in **Python, Statistics, and Machine Learning**
- Apply ML techniques for **regression, classification, clustering, and deep learning**
- Transition into **Generative AI with Transformers**
- Master **LoRA/QLoRA fine-tuning** using open-source frameworks
- Implement **RAG-powered chatbots** with open-source vector DBs
- Gain hands-on experience in **Generative AI projects**

📚 Curriculum Breakdown

Month 1: Foundations & Data Handling

- ✅ Week 1: Python for Data Science (NumPy, Pandas, VS Code, Jupyter)
- ✅ Week 2: Statistics & Probability (SciPy, Statsmodels, Seaborn)
- ✅ Week 3: Data Visualization & EDA (Matplotlib, Seaborn, Plotly, Pandas Profiling)
- ✅ Week 4: SQL & Data Collection (PostgreSQL, APIs, Web Scraping)

Month 2: Machine Learning Core

- ✅ Week 1: Supervised Learning (scikit-learn, XGBoost)
- ✅ Week 2: Feature Engineering & Tuning (Pipelines, Optuna, SHAP)
- ✅ Week 3: Unsupervised Learning (K-Means, DBSCAN, PCA, HDBSCAN)
- ✅ Week 4: Intro to Deep Learning (PyTorch, simple CNNs)

Month 3: Generative AI & Fine-Tuning

✅ Week 1: Transformers & LLM Foundations

Tools: PyTorch, HuggingFace Transformers

- Attention mechanism & Transformer architecture
- Encoder vs Decoder models
- Open-source LLMs (GPT-2, BERT, Mistral, LLaMA)

✅ Week 2: Fine-Tuning with LoRA/QLoRA

Tools: HuggingFace PEFT, Bitsandbytes, Accelerate

- Pretraining vs Fine-tuning vs Prompt Tuning
- LoRA & QLoRA for resource-efficient training
- Domain-specific chatbot fine-tuning project

✅ Week 3: Retrieval-Augmented Generation (RAG)

Tools: FAISS, ChromaDB, LangChain, LlamaIndex

- RAG architecture (Retriever + Generator)
- Embedding generation (`sentence-transformers`)
- Building RAG-based chatbot (with context memory)

✅ Week 4: Capstone Project – Generative AI System

Tools: Streamlit, FastAPI, GitHub, Docker (optional)

- End-to-End project: Fine-tuned + RAG-powered chatbot

- Deploy as a simple web app (Streamlit/FastAPI)
- GitHub project submission & peer review
- Career guidance: DS → ML → GenAI roadmap

🛠️ Project-Based Learning

- ✅ Month 1: EDA & Visualization project on real dataset
- ✅ Month 2: ML project (Classification/Regression with tuning & explainability)
- ✅ Month 3: Generative AI chatbot (Fine-tuned + RAG deployment)

⭐ Suggested Add-ons

- ⚡ Mini hackathons (EDA, ML, GenAI)
- 👨‍💻 GitHub portfolio projects
- 💬 Mentoring sessions
- 🤝 Community networking