

Data Science & AI Mastery: From Foundations to Enterprise

Duration: 24 Hours

Course Overview

This course combines core Data Science principles — including statistics, machine learning, data pipelines, embeddings, and model lifecycle management — with modern Artificial Intelligence and Generative AI applications. Learners will progress from foundational concepts such as probability, deep learning, and data preprocessing through to advanced topics including LLM architectures, Retrieval-Augmented Generation (RAG), AI agents, and responsible AI practices. The programme is designed to equip participants with both the theoretical grounding and the hands-on skills needed to build, evaluate, and deploy AI-powered solutions in real-world environments.

Pre-requisites

- Basic familiarity with programming concepts (Python preferred)
- Foundational understanding of mathematics — algebra and basic calculus
- Comfort working in a command-line or notebook environment
- No prior machine learning or AI experience required

Core Data Science Foundations Covered

- Probability and Statistics for data modeling and AI systems
- Machine Learning concepts and lifecycle
- Deep Learning fundamentals
- Data preprocessing, ingestion, and transformation
- Feature representation and embeddings
- Model evaluation and validation concepts
- Responsible and ethical data usage

Course Agenda

Module 01: Introductory Session

- Machine Learning lifecycle
- Supervised and unsupervised learning concepts
- Data-driven decision making
- Hands-on exposure to ML workflows

Module 02: Introduction to GenAI

- Probability and statistics in generative modeling
- Data distributions and latent space concepts
- Autoencoders for dimensionality reduction
- GANs from a statistical learning perspective

Module 03: Prompt Engineering

- Prompting as feature engineering for LLMs
- Structured inputs and constraints as data controls
- Bias reduction through prompt design

Module 04: LLM Models and Architectures

- Model architecture comparison and selection
- Performance, cost, and scalability trade-offs
- Evaluation of text, image, and audio models

Module 05: AI App Development

- Data ingestion pipelines
- Document preprocessing and chunking strategies
- Embeddings as numerical vector representations
- Information retrieval and data querying systems

Module 06: Customizing LLMs

- Fine-tuning versus embedding-based learning
- Similarity search techniques
- Knowledge graphs and structured data modeling
- Reinforcement Learning from Human Feedback (RLHF)

Module 07: Responsible AI

- Bias detection and mitigation in datasets

- Ethical data usage and governance
- Environmental impact of model training

Module 08: Enterprise Architecture

- End-to-end AI and ML pipelines
- Model lifecycle management
- Role of a Data Scientist in enterprise AI initiatives

Conclusion

This course provides a balanced and integrated approach to Data Science and Artificial Intelligence, ensuring learners leave with both the foundational knowledge and the practical skills to contribute meaningfully to AI-driven projects in any industry context.