KOENIG
step forward

KOENIG SOLUTIONS · COURSE OUTLINE

AI & Data Science: Python, R & Generative AI

Learn to build AI, machine learning and generative AI solutions using Python, R and SQL, from data preprocessing to LLM fine-tuning.

15 Days	37	157	15
120 HOURS	MODULES	TOPICS	HANDS-ON ACTIVITIES

COURSE OVERVIEW

This all-in-one program offers a deep dive into the world of Artificial Intelligence, Data Science, and Generative AI, blending foundational theory with hands-on practice across Python, R, SQL, and cutting-edge AI frameworks. Starting with AI fundamentals and Python programming, learners progress through data analysis, machine learning, deep learning, and advanced neural architectures including CNNs, RNNs, and Transformers. The R programming track complements this journey with robust data manipulation, visualization, and machine learning techniques. The Generative AI specialization introduces Large Language Models (LLMs), prompt engineering, Retrieval-Augmented Generation (RAG), and fine-tuning with open-source tools like LangChain and MLflow. Finally, the SQL module equips learners with essential skills for data querying, transformation, and exploratory analysis. Whether you're aspiring to become a data scientist, AI engineer, or generative AI specialist, this course provides the tools and knowledge to thrive in the AI-driven future.

WHO THIS COURSE IS FOR

- **Data Scientists**
- **AI Engineers**
- **Generative AI Specialists**
- **Machine Learning Engineers**
- **Data Analysts**
- **Python Developers**

LEARNING OBJECTIVES

- Explain core AI, machine learning and deep learning concepts
- Apply Python for data analysis, cleaning and preprocessing
- Build supervised and unsupervised machine learning models
- Implement deep learning architectures including CNNs, RNNs and Transformers
- Apply data manipulation and visualisation techniques in R
- Develop generative AI applications using LLMs and prompt engineering
- Implement RAG and question-answering systems with LangChain
- Query, transform and analyse data using SQL

COURSE SYLLABUS

37 Modules, 157 Topics

PART 1 Complete Artificial Intelligence

01

Introduction to AI and Python Basics

TOPICS

- What is AI?
- History and evolution of AI
- AI applications in various fields
- Introduction to machine learning and deep learning
- Python installation and setup
- Basic syntax and data types
- **Control structures:** if statements, loops
- Functions and modules

02

Python for Data Analysis & Pre-processing

TOPICS

- Introduction to NumPy, Pandas, Matplotlib, Seaborn
- Basic operations with NumPy
- Data manipulation with Pandas
- Data cleaning techniques
- Handling missing and categorical data
- Exploratory Data Analysis (EDA) with visualizations

03

Data Preprocessing and Transformation

TOPICS

- Importance of data preprocessing
- Handling missing values
- Label encoding and one-hot encoding
- **Data normalization techniques:** Standard Scaler, MinMax Scaler
- **Splitting data:** Train, test, and validation sets

04

Introduction to Machine Learning

TOPICS

- Types of machine learning
- Lifecycle of a data science project
- Overview of supervised and unsupervised learning
- Simple Linear Regression
- Multiple Linear Regression
- **Evaluation metrics:** R-squared, RMSE

05

Supervised Learning - Classification

TOPICS

Logistic Regression: Concept and intuition · Confusion matrix and performance metrics · Support Vector Machine (SVM) · Decision Trees · Random Forests

06

Hands-on

Feature Selection and Dimensionality Reduction

TOPICS

Importance of feature selection · Univariate feature selection techniques · Recursive Feature Elimination (RFE) · Principal Component Analysis (PCA)

07

Unsupervised Learning and NLP Basics

TOPICS

- Introduction to clustering
- K-means clustering
- Elbow method
- Introduction to NLP
- Tokenization, stop words, stemming, lemmatization

08

Introduction to Deep Learning

TOPICS

- Need and applications of deep learning
- Working of Artificial Neural Networks (ANN)
- Introduction to TensorFlow and Keras
- **Keras model building:** Construct, compile, evaluate
- Activation functions
- Loss functions and optimization techniques

09

Neural Network Basics and Shallow Networks

TOPICS

- Machine learning problem setup (Neural Networks)
- Neural network mindset
- Vectorization for efficient computation
- Building a neural network with one hidden layer
- Forward propagation and backpropagation

10

Deep Neural Networks

TOPICS

- Building and training deep neural networks
- Computation in deep learning
- Initialization methods
- Regularization techniques to prevent overfitting

11

Optimization and Hyperparameter Tuning (Deep Learning)

TOPICS

- Advanced optimization techniques
- Random minibatching and learning rate decay
- Batch normalization
- Grid search and randomized search (Hyperparameter tuning)

12

Convolutional Neural Networks (CNNs)

TOPICS

- Understanding pooling and convolutional layers
- Building deep CNNs for image classification
- Advanced tricks and methods in deep CNNs
- Transfer learning with pretrained models

13

Computer Vision Applications with CNNs

TOPICS

- Using CNNs for object detection tasks
- Applying CNNs for face recognition
- Implementing neural style transfer for art generation

14

Recurrent Neural Networks (RNNs) and NLP

TOPICS

Introduction to RNNs · Variants of RNNs for sequential data modeling · NLP applications with deep learning models · Word embeddings for text analysis

15

Advanced Sequence Models and Transformers

TOPICS

- Enhancing sequence models with attention mechanisms
- Speech recognition and audio data processing
- Understanding transformer networks
- Applications of transformers in NLP

PART 2 R Programming for Data Science & Machine Learning

01

Introduction To R Basics

TOPICS

- Introduction to R Programming Basics
- Arithmetic, Variables, and Basic Data Types in R
- Vector Basics, Operations, Comparison Operators, Indexing, and Slicing in R

02

R Matrices

TOPICS

- Introduction to R Matrices
- Creating, Arithmetic, Operations, Selection, and Indexing in R Matrices
- Factor and Categorical Matrices in R

03

R Data Frames & Lists

TOPICS

- Introduction to R Data Frames
- Data Frame Basics, Indexing, Selection, and Operations in R
- List and its Operations in R

04

Data Processing With R

TOPICS

Importing data from CSV Files, Excel Files · Connecting to SQL Databases · Web Scraping with R

05

R Programming Concepts

TOPICS

- Introduction to R Programming Concepts
- Logical Operators, Conditional Statements, Loops, and Functions in R
- Built-in R Features, Apply Functions, Math Functions
- Regular Expressions
- Dates and Timestamps in R

06

Data Manipulation & Data Visualization With R

TOPICS

- Data Manipulation Overview with Dplyr, Pipe Operator, and Tidyr in R
- Overview of ggplot2 for Static Visualizations (Histograms, Scatterplots, Bar plots, Box plots, 2 Variable Plotting)
- **Customizing Visualizations:** Coordinates, Faceting, Themes
- Overview of Plotly and Interactive Visualizations in R

07

Supervised Machine Learning With R

TOPICS

- Introduction to Machine Learning Concepts
- **Regression Algorithms:** Simple Linear Regression, Multiple Linear Regression
- Assumptions of Multiple Linear Regression
- **Classification Algorithms:** K-Nearest Neighbours
- **Classification Algorithms:** Decision Trees and Random Forests
- **Classification Algorithms:** Support Vector Machines

08

Unsupervised Machine Learning & Deep Learning With R

TOPICS

- **Unsupervised Learning:** K-Means Clustering
- Natural Language Processing Basics in R
- Deep Learning Concepts using Neural Networks in R

PART 3 Generative AI Speciality

01

Introduction To Generative AI

TOPICS

- Overview of Generative AI
- Architecture of Generative AI
- Applications of Generative AI with the Transformer Library
- Introduction to Generative Adversarial Networks (GANs)

02

Introduction To Large Language Models (LLMs)

TOPICS

- Architecture of Large Language Models
- Types of Large Language Models (LLMs)
- **Task-based Text AI LLMs:** Translation, Summarization, Sentence Similarity, Automatic Speech Recognition, Text-to-Speech, etc.
- **Major Text AI LLMs:** LLaMA, Qwen, Cohere, Falcon LLM
- **Image AI Models & Services:** Object Detection, Image Segmentation, Image Retrieval, Image Captioning, Visual Q&A, Zero-shot Image Classification, etc.

03

Prompt Engineering Using Open-Source Models

TOPICS

Introduction to Prompt Engineering · Prompt Engineering Techniques · Text Prompting with LLaMA (Meta) · Image Prompting with LLaMA (Meta) · Code Prompting with LLaMA (Meta)

04

Basic LLM Systems (RAG) Using Open-Source Models

TOPICS

- Introduction to Retrieval-Augmented Generation (RAG)
- Overview of LangChain
- Concepts of Embeddings, Retrieval, Chains, and Agents with LangChain
- Building simple LLM applications with LangChain (Implied through concepts and practicals)
- Creating chatbots with LangChain (Implied through concepts and practicals)
- Developing vector stores and retrievers using LangChain (Implied through concepts and practicals)
- Building agents with LangChain (Implied through concepts and practicals)
- Creating RAG and conversational RAG applications with LangChain (Implied through concepts and practicals)

05

Advanced LLM Systems (Q&A) Using Open-Source Models

TOPICS

- Difference Between RAG and Question-Answering Systems
- Building a Question-Answering System Over Tabular Data with LangChain
- Building a Question-Answering System Over SQL Data with LangChain

06

Fine-Tuning Techniques Using Open-Source Models

TOPICS

- Introduction to Quantization
- Optimization of Model Weights (Data Types)
- Modes of Quantization
- Fine-Tuning LLMs (e.g., Meta's LLaMA, Alibaba's Qwen, Google's Gemma)

07

Model Evaluation Using Mlflow

TOPICS

- Introduction to MLflow
- Building and Tracking Machine Learning Models with MLflow
- MLflow Deployment Servers
- LLM Evaluation with MLflow

PART 4 SQL for Data Science

01

2 Labs

Introduction to SQL and Data Science

TOPICS

- What is SQL? Why SQL for Data Science?
- Overview of Relational Databases
- **SQL vs NoSQL:** When and Why
- Setting up SQL Environment (SQLite / PostgreSQL / MySQL)

02

2 Labs

Data Retrieval and Filtering

TOPICS

Basic Queries: SELECT, WHERE, ORDER BY · Filtering with AND, OR, IN, BETWEEN, LIKE · Sorting and Aliasing

03

2 Labs

Aggregation and Grouping

TOPICS

- **Aggregate Functions:** COUNT, SUM, AVG, MIN, MAX
- Grouping Data with GROUP BY
- Filtering Groups with HAVING

04

2 Labs

Joins and Subqueries

TOPICS

- **Types of Joins:** INNER, LEFT, RIGHT, FULL OUTER
- Joining Multiple Tables
- **Subqueries:** Scalar, Correlated, and Nested

05

2 Labs

Data Cleaning and Transformation

TOPICS

Handling NULLs and Missing Data · Using CASE, COALESCE, and NULLIF · String and Date Functions

06

2 Labs

SQL for Exploratory Data Analysis (EDA)

TOPICS

Descriptive Statistics in SQL · Creating Derived Columns · Using CTEs (Common Table Expressions)

07

2 Labs

Window Functions and Optimization (Optional/Advanced)

TOPICS

- **Introduction to Window Functions:** ROW_NUMBER, RANK, LEAD, LAG
- **Performance Tips:** Indexing, Query Optimization Basics

HANDS-ON ACTIVITIES

Practise What You Learn

#	MODULE	ACTIVITY
06	**Feature Selection and Dimensionality Reduction**	**Hands-on:** Hands-on PCA implementation
01	**Introduction to SQL and Data Science**	**LAB** ■ Install and connect to a sample database ■ Explore tables using SELECT, LIMIT, and DESCRIBE
02	**Data Retrieval and Filtering**	**LAB** ■ Query a dataset (e.g., Titanic or Iris) to extract specific insights ■ Practice filtering and sorting data
03	**Aggregation and Grouping**	**LAB** ■ Analyze sales or customer data using aggregation ■ Compare grouped statistics
04	**Joins and Subqueries**	**LAB** ■ Join customer and transaction tables to derive insights ■ Use subqueries to filter complex conditions
05	**Data Cleaning and Transformation**	**LAB** ■ Clean and transform messy data for analysis ■ Format dates and standardize text fields
06	**SQL for Exploratory Data Analysis (EDA)**	**LAB** ■ Perform EDA on a real-world dataset (e.g., COVID, e-commerce) ■ Generate summary tables and insights
07	**Window Functions and Optimization (Optional/Advanced)**	**LAB** ■ Use window functions to analyze trends over time ■ Optimize a slow-running query

WHAT IS INCLUDED

Everything You Need to Succeed

- Koenig course material for every module
- Delivery by certified instructors and subject matter experts
- Hands-on exercises throughout the course
- Course completion certificate from Koenig

ABOUT KOENIG

Who We Are

Established in 1993, Koenig Solutions is a globally recognised training provider with 35+ technology partnerships, 5,000+ training programmes and a network of certified instructors and subject matter experts, supporting individuals and enterprise teams worldwide.

FLEXIBLE DELIVERY

Learn the Way That Suits You

Classroom Training | Live Online Classes (ILO) | Fly-Me-A-Trainer (FMAT) | Flexi (Self-Paced) | 1-on-1 Training

Ready to Get Started?

Talk to a learning advisor about schedules, group training and customised delivery.

Visit koenig-solutions.com | info@koenig-solutions.com