

# AI-ML Engineering Training Curriculum Structure (4 months)

## Program Overview

The AI-ML Engineering Program is a 4-month, industry-focused training designed to build practical skills in developing intelligent systems and real-world machine learning applications.

The program starts with Python and data processing, followed by SQL and exploratory data analysis for working with real datasets. It then progresses to core machine learning algorithms, including regression, classification, clustering, and ensemble methods.

Learners are further introduced to deep learning concepts such as neural networks and computer vision. In the final phase, the focus shifts to building and deploying end-to-end ML applications using tools like Streamlit and APIs.

The program follows a hands-on, project-based approach, ensuring learners gain real-world experience and become industry-ready.

**Program Structure:** The 4months sessions are organized into weekly modules balancing theory and practice. Each week ends with project or assessment to reinforce learning.

## Module 1: Python & Data Processing for AI-ML

### Foundations of Programming, Data Handling & Preprocessing

#### Week 1

- Role of Python in AI-ML
- Setting up Anaconda & working environment
- Variables, data types, operators
- Conditional statements & loops
- Basic dataset reading (CSV, Excel)

#### Week 2

- NumPy arrays & vectorization
  - Pandas DataFrames & data manipulation
  - Data cleaning (null values, duplicates)
  - Data transformation techniques
  - Mini Data Processing Project
-

## Module 2: SQL for Data Access & Integration , Database Querying & Data Extraction for ML

### Week 3

- Introduction to relational databases
- SQL basics: SELECT, WHERE, ORDER BY
- JOINS (INNER, LEFT, RIGHT)
- Aggregations (GROUP BY, HAVING)
- Connecting Python with SQL

### Week 4

- Subqueries & nested queries
  - Window functions (ROW\_NUMBER, RANK)
  - Data extraction for ML pipelines
  - Query optimization basics
  - SQL + Python Integration Mini Project
- 

## Module 3: Exploratory Data Analysis (EDA)

### Data Understanding, Visualization & Feature Insights

#### Week 5

- Summary statistics & data profiling
- Handling missing values & outliers
- Feature relationships & correlations
- Data preprocessing techniques
- Preparing dataset for ML models

#### Week 6

- Matplotlib & Seaborn visualizations
  - Distribution plots, boxplots, pairplots
  - Correlation heatmaps & feature importance
  - EDA storytelling techniques
  - Exploratory Data Analysis Project
- 

## Module 4: Machine Learning Fundamentals

### Core ML Algorithms & Model Building

#### Week 7

- ML lifecycle & workflow
- Train-test split & cross-validation

- Linear regression & assumptions
- Evaluation metrics (MAE, RMSE,  $R^2$ )
- Regression Project

### **Week 8**

- Logistic regression
- K-Nearest Neighbors (KNN)
- Decision Trees
- Model evaluation (Accuracy, Precision, Recall, F1)
- Classification Project

### **Week 9**

- Ensemble learning basics
  - Random Forest
  - Gradient Boosting (Intro to XGBoost)
  - Bias-variance tradeoff
  - Ensemble Learning Project
- 

## **Module 5: Unsupervised Learning & Feature Engineering**

### **Clustering, Dimensionality Reduction & Feature Creation**

#### **Week 10**

- K-Means clustering
- Hierarchical clustering
- Cluster evaluation (silhouette score)
- Real-world clustering use cases
- Customer Segmentation Project

#### **Week 11**

- Feature engineering techniques
  - Feature scaling (Normalization, Standardization)
  - PCA (Dimensionality Reduction)
  - Handling categorical data (Encoding techniques)
  - Feature Engineering Project
- 

## **Module 6: Deep Learning Fundamentals**

### **Neural Networks & Computer Vision Basics**

#### **Week 12**

- Neural network intuition
- Activation functions
- Forward & backward propagation
- Building ANN using Keras/TensorFlow
- ANN Classification Project

### **Week 13**

- Convolutional Neural Networks (CNN)
  - Convolution & pooling layers
  - Image classification pipeline
  - Transfer learning basics
  - CNN Image Classification Project
- 

## **Module 7: Model Deployment & MLOps Basics**

### **Deploying Models & Building End-to-End Systems**

#### **Week 14**

- Model saving (pickle, joblib)
  - Building ML apps using Streamlit
  - API basics with Flask/FastAPI
  - Deployment workflow (local/cloud intro)
  - End-to-End Deployment Project
- 

## **Module 8: Advanced Topics, Case Studies & Capstone Project**

### **Real-World Problem Solving & Career Preparation**

#### **Week 15**

- End-to-End ML Case Study (Industry Dataset)
- Multi-model comparison & tuning
- Hyperparameter tuning (GridSearch, RandomSearch)
- Model performance improvement
- Mock Interviews & Resume Building

#### **Week 16**

- Capstone Project (Full ML Pipeline)
- Data collection → preprocessing → modeling → deployment
- Project presentation & storytelling

- GitHub portfolio development
  - Final evaluation & career guidance
- 

### Learning Outcomes & Career Readiness

By the end of the program, learners will be able to:

- Write efficient Python code for data processing and ML workflows
- Perform data cleaning, analysis, and visualization on real datasets
- Use SQL to extract and integrate data for ML pipelines
- Build and evaluate ML models using Scikit-learn
- Apply feature engineering and dimensionality reduction techniques
- Develop deep learning models using TensorFlow and Keras
- Optimize models using evaluation metrics and tuning techniques
- Deploy ML applications using Streamlit and APIs (Flask/FastAPI)
- Build end-to-end ML projects for real-world problem solving
- Create a strong portfolio for AI/ML job roles