Machine Learning – Structured Syllabus

Beginner to Intermediate Level

Unit 0: Prerequisites & Mindset

- What is data and types of data
- Basic statistics intuition
- Why ML depends on data quality
- Python ecosystem overview

Unit 1: Introduction to Machine Learning

- Definition of Machine Learning
- ML vs Traditional Programming
- Applications of ML
- Types of ML (Supervised, Unsupervised, Semi-supervised, Reinforcement)

Unit 2: AI, ML, DL & Data Science

- Artificial Intelligence
- Machine Learning
- Deep Learning
- Generative AI
- Data Science ecosystem

Unit 3: Core ML Concepts

- Algorithms
- Features and labels
- Training vs testing data
- Overfitting and underfitting
- Model lifecycle

Unit 4: Data Preprocessing

- Missing value handling
- Outlier management
- Feature scaling and normalization
- Encoding categorical data
- Data transformation

Unit 5: Scikit-learn Basics

- ML workflow in sklearn
- Train-test split
- Pipelines
- Model fitting and prediction

Unit 6: Regression Algorithms

- Linear regression
- Polynomial regression
- Ridge and Lasso regression
- Regression evaluation metrics

Unit 7: Classification Algorithms

- Logistic regression
- KNN
- SVM
- Naive Bayes
- Decision Trees
- Random Forest
- CatBoost

Unit 8: Distance Metrics & Tree Concepts

- Euclidean distance
- Manhattan distance
- Minkowski distance
- Hamming distance
- Entropy and Gini impurity

Unit 9: Ensemble Learning

- Bagging
- Boosting
- Random Forest
- Boosting vs Neural Networks

Unit 10: Model Evaluation & Improvement

- Cross-validation
- Hyperparameter tuning
- Model comparison
- Best model selection

Unit 11: Automation & Practice

- Kaggle workflow
- End-to-end ML projects
- PyCaret AutoML
- PyCaret for regression and classification