Machine Learning Roadmap (100 Days Plan)

Total Duration: ~100 Days
Prerequisites: Python, Linear Algebra, Probability, Calculus, Statistics

---

Phase 1: Fundamentals (10-15 Days)

Concepts to Learn:

- What is ML? Supervised vs Unsupervised Learning
- Types of ML: Regression, Classification, Clustering, Reinforcement Learning
- Data Preprocessing & Feature Engineering
  - Handling Missing Data (Mean, Median, Mode)
  - Encoding Categorical Data (Label Encoding, One-Hot Encoding)
  - Feature Scaling (Min-Max Scaling, Standardization)
- Exploratory Data Analysis (EDA)
  - Pandas, Matplotlib, Seaborn
  - Data Visualization (Histogram, Box Plot, Scatter Plot)

Hands-On Tasks:

✅ Load and visualize datasets using Pandas, Matplotlib, Seaborn
✅ Apply data cleaning (handling missing values, outliers)
✅ Feature scaling and encoding using Scikit-Learn

Mini-Project:

🔹 **House Price Prediction** – Train a simple regression model on a dataset like **Boston Housing Dataset**

Phase 2: Supervised Learning (25-30 Days)

Concepts to Learn:

- **Regression:**
  - Linear Regression
  - Multiple Linear Regression
  - Polynomial Regression
  - Ridge & Lasso Regression
- **Classification:**
  - Logistic Regression
  - Decision Trees
  - Random Forest

    - Support Vector Machines (SVM)
    - Naive Bayes
    - k-Nearest Neighbors (KNN)
- **Model Evaluation:**
    - Train-Test Split, Cross-Validation
    - Performance Metrics (R^2 Score, MSE, Accuracy, Precision, Recall, F1-score, ROC-AUC)

Hands-On Tasks:

✅ Implement Linear & Logistic Regression from scratch
✅ Use Scikit-Learn for Decision Trees, Random Forest, SVM
✅ Apply Grid Search & Random Search for Hyperparameter tuning

Mini-Project:

🔹 **Titanic Survival Prediction** – Use Logistic Regression to classify survivors
🔹 **Handwritten Digit Classification** – Implement using **MNIST dataset**

Phase 3: Unsupervised Learning (15-20 Days)

Concepts to Learn:

- **Clustering Algorithms:**
    - k-Means Clustering
    - Hierarchical Clustering
    - DBSCAN
- **Dimensionality Reduction:**
    - Principal Component Analysis (PCA)
    - t-SNE
    - Autoencoders (Intro)
- **Anomaly Detection & Outlier Analysis**
    - Isolation Forest, LOF (Local Outlier Factor)

Hands-On Tasks:

✅ Implement k-Means and Hierarchical Clustering
✅ Apply PCA for high-dimensional datasets
✅ Detect fraud transactions using anomaly detection

Mini-Project:

🔹 **Customer Segmentation** – Use k-Means on a shopping dataset
🔹 **Fraud Detection System** – Use Isolation Forest for anomaly detection

---

Phase 4: Advanced ML Topics (20-25 Days)

Concepts to Learn:

- **Ensemble Learning:** Bagging, Boosting (AdaBoost, XGBoost, LightGBM)
- **Time Series Forecasting:** ARIMA, LSTMs, Facebook Prophet
- **Recommendation Systems:** Collaborative Filtering, Matrix Factorization
- **Feature Selection & Engineering:** Mutual Information, Recursive Feature Elimination (RFE)

Hands-On Tasks:

✅ Implement Random Forest & Gradient Boosting
✅ Forecast stock prices using Facebook Prophet
✅ Build a Movie Recommendation System using collaborative filtering

Mini-Project:

🔹 **Stock Price Prediction** – Train models on time-series data
🔹 **Movie Recommendation System** – Collaborative Filtering on MovieLens dataset

---

Phase 5: Practical ML & Model Deployment (10-15 Days)

Concepts to Learn:

- Model Deployment with **Flask / FastAPI**
- Model Monitoring & Scaling (Logging, A/B Testing)
- Working with Big Data (Spark MLlib)
- Introduction to **MLOps** (CI/CD for ML Models)

Hands-On Tasks:

✅ Deploy an ML model using Flask & Streamlit
✅ Integrate models with a web application

Mini-Project:

🔹 **End-to-End ML Model Deployment** – Deploy any previous project on **Render, AWS, or Heroku**

Additional Learning Resources:

- **Books:**
  - "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" – Aurélien Géron
  - "Pattern Recognition and Machine Learning" – Christopher Bishop
- **Online Courses:**

- Andrew Ng's ML Course (Coursera)
- Kaggle Micro-Courses

🚀 Final Capstone Project

After completing all phases, build a **real-world project** such as:
✅ **AI-Powered Chatbot** using NLP
✅ **Fake News Detection Model**
✅ **AI-Based Resume Screener** for hiring

Summary of What You Need to Do:

✅ Learn **Supervised & Unsupervised ML**
✅ Work on **real-world projects**
✅ Learn **model deployment** & **scalability**
✅ Compete in **Kaggle challenges**
