🧠 1-Month Machine Learning Roadmap (Classical ML Focus)

🗓 Duration: 4 Weeks

📏 Focus: Classical Machine Learning only

⚖ Balance: Theory (50%) + Hands-On Practice (50%)

🎓 Outcome: Solid grasp of ML foundations, hands-on with core models, and a completed capstone project.

📍 Week 1: ML Fundamentals & Basic Supervised Models

🎯 Goal: Understand foundational ML concepts and implement basic regression & classification models.

🧠 Concepts

- What is ML? Types of ML: Supervised, Unsupervised
- ML Workflow: Problem definition → Data Collection → Preprocessing → Model Training → Evaluation → Deployment
- Bias-Variance Tradeoff
- Exploratory Data Analysis
- Gradient Descents (SGD, Batch, etc.)
- Overfitting vs Underfitting
- Evaluation Metrics: Accuracy, Precision, Recall, F1-score, MSE, R^2

💻 Hands-On

- Setup: Jupyter Notebook, pandas, numpy, seaborn, scikit-learn
- Dataset Exploration: Iris, Boston Housing
- Linear Regression
- Logistic Regression

☑ Checkpoint Evaluation

- Explain bias-variance tradeoff & evaluation metrics in your own words
- Implement and compare Linear and Logistic Regression (code + visualization)

📍 Week 2: Core Classification Models & Data Preprocessing

🎯 Goal: Explore multiple classification algorithms and practice data preprocessing techniques.

🧠 Concepts

- Decision Trees, k-Nearest Neighbors (kNN), Naive Bayes
- Train/Test Split, Cross-Validation
- Feature Scaling, Handling Missing Data, Encoding Categorical Features
- Feature Selection vs Dimensionality Reduction

💻 Hands-On

- Use Decision Tree, kNN, Naive Bayes on Titanic, Wine, Breast Cancer datasets
- Build Pipelines
- Compare models with metrics and visualizations

☑ Checkpoint Evaluation

- Build full pipeline with preprocessing + 3 classifiers
- Visualize confusion matrices for models
- Write performance comparison summary

📍 Week 3: Unsupervised Learning & Hyperparameter Tuning

🎯 Goal: Implement clustering techniques and use hyperparameter tuning to improve models.

🧠 Concepts

- K-Means, Hierarchical Clustering
- PCA for visualization and feature reduction
- GridSearchCV, RandomizedSearchCV

💻 Hands-On

- K-Means Clustering on Mall   Dataset
- Visualize clusters with PCA
- Use GridSearchCV to tune classifiers

- Apply Elbow Method and Silhouette Score

☑ Checkpoint Evaluation

- Clustering notebook with visualizations
- GridSearch on at least one classifier
- Explanation of tuning results

📍 Week 4: Capstone Project & Final Evaluation

🎯 Goal: Complete and present a full ML project, apply end-to-end knowledge.

🧠 Concepts

- Problem Understanding, Preprocessing, Modeling, Evaluation, Reporting

💻 Hands-On

- Capstone Ideas:
- Titanic Survival (https://www.kaggle.com/code/startupsci/titanic-data-science-solutions)
- House Price Prediction (https://www.kaggle.com/c/house-prices-advanced-regression-techniques)
- Spam Detection with TF-IDF + Naive Bayes
- Customer Segmentation with K-Means
- Use Streamlit/Gradio for final demo (https://docs.streamlit.io/)
- Code push on Github/Bitbucket with self-explanatory README file.

☑ Checkpoint Evaluation

- Capstone Notebook (code + explanation)
- Project Presentation (5–7 mins)
- Written Quiz (theory + code snippets)
- Peer Review

🎁 Final Outcome

By the end of this 1-month roadmap, each student will have:

- ☑ Completed a full ML pipeline project
- ☑ Learned to use scikit-learn, pandas, matplotlib, and seaborn
- ☑ Created a portfolio-ready GitHub repository
- ☑ Gained foundational understanding for advanced ML courses in the future