Machine Learning Algorithms Summary

This document summarizes six essential machine learning algorithms with clear structure including: - Type (Supervised/Unsupervised) - Primary use case - Dataset used - Clean code snippet used to **train** the model only

S.No	Algorithm	Type	Primary Use Case	Dataset Used	Code Snippet (Training Only)
1	Linear Regression	Supervised	Predicting continuous values (e.g., prices, scores)	Boston Housing Dataset	`python model = LinearRegression() mc y_train) `
2	Logistic Regression	Supervised	Binary/ multiclass classification (e.g., cancer detection, spam filter)	Breast Cancer Dataset	`python model = LogisticRegression()  y_train) `
3	Decision Tree	Supervised	Classification and regression with interpretability	Custom CSV Dataset	`python model = DecisionTreeClassifier() y_train) `
4	Random Forest	Supervised	Ensemble model; reduces overfitting, good for complex data	Wine Quality Dataset	`python model = RandomForestClassifier() y_train) `
5	K-Means Clustering	Unsupervised	Clustering unlabelled data (e.g., segmentation, compression)	Custom CSV Dataset	`python model = KMeans(n_clusters=3) model.fit(X_scaled`
6	k-Nearest Neighbors	Supervised	Classification using distance metric (good for small datasets)	Custom CSV Dataset	`python model = KNeighborsClassifier(n_neighbors=5) mod y_train) `

Important Notes: - All algorithms were implemented practically using `scikit-learn`. - Scaling was applied where needed (especially for KNN and K-Means). - Evaluation and model saving were done separately using `joblib`.

Use this summary for interviews, review, and future ML practice.