|  |  |  |  |
| --- | --- | --- | --- |
| **Algorithm** | **Task** | **Best For** | **When to Use** |
| **Linear Regression** | Regression | Predicting continuous values | Linear relationships in small datasets. |
| **Logistic Regression** | Classification | Binary classification problems | Small datasets, simple problems, interpretable. |
| **XGBoost** | Classification/Regression | Complex problems, high accuracy | Large datasets with non-linear relationships. |
| **KMC (K-Means Clustering)** | Clustering | Grouping unlabeled data | Segmenting customers or finding clusters. |
| **KNN (K-Nearest Neighbors)** | Classification/Regression | Simple models for small datasets | When interpretability is not critical and data is well-distributed. |
| **Naive Bayes** | Classification | High-dimensional data, text data | Text classification (spam detection, sentiment analysis). |
| **DT (Decision Tree)** | Classification/Regression | Interpretable models | Tasks requiring clear decision rules. |
| **DTC (Decision Tree Classifier)** | Classification | Categorical target variables | Decision-making problems with interpretable rules. |
| **Random Forest** | Classification/Regression | High accuracy and robustness | Large datasets, noisy data, feature importance. |
| **AdaBoost** | Classification/Regression | Improving weak learners | Medium datasets, clean data, boosting decision stumps or shallow trees. |
| **SVM (Support Vector Machine)** | Classification/Regression | High-dimensional and non-linear data | Small to medium datasets, complex decision boundaries, and high accuracy tasks. |