

Algorithm	Linear data	Non-linear data	Best suited for	Complexity	Problem type	Speciality
Linear regression	Yes	No	Small datasets	Low	Regression	Linear relationship between features and target variable
Ridge regression	Yes	No	Small datasets	Low	Regression	Reduces overfitting by adding a penalty to the model coefficients
Lasso regression	Yes	No	Small datasets	Low	Regression	Reduces overfitting by shrinking the model coefficients
Logistic regression	Yes	Yes	Small and large datasets	Low	Classification	Predicts the probability of a binary outcome
Naive Bayes	Yes	Yes	Small and large datasets	Low	Classification	Classifies data based on the probability of each feature
K-nearest neighbors	No	Yes	Large datasets	High	Regression and classification	Predicts the label of a data point based on the labels of its k nearest neighbors
Decision tree	No	Yes	Small datasets	Medium	Classification and regression	Creates a model by recursively splitting the data
Random forest	Yes	Yes	Large datasets	High	Classification and regression	Combines multiple decision trees to improve accuracy
Support vector machines	Yes	No	Large datasets	High	Classification and regression	Creates a model by finding a hyperplane that separates the data into two classes
Ensemble learning	Yes	Yes	Large datasets	High	Classification and regression	Combines multiple machine learning models to improve accuracy
Clustering	No	Yes	Large datasets	High	Clustering	Groups data points together based on their similarity
K-means clustering	No	Yes	Large datasets	Low	Clustering	Groups data points into k clusters with the smallest within-cluster sum of squares

Hierarchical clustering	No	Yes	Large datasets	Medium	Clustering	Groups data points into clusters hierarchically
DBSCAN	No	Yes	Large datasets	High	Clustering	Groups data points that are close to each other
XGBoost	Yes	Yes	Large datasets	High	Classification and regression	Boosts decision trees to improve accuracy