Machine Learning Algorithms in Scikit-Learn and Python

Supervised Learning Algorithms

- Linear Models:

  Used for regression and classification tasks, e.g., Linear Regression, Logistic Regression.

- Support Vector Machines:

  Useful for classification with the SVM algorithm.

- Decision Trees and Random Forests:

  Decision Trees are used for both classification and regression, while Random Forests improve accuracy by combining multiple trees.

- K-Nearest Neighbors (KNN):

  A simple classification method based on proximity to k nearest points.

- Naive Bayes:

  Probabilistic classifier based on Bayes' theorem.

- Gradient Boosting:

  Boosting method including XGBoost, Gradient Boosting Regressor for better accuracy.

Unsupervised Learning Algorithms

- Clustering:

  Grouping data without labels, e.g., K-Means, DBSCAN.

- Dimensionality Reduction:

  Reducing the feature space, e.g., PCA, t-SNE.

Model Selection and Evaluation

- Cross-Validation:

  Method to validate model stability and performance by partitioning data.

- Hyperparameter Tuning:

  Fine-tuning model parameters for better results, e.g., GridSearchCV.

- Evaluation Metrics:

  Metrics to assess performance, e.g., accuracy, precision, recall, F1-score.