

I. SUPERVISED LEARNING

Regression Models

Predict continuous values. Models: LinearRegression, Ridge, Lasso, DecisionTreeRegressor, RandomForestRegressor, GradientBoostingRegressor, KNeighborsRegressor, SVR, MLPRegressor.

Classification Models

Predict categorical classes. Models: LogisticRegression, DecisionTreeClassifier, RandomForestClassifier, GradientBoostingClassifier, KNeighborsClassifier, SVC, NaiveBayes, MLPClassifier.

Metrics: accuracy_score, precision_score, recall_score, f1_score, roc_auc_score, r2_score, mean_squared_error, mean_absolute_error.

Import Example: from sklearn.linear_model import LinearRegression

II. UNSUPERVISED LEARNING

Clustering: KMeans, DBSCAN, AgglomerativeClustering, MeanShift, SpectralClustering, Birch.

Dimensionality Reduction: PCA, TruncatedSVD, NMF, TSNE, Isomap.

Anomaly Detection: IsolationForest, LocalOutlierFactor, OneClassSVM.

Metrics: silhouette_score, calinski_harabasz_score, davies_bouldin_score.

Import Example: from sklearn.cluster import KMeans

III. SEMI-SUPERVISED LEARNING

Models: LabelPropagation, LabelSpreading.

Used when some data points are labeled, others are not.

Metrics: accuracy_score, f1_score on labeled subset.

Import Example: from sklearn.semi_supervised import LabelSpreading

IV. PREPROCESSING & UTILITIES

Scaling: StandardScaler, MinMaxScaler, RobustScaler.

Encoding: OneHotEncoder, LabelEncoder, OrdinalEncoder.

Missing Values: SimpleImputer, KNNImputer.

Pipelines: Pipeline, ColumnTransformer.

Feature Selection: SelectKBest, RFE.

Data Split: train_test_split.

Import Example: from sklearn.preprocessing import StandardScaler

V. HYPERPARAMETER TUNING

Manual: Set parameters manually.

GridSearchCV: Exhaustive parameter search.

RandomizedSearchCV: Random combinations.

HalvingGridSearchCV: Early stopping for efficiency.

CV Utilities: KFold, cross_val_score, validation_curve.

Import Example: from sklearn.model_selection import GridSearchCV

VI. EXPERIMENT TRACKING

MLflow: Log metrics, parameters, and models.

Weights & Biases: Visual dashboard.

Joblib: Save and load models.

Other Tools: Comet.ml, Neptune.ai.

Integration Example: Combine Pipeline, GridSearchCV, and MLflow for reproducible experiments.