Ensemble methods – Bagging

Comparison of Bagging, Random Forest and Extra Tree algorithms

In this example, we import the necessary libraries, load the Iris dataset, and split it into training and testing sets. We then create instances of Extra-Trees, Random Forest, and Bagging classifiers, specifying the number of estimators (number of trees) as 100 for all. Next, we train each classifier on the training data and make predictions on the test data.

Finally, we calculate the accuracy scores by comparing the predicted classes with the true labels for each classifier. The accuracy scores indicate how well each algorithm performed in classifying the Iris samples. By comparing the accuracies, you can observe the differences in performance between Extra-Trees, Random Forest, and Bagging.

1) Importing libraries

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.ensemble import ExtraTreesClassifier, RandomForestClassifier, BaggingClassifier
from sklearn.metrics import accuracy_score
```

2) Loading data set

```
# Load the Iris dataset
iris = load_iris()
X = iris.data
y = iris.target

n_estimators=10
random_state=42

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=random_state)
```

3) Créate extra tree classifier

```
# Create and train the Extra-Trees classifier
extra_trees = ExtraTreesClassifier(n_estimators=n_estimators, random_state=random_state)
extra_trees.fit(X_train, y_train)
```

4) Create RF classifier

```
# Create and train the Random Forest classifier
random_forest = RandomForestClassifier(n_estimators=n_estimators, random_state=random_state)
random_forest.fit(X_train, y_train)
```

5) Create Bagging

```
# Create and train the Bagging classifier
bagging = BaggingClassifier(n_estimators=n_estimators, random_state=random_state)
bagging.fit(X_train, y_train)
```

6) Doing prediction

```
# Predict the classes for the test data
extra_trees_pred = extra_trees.predict(X_test)
random_forest_pred = random_forest.predict(X_test)
bagging_pred = bagging.predict(X_test)
```

7) Comparing accuracies

```
# Calculate and compare the accuracies
extra_trees_accuracy = accuracy_score(y_test, extra_trees_pred)
random_forest_accuracy = accuracy_score(y_test, random_forest_pred)
bagging_accuracy = accuracy_score(y_test, bagging_pred)

print("Extra-Trees Accuracy:", extra_trees_accuracy)
print("Random Forest Accuracy:", random_forest_accuracy)
print("Bagging Accuracy:", bagging_accuracy)
```