

Machine Learning(Lab)
Assignment 5

Aakanksha Darekar
202200733
09 A1

Problem Statement:

Write a program to implement Boosting and Bagging methods on Iris dataset.

CODE:

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.ensemble import AdaBoostClassifier, BaggingClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score

# Load the dataset
file_path = "/Users/devansh/Downloads/Iris.csv"
data = pd.read_csv(file_path)

# Preprocessing
# Dropping the 'Id' column if present
if 'Id' in data.columns:
    data = data.drop(columns=['Id'])

# Splitting features and target
X = data.drop(columns=['Species'])
y = data['Species']

# Splitting the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Boosting using AdaBoost
adaboost = AdaBoostClassifier(estimator=DecisionTreeClassifier(max_depth=1), n_estimators=50,
random_state=42)
adaboost.fit(X_train, y_train)
boosting_pred = adaboost.predict(X_test)
```

```
boosting_acc = accuracy_score(y_test, boosting_pred)

# Bagging using BaggingClassifier
bagging = BaggingClassifier(estimator=DecisionTreeClassifier(), n_estimators=50, random_state=42)
bagging.fit(X_train, y_train)
bagging_pred = bagging.predict(X_test)
bagging_acc = accuracy_score(y_test, bagging_pred)

# Cross-Validation
boosting_cv = cross_val_score(adaboost, X, y, cv=5)
bagging_cv = cross_val_score(bagging, X, y, cv=5)

# Print results
print(f"Boosting (AdaBoost) Accuracy: {boosting_acc * 100:.2f}%")
print(f"Bagging Accuracy: {bagging_acc * 100:.2f}%")
print(f"Boosting Cross-Validation Accuracy: {boosting_cv.mean() * 100:.2f}%")
print(f"Bagging Cross-Validation Accuracy: {bagging_cv.mean() * 100:.2f}%")
```

OUTPUT:

```
Boosting (AdaBoost) Accuracy: 100.00%
Bagging Accuracy: 100.00%
Boosting Cross-Validation Accuracy: 94.67%
Bagging Cross-Validation Accuracy: 96.67%
```