INTELLIGENT SYSTEMS LAB-9 (25/10/2021)

***NAME**: Harshvardhan Agarwal | **SEC**: A | **REG NO**.: 201800524*

PROBLEM STATEMENT

Write a program to demonstrate the working of decision tree based ID3 algorithms. Use an appropriate dataset for building the decision tree and apply this knowledge to classify a new sample.

1. Submit pdf with code and output.

2. Perform different test analysis by varying the test and train dataset percentage and show output.

PROBLEM SOLUTION

SOURCE CODE AND OUTPUT

```
#importing relevant libraries
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import matplotlib.ticker as ticker
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn import metrics
#importing iris dataset
from sklearn.datasets import load_iris
%matplotlib inline
```

```
#load dataset
data_iris = load_iris()

#Assign features and target labels to respective variables
X, y = data_iris['data'], data_iris['target']
```

```python
for criteria in criterion:
    print(f'\n{criteria}: ')
    values = []
    for test_size in test_sizes:
        X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=test_size)
        tree_clf = DecisionTreeClassifier(criterion=criteria).fit(X_train, y_train)
        yhat = tree_clf.predict(X_test)
        score = metrics.accuracy_score(y_test, yhat)
        values.append(score)
        #Checking accuracy
        print(f'\nFor {test_size*100} % test-set size: ')
        print(f'Train set Accuracy: ', metrics.accuracy_score(y_train, tree_clf.predict(X_train)))
        print("Test set Accuracy: ", score)
```

```
gini:

For 10.0 % test-set size:
Train set Accuracy:  1.0
Test set Accuracy:  0.8666666666666667

For 15.0 % test-set size:
Train set Accuracy:  1.0
Test set Accuracy:  1.0

For 20.0 % test-set size:
Train set Accuracy:  1.0
Test set Accuracy:  0.9666666666666667

For 25.0 % test-set size:
Train set Accuracy:  1.0
Test set Accuracy:  0.9210526315789473

For 30.0 % test-set size:
Train set Accuracy:  1.0
Test set Accuracy:  0.9555555555555556
```

```
entropy:

For 10.0 % test-set size:
Train set Accuracy:  1.0
Test set Accuracy:  1.0

For 15.0 % test-set size:
Train set Accuracy:  1.0
Test set Accuracy:  0.9565217391304348

For 20.0 % test-set size:
Train set Accuracy:  1.0
Test set Accuracy:  1.0

For 25.0 % test-set size:
Train set Accuracy:  1.0
Test set Accuracy:  0.9736842105263158

For 30.0 % test-set size:
Train set Accuracy:  1.0
Test set Accuracy:  0.9777777777777777
```

```
plt.plot(test_sizes, values, 'ro-')
plt.ylabel('Accuracy score')
plt.xlabel('Test size ratio')
plt.show()
```