

Description of three additional features:

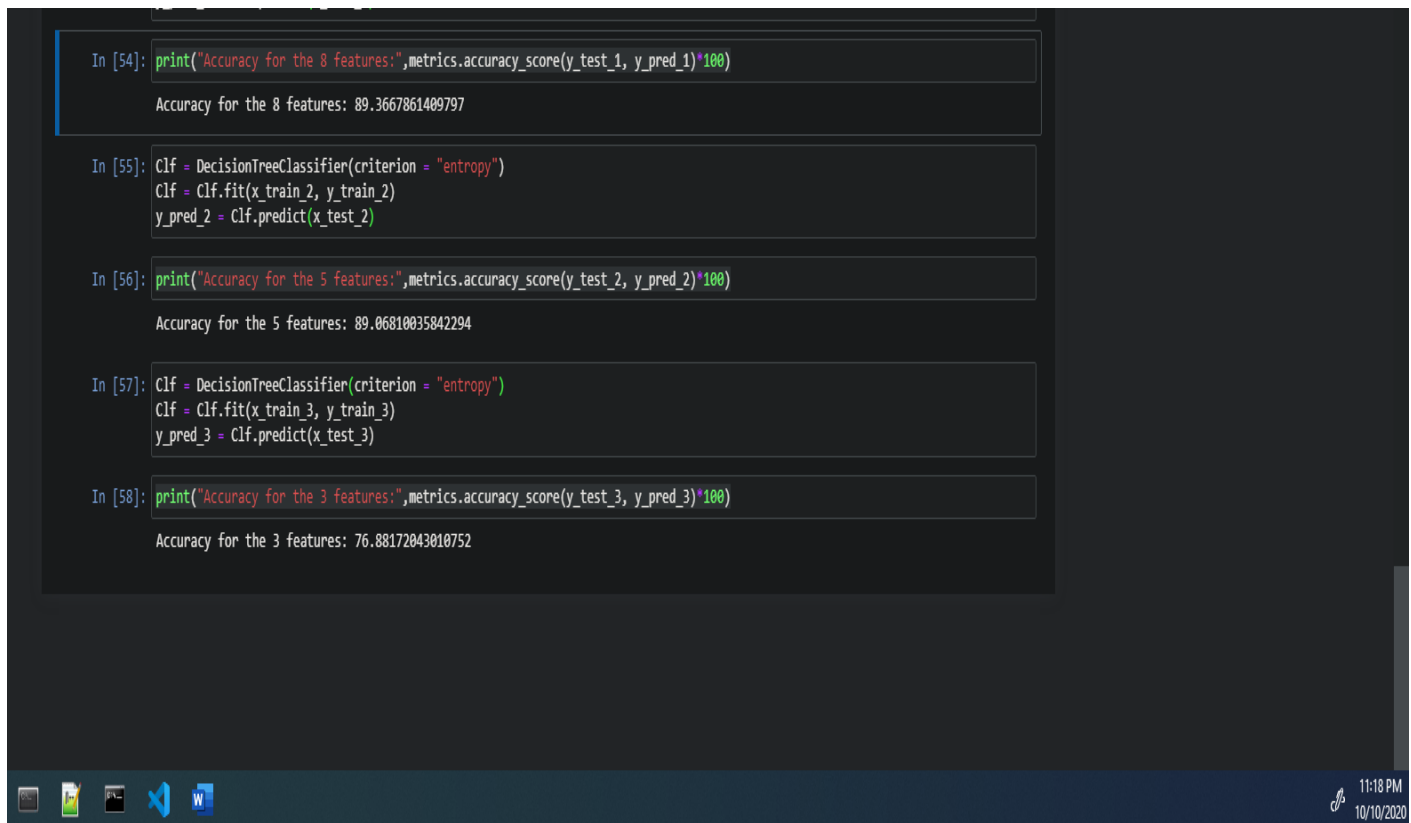
1. "R Quote": This feature is for the data points where a " is followed by a period.
  - a. Example: He didn't like the apple. "Apple is trash"
2. "L Quote": This feature is for the data points where the period is followed by a quote.
  - a. Example: He said, "Apple is trash".
3. "R Digit": When a number starts a new sentence.
  - a. This is a dozen. 12 units make a dozen.

Completion Status:

The code is 100 % completed.

1. It returns the accuracy for 8 features
2. It returns the accuracy for 5 features
3. It returns the accuracy for 3 features

Screenshot:



```
In [54]: print("Accuracy for the 8 features:", metrics.accuracy_score(y_test_1, y_pred_1)*100)

Accuracy for the 8 features: 89.3667861409797

In [55]: Clf = DecisionTreeClassifier(criterion = "entropy")
Clf = Clf.fit(x_train_2, y_train_2)
y_pred_2 = Clf.predict(x_test_2)

In [56]: print("Accuracy for the 5 features:", metrics.accuracy_score(y_test_2, y_pred_2)*100)

Accuracy for the 5 features: 89.06810035842294

In [57]: Clf = DecisionTreeClassifier(criterion = "entropy")
Clf = Clf.fit(x_train_3, y_train_3)
y_pred_3 = Clf.predict(x_test_3)

In [58]: print("Accuracy for the 3 features:", metrics.accuracy_score(y_test_3, y_pred_3)*100)

Accuracy for the 3 features: 76.88172043010752
```