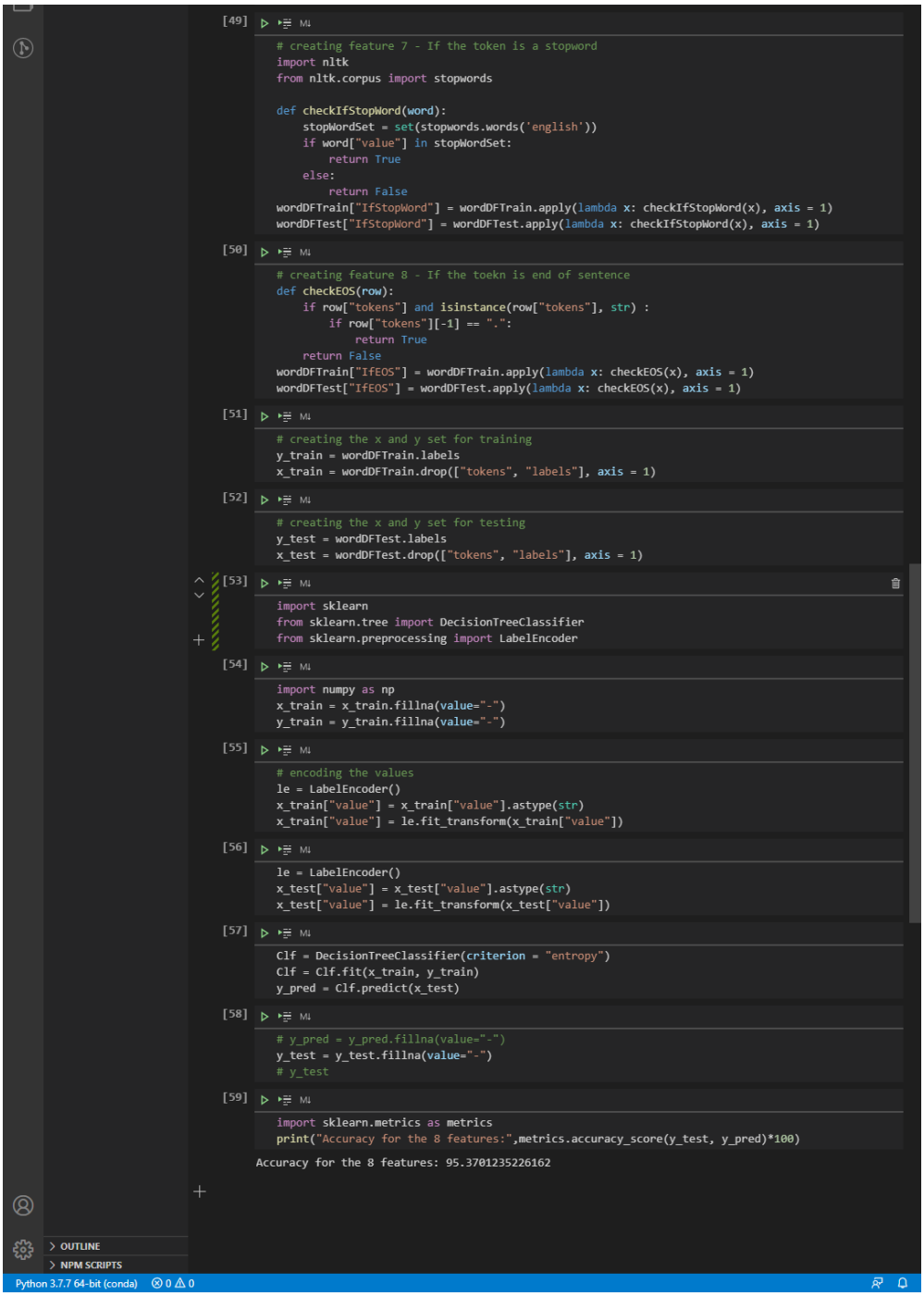
1. The program compiles and is complete.
2. The accuracy is: 95.37 %.



```
[49]
# creating feature 7 - If the token is a stopword
import nltk
from nltk.corpus import stopwords

def checkIfStopWord(word):
    stopWordSet = set(stopwords.words('english'))
    if word["value"] in stopWordSet:
        return True
    else:
        return False
wordDFTrain["IfStopWord"] = wordDFTrain.apply(lambda x: checkIfStopWord(x), axis = 1)
wordDFTest["IfStopWord"] = wordDFTest.apply(lambda x: checkIfStopWord(x), axis = 1)

[50]
# creating feature 8 - If the toekn is end of sentence
def checkEOS(row):
    if row["tokens"] and isinstance(row["tokens"], str) :
        if row["tokens"][-1] == ".":
            return True
    return False
wordDFTrain["IfEOS"] = wordDFTrain.apply(lambda x: checkEOS(x), axis = 1)
wordDFTest["IfEOS"] = wordDFTest.apply(lambda x: checkEOS(x), axis = 1)

[51]
# creating the x and y set for training
y_train = wordDFTrain.labels
x_train = wordDFTrain.drop(["tokens", "labels"], axis = 1)

[52]
# creating the x and y set for testing
y_test = wordDFTest.labels
x_test = wordDFTest.drop(["tokens", "labels"], axis = 1)

[53]
import sklearn
from sklearn.tree import DecisionTreeClassifier
from sklearn.preprocessing import LabelEncoder

[54]
import numpy as np
x_train = x_train.fillna(value="-")
y_train = y_train.fillna(value="-")

[55]
# encoding the values
le = LabelEncoder()
x_train["value"] = x_train["value"].astype(str)
x_train["value"] = le.fit_transform(x_train["value"])

[56]
le = LabelEncoder()
x_test["value"] = x_test["value"].astype(str)
x_test["value"] = le.fit_transform(x_test["value"])

[57]
Clf = DecisionTreeClassifier(criterion = "entropy")
Clf = Clf.fit(x_train, y_train)
y_pred = Clf.predict(x_test)

[58]
# y_pred = y_pred.fillna(value="-")
y_test = y_test.fillna(value="-")
# y_test

[59]
import sklearn.metrics as metrics
print("Accuracy for the 8 features:",metrics.accuracy_score(y_test, y_pred)*100)

Accuracy for the 8 features: 95.3701235226162

OUTLINE
NPM SCRIPTS
Python 3.7.7 64-bit (conda)  0 0
```