Decission Tree Classifier

Submitted by: Musharaf ahsan

Step-1: Import Libraries

In [ ]:
```
import pandas as pd
df = pd.read_csv('mldata.csv')
df.head()
```

Out[ ]:

	age	height	weight	gender	likeness
0	27	170.688	76.0	Male	Biryani
1	41	165.000	70.0	Male	Biryani
2	29	171.000	80.0	Male	Biryani
3	27	173.000	102.0	Male	Biryani
4	29	164.000	67.0	Male	Biryani

Converting Gender into Dummy Values

In [ ]:
```
df['gender'] = df['gender'].replace('Male',1)
df['gender'] = df['gender'].replace('Female',0)
df.head()
```

Out[ ]:

	age	height	weight	gender	likeness
0	27	170.688	76.0	1	Biryani
1	41	165.000	70.0	1	Biryani
2	29	171.000	80.0	1	Biryani
3	27	173.000	102.0	1	Biryani
4	29	164.000	67.0	1	Biryani

Selection of Input and Output Variables

In [ ]:
```
x = df[['weight', 'gender']]
y = df['likeness']
```

In [ ]:
```
x.head()
```

Out[ ]:

	weight	gender
0	76.0	1
1	70.0	1
2	80.0	1
3	102.0	1
4	67.0	1

In [ ]:
```
y.head()
```

Out[ ]:
```
0    Biryani
1    Biryani
2    Biryani
3    Biryani
4    Biryani
Name: likeness, dtype: object
```

Machine Learning Algorithem

In [ ]:
```
from sklearn.tree import DecisionTreeClassifier
```

Create and Fit Model

In [ ]:
```
model = DecisionTreeClassifier().fit(x,y)
```

Prediction

In [ ]:
```
model.predict([[80,1]])
```

```
c:\Users\Dell\miniconda3\envs\pandas_env\lib\site-packages\sklearn\base.py:450: UserW
arning: X does not have valid feature names, but DecisionTreeClassifier was fitted wi
th feature names
  warnings.warn(
```

Out[ ]:
```
array(['Biryani'], dtype=object)
```

How to measure Accuracy of Model?

- Split Data into test and train (80/20) rule
- 80% training Data and 20% testing

In [ ]:
```
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
x_trin, x_test, y_train, y_test = train_test_split(x,y,test_size=0.2)
```

Create a Model

In [ ]:
```
model = DecisionTreeClassifier()
```

Fiting a Model

In [ ]:
```
model = model.fit(x_trin,y_train)

predicted_values = model.predict(x_test)
predicted_values
```

Out[ ]:
```
array(['Biryani', 'Biryani', 'Pakora', 'Biryani', 'Biryani', 'Biryani',
       'Biryani', 'Samosa', 'Biryani', 'Biryani', 'Biryani', 'Biryani',
       'Biryani', 'Pakora', 'Pakora', 'Biryani', 'Biryani', 'Biryani',
       'Biryani', 'Biryani', 'Biryani', 'Biryani', 'Biryani', 'Biryani',
       'Pakora', 'Biryani', 'Biryani', 'Biryani', 'Biryani', 'Pakora',
       'Biryani', 'Biryani', 'Biryani', 'Biryani', 'Biryani', 'Biryani',
       'Biryani', 'Biryani', 'Biryani', 'Pakora', 'Biryani', 'Biryani',
       'Biryani', 'Biryani', 'Biryani', 'Biryani', 'Biryani', 'Biryani',
       'Biryani'], dtype=object)
```

Checking Score

In [ ]:
```
score = accuracy_score(y_test,predicted_values)
score
```

Out[ ]:
```
0.5714285714285714
```

How to train and save your Model

In [ ]:
```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
import joblib

model = DecisionTreeClassifier().fit(x, y)
joblib.dump(model, 'foodie.joblib')
```

Out[ ]:
```
['foodie.joblib']
```

Graph

In [ ]:
```
from sklearn import tree
model = DecisionTreeClassifier().fit(x, y)
```

Graphic Evaluation/ look into what happend

In [ ]:
```
tree.export_graphviz(model, out_file='foodie.dot',
                     feature_names=['age', 'gender'],
                     class_names=sorted(y.unique()),
```

```
                    label = 'all',
                    rounded=True,
                    filled=True)
```