03_Decision_tree_classifier

June 13, 2023

Step-1 Import Data

[ ]:
```
import pandas as pd
df = pd.read_csv("mldata1.csv")
df.head()
```

[ ]:

	age	height	weight	gender	likeness
0	27	170.688	76.0	Male	Biryani
1	41	165	70.0	Male	Biryani
2	29	171	80.0	Male	Biryani
3	27	173	102.0	Male	Biryani
4	29	164	67.0	Male	Biryani

0.0.1 Step-2 Making input and Output Variable

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

[ ]:
```
# selection of input and output variable
X = df[["weight","gender"]]
y = df["likeness"]
```

0.0.2 Step-3 Making Machine Learning Model

[ ]:
```
# Machine learning algorithm
from sklearn.tree import DecisionTreeClassifier
# Create and fit our model
model = DecisionTreeClassifier().fit(X,y)
# predict the result
model.predict([[43,0]])
```

```
c:\Users\Saeed Ahmad\AppData\Local\Programs\Python\Python310\lib\site-
packages\sklearn\base.py:450: UserWarning: X does not have valid feature names,
but DecisionTreeClassifier was fitted with feature names
  warnings.warn(
```

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

0.0.3 Step-4 Checking machine learning model performance

[ ]:
```
# How to measure the accuracy of model
# Split data into test and train(80/20)
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.2)
#Create a model
model = DecisionTreeClassifier().fit(X_train,y_train)
predicted_values = model.predict(X_test)
predicted_values
```

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

[ ]:
```
# checking the score
score = accuracy_score(y_test, predicted_values)
score
```

[ ]:
```
0.6122448979591837
```

0.0.4 Step-5 Making Visualization

[ ]:
```
# Graph
from sklearn import tree
model = DecisionTreeClassifier().fit(X,y)
# Graphic evaluation/look into what happened
tree.export_graphviz(model,out_file= "foodie.dot",
                     feature_names=["age","gender"],
                        class_names=sorted(y.unique()),
                        label="all",rounded=True,filled=True)
```