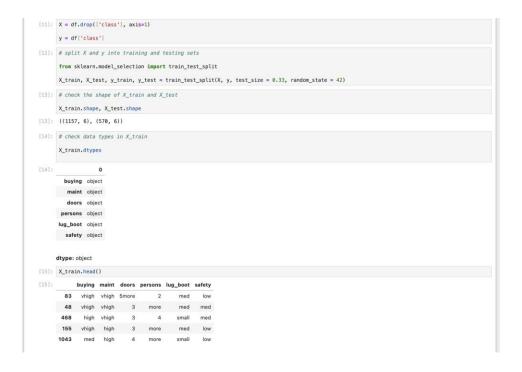
Assignment No. 3

CODE & OUTPUT:

import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_csv("path of dataset")
df.shape

```
[9]: df['class'].value_counts()
       count
      class
     unacc 1209
    acc 384
      good 69
     vgood 65
    dtype: int64
[10]: # check missing values in variables
     df.isnull().sum()
[10]: o
      buying 0
    maint 0
       doors 0
     persons 0
     lug_boot 0
    safety 0
       class 0
    dtype: int64
```



```
[19]: encoder = ce.OrdinalEncoder(cols=['buying', 'maint', 'doors', 'persons', 'lug_boot', 'safety'])
                                                                                                                                                                         ⑥↑↓占♀ⅰ
         X_train = encoder.fit_transform(X_train)
        X_test = encoder.transform(X_test)
[20]: X_train.head()
[20]:
              buying maint doors persons lug_boot safety
           83
       48 1 1 2 2 1 2

        468
        2
        1
        2
        3
        2
        2

        155
        1
        2
        2
        2
        1
        1
        1

        1043 3 2 3 2 2
[21]: X_test.head()
[21]: buying maint doors persons lug_boot safety

        599
        2
        2
        3
        1
        3
        1

        932
        3
        1
        3
        3
        3
        1

        628
        2
        2
        1
        1
        3
        3

        1497
        4
        2
        1
        3
        1
        2

        1262 3 4 3
                                               2
```

```
[22]: from sklearn.tree import DecisionTreeClassifier
[23]: clf_gini = DecisionTreeClassifier(criterion='gini', max_depth=3, random_state=0)
      # fit the model
clf_gini.fit(X_train, y_train)
[23]: DecisionTreeClassifier(max_depth=3, random_state=0)
     In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.
     On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.
     DecisionTreeClassifier
     ?Documentation for DecisionTreeClassifierIFitted
DecisionTreeClassifier(max_depth=3, random_state=0)
[24]: v pred gini = clf gini.predict(X test)
[25]: from sklearn.metrics import accuracy_score
      print('Model accuracy score with criterion gini index: {0:0.4f}'. format(accuracy_score(y_test, y_pred_gini)))
      Model accuracy score with criterion gini index: 0.8053
[26]: y_pred_train_gini = clf_gini.predict(X_train)
     y pred train gini
[27]: print('Training-set accuracy score: {0:0.4f}'. format(accuracy_score(y_train, y_pred_train_gini)))
      Training-set accuracy score: 0.7848
[28]: # print the scores on training and test set
      print('Training set score: {:.4f}'.format(clf_gini.score(X_train, y_train)))
      print('Test set score: {:.4f}'.format(clf_gini.score(X_test, y_test)))
      Training set score: 0.7848
Test set score: 0.8053
```