Assignment day 25

Navie Bayes Classification Results

1 IDV = Pclass Dv =Remaining

Acc Score = 0.578 i.e 57.8%

```
In [97]: x_train,x_test,y_train,y_test = train_test_split(x,y,test_size = 0.2,random_state=0)

In [98]: y_pred=clf.fit(x_train,y_train).predict(x_test)

In [99]: accuracy_score(y_test,y_pred)
Out[99]: 0.5786516853932584

In [100]: confusion_matrix(y_test,y_pred)
Out[100]:
array([[12,  8, 27],
       [ 1,  5, 29],
       [ 6,  4, 86]], dtype=int64)

In [101]:
```

2 IDV =Sex ,DV =Remaining

Acc Score = 0.7191 i.e 71.91%

```
In [101]: y = dataset["Sex"]

In [102]: x = dataset.drop(["Sex","PassengerId"],axis =1)

In [103]: x_train,x_test,y_train,y_test =train_test_split(x,y,test_size=0.2,random_state=0)

In [104]: y_pred=clf.fit(x_train,y_train).predict(x_test)

In [105]: accuracy_score(y_test,y_pred)
Out[105]: 0.7191011235955056

In [106]: confusion_matrix(y_test,y_pred)
Out[106]:
array([[43, 20],
       [30, 85]], dtype=int64)
```

3 IDV = SibSp ,DV = Remaining

Acc Score = 0.6966 i.e 69.66%

```
In [108]: y = dataset["SibSp"]

In [109]: x = dataset.drop(["SibSp","PassengerId"],axis =1)

In [110]: x_train,x_test,y_train,y_test =train_test_split(x,y,test_size=0.2,random_state=0)

In [111]: y_pred=clf.fit(x_train,y_train).predict(x_test)

In [112]: accuracy_score(y_test,y_pred)
Out[112]: 0.6966292134831461

In [113]: confusion_matrix(y_test,y_pred)
Out[113]:
array([[106,  18,   0,   0,   0,   0,   0],
       [ 22,  18,   0,   0,   0,   0,   0],
       [  3,   1,   0,   0,   0,   0,   0],
       [  4,   2,   0,   0,   0,   0,   0],
       [  1,   0,   0,   0,   0,   0,   0],
       [  1,   0,   0,   0,   0,   0,   0],
       [  2,   0,   0,   0,   0,   0,   0]], dtype=int64)

In [114]:
```

3 IDV = Embarked,DV = Remaining

Acc Score = 0.7415 i.e 74.15%

```
In [127]: y = dataset["Embarked"]

In [128]: x = dataset.drop(["Embarked","PassengerId"],axis =1)

In [129]: x_train,x_test,y_train,y_test =train_test_split(x,y,test_size=0.2,random_state=0)

In [130]: y_pred=clf.fit(x_train,y_train).predict(x_test)

In [131]: accuracy_score(y_test,y_pred)
Out[131]: 0.7415730337078652

In [132]: confusion_matrix(y_test,y_pred)
Out[132]:
array([[  0,   0,  30],
       [  0,   0,  16],
       [  0,   0, 132]], dtype=int64)

In [133]:
```