Experiment Title- 2.3

Student Name: Lipakshi **UID: 20BCS5082**
Branch: BE-CSE **Section/Group-20BCSWM_607-B**
Semester:5 **Date of Performance:12/10/2022**
Subject Name: Machine Learning Lab **Subject Code: 20CSP-317**

1. Aim: To Implement K-nearest Neighbour on Classification Problem and Justify the outcome with relevant Parameters.

2. Source Code and Output:

```
importnumpyasnp
importpandasaspd
importseabornassns
importmatplotlib.pyplotasplt

fromsklearn.model_selectionimporttrain_test_split
fromsklearn.metricsimportclassification_report,confusion_matrix
fromsklearn.metricsimportroc_curve,auc
fromsklearn.preprocessingimportlabel_binarize
fromsklearn.multiclassimportOneVsRestClassifier
fromsklearn.metricsimportprecision_recall_curve
fromsklearn.metricsimportroc_auc_score
fromsklearn.neighborsimportKNeighborsClassifier
```

```
data=pd.read_csv('Downloads/iris  data.csv')

data.head()
```

	sepal.length	sepal.width	petal.length	petal.width	variety	Unnamed: 5
0	5.1	3.5	1.4	0.2	Setosa	NaN
1	4.9	3.0	1.4	0.2	Setosa	NaN
2	4.7	3.2	1.3	0.2	Setosa	NaN
3	4.6	3.1	1.5	0.2	Setosa	NaN
4	5.0	3.6	1.4	0.2	Setosa	NaN

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 6 columns):
sepal.length    150 non-null float64
sepal.width     150 non-null float64
petal.length    150 non-null float64
petal.width     150 non-null float64
variety         150 non-null object
Unnamed: 5      0 non-null float64
dtypes: float64(5), object(1)
memory usage: 6.5+ KB
```

```
X=data.drop(['Unnamed: 5','variety'],axis=1) y=data['variety']
# print(X.head())
print(X.shape)
# print(y.head())
print(y.shape)
```

```
(150, 4)
(150,)
```

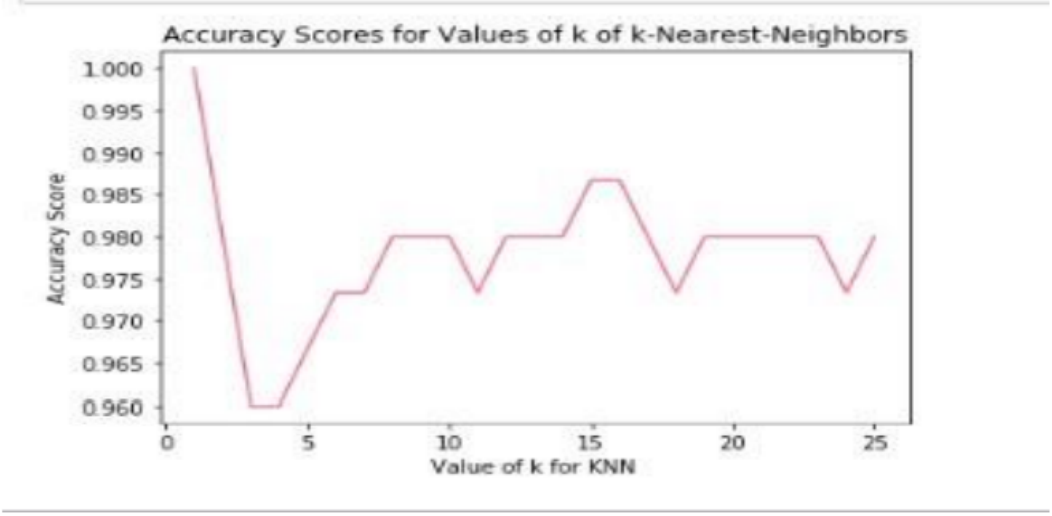
X_train,X_test,y_trai

```
n,y_test=train_test_split(X,y,test_size=0.4,random_state=5)
print(X_train.shape) print(y_train.shape) print(X_test.shape) print(y_test.shape)
```

```
(90, 4)
(90,)
(60, 4)
(60,)
```

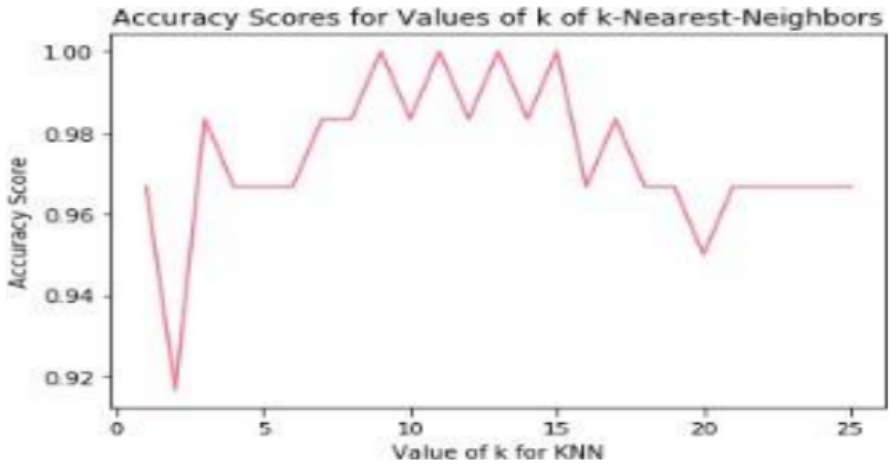
```
k_range=list(range(1,26))
scores=[]
forkink_range:
knn=KNeighborsClassifier(n_neighbors=k)
knn.fit(X,y)
y_pred=knn.predict(X)
scores.append(metrics.accuracy_score(y,y_pred))

plt.plot(k_range,scores)
plt.xlabel('Value of k for KNN')
plt.ylabel('Accuracy Score')
plt.title('Accuracy Scores for Values of k of k-Nearest-Neighbors')
    plt.show()
```



```
k_range=list(range(1,26)) scores=[]
forkink_range: knn=KNeighborsClassifier(n_neighbors=k) knn.fit(X_train,y_train)
y_pred=knn.predict(X_test)
scores.append(metrics.accuracy_score(y_test,y_pred))

plt.plot(k_range,scores) plt.xlabel('Value of k for KNN') plt.ylabel('Accuracy Score')
plt.title('Accuracy Scores for Values of k of k-Nearest-Neighbors') plt.show()
```




```
knn=KNeighborsClassifier(n_neighbors=12) knn.fit(X, y)

# make a prediction for an example of an out-of-sample observation
knn.predict([[6, 3, 4, 2]])
```

```
array(['Versicolor'], dtype=object)
```

Learning outcomes (What I have learnt):

1. I learnt what is K-nearest Neighbor problem.
2. Practical use of K-nearest Neighbor classification.
3. How machine learning helps to analyze datasets.

Evaluation Grid :

Sr. No.	Parameters	Marks Obtained	Maximum Marks

1.	**Student Performance (Conduct of experiment) objectives/Outcomes.**		**12**
2.	**Viva Voce**		**10**
3.	**Submission of Work Sheet (Record)**		**8**
	Total		**30**