

CSBB311: MACHINE LEARNING LAB
ASSIGNMENT 2 :- Classification Using KNN

Submitted By:

Name: Kartik Mittal

Roll No: 221210056

Branch: CSE

Semester: 5th Sem

Group : 2

Submitted To: Dr. Preeti Mehta

Department of Computer Science and Engineering



NATIONAL INSTITUTE OF TECHNOLOGY DELHI

2024

Code :-

```
1  import pandas as pd
2  from sklearn.impute import SimpleImputer
3  from sklearn.preprocessing import LabelEncoder
4  from sklearn.model_selection import train_test_split
5  from sklearn.neighbors import KNeighborsClassifier
6  from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
7  import matplotlib.pyplot as plt
8  import seaborn as sns
9
10 # Preprocessing
11 data = pd.read_csv('titanic.csv')
12 imputer = SimpleImputer(strategy='mean')
13 data['Age'] = imputer.fit_transform(data[['Age']])
14
15 label_encoder = LabelEncoder()
16 data['Sex'] = label_encoder.fit_transform(data['Sex'])
17
18 data = data.drop(['Name', 'Ticket', 'Cabin', 'Embarked'], axis=1, errors='ignore')
19
20 # Splitting the data
21 x = data.drop('Survived', axis=1)
22 y = data['Survived']
23
24 xTrain, xTest, yTrain, yTest = train_test_split(x, y, test_size=0.2, random_state=42)
25
26 # KNN Classifier
27 knn = KNeighborsClassifier(n_neighbors=5)
28 knn.fit(xTrain, yTrain)
29
30 yPred = knn.predict(xTest)
31
32 accuracy = accuracy_score(yTest, yPred)
33 classReport = classification_report(yTest, yPred)
34 confMatrix = confusion_matrix(yTest, yPred)
35
36 print(f'Accuracy: {accuracy:.2f}')
37 print(f'Classification Report:\n{classReport}')
38 print(f'Confusion Matrix:\n{confMatrix}')
39
40 # Plotting the confusion matrix
41 plt.figure(figsize=(8, 6))
42 sns.heatmap(confMatrix, annot=True, fmt='d', cmap='Blues', cbar=False)
43 plt.title('Confusion Matrix')
44 plt.xlabel('Predicted')
45 plt.ylabel('Actual')
46 plt.show()
47
48 # Plotting the accuracy
49 plt.figure(figsize=(4, 4))
50 plt.bar(['Accuracy'], [accuracy], color='green')
51 plt.ylim(0, 1)
52 plt.title('Model Accuracy')
53 plt.show()
```

Output :-

```
Accuracy: 0.76
Classification Report:
              precision    recall  f1-score   support

     0       0.80      0.90      0.85       189
     1       0.61      0.41      0.49        73

   accuracy          0.76       262
  macro avg       0.71      0.66      0.67       262
 weighted avg     0.75      0.76      0.75       262

Confusion Matrix:
[[170  19]
 [ 43  30]]
```

Plots :-



