# **CSBB311: MACHINE LEARNING LAB**

# **ASSIGNMENT 3:- Classification Using SVM**

**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:-

```
import pandas as pd
     from sklearn.model_selection import train_test_split
     from sklearn.svm import SVC
     from sklearn.metrics import accuracy_score, confusion_matrix
     from sklearn.preprocessing import StandardScaler, LabelEncoder
    import matplotlib.pyplot as plt
    import seaborn as sns
10 data = pd.read_csv('tested.csv')
data['Age'].fillna(data['Age'].mean(), inplace=True)
    data['Embarked'].fillna(data['Embarked'].mean(), inplace=True)
     label_encoder = LabelEncoder()
    data['Sex'] = label_encoder.fit_transform(data['Sex'])
    data['Embarked'] = label_encoder.fit_transform(data['Embarked'])
    data = data.drop(columns=['Passengerid'] + [col for col in data.columns if 'zero' in col])
    # Features (X) and Target (y)
25  X = data.drop(columns=['2urvived']) # target column
26  y = data['2urvived']
```

```
X_train, X_temp, y_train, y_temp = train_test_split(X, y, test_size=0.4, random_state=42)
    X_val, X_test, y_val, y_test = train_test_split(X_temp, y_temp, test_size=0.5, random_state=42)
    # Standardize features
    scaler = StandardScaler()
    X_train = scaler.fit_transform(X_train)
    X_val = scaler.transform(X_val)
    X_test = scaler.transform(X_test)
     svm_model = SVC(kernel='linear')
40 svm_model.fit(X_train, y_train)
    # Validate the model
    y_val_pred = svm_model.predict(X_val)
    val_accuracy = accuracy_score(y_val, y_val_pred)
45 print(f'Validation Accuracy: {val_accuracy:.2f}')
    # Test the model
    y_test_pred = svm_model.predict(X_test)
     test_accuracy = accuracy_score(y_test, y_test_pred)
50 print(f'Test Accuracy: {test_accuracy:.2f}')
    conf_matrix = confusion_matrix(y_test, y_test_pred)
    plt.figure(figsize=(6,4))
    sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues', xticklabels=['Did not survive', 'Survived'], yticklabel
     plt.title('Confusion Matrix - Test Data')
```

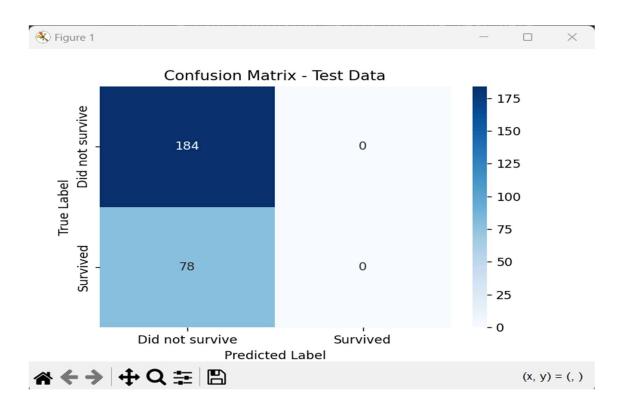
```
plt.title('Confusion Matrix - Test Data')
plt.ylabel('True Label')
plt.xlabel('Predicted Label')
plt.show()

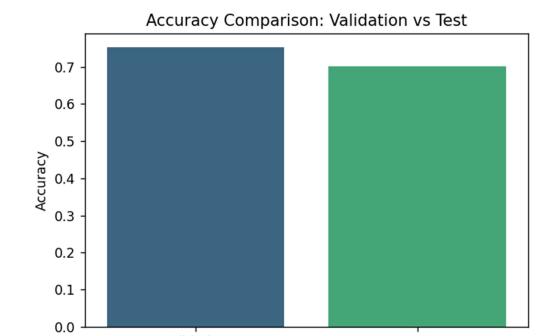
# Plot accuracy comparison
plt.figure(figsize=(6,4))
accuracy_scores = [val_accuracy, test_accuracy]
labels = ['Validation', 'Test']
sns.barplot(x=labels, y=accuracy_scores, palette='viridis')
plt.title('Accuracy Comparison: Validation vs Test')
plt.ylabel('Accuracy')
plt.show()
```

## **Output:-**

```
hine Learning\lab 3 svm\svm.py"
Validation Accuracy: 75%
Test Accuracy: 70%
```

### Plots:-





Test



Validation