NATIONAL UNIVERSITY OF MODERN LANGUAGES <u>ISLAMABAD</u>



MACHINE LEARNING (LAB)

Assignment: 02

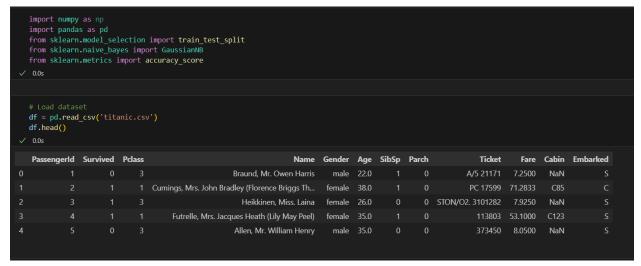
Submitted to Ms. Qurat Raja

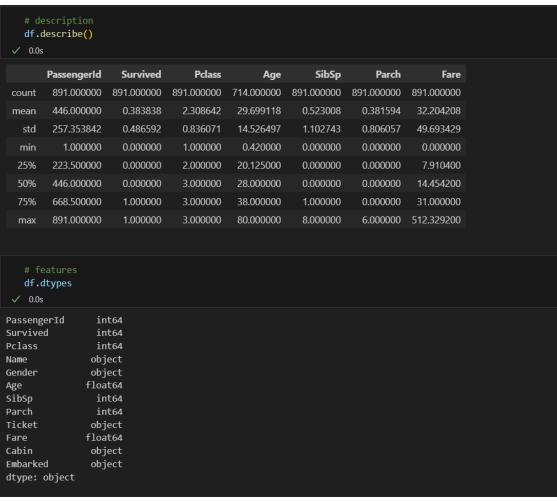
Submitted ByJunaid Asif
(BSAI-144)

Submission Date: October 22, 2024

PART-1

Dataset Import and Exploration





Data Preprocessing

Certainly! Here's how we can properly organize the data preprocessing steps for the Titanic dataset before splitting and training our model.

```
# Dropping irrelevant columns
df = df.drop(['PassengerId', 'Name', 'Ticket', 'Cabin'], axis=1, errors='ignore')

# Handling missing values
df['Age'].fillna(df['Age'].median(), inplace=True)
df['Embarked'].fillna(df['Embarked'].mode()[0], inplace=True)

# Encoding categorical data
df['Gender'] = df['Gender'].map({'male': 0, 'female': 1})
df = pd.get_dummies(df, columns=['Embarked'], drop_first=True) # One-hot encoding for Embarked
```

PART-2

Data Splitting

PART-3

Naïve Bayesian Implementation