NATIONAL UNIVERSITY OF COMPUTER & EMERGING SCIENCES
FAST - PESHAWAR CAMPUS

Subject: AL 2002 - Artificial Intelligence Lab
Instructor: Muhammad Saood Sarwar

Lab Task: K-Nearest Neighbors (KNN)

KNN Classification Using the Titanic Dataset

Download the Titanic dataset from the internet, which contains information about passengers onboard the Titanic. The objective is to predict the survival of passengers using a K-Nearest Neighbors (KNN) machine learning model.
You will use the following features:

- **Pclass** (Passenger Class)
- **Gender**
- **Age**
- **SibSp** (Number of Siblings/Spouses Aboard)
- **Parch** (Number of Parents/Children Aboard)
- **Fare**
- **Embarked** (Port of Embarkation)

The target variable is:

- **Survived**

Question 1: Feature Engineering

Describe how you would convert categorical features like **Gender** and **Embarked** into numeric format so they can be used by the KNN algorithm.

Question 2: Model Training and Evaluation

Perform the following steps:

- Split the dataset into training and testing sets.
- Train the KNN model using the training data.
- Evaluate the model's accuracy on the testing set.
- Repeat the above steps using different random states (e.g., 1, 10, 42) to observe how accuracy varies.

Question 3: Visualization and Interpretation

Plot a graph showing the model accuracy for different random states. Interpret the results and discuss any patterns or variations in accuracy that you observe.