

**CS-4038**  
**Data Mining**  
**Term Project**

**Group Members:**

- 1) **Ayesha Saqib 20F-0360**
- 2) **Mozeb Ahmed 20F-0161**
- 3) **Huma Mukhtar 20F-0139**

# **Phase II Report:**

## **1. Data Preprocessing:**

**Handling Missing Values:** Missing values in the dataset were filled with zeros to ensure data integrity.

**Normalization:** Assignment and quiz scores were normalized based on their total marks and corresponding weights to bring them to a uniform scale for analysis.

## **2. Model Training and Results:**

The analysis was performed using two classification algorithms:

- 1) K Nearest Neighbour (KNN)
- 2) Decision Tree.

### **K Nearest Neighbour (KNN):**

- Accuracy: 92.9%
- Specificity: 87.5%
- Sensitivity: 100%

### **Decision Tree:**

- Accuracy: 78.6%
- Specificity: 75%
- Sensitivity: 83.3%

