Department of Computer Science

Krantiguru Shyamji Krishna Verma Kachchh University

University Campus, Mundra Road, Bhuj, Gujarat - 370 001

http://cs.kskvku.ac.in 02832 235 036 cs@kskvku.ac.in

Course: M.Sc. CA&IT – 6 (NEP)
Subject: Data Mining
Assignment Date: 19-01-2026 **Submission Date:** 02-02-2026, Before 2 PM
Assignment ID: DMASEMMSC61

Instructions:

- Execute the following Programs in your **Jupyter Notebook** and write **all the code and outputs** in your **Practical Notebook.**
- Upload your ***final Jupyter Notebook*** on **GitHub repository**. Stick the **QR Code** of that repository on your blank page (left side) of your practical notebook.
- **Stick the outputs of graph** or dataset on your **blank page** *(left side)* of your **practical notebook.**

Programs (StudentPerformance):

1. Load the Dataset of **StudentPerformance.csv** on Jupyter Notebook and Print the Dataset.
2. Print that rows whose exam score is less than 65 (<65).
3. Print that rows whose Attendance is higher than 90 (90>=).
4. Print that rows whose row number (index) starts with 42 and ends with 54.
5. Print the maximum and minimum values of **Hours_Studied.**
6. Print that row who have minimum **Hours_Studied** value.
7. Print the Dataset displaying which values are null or not through Boolean.
8. Plot a Bar Graph of students to analyse "How *resources are helpful for students"* through **Exam_Score** and **Access_to_Resources** columns using Matplotlib. (Take Exam_Score as "y" axis and Access_to_Resources as "x" axis)
9. Plot a Line Graph of students to analyse "How *motivation level are helpful for students"* through **Motivation_level** and **Exam_Score** columns using Matplotlib.
10. Make a Box Plot of **Exam_Score** column to check for the outliers.
11. Make a Box Plot of **Hours_St_Norm** column to check for the outliers.

12. Write a code in a separate cell to fill the Null/None/NaN values of a student whose **Study Hours** are not present. (Hint: you can use Median to fill null values)

13. Write a code in a separate cell to fill the Null values of a **Teacher Quality** column by **backward fill method** and also, perform **forward fill method** on **Distance from Home** attribute.

14. Write a code in a separate cell to fill the Null values of **Parent Education Level** as "Undefined".

15. Make a new column **Hours_St_Norm** to normalize the **Hours_Studied** column by Min-Max normalization and also make a box plot of **Hours_St_Norm** column.

$$X' = \frac{X - Xmin}{Xmax - Xmin}$$