# CHRIST (Deemed to be University) Data Science Department

**Bachelors of Science (Data Science)**

**Course**: BDS471L – Machine Learning

**Exercise No**: LAB Exercise – 8

**Date:** 18 – 04 – 2024

Use the following platform to demonstrate the following exercises.

Jupyter Notebook, Google Colabs, Any Python Editor

**Topic: Clustering**

1. For the given dataset, Identify the problem statement
2. Apply Imputation/encoding/Scaling techniques as required.
3. Apply the K means algorithm to generate the target variable based on the problem statement,
4. Verify if the number of clusters chosen is inclined with the prediction problem
5. Use the new dataset with a target variable for prediction using any 2 supervised machine learning algorithms.
6. Use appropriate evaluation metrics to check the performance of the model.
7. Modify the clusters and check the performance of the model.
8. Generate AUC-ROC curve to determine which supervised machine learning has high AUC value.

# Evaluation Scheme: (Total 10 Marks)

1. Question 1 , 1 mark
2. Question 2 , 2 mark
3. Question 3 , 1 mark
4. Question 4 , 1 mark
5. Question 5 , 2 mark
6. Question 6 , 1 mark
7. Question 7 , 1 mark
8. Question 8 , 1 mark

# General Instruction:

1. Create a word document and paste all the answers. File name should be your register number followed by lab No: Example: 2048001\_lab1
2. Upload the PDF in Google Classroom on or before the deadline mentioned.