# CHRIST (Deemed to be University) Data Science Department

**Bachelors of Science (Data Science)**

**Course**: BDS471L – Machine Learning

**Exercise No**: LAB Exercise – 6

**Date:** 05 – 04 – 2024

Use the following platform to demonstrate the following exercises.

Jupyter Notebook, Google Colabs, Any Python Editor

**Topic: Linear Regression**

1. For the given dataset, predict the win status of the horse race using logistic regression and a decision tree.
2. Apply the Imputation, encoding and scaling techniques when required.
3. Perform a minimum of 5 cross-validations and test the model.
4. Check for overfitting issues. If overfitting is detected, apply the required overfitting handling methods to the algorithm.
5. User GridsearchCV to check for the best hyperparameters for both the algorithms,
6. Use the Data pipeline to combine the preprocessing and modelling section
7. Generate the pickle file and predict the win status using the pickled file.

# Evaluation Scheme: (Total 10 Marks)

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

**Total 10 marks \*2 = 20 marks, converted to 10 marks**

# 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 answer document in Google Classroom on or before the deadline mentioned.