ITIR626: Assignment 1 (Due date is 26th May 2025)

From the UCI (https://archive.ics.uci.edu/) or Kaggle (https://www.kaggle.com/datasets) datasets or any dataset of your choice, you are to select one classification dataset. A classification dataset has a categorical variable as its Target/dependent variable.

1. Select any one of the right machine learning (ML) algorithms for your classification dataset.
2. Use your selected ML algorithm to train the model on the dataset you have selected and evaluate the results.
3. The evaluation metrics you are to use are as follows: Accuracy, precision, recall, F1-score, ROC, and AUC.
4. Prepare a report which should contain:
    a. A brief background to the problem
    b. Materials and Method
        i. Materials. Describe the model and dataset (how many variables, records, variable types, which is the target variable, and so on).
        ii. Method Describe how you trained your models. What parameters did you use for the algorithms? What was your train/test split ratio? Provide flowcharts/diagrams of your models.
    c. Results and Discussion
        i. Present the results of your experiments.
        ii. Briefly discuss your assessment of the model performance.
5. Place your code, report, and datasets into a folder, zip it, and submit only the zip file. Make sure your code is fully functional, as it will simply have executed them to confirm if it works. Feel free to use either a notebook format (Jupyter Notebook) or Python scripts for your coding.