**Institute of Business Administration**

**CSE-473: Introduction to Machine Learning**

**Assignment 01**

**Due Date: 10th November 2024 (11:55 PM)**

**Instructions**

* You are required to individually participate in the challenge posted on the Kaggle website. The link of the Kaggle challenge is copied below:

<https://www.kaggle.com/t/5dfe8589017741f4a9457337d6c6825e>

* The challenge must be implemented using Python.
* You are expected to use your proper full name. This is important because many people have similar first names. Any entry with an indistinguishable name/username (like warrior, beast, blackbeauty, etc.) would be removed.
* You are allowed to post 10 entries per day. This is to ensure you start participating from the very first day. Being on the top of the leaderboard is important but more important is the breadth of your experiments and making a consistent effort.
* You are required to submit/fill a small google form after every entry to ensure that we have a detailed record of your submissions like which model you used, what data transformations you applied, how you handled the missing data, what model parameters were used, etc. The link to the form submission is copied below.

<https://forms.gle/8mebDQFjjERBixVK7>

* Your task is to primarily try following classification algorithms along with the data preprocessing, cleaning, and transformation techniques to increase the score.
  + Decision Tree
  + Naive Bayes
  + K-Nearest Neighbor
  + Random Forest
  + Gradient Boosting
  + Adaptive Boosting
  + Light GBM
  + XGBoost
  + CatBoost
  + BaggingClassifier
  + ExtraTree Classifier (Extremely Randomized Tree)
  + Voting
  + Stacking
* You must also submit the python code notebook (on LMS), demonstrating your best entry on Kaggle. Please note that the submitted file on execution must show the same score as visible on Kaggle. The python notebook must also include your overall findings and insights as a markdown cell.