**Course Project – IME692, Fall 2020**

Use the datasets uploaded with this document to complete this assignment. There are two files: train.csv and test.csv. Use train.csv to train your prediction model and test.csv to only test the performance of the prediction model using misclassification error rate. There are twenty predictors labeled $x1, x2, \ldots, x20$ and the dependent variable y has binary class labels (0 and 1). Your project report should contain following information.

(a) Develop a classification model that has lowest misclassification error rate on the test data. What is the training error rate for this model? Also, provide a brief explanation for the good performance of this model. (Hint: a good classification model should have test misclassification error rate below 0.20).

(b) Document all models you experimented with to evaluate the performance on the test data. The code used for these models should be uploaded along with your submission. The report should include a brief description of all classification models you have used to complete the assignment. If you are using a classification model that is not covered in the course, then provide an appropriate reference for this model.

(c) Generate ROC curve for the best performing classification model. How would you interpret the findings from the ROC curve?

(d) Briefly mention the contribution of each team member in the project.