

## Machine Learning (BIA-5302) Fall 2023 Assignment 02 Deadline, Friday November 24<sup>th</sup>, 2023

The dataset Cars\_Data\_full.xls required for this assignment can be found along with the Assignment 02.

In this assignment, we will focus on implementing centralized processing of ML models.

**Goal:** After working through the assignment questions, students will be able to do the following:

- Understand Regression and Random Forest models.
- Compare and contrast the advantages and disadvantages of different prediction models and determine which one is most suitable for car pricing.

### Assignment Questions

You need to submit a report that addresses the following questions.

**Q1.** Discuss a conceptual model to predict the MSRP including significant variables, correlation, your plan for training and testing.

**Q2.** Build both Multi-Linear Regression model and Random Forest model to predict the MSRP of a car. (*P.S. do not forget to upload your code*).

**Q3.** Validate and interpret the model using a different number of significant variables.

**Q4.** Compare results of random forest and linear regression, support each model by validating the results.

**Q5.** Make a decision and offer your recommendations to the professor. Justify your decision based on your analytics results.

### Submission Guidelines

Please note that this is group submission, any copy of code, or answers will be considered as plagiarism. To Submit: Zip all of your **word/pdf/code** files into a folder groupnumberXX.zip and upload to folder “Assignment 02”.