The dataset was on graduate admissions statistics. The variables and features of the dataset include, amongst other things, ‘GRE score’, ‘TOEFL score’, ‘University Rating’… so on. The target would be the chance to admit expressed as a probability.

My objective is to see which would be the best model for regression and classification applied to better understand the interrelations and interactions of the features and the target. The model selected based on the lowest error and highest accuracy score would be:

For Regression: Linear Regression (rms of ~ 0.0770)

For Classification: Extra Tree Classifier (achieved the accuracy of 0.95 (95%))

In relation to the Chance of Admit, CGPA has the highest score in terms of features importance, followed by GRE score.