Client: *Donegal Community Bank*

Consultant: *Teejay Data Analytics*

Annual Report on Customers Loan-worthiness

This report is based on the data provided by Donegal Community Bank on finding a way to decide whether an applicant for a loan is likely to default.

Method: We carried out data cleaning on the raw data. To predict an outcome (target) variable based on the values of other "predictor" variables, we split the cleaned data into training, validation, and test (60%, 20%, 20%). We applied fitting on four models; a logistic regression model; a Ridge regression model; a LASSO regression model; and an Elastic Net model.

Result: To evaluate the performance of our model, we applied a classical tool: the Receiver Operating Characteristic (ROC) curve. The ROC curves for the models considered are shown in Fig 1. The ROC curve is generated by plotting all possible cutoff values, which are the probabilities assigned to each observation. Selecting a different cutoff value will alter the sensitivity and specificity of the prediction tool [1].

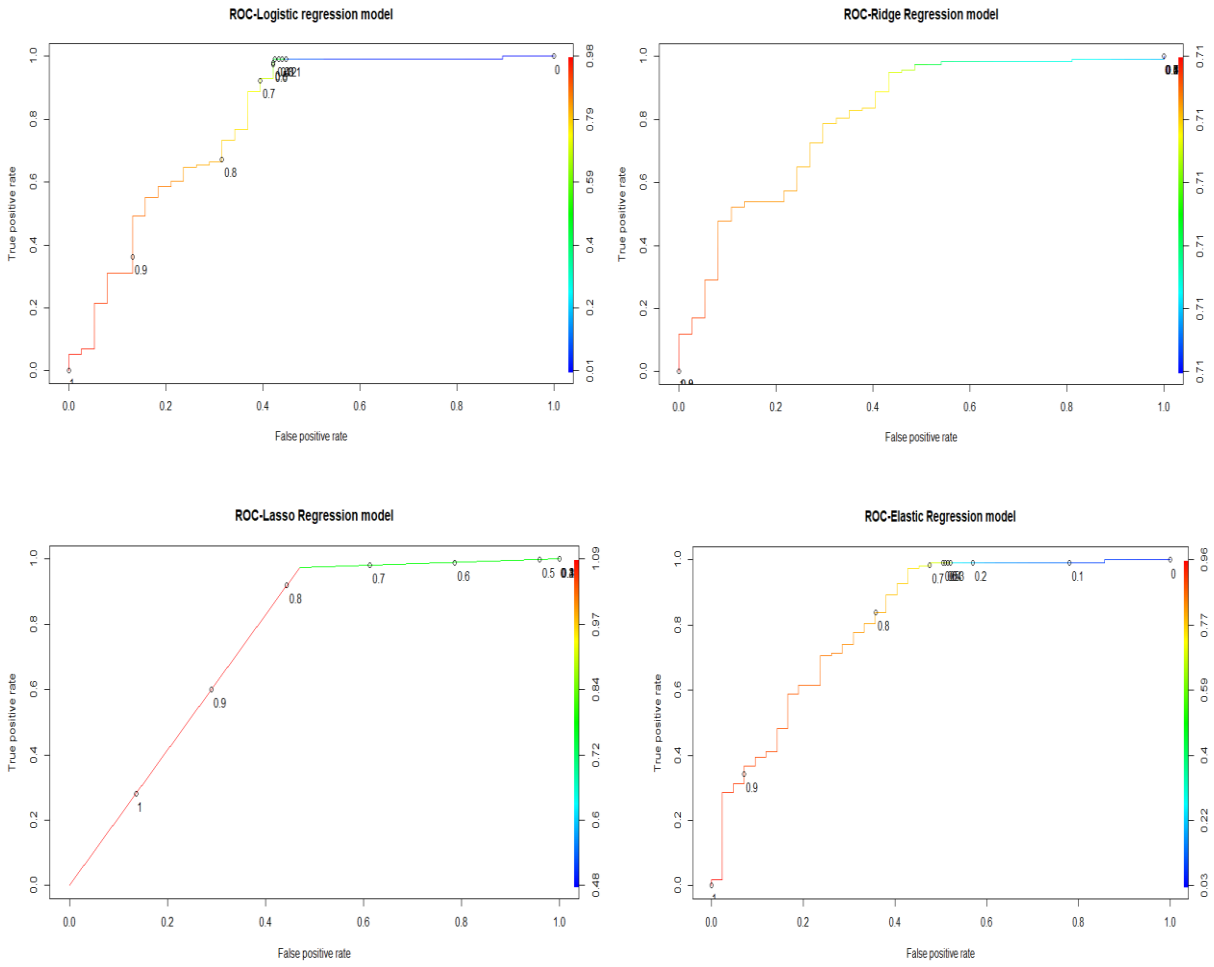

Fig 1: ROC Curves for Different Regression Models

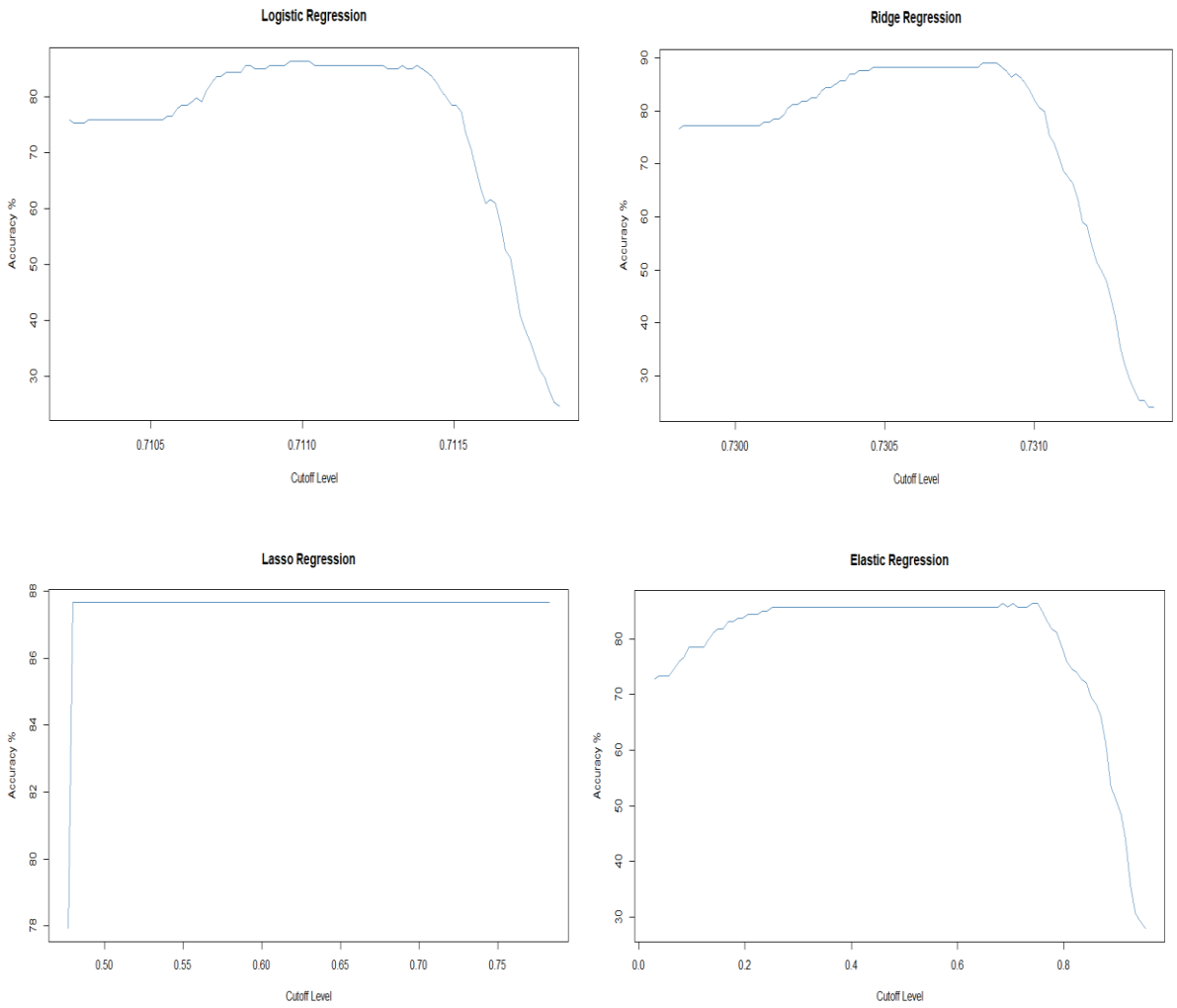

Fig 2: Accuracy vs. Cuttoffs for Different Models

Recommendation: From Fig 2, It can be observed that Lasso give higher level of accuracy over a wider range of cutoff values. The logistic regression is used as our baseline since it is easy to implement, however, its accuracy is about 85%. Similarly, the ridge and elastic regression provided an accuracy of about 85% for only a limited range of cutoff value. This make is difficult to achieve high accuracy especially in situations where the optimal threshold value is unknown apriori. I hereby recommend the use of Lasso for the accurate prediction of defaulters of loan at Donegal Community bank, Ireland. However, Lasso selects only based on Predictive power of variables. If N variables are correlated with one another, it retains only 1 of them.

Reference:

1. https://towardsdatascience.com/illustrating-predictive-models-with-the-roc-curve-67e7b3aa8914