

Predicting Credit Card Customer Churn

Joseph Gallegos, Ajay Kallepalli,
Lyndon Liang, Charles Liu, Anshuman
Mahalley

Introduction

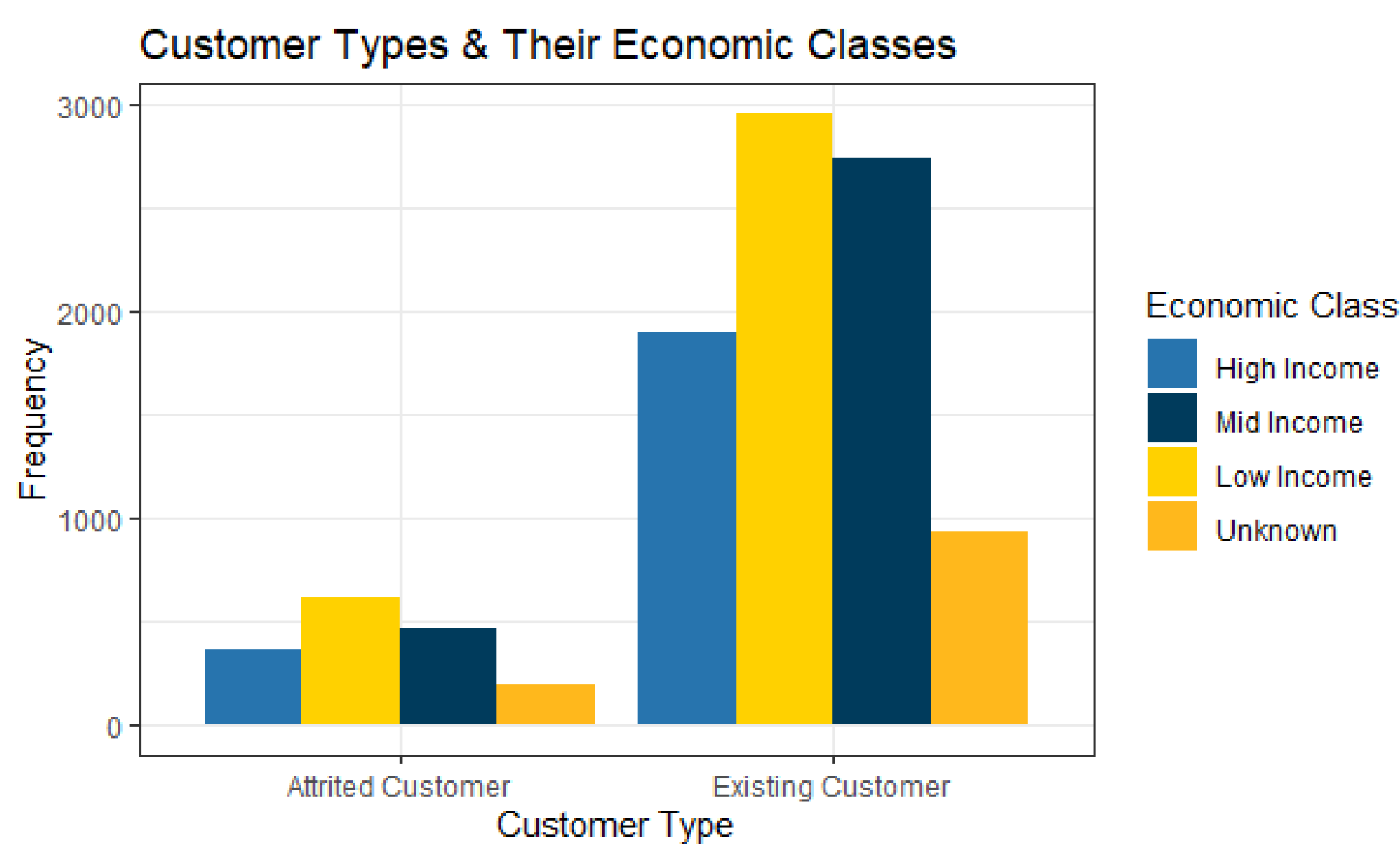
As a competitive business, a bank would like to retain their existing customers and grow their customer base. To reduce the loss of customers (or churning), we seek to investigate variables that potentially influence customer behaviors and determine how the bank can minimize losses and optimize performance.

Exploratory Data Analysis

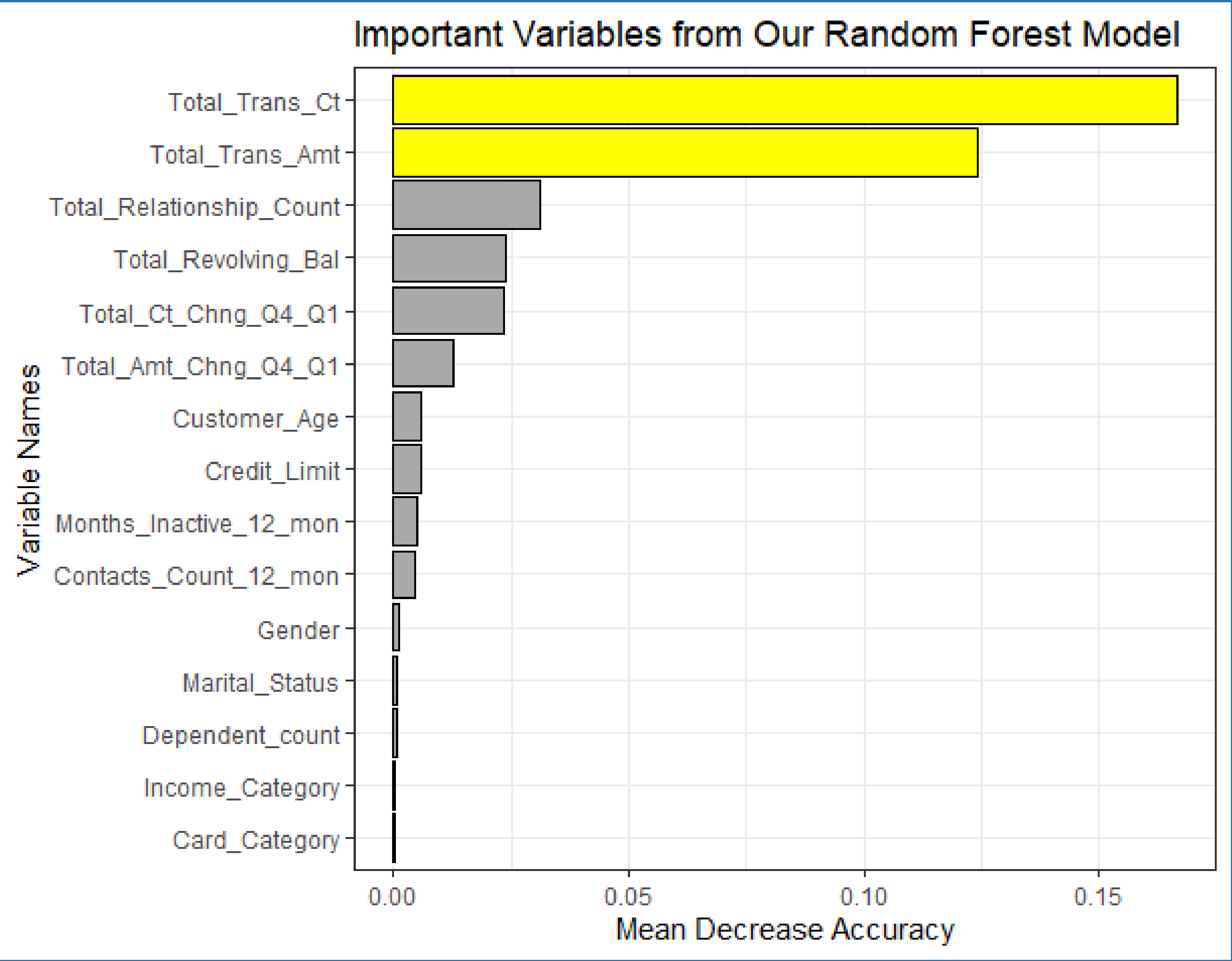
We started by modifying the character variables into factors to help with modeling. We also decided to modify the economic class variable to simplify the process of determining people's Social Economic Status. We decided that lower income status would be "Less than \$40K", middle income would be "\$40K - \$60K" "\$60K - \$80K", and high income would be "\$80K - \$120K" "\$120K +". We then plotted this variable to see if there is significant difference between the economic classes from the attrited customers and existing customers. (seen below)

Plot

Plot for Economic class with attrited vs existing customers



Banks should focus on analyzing customer transaction history and activity to accurately predict churn rather than user demographics.



Model Proposal

We considered various classification methods like logistic regression, K Nearest Neighbors and random forest to predict which customers would churn. Our final model was a random forest model with the following variables listed in decreasing order of their MeanDecreaseAccuracy: Total_Trans_Ct, Total_Trans_Amt, Total_Revolving_Bal, Total_Relationship_Count, Avg_Utilization_Ratio etc. Hyperparameter tuning was done using grid search and our best performing model had parameters mtry = 15, ntree = 500 with an accuracy of 96.45

Conclusion

We see from the model that there are certain variables that correspond to a customer being more likely to cancel their subscription. To maximize profit it would be best for a company to use this model to identify those customers that are most likely to churn and proactively try and improve their experience through better support or various other direct incentives. From our analysis we see that this prediction is best done by using customers credit card usage statistics instead of their demographics data. Ultimately there are numerous valuable and actionable insights that a company can glean from our presentation.



Customer's credit cards before and after the model!

References

Data: <https://www.kaggle.com/sakshigoyal7/credit-card-customers>
Website: <https://jfin-swufe.springeropen.com/articles/10.1186/s40854-016-0029-6>