Predictive Donor Churn Models for Duke Athletics

Hannah Marchuk, Ronald Nhondova, Sergio Wallace-Vera, Andrew Carr, Larry Cleaver, John Haws

Duke University, Duke University Office of Information Technology, Duke Athletics, Data+

Abstract

➢ Donations to Duke Athletics make a significant portion of their annual budget, so determining which donors are more likely to discontinue their donation, i.e. churn, serves financial importance.

➢ This project aimed to understand who is more likely to churn and why donors decide to churn by using internal and supplementary data to make predictive models.

➢ Using these models is useful for predicting donors that will churn, so that Duke Athletics can then target those donors with promotions to discourage churn.

Introduction

➢ Duke Athletics provided files detailing account profiles, donor profiles, season ticket, and single ticket sales which included variables such as: addresses, total years active, lifetime amount donated, ticket amount spent, and others.

➢ Other sources were included such as median income per zip code, unemployment rate per zip code, sport recruitment data, and tax rates.
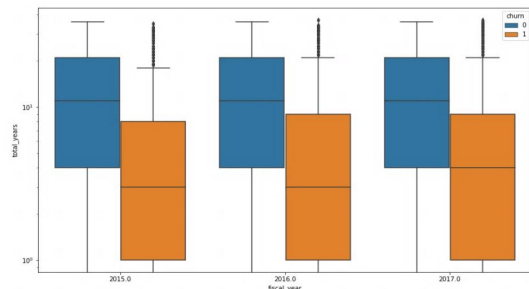
Results

Figure 1. Distribution of donor churn for total years donated from 2015-2017.

➢ Donors who donated over a longer period of time are less likely to churn (Figure 1).

➢ The Random Forest model had 84% accuracy, while the LightGBM had 86% accuracy.

➢ The Random Forest model had the following precision, recall, and f1-scores for those predicted to not have churned: 0.90, 0.92, 0.91 and the following scores for those predicted to have churned, 0.46, 0.37, 0.41.

➢ The LightGBM model had the following precision, recall and f1-scores for those predicted to not churn: 0.91, 0.92, 0.92 and the following scores for those predicted to have churned: 0.47, 0.43, 0.45.
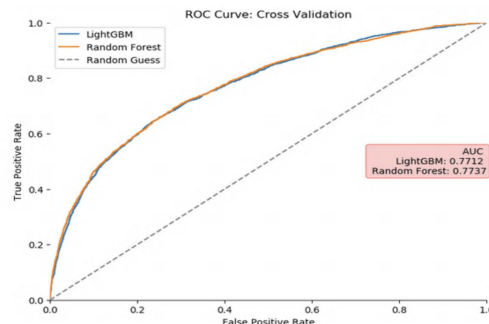

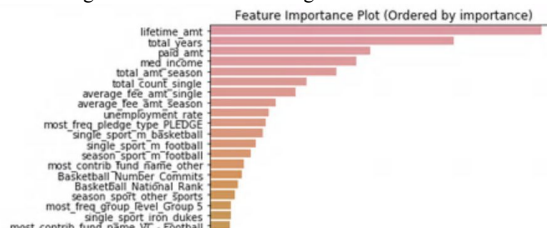Figure 2. ROC Curve for LightGBM and Random Forest

Figure 3. Feature Importance Plot

Conclusion

➢ Total years as active donor and lifetime amount donated to date are significant determinants of churn.

➢ Analysis of the dataset suggests higher churn rates for donors with low lifetime donations and low total years as an active donor

➢ The predictive model that we found with the best overall performance used LightGBM and was able to attain 86% accuracy and 68% F1-score

➢ Given more time to analyze and tune the model, there would be an emphasis on reducing false negatives and false positives. These false outcomes from the model are significant because they would result in Duke Athletics not identifying a donor that is likely to churn or use promotions on a donor that isn't likely to churn.

➢ Possible next step to this project would be analysis of churn split by voluntary and involuntary (e.g. death, financial reasons, etc.).

➢ Possibly look into gathering data from churned donors about the reasons why they stopped donating.

References

1. 2020 Basketball Team Rankings. (n.d.). Retrieved July 29, 2020, fromhttps://247sports.com/Season/2020-Basketball/CompositeTeamRankings/

2. El-Sibaie, A. (2020, June 16). 2018 Tax Brackets, Rates & Credits: Standard Deduction. Retrieved July 29, 2020, from https://taxfoundation.org/2018-tax-brackets/

3. State Individual Income Tax Rates 2000-2020. (2020, February 25). Retrieved July 29, 2020, from https://www.taxpolicycenter.org/statistics/state-individual-income-tax-rates-2000-2020

4. IPUMS NHGIS. (n.d.). Retrieved July 20, 2020, from https://data2.nhgis.org/main