Waze User Churn Project: Logistic Regression Modeling

Prepared for Waze Leadership Team

OVERVIEW

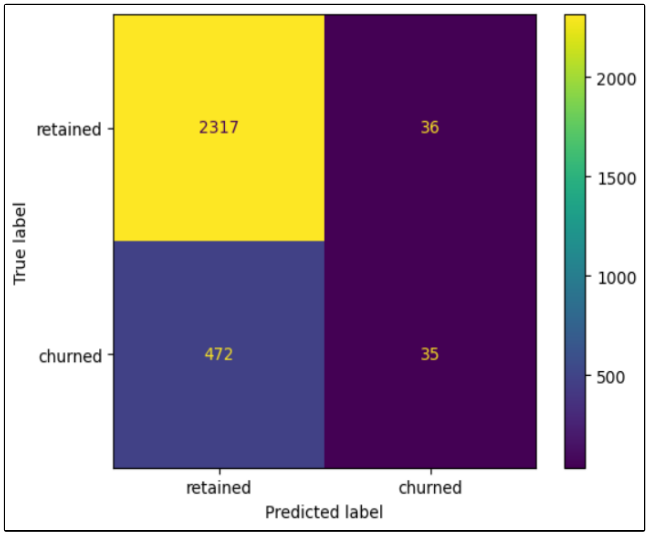
The Waze data team has conducted a data analytics project in predicting user churn and increasing the user retention rate of the app. Binomial logistic regression is useful in predicting the likelihood of user churn based on the predictive factors. For Milestone 5 of the project, this report summarizes the findings and insights made for binomial logistic regression, and how it may impact future development of the project.

PROJECT STATUS

- The goal of this project is to build a binomial logistic regression model to predict user churn.
- Binomial logistic regression is effective in making informed decisions by revealing any significant relationships between the variables as well as predicting the binary outcomes of the target variable.
- The dataset is prepared by tidying, cleaning, and organizing existing data. Following that, the logistic regression model is built and the model performance is evaluated.
- For training and testing purposes, 80% of the dataset is set as training and 20% is set as test.

KEY INSIGHTS

- Based on the confusion matrix below, the precision for churned users is average (49%), but the recall is lower than usual (7%). This means that there is a significant number of false negatives present, and the model fails to capture users who main churn effectively.
- Out of all features, it is the number of activity days that greatly influences the likelihood of user churn.



NEXT STEPS

- Given with the key insights obtained, it is not reliable to use the model in making business decisions on predicting user churn due to the low recall score metric. However, it can be useful to go further in explanatory efforts to gain additional insights by listing additional factors that may lead to user churn, which in turn can lead to greater user retention. Perhaps further study should be taken into account by trying out different features like professional drivers and total distance driven per day.