Waze User Churn Project – Machine Learning Model

Executive Summary Report V

Milestone 6

Project Overview

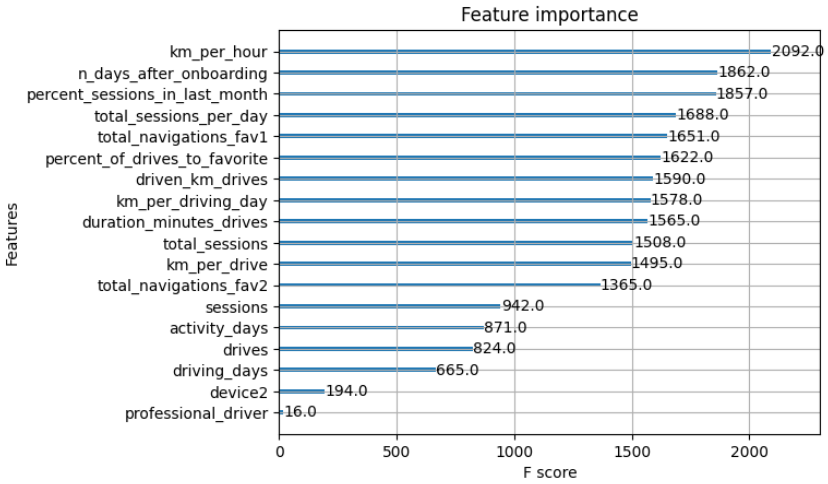
The project focuses on boosting growth by reducing the rate of monthly user churn in the Waze app. Churn refers to users who have either uninstalled the app or ceased using it. Our objective is to develop a machine learning model to predict user churn using the provided data.

Key Insights

- Engineered features made up 6 of the top 10 predictors, including speed, percentage of sessions in the last month, total sessions per day and distance per driving day.
- The XGBoost model outperformed the random forest model in data fitting. Additionally, the model's recall score (18%) nearly doubled compared to the previous logistic regression model, while maintaining similar accuracy and precision.
- The tree-based model ensembles performed better across evaluation metrics than a single logistic regression model, requiring less data preprocessing, though they are less interpretable.

Details

To build a model with optimal predictive capabilities, the Waze data team created and compared two models: a random forest and XGBoost. The data was divided into training, validation and test sets for this analysis.



Next Steps

1. Collect data at the individual drive level (e.g. driving times, locations), as well as more detailed user interaction data (e.g. frequency of reporting road hazards.
2. Track the number of unique starting and ending points per user each month.
3. Perform a second iteration of the User Churn Project with more refined data.