# Waze User Churn Project - Machine Learning Models

Executive Summary Report

## ISSUE / PROBLEM

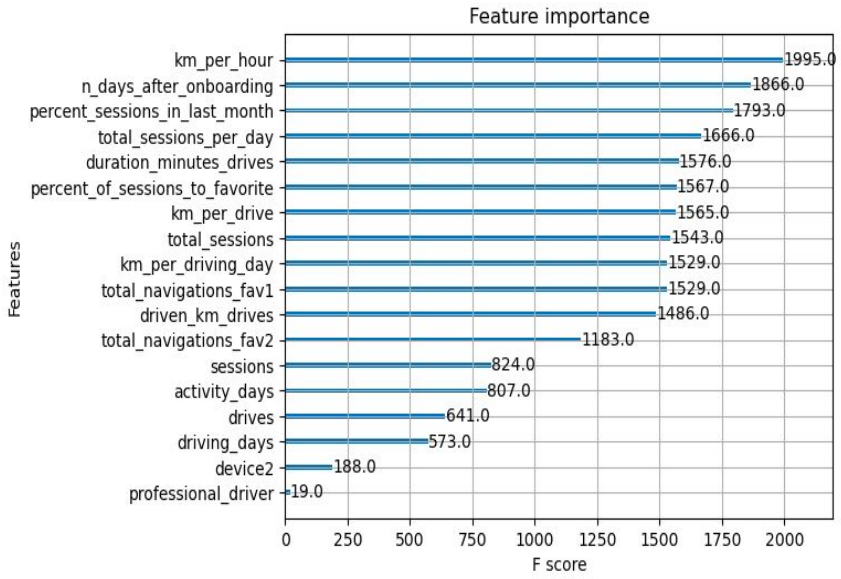
The Waze data team is looking for ways to reduce user churn rates. This stage of the project consisted of developing machine learning models to predict whether or not a user would churn. Churning is classified as when a user uninstalls / stops using the Waze app. This stage of the project is part of an overall company objective of increasing growth.

## RESPONSE

In order to create a model that would best predict whether or not a Waze user would churn we developed two different machine learning models. A random forest model and a XGBoost model. We compared the scoring metrics of both models to choose a champion model.

The data was split into 3 different sets, training, testing and validation. Splitting the data 3-ways comes with a trade off, this means there is less data to utilize for the testing of the champion model but we are able to allow all the testing data to be used by the champion model which provides a better prediction of future performance.

## KEY INSIGHTS




- 6 of the top 10 predictive features were created during the feature engineering portion of this project
- The XGBoost model fit the the data better than the random forest model did
- When compared to linear regression models, the ensemble of tree based machine learning models are more valuable when predicting user churn rates. This is due to the fact that they achieved higher scores on each scoring metric utilized and required less preprocessing of the data.

## IMPACT

The developed machine learning models would be suitable for supplemental usage but would not be recommended to be used as a user churn predictor as the model currently stands.

It is evident that there is a need for additional data in order to improve the reliability of the model. Data linked to how users are actively using the Waze application such as whether a user actively participates in confirming / reporting road hazards and warnings could be helpful in improving the model.

As can be seen in the graphic to the right, engineered features played a significant role in the prediction of whether a user will churn or not so looking to rework the project and engineer more features with newer data is recommended.