Predicting User Churn | Machine Learning Results

Executive Summary for Waze Data Team

Prepared by: Rebecca Iglesias-Flores

ISSUE / PROBLEM

The **Waze Data Team** is developing a **machine learning model** to predict user churn, which measures the number of users who stop using or uninstall the app. This initiative is part of a broader strategy to drive growth at Waze.

At **Milestone 6**, the model has been **built and evaluated**, yielding key insights that could influence the project's next steps.

IMPACT

The **Milestone 6 models** highlight a critical need for **additional data** to improve churn prediction accuracy. Current data alone is insufficient for consistently predicting user churn.

Incorporating **drive-level data** (e.g., drive times, geographic locations) and **more granular user interaction data**—such as road hazard reports and confirmations—could enhance model performance. Additionally, tracking the **monthly count of unique starting and ending locations** per driver may provide further insights.

Given the demonstrated impact of **engineered features**, the Waze team recommends a **second iteration** of the **User Churn Project** to refine predictions with enriched data.

RESPONSE

	model	precision	recall	F1	accuracy
0	Best XGB test	0.418502	0.187377	0.258856	0.809790
0	XGBoost Classifier CV	0.417392	0.168856	0.240105	0.810700
0	Best XGB val	0.386473	0.157791	0.224090	0.806294
0	Best RF val	0.450000	0.124260	0.194745	0.817832
0	RandomForestClassifier CV	0.464125	0.122185	0.193340	0.819442

We conducted **cross-validation** on two models—**Random Forest Classifier** and **XGBoost Classifier**—to select the one with the highest predictive power.

While both models achieved **high accuracy**, they underperformed on our **preferred evaluation metric**. The best recall score came from the **XGBoost classifier**, correctly identifying **18.7% of users who actually churned**.

KEY INSIGHTS

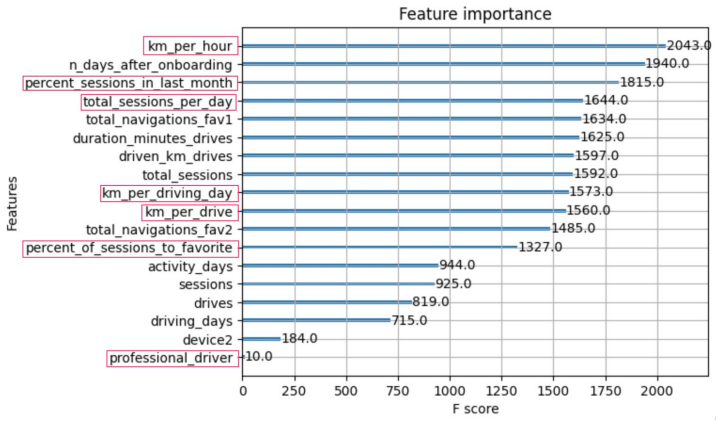

Figure 2. Feature importance of XGBoost best extimator

Six of the top ten predictive features were engineered, including **km_per_hour, percent_sessions_in_last_month, total_sessions_per_day, percent_of_drives_to_favorite, km_per_drive, and km_per_driving_day**. These features played a critical role in model performance, as highlighted in the accompanying visual.

The **XGBoost model outperformed the random forest model**, demonstrating a better fit to the data. Notably, the **recall score (17%) nearly doubled** compared to the previous logistic regression model in Milestone 5, while maintaining **comparable accuracy and precision**.

The ensemble-based tree models used in this milestone proved to be more effective than a standalone logistic regression model, yielding **higher performance across all evaluation metrics** while requiring **less data preprocessing**. However, the tradeoff is reduced interpretability, making it more challenging to understand how individual predictions are made.