Evaluating Machine Learning Models to Predict User Churn

Milestone 6

❯ ISSUE / PROBLEM

The Waze data team is working to improve user retention by predicting which users are likely to churn—defined as users who uninstall or stop using the app. Understanding churn behavior is essential to help Waze take proactive steps in retaining its user base. In this phase of the project, the team built and tested machine learning models to identify churn patterns based on user activity data. This report summarizes the model development, performance outcomes, and recommendations for future improvements.

❯ IMPACT

The ML models developed in Milestone 6 highlight the potential of using user activity data to predict churn, but also reveal several limitations that impact performance. While the Random Forest model achieved reasonable recall, precision and overall predictive power remain moderate. This suggests that additional, more detailed data could improve model accuracy. For example, drive-level metrics such as route frequency, drive duration variability, or user engagement with app features (e.g., reporting traffic issues or interacting with suggested routes) may offer better predictive signals. The Waze team recommends a second phase of the User Churn Project focused on collecting and integrating richer behavioral data and expanding feature engineering efforts to refine and improve model performance.

❯ RESPONSE

To build a robust prediction model, the Waze data team developed and tested both Random Forest and XGBoost classifiers. The dataset was divided into training, validation, and test sets to ensure reliable model evaluation. This three-way split allowed for unbiased model selection using the validation set and an accurate estimate of final performance using the test set. Both models were tuned using grid search with cross-validation and evaluated using multiple metrics, with recall prioritized. Random Forest showed stronger performance in recall and generalization across all datasets. Based on these results, Random Forest was selected as the champion model. Its test performance closely matched validation results, indicating stable and consistent behavior.

❯ KEY INSIGHTS

- Engineered features accounted for six of the top 10 most important predictors: `activity_days`, `km_per_driving_day`, `total_sessions_per_day`, `km_per_hour`, `km_per_drive`, and `percent_of_sessions_to_favorite`.
- The Random Forest model outperformed XGBoost in terms of recall, making it the preferred model for predicting user churn.
- On the test set, the Random Forest model achieved a **recall of 65.4%**, **precision of 33.7%**, **F1 score of 44.5%**, and **accuracy of 70.1%**.
- The confusion matrix revealed that the model correctly predicted **342** of the **523** users who actually churned, with **181 false negatives** and **673 false positives**.
- Despite the lower precision, the high recall is aligned with business priorities, as missing actual churners is riskier than incorrectly targeting retained users.
- Feature importance analysis showed that user activity metrics such as `activity_days`, `driving_days`, and `km_per_driving_day` are the strongest indicators of churn.