MACHINE LEARNING MODEL THAT PREDICTS USER CHURN

Regression Modeling analysis

OVERVIEW

The Waze Data Team is developing an analytics project to reduce monthly user churn and support app growth. Churn refers to users who uninstall or stop using the app. To address this, the team applied binomial logistic regression, a predictive model aiding business decisions.

🎯 Goal: Use user data to build and analyze a binomial logistic regression model.

PROJECT STATUS

Milestone 5 - Regression Modeling

🎯 Methods: Created relevant features, checked for multicollinearity, built the model, and evaluated its performance.

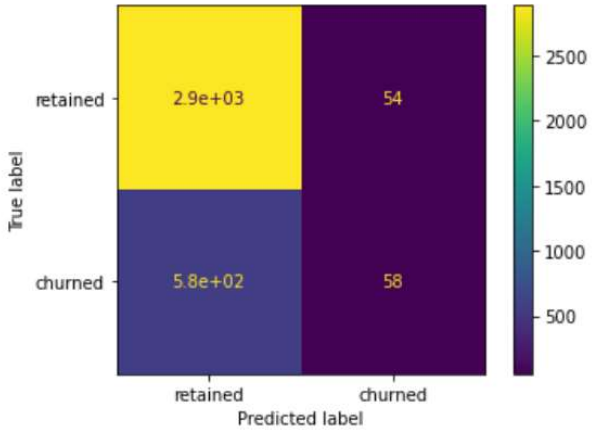
🎯 Impact: With adequate data, logistic regression helps identify key variable relationships and predict binary outcomes, supporting strategic decisions in areas like marketing and product development.

NEXT STEPS

*Based on the model's results, we recommend using insights from this milestone to guide further analysis.

*While not yet suitable for major business decisions, the model highlights the need for richer feature data linked to churn and suggests refining the target user profile. These steps are crucial to support Waze's goal of reducing monthly churn and driving user growth.

KEY INSIGHTS



The effectiveness of the binomial logistic regression model was assessed using accuracy, precision, and recall.

While the model achieved a moderate precision of 53%, it exhibited very low recall, identifying only 9% of actual churned users.

Among all features, activity_days emerged as the most significant predictor, showing a negative correlation with churn—suggesting that users who are more active are less likely to leave.

In contrast, although exploratory data analysis (EDA) previously indicated that churn increased with higher values of km_per_driving_day, this variable turned out to be among the least influential in the final model.