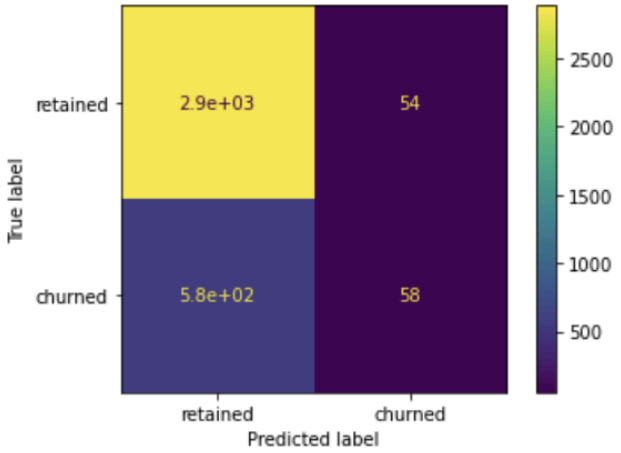
Project: Waze Monthly User Churn Predictive Model

Regression Modeling

Overall Project Goal: Increase app growth by creating a predictive model that predicts Waze monthly user churn and accurately identifies who, when, and why users churn.

Memo Objective: This report offers an evaluation and insights derived from a binomial logistic regression model on provided user data. Of note, a binomial logistic regression model offers predictive power for categorical dependent variables, in this case user churn or retention status.

Results:



The figure on the left displays the confusion matrix, which represents how accurate the logistic model is at predicting user churn and retention rates based on predictor variables.

Key Insights

- The variable *Activity_days* was the most important feature in the model, as it had the largest negative correlation with user churn.
- Previous EDA indicated that user churn rates increased as the values in the variable *km_per_driving_day* increased. In the model, *km_per_driving_day* was the second-least-important variable.
- The model has mediocre precision (53% of its positive predictions are correct) and very low recall, with only 9% of churned users identified. Thus the model makes a lot of false negative predictions and fails to capture users who will churn.

Next Steps

Given the model's low efficacy in recall and precision, this model should not be used to make significant business decisions. However, its low predictive performance implies a need for additional data that may correlate with user churn.