Fraud Detection Model Performance Report

An interactive analysis of Poundbank's model degradation.

Summary Performance Analysis Root Cause Recommendations

Executive Summary

This analysis reveals a significant drop in the fraud detection model's accuracy, primarily due to an evolution in fraudulent transaction patterns. The model's performance degraded sharply in the second quarter of 2019 as fraudsters altered their behavior regarding login times and transaction amounts. This report breaks down when the performance dropped, why it happened, and what steps should be taken next.

Alert Period

Apr - Jun 2019

Months with confirmed performance degradation.

Primary Cause

Data Drift

A shift in user behavior patterns.

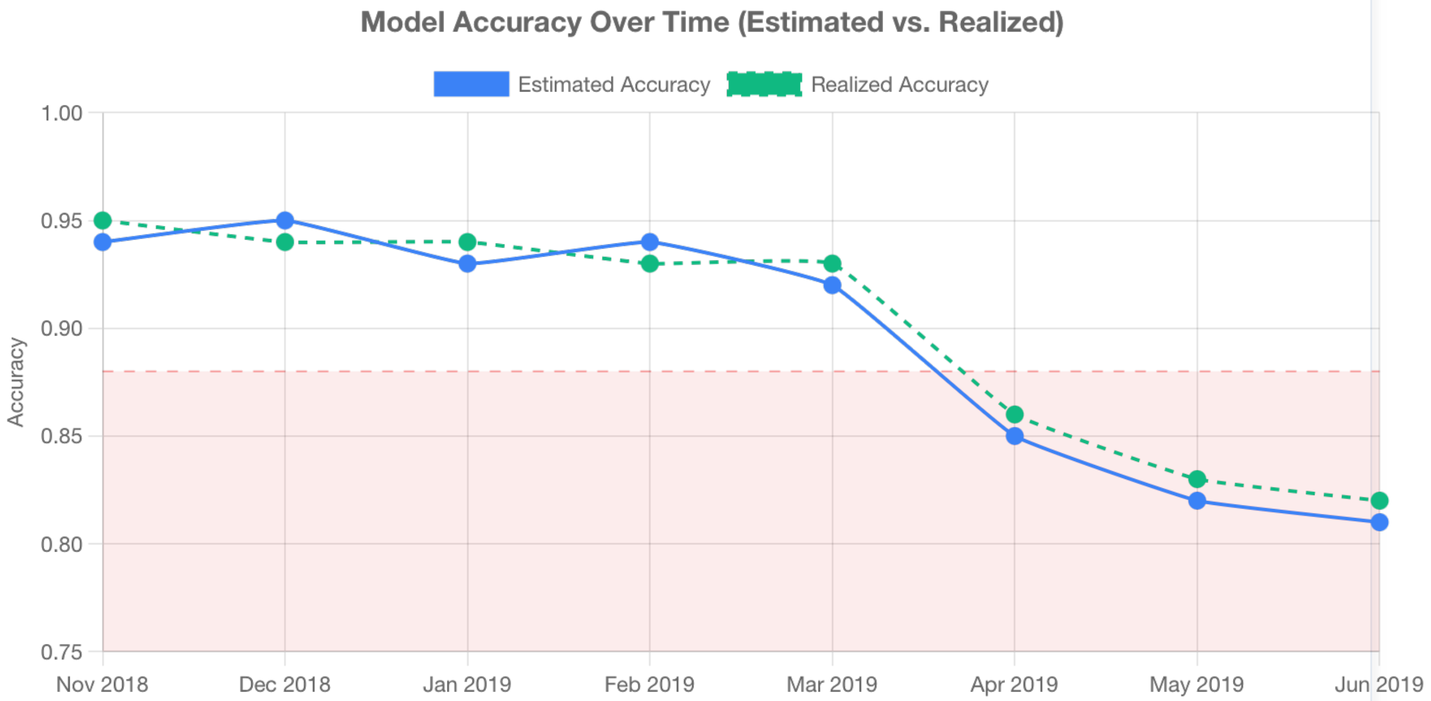
Top Correlated Feature

time_since...

The strongest indicator of performance drop.

Model Performance Analysis

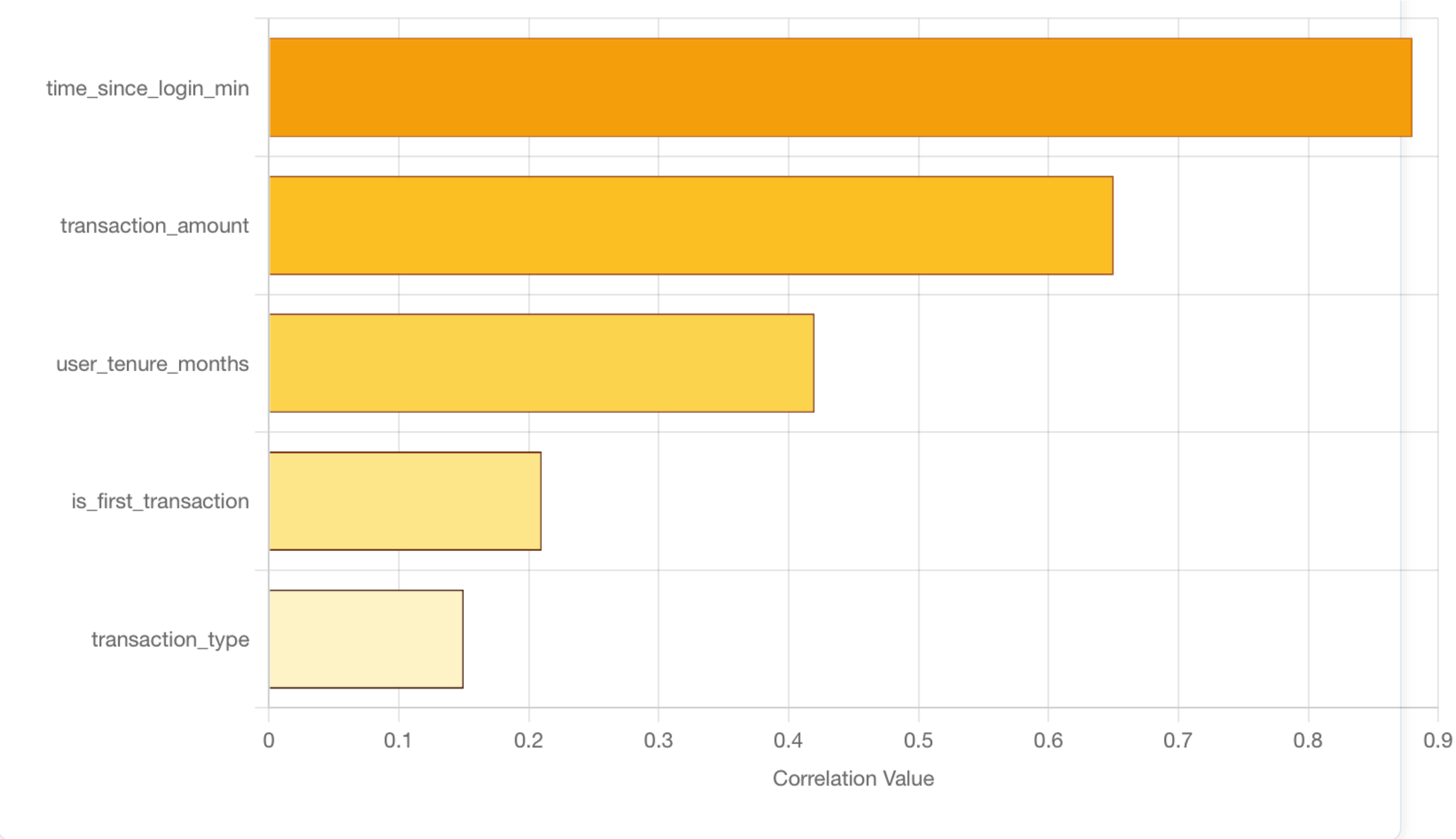
The model's accuracy was tracked using two methods: estimated performance (CBPE), which provides an early warning, and realized performance, calculated after true fraud labels are available. The chart below shows a clear alignment between both metrics, confirming a severe performance drop starting in April 2019. The red shaded area highlights the period where accuracy fell below the acceptable threshold, triggering alerts.



Root Cause: Data Drift

Data drift occurs when the patterns in the live data diverge from the data the model was trained on. Our analysis ranked features by how strongly their drift correlated with the performance drop. `time_since_login_min` was the clear primary driver. This indicates that fraudsters are no longer making transactions immediately after logging in, a pattern the original model relied on for detection.

Feature Drift Correlation with Performance



Conclusion & Recommendations

The investigation confirms that the model's declining accuracy is a direct result of evolving fraud tactics. To restore and maintain the model's effectiveness, we propose the following actionable steps.

1. Model Retraining

The model must be retrained using recent data that includes these new fraudulent patterns. This will allow it to learn the evolved behaviors related to login times and transaction amounts, restoring its predictive accuracy.

2. Continuous Monitoring

Implement a permanent data drift and performance monitoring system. This will provide early warnings of future pattern shifts, enabling proactive model maintenance instead of reactive fixes after performance has already degraded.