

HDBSCAN-based SecureRide Fraud Detection System

In the dataset collected between spring 2022 and spring 2023, data analysts using Tableau dashboards identified a concerning trend: the cancellation rate had risen from 5 percent of accepted proposals to 10 percent. Subsequently, through diligent monitoring of the location tracker service responsible for relaying drivers' locations, it was revealed that certain drivers who canceled trips were surreptitiously completing the journey to the canceled destination. The motive behind this behavior was an attempt to evade commission fees imposed by the ride-hailing company. To address this issue and distinguish between drivers canceling trips for legitimate reasons and those engaging in fraudulent activities, a dedicated project has been initiated.

- Data collection and processing focused on cancellation rates, gross income, and average ride time to identify potential fraud indicators.
- Implementation of HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) replaced traditional k-means clustering for grouping drivers and distinguishing fraudulent ones.
- Utilization of domain analysts led to the identification and blocking of fraudulent driver groups, resulting in the prevention of further fraudulent activities.

This transformative project has yielded remarkable results, successfully identifying and blocking a thousand fraudulent drivers and concurrently reducing the cancellation rate from 10 percent to an impressive 6 percent. The strategic implementation of anti-fraud measures, coupled with public promotion of these features within the application significantly contributed to a boost in the company's income.