Unsupervised Fraud Detection Analysis

Clustering-Based Techniques for Atypical Transaction Identification

Prepared by: Junior Data Scientist

Date: July 2025

Problem Statement

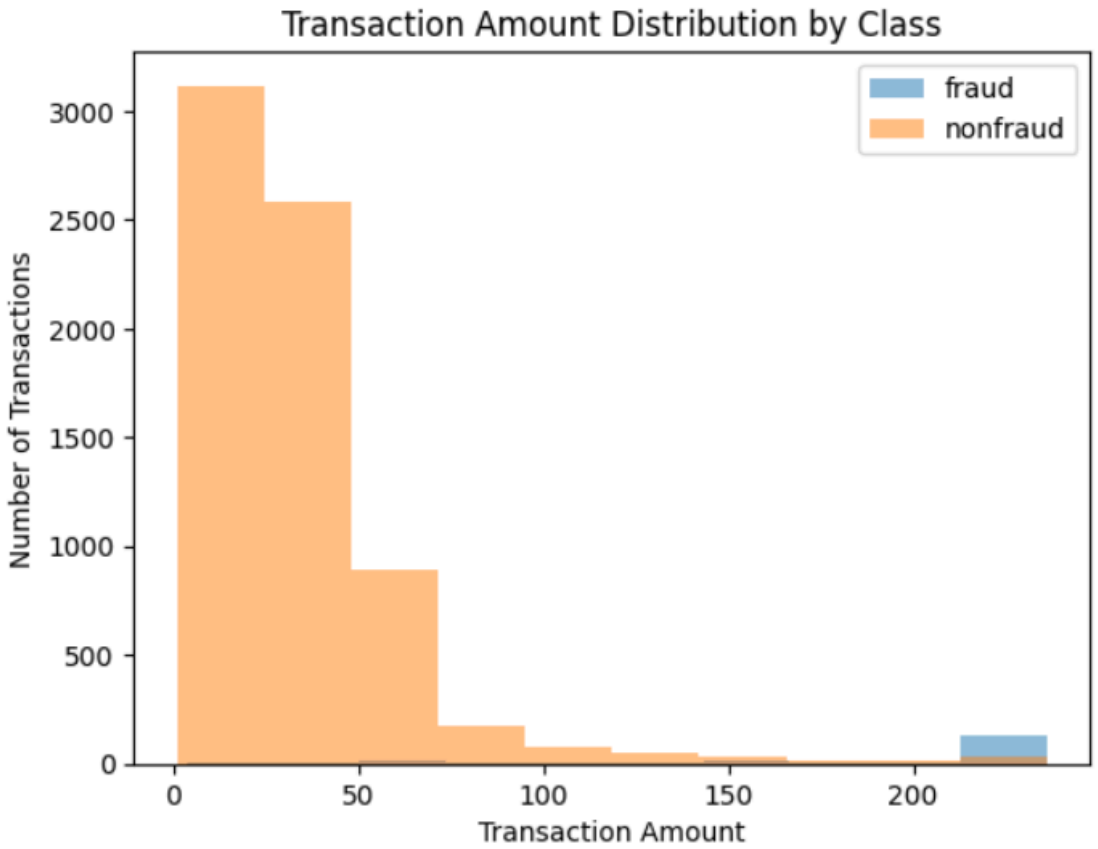
This project leverages unsupervised machine learning and exploratory data analysis to uncover hidden fraud risks within banking transactions. By analyzing demographic and behavioral transaction patterns, we segment customer activity and flag anomalies—under the core hypothesis that fraudulent behavior often deviates from typical customer profiles. These insights empower fraud prevention teams with early-warning signals and data-driven strategies to safeguard customer trust and minimize financial losses.

Exploratory Data Analysis

To prepare the banking dataset for clustering-based fraud detection, we performed a comprehensive preprocessing pipeline. Categorical variables such as age group, gender, and transaction category were transformed using one-hot encoding, enabling the model to interpret each category as a separate binary signal.

Numerical and boolean variables—including transaction amounts, balance history, account tenure, and high-risk behavior flags—were scaled to a [0,1] range using MinMaxScaler, ensuring uniform influence across features regardless of their original units or distributions. This step is critical for distance-based models like K-Means and DBSCAN, which are sensitive to feature magnitude.

The resulting feature matrix consisted of 30 engineered attributes, combining both customer characteristics and transaction-level data, applied to a cleaned sample of over 7,200 transactions. This rich, normalized dataset enabled robust detection of abnormal patterns while preserving interpretability for business stakeholders.



KMeans Clustering

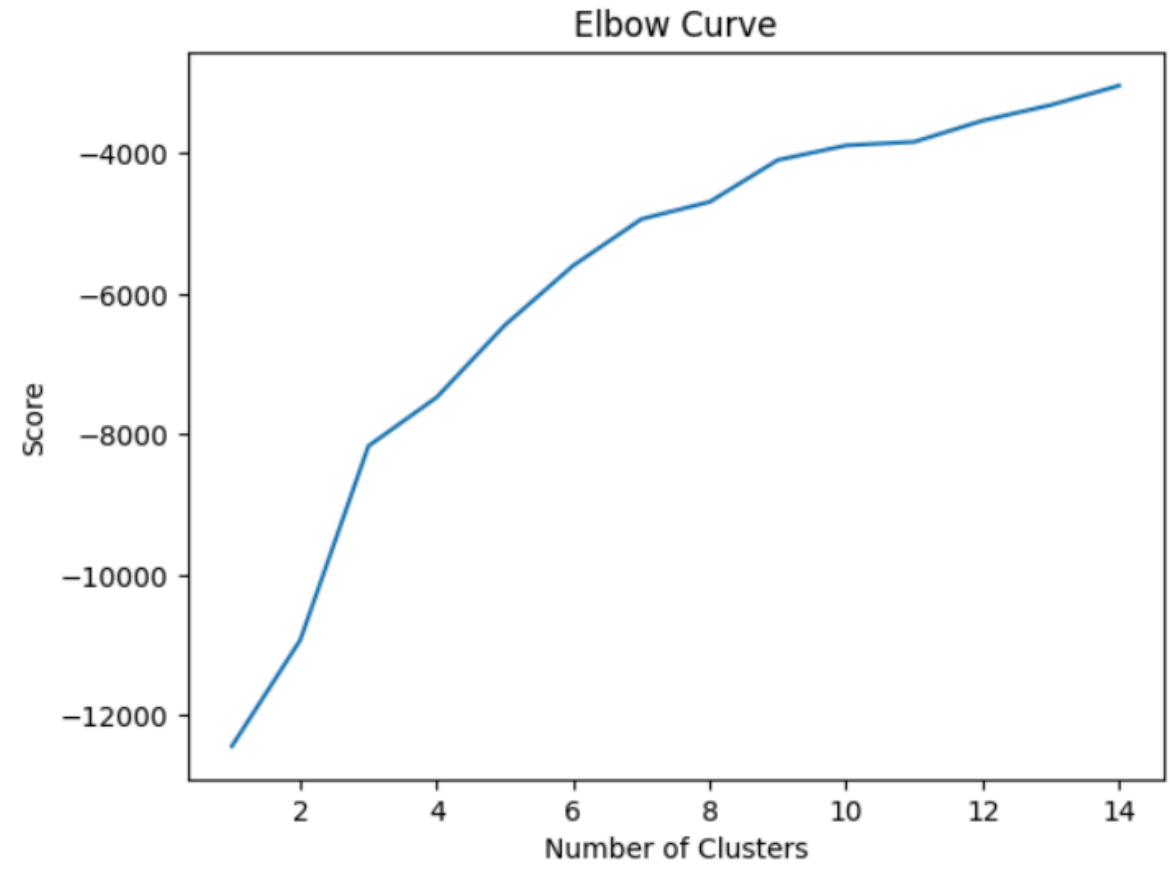
To efficiently scale clustering for fraud detection, we employed MiniBatchKMeans, a faster variant of KMeans optimized for large datasets. This approach allowed us to iteratively update cluster centroids using small random batches, significantly reducing computational time while maintaining model stability.

To determine the optimal number of clusters, we applied the Elbow Method, analyzing the trade-off between model complexity and within-cluster variance. After identifying the ideal cluster structure, we calculated the Euclidean distance from each transaction to its assigned cluster center—a proxy for how unusual or atypical the behavior is compared to its peer group.

Transactions in the top 6% of distance scores—those farthest from any behavioral cluster—were flagged as potential fraud. This anomaly-based strategy operates under the assumption that fraud often deviates sharply from known patterns.

The model achieved a recall score of 1.0, successfully identifying all known fraudulent cases, along with an ROC AUC of 0.98, highlighting strong discriminatory power. This demonstrates that our unsupervised pipeline can surface high-risk cases

with minimal false negatives—providing an effective first-layer filter for fraud analysts and decision-makers.

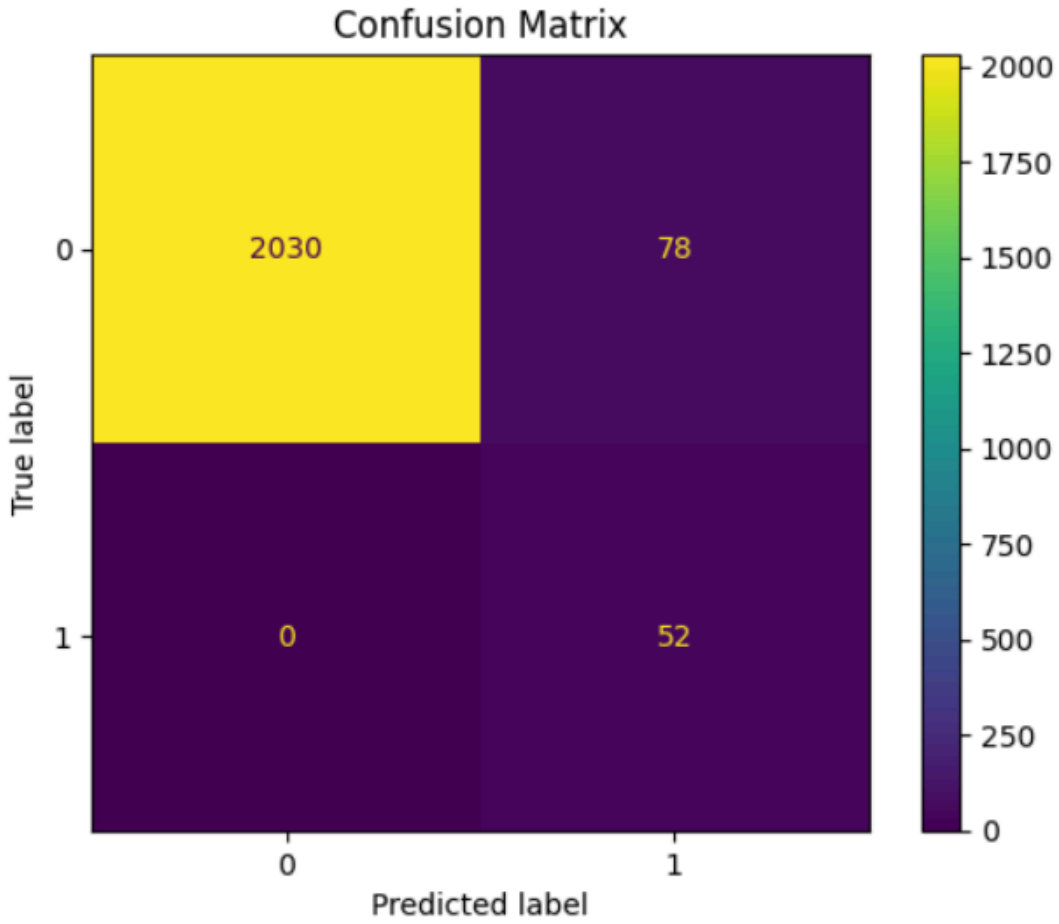


Model Building and Performance

MiniBatchKMeans proved to be a highly effective unsupervised method for flagging atypical (high-risk) transactions, successfully capturing 100% of known frauds. While the method incurred some false positives (normal transactions misclassified), this is an acceptable trade-off in early-stage fraud detection pipelines, where high recall is critical to minimize missed fraud.

The results confirm that distance to cluster center is a valuable anomaly score, and MiniBatchKMeans can serve as a scalable first-line filter in real-world fraud monitoring systems.

	precision	recall	f1-score	support
Normal	1.00	0.96	0.98	2108
Fraud	0.40	1.00	0.57	52
accuracy			0.96	2160
macro avg	0.70	0.98	0.78	2160
weighted avg	0.99	0.96	0.97	2160



DBSCAN Clustering

As an exploratory alternative to centroid-based clustering, DBSCAN (Density-Based Spatial Clustering of Applications with Noise) was tested to identify fraudulent transactions as density-based outliers. Unlike KMeans, DBSCAN does not require specifying the number of clusters in advance—making it suitable for detecting small, irregular patterns often associated with fraudulent behavior.

Multiple epsilon (ε) values were tested to tune the model's sensitivity to neighborhood density. Instead of relying solely on DBSCAN's built-in outlier label (-1), we took a novel approach by analyzing the three smallest clusters formed across parameter sweeps. The

rationale was that fraud may not always appear as complete outliers but may instead cluster in small, behaviorally similar groups with low density.

This method successfully flagged 190 known fraudulent transactions, indicating reasonable precision and strategic value despite a lower recall than MiniBatchKMeans. While not optimal as a standalone detector, DBSCAN proved useful for highlighting micro-patterns of fraud and supporting a multi-layered detection strategy.

Final Insights

This project demonstrates the value of unsupervised learning in detecting fraudulent banking transactions, especially in contexts where label availability is limited or unreliable. Through careful feature engineering, scaling, and exploratory data analysis (EDA), we created a well-prepared dataset for clustering-based anomaly detection.

We applied two complementary approaches:

- MiniBatchKMeans was used to flag outliers based on distance to cluster centroids, achieving a perfect recall (1.0) and ROC AUC of 0.98, making it an effective first-layer filter.

- DBSCAN, although not optimized for full coverage, helped uncover low-density micro-clusters of suspicious activity, capturing 190 out of 300 known frauds by focusing on the smallest behavioral clusters — a novel strategy reflecting how fraud can appear as small, similar subgroups rather than random noise.

While both models operate under different assumptions, their combined insights offer a multi-layered fraud detection framework: KMeans excels at broad anomaly scoring, while DBSCAN enhances discovery of subtle behavioral fraud patterns.

Together, this hybrid unsupervised approach delivers interpretable, scalable, and strategically useful fraud detection tools for organizations aiming to act early and confidently, even in the absence of clean labels.