Anomaly Detection in Fraud

1. Overview

Fraud could be considered as an anomaly problem to solve, and we could leverage probability and statistical approaches as well as ML/DL approaches for such problem. In the past several years, graph-based fraud detection is also getting more and more popular.

The rationale behind using anomaly detection to solve fraud detection problem is because fraudulent behaviors are usually not common and are concealed.

The challenges in fraud detection is that fraud is constantly changing and adapting and even organized, so as data scientists/researchers, we should remain vigilant, continuously update our techniques/knowledge to identify current anomalies/fraud in the system, and collaborate across industries to stay ahead of these evolving threats (prevention).

2. Summary of Techniques in Anomaly Detection (this list will keep updating)

Probability and Statistical Approaches	Machine Learning Approaches
• Z-scores/Robust Z-scores • IQR Rule • Mahalanobis Distances/Robust Mahalanobis • Benford's Law • Gaussian mixture models	• Local Outlier Factor (LOF) • Isolation Forests • Classifier-Adjusted Density Estimation (CADE) • One-Class Support Vector Machine (SVM) • k-Nearest Neighbors (kNN) • Neural Networks

Supervised	Unsupervised
• Supervised Neural Networks • Support Vector Machine learning • K-Nearest Neighbors Classifier	• Mahalanobis Distances/Robust • Cluster Analysis (e.g. density-based clustering algorithm, K-Means)

3. Implementation with Python