

Enhancing Financial Fraud Detection Using Advanced Anomaly Detection Models in Banking Transactions

Presented by: Vishwas Anand

Jain University, MBA (Business Intelligence and Analytics)

Faculty Guide: Dr. Shalini



by Vishwas Anand



Executive Summary: Project Overview



Purpose

Investigate and create next-generation anomaly detection models to improve financial fraud prevention and detection in banking transactions.



Significance

Addresses the evolving landscape of financial fraud, where perpetrators leverage AI and automation, making traditional methods less effective.



Impact

Aims to reduce financial losses, improve operational effectiveness, and safeguard customers from sophisticated fraud attacks in a digital environment.



Introduction & Background: The Evolving Threat

Importance of Fraud Detection

Crucial for safeguarding customer assets and upholding trust in the banking sector amidst increasing regulatory demands and operational limitations.

Challenges with Traditional Systems

Legacy rule-based systems suffer from high false positives, slow detection, and lack of adaptability to new fraud patterns, leading to inefficiency and increased costs.

Rise of Digital Banking & New Threats

Booming digital transaction volumes increase attack surfaces. Scammers use AI, generative AI, and deep fakes for sophisticated scams, including synthetic identity fraud (expected to cause over \$23 billion in U.S. losses by 2030).

Faster payment systems introduce vulnerabilities, requiring real-time detection tools to flag suspicious transactions within milliseconds.



Objectives of the Study: Key Goals



Develop & Evaluate Models

Create sophisticated anomaly detection models, including machine learning (isolation forests, autoencoders, gradient boosting) and graph neural networks, for efficient fraud detection.



Address Key Challenges

Tackle issues like data imbalance, real-time detection, model interpretability, and privacy regulations (e.g., GDPR compliance).



Review of Literature: ML & Anomaly Detection

Machine Learning Dominance

Systematic reviews (Polak et al., 2024; Mustika et al., 2025) show ML's prevalence, especially supervised learning for credit card fraud. Ensemble and deep learning models (LSTMs) excel with temporal data.

Deep Learning & Unsupervised Methods

Advanced architectures like CNNs and LSTMs achieve high accuracy. Unsupervised techniques (autoencoders, isolation forests) identify new fraud patterns without labeled data.

Hybrid & Graph-Based Models

Combining multiple ML models enhances detection. Graph Neural Networks (GNNs) detect complex relationships in fraud rings, revealing patterns feature-based models miss.

Explainable AI (XAI)

XAI methods (SHAP, LIME) provide transparency for regulatory compliance and trust by explaining model predictions, transforming "black-box" models.

Research Methodology: Approach & Data



Dataset Description

Utilized a random sample of 200 banking transactions from a larger Kaggle database, ensuring diversity in transaction attributes (value, type, location, device, customer behavior).



Data Analysis Approach

Quantitative, experimental design involving data preprocessing, exploratory data analysis (EDA), and feature engineering.



Models & Techniques

Employed logistic regression, random forests, isolation forests, autoencoders, and graph neural networks. Evaluated performance using accuracy, precision, recall, F1-score, AUC, and detection latency.

RESEARCH METHODOLOGY



data load

Intoe recial sahn ancerdars.



Datals training



①+

data proal cussing

Strore cefcerden an sthotall pmaite and pericuring the ssedate.

②+

modelevaluation

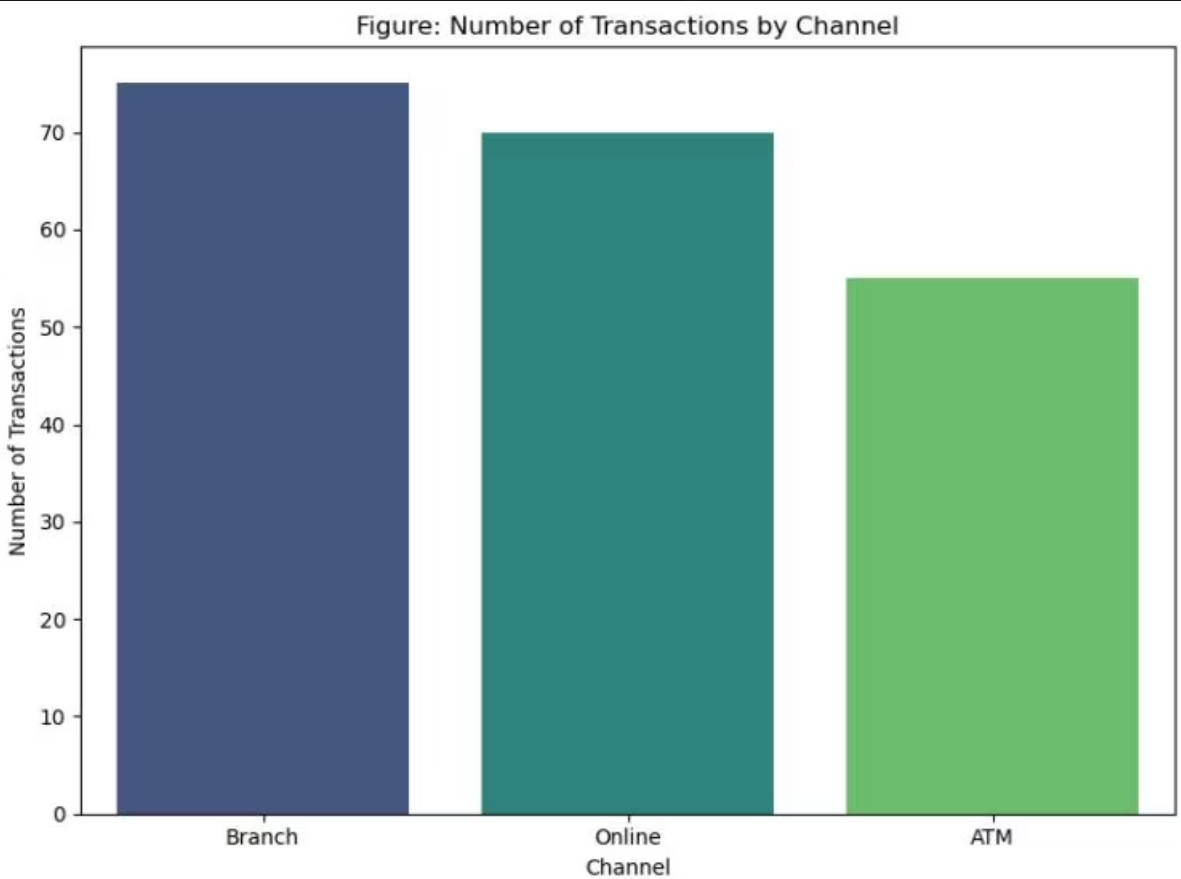
Not the polduar efftestunces dom essonathe, at the legerings evahcofits.



Data Analysis & Key Findings

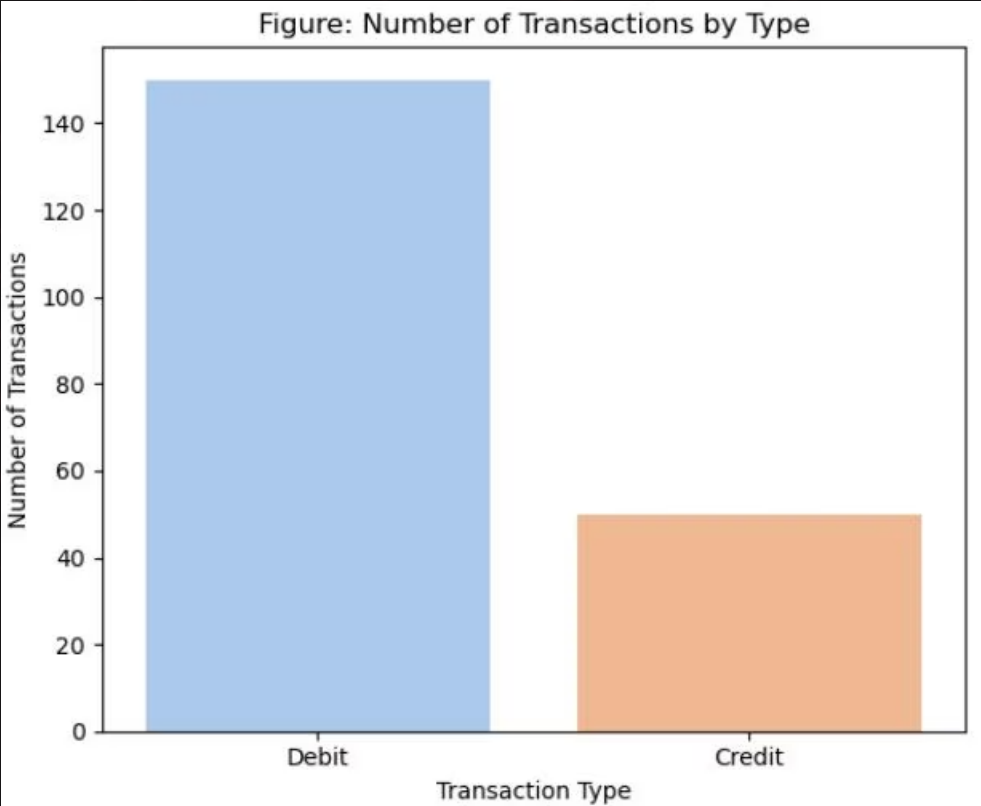
Transaction Distribution by Channel

Branch: 75, Online: 70, ATM: 55. Digital channels are nearly on par with traditional banking, emphasizing the need for robust online fraud detection.



Transaction Type Distribution

Debit: 150 (75%), Credit: 50 (25%). Debit transactions dominate, requiring models sensitive to anomalous patterns in outgoing payments.

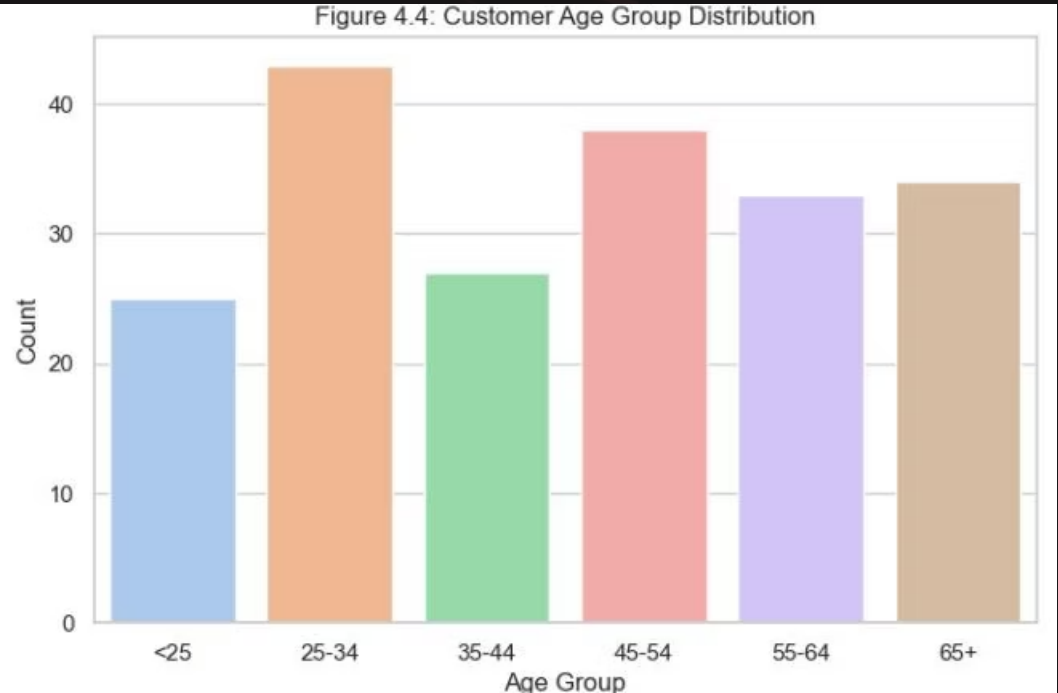


The analysis of 200 sampled banking transactions revealed key patterns. Traditional in-person banking remains significant, but digital channels are rapidly catching up. The prevalence of debit transactions highlights a critical area for fraud detection focus.

Data Analysis & Key Findings(Continued)

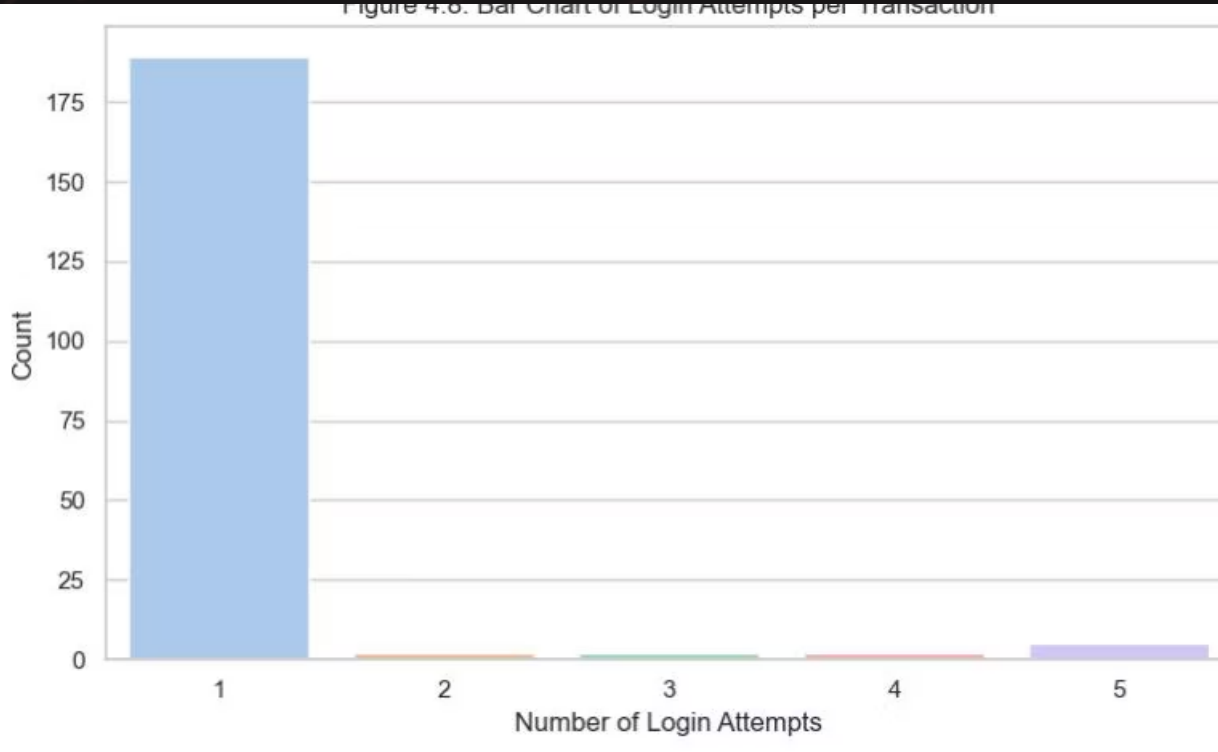
Customer Demographics

Diverse age range (18-80) and occupations (engineers, doctors, retirees, students). This supports developing fraud models robust to varied customer profiles.



Login Attempts

70% of transactions had one login attempt, but multiple attempts (up to 5) were observed. These can be red flags for brute-force attacks or suspicious access.



The data shows a broad customer base, with varying ages and occupations. Behavioral indicators like multiple login attempts are crucial for real-time fraud detection, signaling potential security risks.

Recommendations & Improvements



Enhance Digital Monitoring

Prioritize real-time anomaly detection for online and ATM channels.



Develop Multi-Feature Models

Incorporate transaction amount, channel, location, age, occupation, duration, and login attempts.



Implement Behavioral Analytics

Use multiple login attempts and unusual transaction durations as fraud triggers.



Geographic Risk Profiling

Deploy location-based risk scoring for high-activity cities.



Address Data Imbalance

Use advanced sampling techniques (e.g., SMOTE) to improve model sensitivity to rare fraud cases.

These recommendations aim to strengthen fraud prevention strategies by leveraging diverse data features and advanced analytical techniques.

Conclusion & Future Directions

Project Impact

Demonstrated the feasibility of advanced anomaly detection models using sampled banking data, highlighting the complexity of modern banking and the need for sophisticated systems.

References

Detailed references are available in the full report.



Future Research

Explore larger datasets, graph neural networks for complex relationships, and NLP for unstructured data. Focus on real-time, adaptive, and explainable AI systems.

Acknowledgements

Thanks to Jain University, Dr. Shalini, and all supporters for their invaluable guidance and assistance.