Capstone Research Proposal: Advanced Fraud Detection System

Integrating Anomaly Detection and Linguistic Modeling for Enhanced Financial Security

*Author: Komal Shahid*

*DSC 680: Final Capstone Project*

*June 2025*

# Abstract

Financial fraud in the modern digital economy represents a pervasive and rapidly evolving threat, jeopardizing individuals and institutions alike. This research proposal outlines an innovative approach to fraud detection, integrating autoencoder-based anomaly detection with fine-tuned BERT-derived linguistic modeling. The system is designed to analyze both structured transaction data and unstructured textual descriptions, enabling a comprehensive and ethically aligned fraud detection strategy. By introducing underexplored attributes such as behavioral signatures and inter-transaction temporal dynamics, this project aims to advance beyond conventional fraud detection methodologies, capturing nuanced, non-obvious indicators of fraudulent activity that traditional models may overlook. Emphasizing ethical AI practices, including fairness auditing, transparency, and explainability, the research seeks to deliver a solution that is not only statistically robust but also poised for practical application.

# Introduction

Financial fraud has become an omnipresent challenge in the digital era, with its consequences ranging from personal financial losses to large-scale institutional crises. For example, in 2024, global fraud losses amounted to over $42 billion, primarily targeting online banking and digital payment platforms (Olowu et al., 2024). Although progress has been made in fraud prevention, conventional detection systems often fall short when facing the increasingly sophisticated strategies of malicious actors.

Existing fraud detection methods typically rely on rule-based systems or singular AI models, which lack adaptability and fail to account for nuanced data patterns. Moreover, ethical concerns surrounding bias, accountability, and transparency plague numerous AI-driven solutions. Building on this foundation, the project also explores inter-transaction timing patterns and behavioral signals—dimensions often overlooked in conventional models. By examining how transactions relate over time and across contexts, the system aims to better reflect the complexity of real-world financial fraud, especially in sectors where trust, speed, and regulatory compliance are paramount.

# Problem Statement

Despite advancements in AI-driven fraud detection, several critical gaps persist:

* Precision vs. False Positives: Financial institutions risk alienating legitimate customers due to overly sensitive models that flag valid transactions as fraudulent.
* Adaptability: Static rule-based systems struggle to keep pace with the dynamic strategies employed by fraudsters.
* Transparency: Many systems operate as "black boxes," offering limited explainability, which raises ethical concerns.
* Overlooked Attributes: Existing literature and implementations often neglect transaction interdependencies and the broader temporal dynamics of fraudulent behavior, missing valuable signals.

This research aims to address these challenges comprehensively by developing a hybrid system that combines anomaly detection and linguistic modeling, while introducing unique attributes—such as behavioral frequency patterns—that have been historically underutilized in fraud detection.

# Research Objectives

The primary objectives of this project include:

* Designing an autoencoder model to identify anomalies in structured transaction data, with an emphasis on inter-transaction patterns and temporal dynamics.
* Fine-tuning a BERT-derived model (e.g., FinBERT) to classify fraud-related descriptions and identify subtle linguistic cues indicative of fraudulent behavior.
* Exploring underutilized attributes such as behavioral frequency, geospatial patterns, and contextual data enrichment to enhance fraud detection capabilities.
* Evaluating model performance using robust statistical metrics, including ROC-AUC, precision, recall, and F1-score, ensuring a balance between accuracy and practical applicability.
* Implementing fairness auditing and explainability frameworks to align the system with ethical AI principles.

# Methodology

## Data Collection

The project will leverage publicly available datasets that encompass both structured transaction data and unstructured textual descriptions, such as those from Kaggle or government fraud reports. Additionally, synthetic data generation methods will be employed to simulate complex fraud scenarios, including temporal patterns and behavioral anomalies.

## Data Sources

The primary training and evaluation rely on IEEE-CIS and Talha (2025) datasets, both of which provide labeled, structured, and unstructured fraud data sufficient for the initial scope of this system. Additional datasets were considered for external benchmarking but not used in core training, to maintain fidelity to real-world institutional fraud contexts.

1. **IEEE-CIS Fraud Detection Dataset** (Kaggle, 2020): A large-scale dataset containing anonymized transaction records, including numerical and categorical features such as transaction amount, device type, and card usage patterns. This dataset enables training and evaluation of the autoencoder model.
2. **AI-Powered Banking Fraud Detection Dataset (2025)** (Talha, 2025): A synthetic dataset that includes transaction descriptions and fraud labels, suitable for fine-tuning FinBERT. The text data allows exploration of semantic patterns in fraudulent behavior.

These datasets provide a rich foundation for identifying behavioral and linguistic characteristics of fraud, such as transaction frequency, velocity, and deceptive language cues.

We also use, Synthetic Minority Over-Sampling Technique (SMOTE), which is also available in python as a package to address the class imbalance in the datasets. It generates samples of minority classes to make the class distribution balanced by creating synthetic examples in the feature space of the minority class.

## Model Development

* Autoencoder-Based Anomaly Detection: This model will focus on identifying irregularities in transaction data by learning typical patterns and flagging deviations. Unique features such as time-based transaction clustering and behavioral frequency analysis will be incorporated.
* Fine-Tuned BERT Model: A pretrained FinBERT model will be adapted to identify fraud-related language in transaction descriptions. Advanced tokenization techniques will be used to capture context-sensitive signals.

## Exploration of Overlooked Attributes

To ensure a unique and impactful approach, the project will integrate attributes less commonly addressed in existing solutions, such as:

* Temporal Dynamics: Analyzing the timing and frequency of transactions to identify abnormal behavioral patterns.
* Geospatial Patterns: Incorporating location-based data to detect suspicious activity across regions.
* Contextual Enrichment: Combining transaction metadata with external factors like market trends to better contextualize fraudulent behavior.

## Evaluation and Ethical AI Practices

While AI presents unprecedented opportunities for detecting financial fraud, it also introduces profound ethical challenges—particularly in high-stakes domains like banking, where misclassifications can result in real financial harm or unfair treatment of individuals. One core concern is algorithmic bias: if models learn patterns from imbalanced datasets, they risk amplifying disparities across age, income, geography, or other protected attributes. Another concern is opacity—many deep learning models are difficult to interpret, making it hard for financial institutions to justify decisions to regulators or customers. This project acknowledges those risks and directly addresses them through balanced data, fairness auditing, transparent documentation, and model explainability. Tools like SHAP and LIME will be applied to surface feature contributions behind each prediction, while evaluation metrics will be disaggregated by demographic segments to monitor equity across user groups. Additionally, synthetic examples may be used for stress-testing, but real-world decisions will only be based on validated, authentic datasets to prevent hallucinated or unrepresentative inference. In doing so, the system aims to uphold not only technical accuracy, but also the ethical accountability increasingly expected in AI governance.

# Expected Outcomes

The project aims to deliver a fraud detection system that:

* Outperforms traditional systems by utilizing underexplored attributes and hybrid modeling approaches.
* Minimizes false positives, enhancing customer experience and institutional trust.
* Aligns with ethical AI practices, ensuring fairness and transparency in decision-making processes.

# Significance

By combining deep learning and NLP, this project offers a holistic approach to fraud detection that reflects the complexity of real-world financial behavior. It also contributes to the growing field of ethical AI by embedding fairness and interpretability into the model development process.

Given the strict regulatory landscape governing financial institutions—such as GDPR, the Fair Credit Reporting Act (FCRA), and proposed legislation like the EU’s AI Act—any AI-powered fraud detection system must adhere to principles of fairness, accountability, and transparency. This includes ensuring that models provide explainable inferences, are free from systemic bias, and meet the evolving standards of “responsible AI.” In this project, fairness audits and explainability tools will be used to evaluate model behavior across demographic segments, providing both ethical integrity and regulatory defensibility.

The resulting system could serve as a prototype for financial institutions seeking to modernize fraud prevention efforts without compromising user trust or legal compliance.

**References**

Corcoran, S. M., & Soscia, R. J. (2024, December 17). *AI and other decision-making tools: Does the Fair Credit Reporting Act apply?* Workplace Privacy, Data Management & Security Report. https://www.workplaceprivacyreport.com/2024/12/articles/artificial-intelligence/ai-and-other-decision-making-tools-does-the-fair-credit-reporting-act-apply/

IBM. (2024). *AI compliance: What it is, why it matters and how to get started.* IBM Think Insights. https://www.ibm.com/think/insights/ai-compliance

Olowu, O., Adeleye, A. O., Omokanye, A. O., Ajayi, A. M., Adepoju, A. O., Omole, O. M., & Chianumba, E. C. (2024). *AI-driven fraud detection in banking: A systematic review of data science approaches to enhancing cybersecurity*. GSC Advanced Research and Reviews, 21(2), 227–237. https://doi.org/10.30574/gscarr.2024.21.2.0418

Talha, M. (2025). *AI-powered banking fraud detection dataset*. Kaggle. https://www.kaggle.com/datasets/mdtalhask/ai-powered-banking-fraud-detection-dataset-2025