

Explainable AI for Anti-money laundering in banking.

Course Title: Research Methodology in Computer Science

Student ID: 253001288

Student name: Aishah Abdulla Saeed Ali Aldhanhani

Problem Statement

These days in the banking and financial sector, many accounts are involved in money laundering, and identifying these suspicious transactions has become a major challenge, especially with traditional methods. The solution was to introduce Explainable Artificial Intelligence to become one of the foundations of banking transactions, making it easier to detect abnormal transactions that can be classified as money laundering. Although there are some methods used by certain banks that fall under the term “black box,” and the results are often satisfactory and accurate, they lack transparency and provide no reasons or explanations for the outcome (Kumar et al., 2023). The lack of transparency shakes the customer’s trust in the bank and creates fear when carrying out or initiating high-value transactions, fearing that they might be considered suspicious or related to money laundering.

The current models assist only in providing decisions, but they are not supported by explanations; in other words, they lack transparency. This represents a major weakness in the existing models, as the limited interpretability and the absence of any method to trace the reasons or outcomes produced by the system have become a significant challenge between customers and the banking sector. On the other hand, this also poses a risk to regulatory compliance and consumer protection.

In summary, the need to develop a model that provides decisions along with explanations has become a fundamental requirement in the banking sector. focusing on model interpretation and regulatory compliance. It emphasizes the importance of employing certain techniques (Zhou et

al., 2024) Employing Explainable Artificial Intelligence is now a necessity not only in name, but also as a means to enhance regulatory compliance and strengthen trust with customers. Implementing such a system will bridge a significant gap within the banking industry.

Research Objectives

Primary Goal:

To build a model based on Explainable Artificial Intelligence, enabling the detection of suspicious transactions with full transparency for customers.

Sub-Objectives:

- To analyze whether the currently used systems lack transparency, clarity, or the ability to provide explanations for the results they generate.
- To review whether modern approaches such as Explainable Artificial Intelligence can address this problem, and whether identifying the strengths and weaknesses of the existing model will enable the new system to overcome these limitations.
- To design an ideal model that provides both results and their underlying reasons, using advanced AI tools, and to determine whether this will improve decisions related to detecting suspicious transactions or risks associated with credit reports.

- To evaluate whether the proposed advanced system produces accurate predictions, whether its explanations are clear, and those explanations align with and comply with regulatory requirements.
- To compare the results of the newly proposed system with those of the currently used systems,

Literature Review:

Several previous studies have addressed the use of artificial intelligence in the banking sector for

fraud detection and risk management, as reviewed in the research paper by Kumar et al. (2023). The study emphasized the importance of interpretability in strengthening customer trust, ensuring regulatory compliance, and enhancing consumer protection. It also highlighted that most models currently used in the banking sector rely heavily on black-box algorithms, The difficulties and challenges faced were the inaccuracy and imbalance of information and the results of false positives. (Han et al., 2020) which creates significant difficulty in understanding the reasons behind classifying certain transactions as suspicious. Many studies have previously been presented that addressed the use of artificial intelligence

in the banking sector for fraud detection and risk management, and this was reviewed in the research paper by Kumar et al. (2023). It confirmed the importance of

interpretability in enhancing customer trust, regulatory compliance, and consumer protection. It also confirmed that most of the models used in the banking sector rely entirely on black-box algorithms, which created difficulty and an obstacle in understanding the reasons behind classifying suspicious transactions. One of the challenges highlighted by the paper is the limited number of practical trials conducted and the complete reliance on synthetic data simulations. (Kumar et al., 2023) .

In summary of what was mentioned in previous studies, integrating these technologies contributes to reducing bias and enhancing transparency, especially when the customers are distinguished and the bank does not want to lose this trust. These technologies will also contribute to strengthening the reasons and will support the alerts issued by the system in case suspicious transactions are detected. It was also mentioned in the context of previous studies that introducing these models after development that removes the obstacles and at the same time meets all regulatory requirements and enhancing trust in the banking sectors, not only between customers. There was a focus on future plans and directions to allow the XAI system time to expand across many expects. (Liet al., 2024)

Reference

- Han, Y., Liu, Z., & Chen, X. (2020). Artificial Intelligence for Anti-Money Laundering: A Review and Extension. SSRN. Retrieved from <https://papers.ssrn.com/>
- Han, Z., Wang, L., & Li, J. (2023). Interpretable Machine Learning Models for Financial Fraud Detection. AIP Publishing. Retrieved from <https://aip.scitation.org/>
- Kumar, R., Sharma, A., & Verma, P. (2023). Applications of Explainable Artificial Intelligence in Finance — A Review. Springer. Retrieved from <https://link.springer.com/>
- Li, M., Zhang, Q., & Huang, Y. (2024). Explainable AI in Financial Services: Transparency and Trust in Risk Management. Springer. Retrieved from <https://link.springer.com/>
- Nguyen, T., Tran, H., & Pham, K. (2025). Comparative Analysis of Machine Learning Algorithms for Money Laundering Detection. Springer. Retrieved from <https://link.springer.com/>
- Zhou, X., Chen, Y., & Li, D. (2024). Explainable Artificial Intelligence (XAI) in Finance: A Systematic Literature Review. Springer. Retrieved from <https://link.springer.com/>