

ABSTRACT

The rapid development of Decentralized Finance (DeFi) has introduced transformative financial innovations, offering open, transparent, and intermediary-free financial services. However, this growth has also increased exposure to sophisticated fraudulent activities. Existing fraud detection approaches on benchmark datasets have achieved high accuracy, but most operate purely on a statistical paradigm, focusing on numerical feature correlations while overlooking the deeper contextual meaning of transaction behavior. This research addresses that methodological gap by exploring whether a semantic-based paradigm, capable of understanding contextual and narrative patterns in transaction data, can provide a competitive alternative for detecting fraudulent activities in the Ethereum ecosystem.

This thesis proposes and evaluates a novel framework based on the RoBERTa transformer model, optimized to process blockchain transaction data in descriptive sentence form. By converting structured transaction records into narrative inputs, the model leverages its deep contextual understanding to detect anomalous patterns. To address RoBERTa’s high computational demands, Gradient Accumulation (GA) is implemented to improve training efficiency without sacrificing performance. The framework is evaluated on the public “Ethereum Fraud Detection” benchmark dataset, using only seven core behavioral features and without statistical oversampling, ensuring robustness against real-world class imbalance.

Experimental results show that the proposed RoBERTa+GA model achieves an F1-score of 79.8% and a recall of 75.0%, demonstrating competitive performance under more challenging conditions compared to traditional methods. However, a false negative rate of 25.0% highlights the limitations of fully automated detection, reinforcing the need for a layered defense system combining AI models with human analyst verification. The findings also reveal a persistent gap between advanced technological capabilities and current regulatory frameworks, offering valuable insights for the evolution of more adaptive, data-driven oversight mechanisms in blockchain-based financial systems.

Keyword: DeFi, Blockchain, Fraud Detection, RoBERTa, Ethereum Transactions