



Contribution ID: 34

Type: **Physical Poster shown at the Meeting**

A Reinforcement Learning Framework for Fraud Detection in Highly Imbalanced Financial Data

Anomaly detection in financial transactions is a challenging task, primarily due to severe class imbalance and the adaptive behavior of fraudulent activities. This paper presents a reinforcement learning for fraud detection (RLFD) framework to address this problem. We train a deep Q-network (DQN) agent with a long short-term memory (LSTM) encoder to process sequences of financial events and identify anomalies. On a proprietary, highly imbalanced dataset, 10-fold cross-validation highlights a distinct trade-off in performance. While a gradient boosted trees (GBT) baseline demonstrates superior global ranking capabilities (higher ROC and PR AUC), the RLFD agent successfully learns a high-recall policy directly from the reward signal, meeting operational needs for rare event detection. Importantly, a dynamic orthogonality analysis proves that the two models detect distinct subsets of fraudulent activity. The RLFD agent consistently identifies unique fraudulent transactions that the tree-based model misses, regardless of the decision threshold. Even at high-confidence operating points, the RLFD agent accounts for nearly 30\% of the detected anomalies. These results suggest that while tree-based models offer high precision for static patterns, RL-based agents capture sequential anomalies that are otherwise missed, supporting for a hybrid, parallel deployment strategy.

INFN OpenAccess Repository link

Authors: PAPANASTASSIOU, Alkis (Istituto Nazionale di Fisica Nucleare); CAMAIANI, Benedetta; LENZI, Piergiulio (Istituto Nazionale di Fisica Nucleare); Dr CRUPI, Riccardo (Intesa Sanpaolo S.p.A)

Session Classification: POSTER AND VIDEO UPLOAD