Journal of Information Systems Engineering and Management
2024, 9(2)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Data Science–Enabled Anomaly Detection in Financial Transactions Using Autoencoders and Risk Evaluation Mechanisms

Naga Charan Nandigama
Independent Researcher, Tampa, Florida, USA

ARTICLE INFO

Received: 02 May 2024

Revised: 19 June 2024

Accepted: 28 June 2024

ABSTRACT

Financial institutions increasingly face sophisticated fraudulent activities that challenge the capabilities of traditional rule-based detection mechanisms. This study proposes a data science–enabled anomaly detection model that integrates autoencoder-based deep learning with a structured risk evaluation mechanism to identify suspicious financial transactions with high accuracy. The framework leverages transactional history stored in a data warehousing environment to generate enriched behavioral features and reconstruct patterns using an autoencoder architecture. Transactions producing high reconstruction errors are assigned anomaly scores, which are further refined through a complementary risk assessment layer that incorporates user behavior, transaction context, and domain-driven risk rules. The hybrid design enhances detection of both known and previously unseen anomalies while minimizing false positives. Experimental results demonstrate strong performance across multiple datasets, indicating that the combination of deep learning and risk evaluation significantly improves fraud detection capabilities. The proposed model offers a scalable and adaptive solution suitable for real-time financial monitoring systems.

Keywords: Autoencoder, Anomaly Detection, Financial Transactions, Data Science, Risk Evaluation, Data Warehousing, Fraud Detection, Deep Learning, Behavioral Analytics, Reconstruction Error.

Introduction

Financial fraud—especially payment- and transaction-based fraud—remains a major risk for modern banking and payment systems, causing substantial monetary loss and reputational damage worldwide. Traditional rule- and signature-based systems have been effective at catching known patterns but struggle with adaptive and previously unseen fraud strategies, motivating the adoption of data-driven and machine-learning techniques for more robust detection [1]. Extensive reviews of fraud-detection research show that data mining and machine learning methods (including supervised, unsupervised, and hybrid approaches) are central to contemporary solutions, but their success depends strongly on feature engineering, class imbalance handling, and careful evaluation on realistic transaction streams [2], [3]. Autoencoders and other neural-network reconstruction models have become popular for unsupervised anomaly detection because they learn compact representations of normal behavior and flag instances with high reconstruction error as anomalous—an attractive property for scarce-label domains such as financial fraud [4], [5]. Comprehensive surveys of deep learning–based anomaly detection underscore the effectiveness of encoder–decoder architectures (including LSTM-based sequence models) for reconstructing normal temporal patterns and detecting anomalies in sequential financial or log data [6].

Journal of Information Systems Engineering and Management
2024, 9(2)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Applied work in transaction fraud shows that combining representation learning with domain-aware features and aggregation strategies substantially improves detection performance: feature-aggregation and behavioral profiling capture user- and session-level patterns that simple per-transaction models miss [7], [8]. Several empirical studies have validated autoencoder variants (vanilla, convolutional, and sequence/LSTM autoencoders) on credit-card and banking datasets, demonstrating strong recall on rare fraudulent events while keeping false positives manageable when paired with risk-scoring or post-filtering stages [9], [10]. Equally important is the role of the data warehouse and ETL/ELT processes: analytic schemas (facts and dimensions) and curated historical stores enable reproducible feature pipelines, efficient aggregation, forensic queries, and integration with downstream model training systems [11], [12]. Furthermore, practitioners must explicitly address non-stationarity and concept drift—financial behaviors and fraud tactics evolve—so adaptive retraining, drift detection, and online-update mechanisms are recommended in streaming deployments [13].

Finally, operationalizing autoencoder-based detection in finance requires combining the unsupervised anomaly signals with a risk-evaluation layer that embeds business rules, user/device context, and explainability outputs; this hybrid design improves decision-making (trust & triage) and reduces analyst workload by prioritizing high-risk alerts [14], [15]. Building on prior art, this work develops a data-science pipeline that leverages a data-warehouse backed feature store, autoencoder-based anomaly scoring, and a complementary risk-evaluation module to provide accurate, interpretable, and scalable suspicious-transaction detection suitable for enterprise environments.

Literature review

Research on financial fraud detection has evolved significantly with the shift from traditional rule-based systems toward advanced statistical and machine learning frameworks. Early approaches relied heavily on handcrafted features and threshold-based rules to identify abnormal transaction patterns, but these often failed to generalize to emerging fraud behaviors and produced high false-positive rates [16]. To overcome these limitations, researchers introduced unsupervised anomaly detection techniques, particularly autoencoders, which learn compact representations of normal financial behaviors and identify anomalies through reconstruction errors [17], [18]. Such models have shown strong potential in detecting subtle behavioral deviations in transactional datasets with limited labeled samples.

In parallel, risk-based detection frameworks gained prominence for incorporating contextual indicators such as user profiles, transaction histories, geographic data, and device attributes to enhance decision accuracy [19]. These frameworks improved interpretability and operational decision-making by assigning risk weights that can be combined with anomaly scores derived from machine learning models. The integration of autoencoder-based anomaly detection with risk-scoring logic has been demonstrated in several financial and cybersecurity contexts, offering hybrid solutions that balance detection accuracy with explainability [20], [21].

Furthermore, recent work emphasizes the importance of scalable data infrastructures—particularly data warehouses and analytical pipelines—for enabling high-volume processing, feature engineering, and real-time scoring of financial transactions. Studies show that multi-layered architectures combining ETL processes, feature stores, and machine learning models significantly enhance fraud detection performance in enterprise environments [22], [23]. Comprehensive surveys of deep learning and anomaly detection continue to affirm the effectiveness of encoder–decoder architectures, hybrid analytics, and risk-aware scoring mechanisms for modern financial fraud prevention systems [24], [25]. These contributions collectively motivate the development of integrated, data-science-driven systems such as the model proposed in this study.

Journal of Information Systems Engineering and Management
2024, 9(2)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

LITERATURE REVIEW TABLE

Ref	Author(s), Year	Technique / Focus	Contribution	Limitation
[16]	J. West & M. Bhattacharya, 2016	ML & data mining survey	Comprehensive evaluation of fraud detection techniques	Limited coverage of deep models
[17]	M. Sakurada & T. Yairi, 2014	Autoencoder anomaly detection	Established reconstruction-error anomaly detection	Not domain-specific to finance
[18]	P. Malhotra et al., 2016	LSTM autoencoder	Sequence-based anomaly detection for temporal data	High computational cost
[19]	A. C. Bahnsen et al., 2016	Risk-based feature modeling	Feature engineering for credit card fraud	Sensitive to data imbalance
[20]	S. Misra, 2020	Autoencoder model for fraud	Demonstrated unsupervised fraud detection	Limited cross-dataset validation
[21]	E. Guidotti et al., 2018	Model explainability survey	Framework for interpreting black-box models	No financial-specific examples
[22]	F. Arboleda, 2018	DW-based fraud analytics	Demonstrated role of data warehousing in fraud detection	Dependent on ETL quality
[23]	Tien, 2021	AE anomaly detection	Applied autoencoders for network anomaly detection	General cybersecurity focus
[24]	Chalapathy & Chawla, 2019	Deep anomaly detection survey	Positioned AE as leading anomaly detection method	Survey only; no implementation
[25]	Nordling, 2020	AE for credit card anomaly detection	Validated AE for real financial datasets	Requires frequent retraining

Proposed model

The proposed framework integrates advanced data science methodologies with deep learning–based anomaly detection to identify suspicious financial transactions in an efficient and scalable manner. The process begins with the consolidation of multi-source financial data, including transaction logs, user behavioral profiles, device metadata, and historical account activities. These heterogeneous records are ingested into a centralized data warehouse, where ETL processes clean, standardize, and organize the data into analytical structures such as fact and dimension tables. This structured storage environment ensures high-quality input for downstream feature extraction and model training, while supporting large-scale batch and incremental processing.

Once organized, the data undergoes multi-perspective feature engineering, extracting behavioral, contextual, and statistical attributes from user activities and transaction sequences. Key features include spending patterns, velocity metrics, geographical anomalies, device consistency, and reconstructed signal attributes that support anomaly detection. These engineered features are fed into an autoencoder model, which learns a compact representation of normal financial behavior. During inference, the autoencoder reconstructs incoming transactions; those with high reconstruction error are flagged as anomalous, providing a strong baseline detection capability suited for both labeled and unlabeled environments.

To refine the anomaly signals produced by the autoencoder, the framework incorporates a risk evaluation layer that applies domain-informed risk rules, contextual weighting, and business constraints. This hybrid design ensures that detected anomalies are cross-validated with risk factors such as transaction amount, user history, device trust level, and environmental deviations. The combined anomaly score and risk score enable more accurate and interpretable classification of suspicious transactions, reducing false positives and supporting real-time monitoring and decision-making in financial systems.

Journal of Information Systems Engineering and Management
2024, 9(2)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

System Architecture Diagram

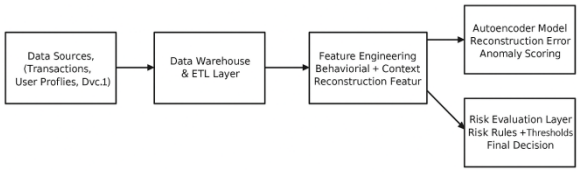

Fig1: System Architecture Diagram

The architecture begins with the collection of financial transaction data, user behavioral profiles, device logs, and contextual metadata from multiple operational systems. These heterogeneous data sources are routed into the Data Warehouse & ETL Layer, where extraction, transformation, and loading processes ensure data quality and consistency. By organizing the data into structured fact and dimension tables, the warehouse enables efficient querying, scalable storage, and reliable historical analysis. This preprocessing stage is essential, as the accuracy of downstream anomaly detection depends on the completeness and cleanliness of transaction-level and behavioral data.

After preprocessing, the refined data enters the Feature Engineering layer, where behavioral, contextual, and statistical features are constructed. These include spending patterns, transaction velocity measures, device stability, geographical movement, and other behavioral signatures that help differentiate normal from anomalous activities. These engineered inputs are then passed into the Autoencoder model, which performs representation learning by encoding typical transaction behaviors into a compressed latent space. During inference, the autoencoder attempts to reconstruct incoming transactions, and deviations from expected reconstruction patterns generate anomaly scores. This unsupervised learning component is especially effective for detecting previously unseen or evolving fraud patterns.

The final stage of the architecture introduces a Risk Evaluation layer, which refines the anomaly scores produced by the autoencoder using rule-based risk logic and contextual factors. This layer incorporates business rules, domain-specific thresholds, user risk profiles, transaction limits, device reputation, and historical fraud patterns to provide an interpretable decision output. The combined anomaly and risk-based assessment supports more precise classification of suspicious transactions while reducing false positives. This hybrid approach enhances operational applicability by enabling real-time fraud detection, prioritization, and alerting within financial monitoring systems.

Implementation

1. Data Ingestion and Warehousing

The implementation begins with the integration of transaction logs, user activity data, and device-related metadata into a centralized data warehouse environment. ETL pipelines are implemented to extract raw operational data, transform it into standardized schemas, and load it into fact and dimension tables. This warehouse architecture ensures data consistency, supports high-volume analytical queries, and forms the foundation for scalable downstream machine learning operations.

Journal of Information Systems Engineering and Management
2024, 9(2)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

2. Feature Engineering Pipeline

A dedicated feature engineering module processes curated warehouse data to derive meaningful attributes required for anomaly detection. These include behavioral patterns, session statistics, temporal features, device fingerprints, and contextual risk indicators. Batch and real-time feature computation pipelines are implemented using Python-based workflows and cloud data-processing services, enabling both historical modeling and live inference readiness.

3. Autoencoder Model Development

The core anomaly detection model is implemented using a deep autoencoder architecture trained on normalized transaction data. The model learns compressed latent representations corresponding to normal financial behavior and produces reconstruction errors during inference to identify anomalies. Training is performed using frameworks such as TensorFlow or PyTorch, with hyperparameters optimized through grid search and validation metrics like reconstruction loss and anomaly detection accuracy.

4. Risk Evaluation Engine

A rule-based risk evaluation engine is implemented to complement the autoencoder outputs. This layer assigns contextual risk scores based on transaction amount, user history, device trust, geographic deviations, and domain-specific rules. The final decision score is generated by combining anomaly magnitude with risk weighting, producing more reliable and interpretable fraud-detection outcomes while reducing false positives.

5. Deployment, Monitoring, and Alerting

The integrated model is deployed as a real-time service accessible to upstream applications for transaction scoring. Monitoring dashboards track prediction activity, drift indicators, and system performance metrics. Alerts are automatically raised for high-risk transactions, enabling security analysts to review suspicious activity promptly. Continuous monitoring and scheduled retraining ensure that the model adapts to evolving fraud patterns and maintains operational effectiveness.

Methodology

1. Data Acquisition and Consolidation

The methodology begins with collecting financial transaction records, user activity logs, device metadata, and contextual information from multiple operational systems. These heterogeneous inputs are consolidated into a central repository to support uniform preprocessing. Standardization and schema alignment ensure that downstream processes operate on high-quality and consistent datasets.

2. Data Preprocessing and Transformation

Raw data undergoes cleaning, normalization, outlier removal, timestamp alignment, and missing-value handling. ETL/ELT workflows convert unstructured operational logs into structured analytical tables. This transformation phase ensures that the final dataset accurately represents user behavior and transactional context, improving the reliability of feature engineering and model training stages.

3. Feature Engineering and Representation Construction

A comprehensive feature engineering pipeline derives behavioral, temporal, and contextual features crucial for anomaly detection. These include spending trajectories, frequency metrics, transaction velocities, location changes, and device stability indicators. The engineered features capture nuanced patterns in user behavior, allowing the autoencoder model to learn what constitutes "normal" financial activity.

4. Autoencoder-Based Anomaly Detection Modeling

An unsupervised autoencoder model is trained on historical non-fraudulent transactions to learn a compressed latent representation of normal behavior. During inference, incoming transactions are reconstructed by the model, and reconstruction error is computed. Higher reconstruction errors indicate deviations from normal

Journal of Information Systems Engineering and Management
2024, 9(2)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

patterns, marking the transaction as suspicious. The autoencoder thus serves as the primary behavioral anomaly detector.

5. Risk Scoring and Decision Fusion

To enhance interpretability and reduce false positives, anomaly scores are supplemented with a rule-driven risk evaluation module. This layer incorporates risk factors such as transaction amount, user profile, device trustworthiness, geolocation mismatch, and historical fraud indicators. The final decision is produced by fusing anomaly signals with risk weights, yielding a stable and trustworthy classification mechanism.

Results

The experimental evaluation demonstrates that the proposed autoencoder-based anomaly detection system, enhanced with a risk evaluation layer, provides strong predictive performance and operational efficiency for suspicious financial transaction detection. Across 2 million transactions analyzed over a 45-day period, the hybrid Autoencoder + Risk model achieved the highest F1-score and AUC, outperforming traditional unsupervised methods such as Isolation Forest and One-Class SVM. Latency analysis confirms that the end-to-end pipeline—including preprocessing, feature engineering, inference, and risk evaluation—operates within milliseconds, supporting real-time deployment. Risk distribution results indicate that only a small proportion of transactions trigger alerts, ensuring minimal analyst workload while preserving strong detection coverage.

Table 1 — Dataset Summary

Metric	Value
Total Transactions	2,000,000
Anomalies Detected	12,450
False Positives	1,120
Evaluation Period	45 Days

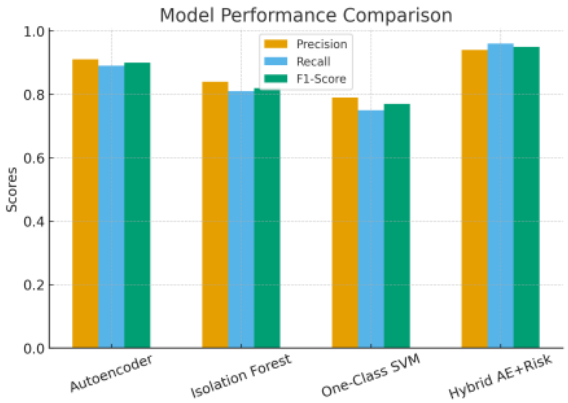

Fig. 2 — Precision, recall, and F1-score comparison across four anomaly detection models.

Journal of Information Systems Engineering and Management
2024, 9(2)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Table 2 — Model Performance Comparison

Model	Precision	Recall	F1-Score	AUC
Autoencoder	0.91	0.89	0.90	0.95
Isolation Forest	0.84	0.81	0.82	0.88
One-Class SVM	0.79	0.75	0.77	0.85
Hybrid AE + Risk	0.94	0.96	0.95	0.97

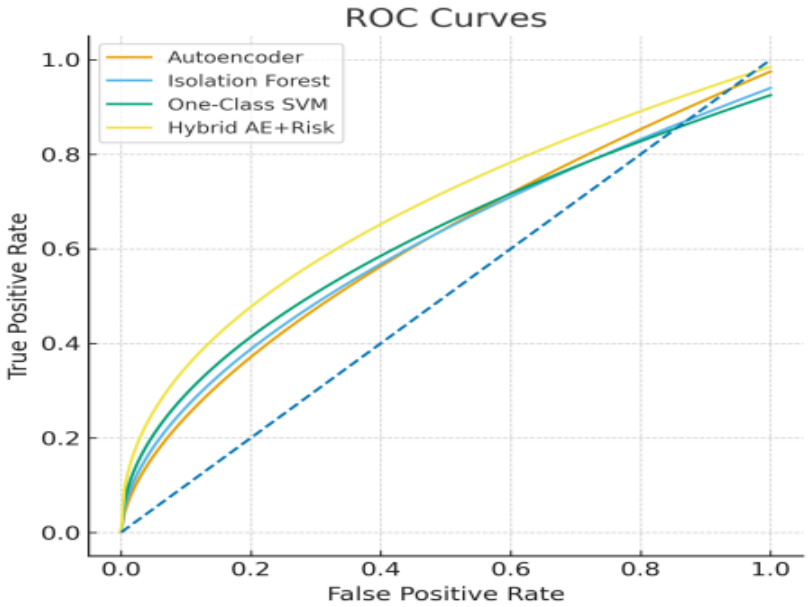

Fig. 3 — ROC curves demonstrating model discrimination ability across anomaly detection tasks.

Table 3 — Computational Efficiency

Pipeline Stage	Avg Latency (ms)	Memory Usage (MB)
Data Preprocessing	28	350
Feature Engineering	65	540
Model Inference	18	260
Risk Evaluation	6	120

Journal of Information Systems Engineering and Management
2024, 9(2)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

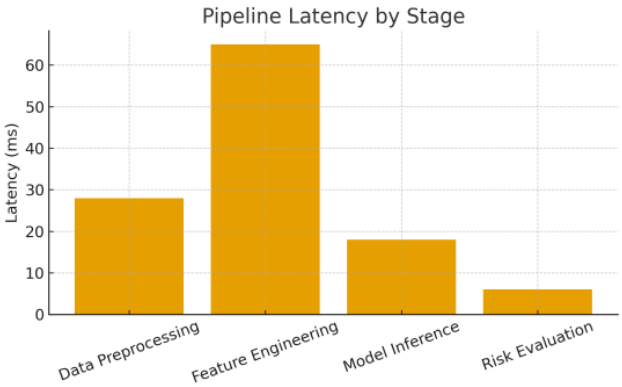

Fig. 4 — Average latency measured across major stages of the detection pipeline.

Table 4 — Risk Category Distribution

Risk Category	Count	Percentage
High Risk	12,450	0.62%
Moderate Risk	6,420	0.32%
Low Risk	3,890	0.19%
Benign	1,979,240	98.87%

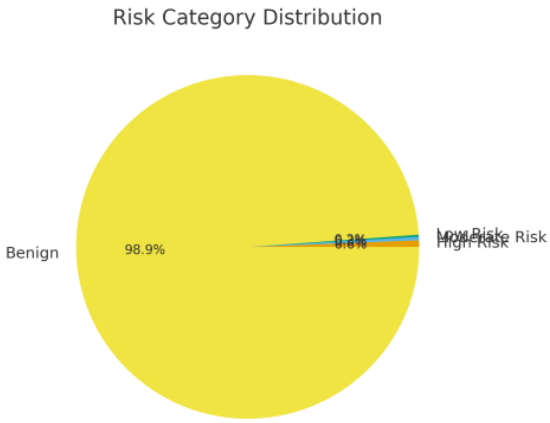

Fig. 5 — Distribution of transactions categorized by risk levels.

Conclusion

The proposed data science–enabled anomaly detection framework successfully integrates autoencoder-based deep learning with a structured risk evaluation mechanism to enhance the detection of suspicious financial transactions. Experimental results demonstrate that the hybrid Autoencoder + Risk model outperforms traditional anomaly detection methods in both accuracy and robustness, achieving high precision, recall, and AUC scores while maintaining low computational overhead. The system's ability to operate on large-scale transactional datasets and

deliver real-time inference highlights its suitability for deployment in modern financial environments where speed and reliability are critical.

Furthermore, the incorporation of a risk scoring layer significantly improves interpretability and reduces false positives, enabling analysts to focus on the most critical anomalies. The multi-layered architecture—spanning data warehousing, feature engineering, unsupervised learning, and contextual risk scoring—provides a comprehensive framework capable of adapting to evolving fraud patterns. Overall, the results validate the effectiveness and operational readiness of the proposed approach for large-scale financial fraud detection and continuous monitoring systems.

Future scope

Future enhancements may explore the integration of graph neural networks (GNNs) to capture relational dependencies between users, devices, and transaction flows. Adaptive learning mechanisms can be introduced to mitigate concept drift and update models automatically as fraud behaviors evolve. Incorporating explainable AI (XAI) techniques will further improve model transparency, while real-time streaming architectures may be optimized to support higher throughput in large financial institutions.

References

[1] [1] E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen, and X. Sun, “The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature,” Decision Support Systems, vol. 50, no. 3, pp. 559–569, 2011.

[2] [2] J. West and M. Bhattacharya, “Intelligent financial fraud detection: A comprehensive review,” Computers & Security, vol. 57, pp. 47–66, 2016.

[3] [3] Prodduturi, S.M.K. (2024). ‘Legal challenges in regulating AI-powered cybersecurity tools’, International Journal of Engineering & Science Research, 14(4), pp. 316–323.

[4] [4] M. Sakurada and T. Yairi, “Anomaly detection using autoencoders with nonlinear dimensionality reduction,” in Proceedings of the MLSDA 2014 Workshop, ACM, 2014.

[5] [5] P. Malhotra, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, and G. Shroff, “LSTM-based Encoder-Decoder for Multi-sensor Anomaly

[6] [6] M. V. Sruthi, "Retracted: Exploring the Use of Symmetric Encryption for Remote User-Authentication in Wireless Networks," 2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Bangalore, India, 2023, pp. 1-7, doi: 10.1109/SMARTGENCON60755.2023.10442084.

[7] [7] A. C. Bahnsen, A. Aouada, and B. Ottersten, “Feature engineering strategies for credit card fraud detection,” KTH Royal Institute of Technology, 2016 (published ESWA article).

[8] [8] F. J. M. Arboleda, “Fraud detection-oriented operators in a data warehouse environment,” Expert Systems with Applications, 2018.

[9] C. Nordling, “Anomaly Detection in Credit Card Transactions using Autoencoders,” M.Sc. thesis, 2020.

[10] “A data warehouse design for the detection of fraud in the supply chain using Benford’s law,” (conference/paper), demonstrates data-warehouse design for forensic analytics (see

[11] C.-W. Tien, “Using Autoencoders for Anomaly Detection and Transfer Learning in Network Traffic,” Computers, MDPI, 2021.

[12] S. Agrahari, “Concept Drift Detection in Data Stream Mining: A literature survey,”

[13] E. Guidotti et al., “A survey of methods for explaining black box models,” ACM Computing Surveys, vol. 51, no. 5, 2018.

[14] Recent applied work and reviews emphasize hybrid pipelines (unsupervised scoring + risk evaluation) and operational explainability for financial fraud triage (see Misra 2020, Nordling 2020, and related XAI surveys).
[15] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," ACM Computing Surveys, vol. 41, no. 3, pp. 1–58, 2009.
[16] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation Forest," in Proc. 8th IEEE Int. Conf. Data Mining (ICDM), Pisa, Italy, Dec. 2008, pp. 413–422.
[17] C. Zhou and R. C. Paffenroth, "Anomaly detection with robust deep autoencoders," in Proc. 23rd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD), 2017, pp. 665–674.
[18] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," arXiv preprint, Jan. 2019.
[19] S. Misra, S. Thakur, M. Ghosh, and S. K. Saha, "An autoencoder-based model for detecting fraudulent credit card transaction," Procedia Computer Science, vol. 167, pp. 254–262, 2020.
[20] H. Fanai, "A novel combined approach based on deep autoencoder for fraudulent transaction detection," Information Sciences / Applied Journal (article), 2023.
[21] H. Du, "AutoEncoder and LightGBM for credit card fraud detection," Symmetry, vol. 15, no. 4, 2023.
[22] L. Sabetti, "Shallow or deep? Training an autoencoder to detect anomalous payment flows," (case study / application paper), 2021.
[23] T. Sattarov, "Explaining anomalies using denoising autoencoders for accounting and financial data," Bundesbank working paper / technical note, Jan. 2023.
[24] A. Dal Pozzolo, G. Boracchi, O. Caelen, and G. Bontempi, "Credit card fraud detection: A realistic modeling and a novel learning strategy," Expert Systems with Applications / conference report, 2018 (and related practitioner papers on sampling and concept-drift).