© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715472

# Anomaly Detection in E-commerce Transactions Using Hybrid Deep Learning Models

SHIVAM GUPTA[1], DR. UJWALA SAV[2]
[1, 2]*Vidyalankar School of Information Technology, University of Mumbai*

***Abstract-*** ***As online shopping continues to explode, so do the tactics used by fraudsters. The old-school method of using "rule-based" systems essentially a rigid checklist of "if this, then that" is falling behind because modern fraud is constantly evolving. To stay ahead, we've developed a hybrid framework that acts like a digital detective, combining two powerful tools: Autoencoders and Isolation Forests.***

***Index Terms-*** ***Anomaly Detection, E-commerce, Autoencoder, Isolation Forest, Fraud Detection, Deep Learning***

## I. INTRODUCTION

As online shopping continues to explode, e-commerce platforms are handling millions of transactions daily, making them a prime target for increasingly sophisticated fraud. Traditional security relies on rigid rules or pre-labelled data, but these "old-school" methods often fail to catch new, creative scams and struggle with the fact that genuine fraud examples are rare and hard to document. To bridge this gap, our study moves away from the "rulebook" approach and instead uses a hybrid anomaly detection system. By combining the pattern-recognition power of deep learning with smart statistical techniques, our system learns to identify "weird" behaviour on its own—without needing a human to label it first. This creates a more flexible, proactive shield that can spot hidden threats in massive datasets and protect both a company's bottom line and its customers' trust.

## II. PROBLEM STATEMENT

E-commerce platforms face multiple challenges in identifying:

- Fraudulent transactions (e.g., fake payments, repeated orders)
- Unusual buying behaviour (e.g., sudden bulk purchases, abnormal frequency)

The Problem: Why Current Systems Fail

- A "Data Desert": There is a chronic lack of accurately labelled fraud data, making it hard to "train" traditional AI on what a scam actually looks like.
- Moving Targets: Fraud patterns are dynamic and constantly evolving; by the time a rule is written to stop one trick, fraudsters have already moved on to the next.
- The Firehose Effect: The sheer volume and velocity of modern transactions make it nearly impossible for manual or rigid systems to keep up without lagging.

The Solution: A Smarter Hybrid Model

- Beyond the Rulebook: Our paper proposes a model that doesn't just check boxes; it identifies anomalies in real-time by understanding the "DNA" of a normal transaction.
- Self-Learning: Because it doesn't rely entirely on predefined rules, the system can spot suspicious behaviour it has never seen before.
- Proactive Defence: By combining deep learning with statistical scanning, the framework creates a flexible shield that adapts as quickly as the scammers do.

## III. LITERATURE REVIEW

The Traditional Approach: For years, research has leaned on statistical models, clustering, and distance-based tracking to find outliers. While these worked for simpler datasets, they often "hit a wall" when faced with the massive, high-dimensional data found in modern e-commerce.

The Power of Autoencoders: Deep learning has introduced a smarter way to handle this complexity.

Autoencoders are particularly effective because they learn to "compress" and then reconstruct normal data; if a piece of data doesn't fit the usual pattern, the model can't reconstruct it accurately, signaling a potential anomaly.

The Precision of Isolation Forest: On the more classical side, Isolation Forest stands out as a highly efficient tool. Instead of defining what "normal" looks like, it works by randomly partitioning data to isolate the "loners"—the anomalies—making it incredibly fast and effective for massive datasets.

The Future is Hybrid: Recent breakthroughs suggest that we don't have to choose between deep learning and classical math. By combining the pattern-recognition of models like Autoencoders with the outlier-spotting speed of Isolation Forest, we can create a system that is far more accurate and adaptable in unsupervised environments.

## IV. PROPOSED METHODOLOGY

4.1 System Overview
The proposed system uses a hybrid approach:
1. Autoencoder for feature learning
2. Isolation Forest for anomaly scoring

4.2 Data Pre-processing
- Transaction data is cleaned and normalized
- Features include:
- Transaction amount
- Frequency of purchase
- Time intervals
- Product categories

4.3 Autoencoder Model
The Autoencoder is trained on normal transaction data:
- Encoder compresses input into a latent representation
- The original input is reconstructs decoder
- Reconstruction error is calculated

Higher reconstruction error indicates an anomaly.

4.4 Isolation Forest
The latent features generated by the Autoencoder are passed to Isolation Forest:
- Random trees isolate anomalies quickly
- Shorter path length → higher anomaly score

4.5 Hybrid Detection
The final anomaly score is computed by combining:
- Reconstruction error (Autoencoder)
- Isolation score (Isolation Forest)

## V. SYSTEM ARCHITECTURE

Flow:
1. Input Transaction Data
2. Data Pre-processing
3. Autoencoder Training
4. Feature Extraction
5. Isolation Forest
6. Anomaly Detection Output

## VI. IMPLEMENTATION DETAILS

- Programming Language: Python / PHP integration possible
- Libraries:
- TensorFlow / PyTorch (Autoencoder)
- Scikit-learn (Isolation Forest)
- Dataset:
- Simulated e-commerce transactions
- Real-world datasets (if available)

## VII. RESULTS AND DISCUSSION

The hybrid model was evaluated on transaction datasets with injected anomalies.
Performance Metrics:
- Precision
- Recall
- F1-score

Observations:
- Autoencoder effectively captured normal behaviour patterns
- Isolation Forest improved detection of rare anomalies
- Hybrid model outperformed standalone models

Advantages:

- Works without labeled data
- Detects new fraud patterns
- Scalable for large datasets

## VIII. APPLICATIONS

- Fraud detection in online shopping
- Payment security monitoring
- Customer behaviour analysis
- Inventory anomaly detection

## IX. LIMITATIONS

- Tuning of hyperparameters required
- In rare cases it may produce false positives
- Performance depends on quality of data provided

## X. FUTURE WORK

- Integration can be done with real-time streaming systems
- Usage of advanced models like Transformers
- Incorporation of user profiling in the system
- Deployment should be done oncloud platforms

## CONCLUSION

By combining the pattern-recognition power of Autoencoders with the outlier-spotting efficiency of the Isolation Forest algorithm, this paper introduces a hybrid framework designed to catch sophisticated e-commerce fraud in real-time. Unlike traditional security that relies on a rigid "checklist" of known scams, this unsupervised approach learns to identify "normal" behaviour on its own, allowing it to flag unusual transactions without needing a massive library of pre-labelled data. The results show that blending deep learning with classical statistical methods creates a robust, scalable shield that can adapt to new, creative fraud tactics as they emerge, effectively protecting both a company's bottom line and its customers' trust in an increasingly high-velocity digital market.

## REFERENCES

[1] Liu, F. T., Ting, K. M., & Zhou, Z. H., "Isolation Forest," IEEE, 2008.

[2] Goodfellow, I., Bengio, Y., & Courville, A., Deep Learning, MIT Press, 2016.

[3] Chandola, V., Banerjee, A., & Kumar, V., "Anomaly Detection: A Survey," ACM Computing Surveys, 2009.

[4] Aggarwal, C. C., Outlier Analysis, Springer, 2017.

[5] Chalapathy, R., & Chawla, S., "Deep Learning for Anomaly Detection," arXiv, 2019