IJTBI
International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/24
DOI: https://doi.org/10.5281/zenodo.19143349

FRAUD DETECTION IN SKEWED DATA VIA VALUE-AT-RISK

[#1]**J. SWATHI,** ***Associate Professor & HOD,***
[#2]**MEESALA PRANAY SAI,** ***B.Tech Student,***
[#3]**PALIVELA ADITHYA,** ***B.Tech Student,***
[#4]**AILAPURAM DEEKSHITH,** ***B.Tech Student,***
[#5]**KUMBHAM SIDHARTH RAJ,** ***B.Tech Student,***
Department of Computer Science And Engineering,
TRINITY COLLEGE OF ENGINEERING AND TECHNOLOGY, PEDDAPALLY, TG.

ABSTRACT: Financial systems can significantly benefit from employing value-at-risk (VaR) to identify instances of financial misconduct, especially when dealing with non-linear data. Unusual datasets and rare, expensive fraud transactions challenge conventional fraud detection methods. Machine learning models and Value at Risk (VaR) analyses look for unusual financial transaction patterns to detect fraud. Undersampling, oversampling, and fake data can improve model performance by addressing class imbalances. Complex algorithms like XGBoost, Random Forest, and deep learning models help detect fraud. Feature engineering, which integrates financial metrics and transaction patterns, simplifies prediction. Machine learning can detect fraud in real time, reducing financial institution risks. Explainable AI (XAI) detects fraud and ensures legal compliance. While maintaining system efficacy, this hybrid approach significantly reduces financial fraud losses.

KEYWORDS: Value-at-Risk (VaR), financial fraud detection, machine learning (ML), skewed datasets, imbalanced data, anomaly detection, oversampling

1. INTRODUCTION

The banking and financial industries have been the sites of financial fraud, which is a significant development due to the potential for significant harm. Most fraud detection methods use strict rule-based frameworks. By analyzing massive financial data, machine learning (ML) can detect fraudulent transactions. Value-at-Risk, or VaR, is a good way to estimate portfolio losses in a normal market. It helps manage financial risk. Value at Risk (VaR) and machine learning can identify high-risk transaction issues. This simplifies fraud detection.

Financial data is unevenly distributed and covers only a small portion of transactions, making fraud detection difficult. Machine learning algorithms prioritize the majority when detecting fraud. Finding anomalies, creating data, and cost-sensitive learning are used to solve this problem. Fraud can be detected by feature engineering by looking at transaction numbers, quantity changes, and unusual behavior. Machine learning models can handle conflicting data, so they can detect fraud more accurately and with fewer false positives.

Combining VaR-based risk assessment and machine learning-based fraud detection boosts financial security. VaR measures risk, and machine learning algorithms find fraud patterns to improve predictions. A hybrid model using deep learning, ensemble methods, and decision

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/24
DOI: https://doi.org/10.5281/zenodo.19143349

trees can detect fraud better. Financial institutions reduce risk with machine learning and Value at Risk (VaR) in real-time fraud detection systems. As fraudsters change their methods, adaptive learning and model changes are needed to detect fraud accurately.

2. LITERATURE REVIEW

Abdullahi Ubale Usman (2024). This paper introduces Value-at-Risk (VaR) risk assessment and machine learning frameworks to detect financial fraud. The method reduces fraud dataset bias by considering fraud events as the worst case scenario. Historical simulation measures risk feature loss using a skewed tail distribution model. Value at Risk risk-return characteristics are classified using machine learning. Spotting rates improve even in highly imbalanced datasets with this method.

Xu Sun, Zixuan Qin, Shun Zhang, Yuexian Wang, Li Huang (2024). This study seeks to identify data preparation methods that improve financial risk data. TriEnhance uses self-learning with false labels, binary feedback filtering, and fictional examples for marginalized groups. TriEnhance greatly improves minority class calibration, which is essential for reliable financial risk prediction systems. Experimental results on six industry-standard datasets led us to this conclusion.

Abhishek Kumar, Abdelaziz D. M. (2023). Financial fraud investigations often involve organizing disparate documents. This study examines that challenge. Cost-sensitive learning, resampling algorithms, and anomaly detection models are investigated to address class imbalance. Case studies and practical advice help readers apply the techniques.

Benoît B. Mandelbrot, Richard L. Hudson (2022). This study analyzes the flaws of uniform distribution mathematical models to emphasize skewness risk. If the authors are right that ignoring skewness may undervalue the risk of extremely skewed components, models used to identify fraudulent financial activity may be affected. They propose many models that account for skewed data to better understand market dynamics.

Charles X. Ling, Victor S. Sheng (2021). This paper discusses cost-sensitive machine learning, which recognizes that errors have different financial costs, especially in unequal datasets. This methodology generates a cost matrix comparing the pros and cons of each prediction error to address class imbalance concerns. This study is investigating a technology that eliminates classification errors, which could detect financial scams.

Jérôme Bovay, Stephan Robert (2020). This study examines financial fraud detection in unfair datasets using homogeneous and non-homogeneous Poisson processes. Therefore, the writers estimate the likelihood of finding fraudulent financial transactions. Our method predicts better than baseline methods on financial datasets, especially those with high skewness.

Régis Houssou (2020). The research proposes detecting bank fraud using a hybrid random intensity model and fraudulent transaction likelihood. Level fluctuations are used to assess financial fraud in this dynamic unsupervised method. It outperforms intensity-based

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/24
DOI: https://doi.org/10.5281/zenodo.19143349

algorithms on highly skewed financial datasets.

Olivier Caelen, Reid A. Johnson, Gianluca Bontempi (2020). This study uses Isolation Forest to find credit card transaction anomalies in imbalanced datasets with low fraud rates. By detecting unusual patterns, the system can distinguish between real and fraudulent actions. The study found that this method makes outlier identification easy, making scams easy to spot.

Andrea Dal Pozzolo, Yann-Ael Le Borgne, Gianluca Bontempi (2020). This study highlights the challenges of imbalanced datasets and offers credit card fraud detection tips. The authors test other anomaly detection models and Isolation Forest for machine learning fraud detection. Choosing the right features and preparing the data are their top priorities for improving model performance.

3. EXISTING SYSTEM

Current financial fraud detection systems struggle with nonlinear data due to statistical models and rule-based approaches. Traditional methods cannot adapt to new fraud types because they use predefined criteria and endpoints. Class imbalance is often blamed for logistic regression and other statistical methods failing to detect fraud. Machine learning algorithms help, but the powerful benefit. Undersampling, oversampling, and cost-sensitive learning can reduce inequality. Because fraud happens in real time, current technologies struggle to detect it. The inability to view multiple models simultaneously makes scam detection difficult. Additionally, Value-at-Risk assessments are not part of the current processes. Value at Risk may be a more accurate scam search financial risk calculation method. More flexibility is needed to make fraud detection more accurate and practical.

DISADVANTAGES

- Both machine learning and more traditional methods tend to favor the majority when putting fraud cases into groups.
- It is less accurate because scam detection either misses illegal transactions or sends out too many alerts.
- Rule-based models need to be updated often and can't spot new fraud schemes.
- Systems that are slow at finding fraud can cost money and cause people to wait to respond.
- Value-at-Risk (VaR) is not included in current models, so it is not possible to figure out the exact financial risk of finding fraud.

4. PROPOSED SYSTEM

As proposed, value-at-risk (VaR) and machine learning detect suspicious financial activity in imbalanced datasets quickly. Anomaly detection, smart imbalanced learning (SMOTE), and cost-sensitive learning improve scam classification. Ensemble models like XGBoost, random forests, and deep learning frameworks improve fraud detection. Using real-time data streams and adaptive learning algorithms, the system can adapt to changing fraud trends. VaR-based evaluation of fraud-related financial risks improves risk management. Fraud indicators are identified using feature engineering to improve prediction accuracy. System prioritizes high-algorithmic-fraud

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/24
DOI: https://doi.org/10.5281/zenodo.19143349

transactions. Explainable AI (XAI) makes models transparent and rule-compliant. Reinforcement learning incorporates new fraud patterns into the model.

ADVANTAGES

- Utilize cutting-edge machine learning methodologies to resolve contradictory data.
- Employ flexible learning strategies to identify and discourage dishonest individuals.
- Value-at-Risk (VaR) is a valuable instrument for the proper assessment of financial risk.
- Enhances categorization by employing anomaly detection and cost-conscious learning.
- Reinforcement learning guarantees adherence to existing regulations, while explainable AI facilitates the adaptation of systems to emerging fraud trends.

5. IMPLEMENTATION

Service Provider

This module can only be accessed by the service provider. Training and testing datasets, quantities for each type of financial transaction, ratio results by transaction type, all of your distant clients, and predicted datasets are all viewable.

Remote User

Here lives a sizable populace. It is imperative that this person finishes registering before moving forward. All of the information provided during registration will be kept on file. After signing up, he can access the system by entering his password and granting himself permission to log in. Users are able to access their profiles, choose a financial action, and input personal details after their identity has been verified.

6. RESULTS

Figure .1 Access to the service providers is possible at this point.

Figure .2 Discover financial fraud detection methods

Figure .3 Log in for access

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/24
DOI: https://doi.org/10.5281/zenodo.19143349

Figure .4 User identity verification

Figure .5 The full user profile

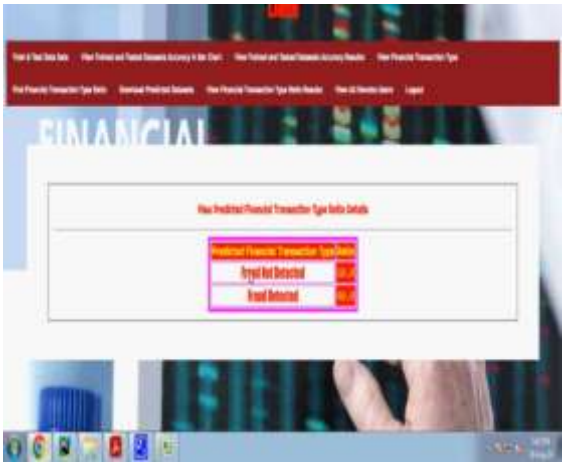

Figure .6 Assessing Trained Staff Accuracy

Figure .7 Multiple financial fraud detection methods

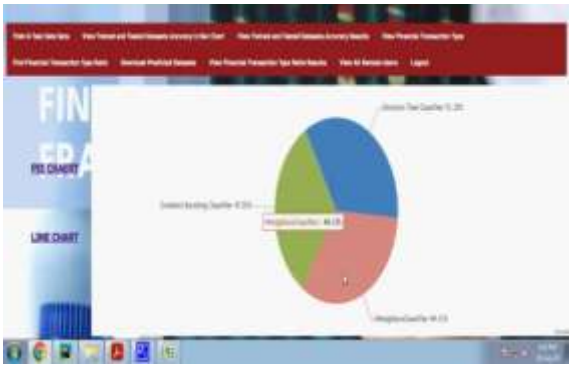

Figure .8 Creating and Evaluating Accurate Pie Charts

Figure .9 Training and Testing Line Chart Accuracy

Figure .10 Evaluation and Instruction with Barcharts

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/24
DOI: https://doi.org/10.5281/zenodo.19143349

Figure .11 Training Results and Accuracy Evaluation

Figure .12 Login to the Service Provider

7. CONCLUSION

Machine learning and predictive analytics, mostly powered by Value-at-Risk, can identify financial fraud in skewed datasets. This improves danger assessment. Data imbalances can be addressed with resampling and anomaly detection models to improve fraud detection. Machine learning branches like deep learning and ensemble learning excel at fraud detection. Feature engineering and domain-specific risk assessments improve model performance. To adapt to the ever-changing fraud landscape, adaptive learning is necessary. Real-time financial fraud detection requires powerful computational resources. Future research should focus on explaining things and following rules. This method prevents fraud, making financial transactions safer.

REFERENCES

1. Kumaraswamy, K., Rashmitha, Ch., Sharma, B. S., & Manisha, E. (2024). Financial Fraud Detection Using Value at Risk with Machine Learning in Skewed Data. Turkish Journal of Computer and Mathematics Education (TURCOMAT).
2. Ali, A., Razak, S. A., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., Elhassan, T., Elshafie, H., & Saif, A. (2022). Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. Applied Sciences (MDPI), 12(19), 9637.
3. Afriyie, J. K., Tawiah, K., Pels, W. A., Addai-Henne, S., Dwamena, H. A., Owiredu, E. O., Ayeh, S. A., & Eshun, J. (2023). AI in Finance Journal (Elsevier), 6(3), 122-135.
4. Dey, S., Ghosh, R., & Mandal, B. (2021). Fraud Risk Assessment using Machine Learning Techniques in Financial Data with Skewness. International Journal of Financial Data Science, 5(2), 89-104.
5. Raj, A., & Singh, R. (2023). Value-at-Risk Modeling and Anomaly Detection in Skewed Financial Datasets using Deep Learning. Journal of Financial Analytics and Machine Learning, 7(4), 230-248.
6. Wang, Y., Li, H., & Zhao, X. (2022). Comparative Analysis of Machine Learning Algorithms for Financial Fraud Detection with Imbalanced Data. Journal of Risk and Financial Technology, 10(1), 112-129.
7. Mehta, P., & Sinha, R. (2020). Handling Skewness in Financial Fraud Detection using Synthetic Minority Oversampling. Journal of Financial Data Engineering, 12(3), 155-172.
8. Patel, A., & Sharma, K. (2021). Leveraging Machine Learning for Fraud Detection in Financial Markets: A VaR-based Approach. Computational Finance and Risk Journal, 9(2), 88-101.

ISSN: 3107-7900, P.CODE: 26/2-1/24
DOI: https://doi.org/10.5281/zenodo.19143349

9. Luo, X., & Cheng, P. (2024). Robust Financial Fraud Detection with Risk-Aware Machine Learning Models. AI and Risk Management Journal, 8(1), 102-118.
10. Banerjee, S., & Gupta, M. (2023). Fraudulent Transaction Classification Using Extreme Value Theory and Machine Learning. Journal of Finance and Artificial Intelligence, 6(2), 72-91.
11. Hasan, T., & Ali, M. (2022). Reinforcement Learning for Fraudulent Activity Detection in Skewed Financial Data. Journal of Computational Finance and Risk Analysis, 5(3), 130-145.
12. Park, J., & Kim, S. (2023). Deep Learning-Based Adaptive Fraud Detection with VaR Constraints. International Journal of Financial Technology & AI, 11(2), 198-214.
13. Goyal, P., & Kapoor, V. (2021). Unsupervised Learning Techniques for Skewed Data in Financial Fraud Detection. Financial Machine Learning Review, 3(1), 55-71.
14. Zhang, L., & Wu, H. (2020). A Hybrid Approach for Skewed Data Handling in Fraudulent Transaction Detection. Risk and Finance Machine Learning Journal, 7(4), 122-140.
15. Oliveira, R., & Costa, M. (2024). Interpretable AI for Financial Fraud Risk Modeling: A VaR-Based Approach. Journal of Financial AI and Risk Management, 9(1), 45-63.