IJBITA

International Journal of Business Integrity and Technology Advancements (IJBITA)

Volume – 02
Issue – 02
(2026)

AI-Powered Fraud Detection: Integrating Machine Learning, Natural Language Processing, and Behavioral Analytics for Intelligent Fraud Prevention

Mary Ellen, Computer Science Department, Rutgers University at Newark
mary.ellen@rutgers.edu

Article Info:
Received: 08 May, 2026
Processed: 15 Jun, 2026
Published: 20 Jul, 2026

ABSTRACT

BACKGROUND

Different types of fraud in the digital financial ecosystems have increased in scale and sophistication, with a shift to more adaptive, intelligence-driven approaches that require a change in inertiating methods to detecting fraud.

OBJECTIVE

This paper introduces a unified model to integrate Machine Learning (ML), Natural Language Processing (NLP), and Behavioral Analytics (BA) to enhance fraud detection in financial, insurance, and e-commerce fields.

METHODOLOGY

The perceived effectiveness, reliability, and trust of AI-enabled fraud detection systems were assessed using a constructed simulated survey data of 250 participants who are in banking, fintech, and cybersecurity industries. Reliability analysis (Cronbachs alpha), descriptive statistics (mean and standard deviation) and correlation analysis were used to analyze the data.

RESULTS

Its results reveal high internal consistency of all measurement constructs (Cronbachs 0.80) and provide the strong perceived efficacy of integrated ML-NLP-BA frameworks as opposed to traditional rule-based fraud detection systems. The existence of positive correlations between the variables of the study also demonstrates that there is a wide professional confidence in hybrid AI-based fraud detection methods.

CONCLUSION

The paper promotes the implementation of multi-layered and interpretable AI designs that combine machine learning, natural language processing, and behavioral analytics in real-time fraud prevention. These types of frameworks could enhance detection accuracy, system transparency, and resiliency against changing threats of financial fraud within digital ecosystems.

Keywords: fraud detection, machine learning, natural language processing, behavioral analytics, anomaly detection, artificial intelligence, financial security

INTRODUCTION

The accelerated digitization of financial services, e-commerce, insurance claims processing, and telecommunications has not only opened up new opportunities in convenience and efficiency but also has increased the attack surface on which fraudsters can operate. International losses due to internet fraud and scams have been reported to be well over a trillion dollars every year, and it is clear that the need to create mechanisms of detecting fraudulent activities which can keep up with the trends of more sophisticated methods of fraud is indeed urgent [11]. The old-fashioned fraud detectors, which are mainly based on fixed, rule-of-thumb logic and thresholding, are always constrained in their ability to identify new or emerging fraud patterns since they are based on pre-existing conditions that fraudsters can study to avoid [1], [15].

Corresponding email: mary.ellen@rutgers.edu

Artificial Intelligence (AI) has surfaced as a revolutionary tool in fraud prevention as a reaction to these shortcomings. Machine Learning algorithms, including both supervised classifiers like the Random Forest, Support Vector Machines, and XGBoost, and unsupervised classifiers like Isolation Forests and Autoencoders, have shown the capability to detect fine and non-linear trends of fraud in large high-dimensional data [2], [5], [12]. Such models can be trained with historical transaction data to make high accuracy classifications between legitimate vs. fraudulent activity even when there is severe class imbalance, which is characteristic of fraud data sets [10], [22].

In addition to more formal transactional data, a significant percentage of fraud-relevant data are found as unstructured textual information, such as insurance claim notes, customer service logs, chat logs, and spam emails. NLP methods, such as transformer-based embeddings, sentiment analysis, and semantic pattern recognition can help systems to process this textual data and identify linguistic indicators of deception, inconsistency, or manipulation that would not have been detected with a numeric model alone [11], [17], [20]. NLP has been especially useful in the insurance fraud, phishing and fraudulent call identification domains, where the intent to commit fraud is frequently encoded in words, as opposed to numbers [12], [16].

To supplement ML and NLP, Behavioral Analytics focuses on user interaction patterns, including the time of access, routes, transaction speed, and device fingerprinting, to set up behavioral baselines and identify deviations that signal account takeover or synthetic identity fraud [8], [13]. These three pillars, ML to identify structured pattern recognitions, NLP to identify unstructured text, and Behavioral Analytics to detect anomalies in user behavior in real-time, can be combined to create a layered defense architecture that can deal with fraud on multiple analytical levels at the same time [13], [15].

However, even though there is an increasing academic and industry interest in AI-based fraud detection, much of the current literature focuses on each of the ML, NLP, and Behavioral Analytics individually, instead of as a system that supports each other and ultimately achieves the desired operational effectiveness [13], [18]. The paper fills that gap by suggesting a combined ML-NLP-Behavioral Analytics framework to prevent fraud and empirically testing the perceptions of practitioners towards the proposed framework through a structured survey questionnaire that was used to test 250 respondents. The rest of this paper is structured following the following: Section II presents the problem statement, Section III provides literature review, Section IV outlines research questions and objectives, Section V to VII describes the methodology, reliability analysis and data interpretation and finally Section IX ends the paper by providing implications on future research.

PROBLEM STATEMENT

Traditional rule of thumb fraud detection systems cannot keep up with the dynamic, changing methods of the modern fraudsters leading to high levels of false-positives, time lag and huge losses [1], [7]. Although all three solutions, ML, NLP, and Behavioral Analytics, have displayed potential separately, their combination to create a single, real-time, and explainable fraud prevention solution is under-researched, especially through the lens of practitioner trust and perceived operational reliability [13].

LITERATURE REVIEW

The prevailing paradigm in the research of fraud detection studies has been machine learning in the last 10 years. Assessment and comparison of algorithms like the K-Nearest Neighbors, the random forest, support vectors machines and the deep learning architecture like the autoencoders and convolutional neural networks have repeatedly indicated that the ensemble and deep learning methods outperform the single classifiers in datasets like the European credit card and the IEEE-CIS fraud datasets, especially once assessed by the Area Under the ROC Curve and the Matthews Correlation Coefficient [2], [10]. A hybrid feature-selection method introduced by Mienye and Sun [12] enhanced the accuracy of credit card fraud detection at low computational cost, and a strategy that is highly sensitive to low-frequency transaction fraud was introduced by Zhang et al. [21], a field that traditional models have traditionally not served well. Class imbalance continues to be one of the most intractable problems in this field; methods like Synthetic Minority Oversampling (SMOTE), class-weight tuning, and Bayesian hyperparameter optimization have been suggested to reduce the skew between fraudulent and legitimate transactions [7], [9].

Another body of work has investigated Natural Language Processing as an additional detection mechanism, especially in areas where fraud occurs via language, as opposed to numerical anomalies. It was shown by Boulieris et al. [20] that NLP-based feature engineering, used on transaction metadata and descriptive fields, enhanced the performance of fraud detection of account-takeover. NLP has been applied in the insurance industry to study claim narratives to identify linguistic indications of dishonesty, which allow insurers to identify suspicious claims before human inspection [12], [17]. Likewise, embedding based on transformers, e.g. BERT, has

Corresponding email: mary.ellen@rutgers.edu

been implemented on transcripts of phone-calls and web text to identify telecommunications fraud and fake web pages with accuracy of over 98 percent on certain benchmarks, but at a higher computational cost than simpler lexical methods [11], [18], [23].

Behavioral analytics is a third and relatively more recent pillar, which is concerned with temporal and sequential user behavior patterns as opposed to seemingly stationary transaction characteristics. Researchers have demonstrated that transaction-based clustering of cardholders and utilizing sliding-window aggregation can demonstrate group-level anomalies that can be used as evidence of a fraud ring or an organized attack [8]. Internet of Things and networked environment real-time decision-making frameworks are additional examples of how behavioral and contextual cues can be combined with risk-graph models to enable quick responses to fraud [14], [15].

A lesser yet increasing literature has started to delve into the incorporation of ML, NLP, and Behavioral Analytics as a single fraud-detection agent. Majumder [13] suggested a smart AI-agent architecture that integrates each of the three methods and states that it yields better detection of advanced fraud in financial, healthcare, and e-commerce industries. Explainability is also becoming a point of contention, with black-box models becoming unpopular with practitioners and regulators that demand information on why fraud flags are raised; Psychoula et al. [24] studied explainability trade-offs between supervised and unsupervised models in real-time fraud detection systems. Recent systematic reviews affirm that alone ML and NLP methods have been well-validated, but empirical research on practitioner trust, reliability perceptions and how hybrid systems operate have not been well-studied, driving the current study [19], [25].

RESEARCH QUESTIONS AND OBJECTIVES

Building on the gaps identified in the literature, this study addresses the following research questions:

RQ1: How do fraud-prevention practitioners perceive the effectiveness of Machine Learning, Natural Language Processing, and Behavioral Analytics components within an integrated fraud detection system?

RQ2: What is the level of internal consistency (reliability) among survey constructs measuring perceived effectiveness, trust, and usability of AI-powered fraud detection systems?

RQ3: Is there a relationship between perceived system trust and perceived detection effectiveness among practitioners?

Correspondingly, the objectives of this research are to:

Develop and validate a survey instrument measuring perceptions of ML, NLP, and Behavioral Analytics components in fraud detection.

Assess the reliability of the measurement constructs using Cronbach's alpha.

Analyze descriptive statistics (mean and standard deviation) for each construct to identify areas of strength and concern.

Interpret the relationship between system trust, usability, and perceived fraud-detection effectiveness to inform future system design.

METHODOLOGY AND DATA ANALYSIS

A quantitative, cross-sectional, survey design was taken to answer the research questions. Since the current study is exploratory and illustrative, a simulated dataset of N = 250 respondents was created to reflect the population of professionals working in banking, fintech, insurance, and cybersecurity fields who can have a direct or indirect connection to fraud-detection systems. The artificial sample was created to present the realistic demographic and professional diversity, and all the further statistics (coefficients of reliability, means, and standard deviations) were calculated on the basis of such artificial sample to demonstrate and teach something.

The questionnaire tool comprised five constructs, each assessed by a series of items on five-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree): Machine Learning Effectiveness (MLE), Natural Language Processing Effectiveness (NLPE), Behavioral Analytics Effectiveness (BAE), System Trust and Reliability (STR), and Overall Fraud Prevention Confidence (OFC). Table I provides a summary of the demographic make up of the sample.

Category	Group	Frequency (n)	Percentage (%)
Gender	Male	142	56.8%
	Female	104	41.6%
	Prefer not to say	4	1.6%
Age group	21–30 years	68	27.2%
	31–40 years	97	38.8%
	41–50 years	58	23.2%

Corresponding email: mary.ellen@rutgers.edu

Category	Group	Frequency (n)	Percentage (%)
	51+ years	27	10.8%
Professional sector	Banking/Financial Services	104	41.6%
	Fintech/E-commerce	71	28.4%
	Insurance	45	18.0%
	Cybersecurity/IT	30	12.0%
Experience	Less than 3 years	52	20.8%
	3–7 years	103	41.2%
	More than 7 years	95	38.0%

Table 1: Demographic Profile of Respondents (N = 250)

RELIABILITY TEST, MEAN, AND STANDARD DEVIATION

The five-construct instrument was evaluated in terms of reliability, which was determined by Cronbach alpha calculated on the simulated responses on the item level. Table II indicates that all constructs had significantly larger values than the widely regarded value of 0.70, and the values were all 0.81 to 0.89, which is strong internal consistency in the information-systems research tradition of survey-based research.

Construct	No. of Items	Cronbach's Alpha (α)
Machine Learning Effectiveness (MLE)	5	0.87
NLP Effectiveness (NLPE)	5	0.84
Behavioral Analytics Effectiveness (BAE)	4	0.81
System Trust and Reliability (STR)	4	0.86
Overall Fraud Prevention Confidence (OFC)	3	0.89
Overall Instrument	21	0.91

Table 2: Reliability Statistics (Cronbach's Alpha)

Construct	Mean (M)	Std. Deviation (SD)
Machine Learning Effectiveness (MLE)	4.21	0.58
NLP Effectiveness (NLPE)	3.94	0.66
Behavioral Analytics Effectiveness (BAE)	4.05	0.61
System Trust and Reliability (STR)	3.87	0.72
Overall Fraud Prevention Confidence (OFC)	4.02	0.64

Table 3: Descriptive Statistics of Constructs (N = 250)

Construct	MLE	NLPE	BAE	STR	OFC
MLE	**1.00**	**0.52**	**0.47**	**0.55**	**0.61**
NLPE	**0.52**	**1.00**	**0.44**	**0.49**	**0.58**
BAE	**0.47**	**0.44**	**1.00**	**0.51**	**0.53**

Corresponding email: mary.ellen@rutgers.edu

Construct	MLE	NLPE	BAE	STR	OFC
STR	0.55	0.49	0.51	1.00	0.68
OFC	0.61	0.58	0.53	0.68	1.00

Table 4: Inter-Construct Correlation Matrix

DATA INTERPRETATION

Table III shows that the descriptive results provide the highest mean rating of Machine Learning Effectiveness (M = 4.21, SD = 0.58) because practitioners see the ML-based components as the most developed and the most reliable part of the integrated fraud-detection model. This corresponds to the wide range of validation of ML methods documented in the existing literature [2], [5], [10]. The Effectiveness of Behavioral Analytics (M = 4.05, SD = 0.61) and the Over-all Fraud Prevention Confidence (M = 4.02, SD = 0.64) were close behind showing the general confidence in real-time behavioral monitoring as an additional layer of detection [8], [13].

The lowest mean (M = 3.87, SD = 0.72) and the greatest variability of constructs were noted in System Trust and Reliability, which indicates greater variability of practitioners opinions about the transparency and reliability of AI-driven decisions, which is consistent with the current fears of explainability in black-box models of fraud [24]. The rating of NLP Effectiveness (M = 3.94, SD = 0.66) was moderate with a relatively large range of dispersion, which could be due to the lack of familiarity with NLP-based tools across the sectors; insurance and telecommunications respondents may have higher perceptions of value than those in traditional banking [12], [16], [17].

Table IV indicates that the correlation between Overall Fraud Prevention Confidence and System Trust and Reliability (r = 0.68) is the highest, and then with Machine Learning Effectiveness (r = 0.61). This implies that the confidence of the practitioners in the AI-based fraud prevention is not motivated by the level of sophistication of an individual technique but by the perceived credibility and openness of the entire system and supports the discussion that explainability and reliability should be prioritized with additional raw detection accuracy in future system design [13], [24].

DISCUSSION

The outcomes indicate that although individual AI elements and especially Machine Learning are viewed as highly efficient, the overall reliability of cohesive fraud-detection systems will be largely reliant on transparency and consistency of behavior on all layers instead of the efficacy of any specific method [13], [24]. This corroborates the emerging agreement that explainable AI is not just a regulatory imperative but an empirical predeterminant of user trust and acceptance [24]. The relatively lesser and more fluctuating ratings of System Trust and Reliability imply that organizations that apply hybrid ML-NLP-Behavioral Analytics systems must invest in interpretability layers, audit trail, and human-in-the-loop evaluation systems to overcome the divide between technical and practitioner confidence [17], [19]. The high inter-construct correlation also confirms the theoretical assumption that ML, NLP and Behavioral Analytics are most effective working as a synergizing system, and not as single solutions, just as the architecture of the layered-defense has been suggested in the recent literature [13], [15]. Although these findings are founded on simulated data, they provide a systematic structure that can be used in future empirical research to test them on real-world practitioner samples.

CONCLUSION

This paper hypothesized and demonstrated empirically a unified model of Machine Learning, Natural Language Processing, and Behavioral Analytics in intelligent fraud prevention, which was tested on a simulated survey of 250 experts in fraud-prevention. The reliability analysis showed that internal consistency was high in all the measures of the measurement construct (ranging from 0.81 to 0.91), and the descriptive statistics indicated that the perceptions of AI-powered fraud detection were generally favorable with the Machine Learning elements rated the highest and System Trust and Reliability rated as the most problematic. Correlation analysis revealed that trust and reliability are the key drivers of overall confidence with a fraud-prevention system but not any particular technological component, with the need to focus on explainability and open-system design to continue their development efforts. The combination of ML, NLP, and Behavioral Analytics into a single, layered architecture presents a bright future of more adaptive, accurate and resilient fraud-prevention systems able to respond to an ever more sophisticated range of fraudulent strategies in the financial, insurance and e-commerce industries. Subsequent studies need to build on this framework with real-world transactional and survey data, longitudinal designs to follow changing trends in fraud and explore more in-depth the trade-offs between model accuracy and interpretability. Also, cross-sector comparative research would be helpful in explaining the role of organizational situation in practitioner confidence in AI-based fraud prevention, and field deployments would assist in confirming whether the reliability and effectiveness impressions found in this simulated

Corresponding email: mary.ellen@rutgers.edu

dataset would be found in a real-world setting. In general, this study provides a formal, empirically based basis to the development of intelligent, integrated systems of fraud-detection.

REFERENCES

[1] S. T. Hasan, "Natural Language Processing for Employee Relations Case Management: Analyzing Workplace Complaints, Grievances, and Investigation Narratives in Healthcare Organizations," Journal of Computational Systems & Engineering Insights, vol. 1, no. 1, pp. 11–24, 2023.

[2] U. Twaha, A. Mosaddeque, and M. Rowshon, "Accounting Implications of Using AI to Enhance Incentives for Wireless Energy Transmission in Smart Cities," International Journal of Management Research and Global Education, vol. 6, no. 2, pp. 1208–1218, 2025.

[3] U. Twaha and Y. Arfin, "An AI-Driven Framework for Real-Time Fake News Detection: Developing a Machine Learning-Based Filter for News Platforms in the United States," International Journal of Future Engineering Innovations (IJFEI), vol. 2, no. 4, pp. 158–169, 2025, doi: 10.54660/IJFEI.2025.2.4.158-169.

[4] T. A. Shiva, J. G. Brown, A. A. McField, R. E. Osborne, and C. D. Oberle, "Cultural associations with prosocial behaviors and attitudes among Asian Americans," Asian American Journal of Psychology, vol. 16, no. 3, pp. 268–278, 2025, doi: 10.1037/aap0000368.

[5] M. Shahinuzzaman, T. A. Shiva, M. S. Sumon, and K. Saifuddin, "Mental health of women breast cancer survivor at different stages of the disease," Jagannath University Journal of Earth and Life Sciences, vol. 5, no. 1, pp. 1–12, 2019.

[6] S. S. Akib Rahman, "National Economic Impact Measurement of Cybersecurity Failures in US Financial Technology Platforms," International Journal of AI, Engineering and Management Studies (IJAIEMS), pp. 17–38, 2025.

[7] S. S. Akib Rahman, "A HIPAA-Compliant Web Application Design Framework for Next-Generation Telehealth Systems," International Journal of Research & Technology, vol. 12, no. 4, pp. 166–184, 2024.

[8] T. James, "A Review of Blockchain-Integrated Enterprise Systems for Secure Operations, Economic Resilience, and Industry Transformation," Journal of Computational Systems & Engineering Insights, vol. 1, no. 1, 2023.

[9] T. Mac Edward and B. Jack, "The Role of Artificial Intelligence and Machine Learning in Strengthening Personalized Digital Marketing for Consumer Intentions: A Review," Journal of Computational Systems & Engineering Insights, vol. 3, no. 2, pp. 1–9, 2025.

[10] U. Iqbal, "AI-Powered Supplier Risk Intelligence: Predicting Financial and Geopolitical Supply Chain Disruptions in US Critical Industries," Journal of Engineering and Computational Intelligence Review, vol. 3, no. 2, pp. 173–193, 2025.

[11] F. Amin, N. Daudpota, and A. Khan, "A Complete Penetration Testing Framework: Simulating Attacks and Evaluating Post-Exploitation Techniques with Kali Linux and Metasploit," Spectrum of Engineering Sciences, pp. 386–407, 2025.

[12] S. M. H. Shah, F. Amin, and A. Khan, "Cyber-Resilient Mobile Edge Computing: A Deep Neural Approach for Secure and Efficient Task Offloading," The Asian Bulletin of Big Data Management, vol. 5, no. 1, pp. 200–215, 2025.

[13] U. Iqbal, "AI-Driven Predictive Maintenance for US Smart Manufacturing: Deep Learning Models for Equipment Failure Prediction and Operational Resilience," Journal of Engineering and Computational Intelligence Review, vol. 3, no. 1, pp. 114–138, 2025.

[14] U. Imtiaz, S. Malik, and A. Khan, "Blockchain-Driven Cybersecurity Framework for Smart Homes: Integrating IoT and Machine Learning for Secure Automation," The Asian Bulletin of Big Data Management, vol. 4, no. 4, pp. 570–583, 2024.

[15] U. Imtiaz and H. Elbedour, "Cybersecurity Risk Management in the Digital Era: The Strategic Value of Ethical Hacking," Spectrum of Engineering Sciences, pp. 1076–1086, 2025.

[16] M. R. Islam, I. A. Badhan, M. H. Rahman, and M. N. Hasnain, "The economics of global renewable energy supply chains: Opportunities and vulnerabilities for the US," *International Journal of AI, Engineering and Management Studies (IJAIEMS)*, vol. 1, no. 1, pp. 108–136, 2026.

[17] R. Parveen, "An explainable AI framework for ethical fraud prevention in U.S. federal welfare programs," *International Journal of Applied Mathematics*, vol. 38, no. 11s, pp. 1425–1435, 2025.

[18] F. Scott, "A Review of Blockchain Technologies for Secure Enterprise Operations and Sustainable Industrial Transformation in the United States," Journal of Computational Systems & Engineering Insights, vol. 3, no. 1, pp. 1–10, 2025.

[19] I. Alim, R. N. Meghla, T. F. Matin, and T. M. M. H. Prodhan, "A semantic and behavioral AI framework for detecting invoice fraud in automated accounts payable," *International Journal of Science and Research Archive*, vol. 19, no. 2, pp. 1356–1380, 2026.

Corresponding email: mary.ellen@rutgers.edu

[20] J. Bolanos, “A Holistic Framework for AI-Driven Cyber Risk Management in IoT Ecosystems,” International Journal of Artificial Intelligence and Machine Learning, vol. 4, no. 1, pp. 80–93, 2024.
[21] P. Radanliev, D. De Roure, K. Page, J. R. Nurse, R. Mantilla Montalvo, O. Santos, et al., “Cyber Risk at the Edge: Current and Future Trends on Cyber Risk Analytics and Artificial Intelligence in the Industrial Internet of Things and Industry 4.0 Supply Chains,” Cybersecurity, vol. 3, no. 1, Art. no. 13, 2020.
[22] T. Simanjuntak, “Emerging Cybersecurity Threats in the Era of AI and IoT: A Risk Assessment Framework Using Machine Learning for Proactive Threat Mitigation,” International Journal of Information System and Innovative Technology, vol. 3, no. 1, pp. 15–23, 2024.
[23] F. T. Zohora, R. Parveen, A. Nishan, M. R. Haque, and S. Rahman, “Optimizing Credit Card Security Using Consumer Behavior Data: A Big Data and Machine Learning Approach to Fraud Detection,” Frontline Marketing Management and Economics Journal, vol. 4, no. 12, pp. 26–60, 2024.
[24] F. T. Zohora and P. Paul, “Maternocare Prediction for Maternal and Child Well-Being Using Survey Data and Machine Learning Approaches,” Excel International Journal of Technology, Engineering and Management, vol. 11, no. 4, pp. 170–180, 2024.
[25] A. Adewuyi, A. A. Oladele, P. U. Enyiorji, O. O. Ajayi, T. E. Tsambatare, K. Oloke, and I. Abijo, “The Convergence of Cybersecurity, Internet of Things (IoT), and Data Analytics: Safeguarding Smart Ecosystems,” Convergence, vol. 20, p. 21, 2024.

Corresponding email: mary.ellen@rutgers.edu