Vol 5, Issue 4, April 2025 E-ISSN: 2582-9734
International Journal of Engineering, Science, Technology and Innovation (IJESTI)

Machine Learning-Enhanced Digital Forensics: A Multi-Dataset Analysis of Cybercrime Detection, Classification, and Judicial Admissibility

Amtabh Srivastava

Research Scholar, Computer Science, Sunrise University, Alwar.

Dr. Jitender Rai

Computer Science, Sunrise University, Alwar.

***Email**: Amitabh7500@Yahoo.Com*

ABSTRACT

The rapid growth of cybercrime has challenged traditional forensic methods, necessitating advanced approaches for analyzing complex digital evidence. This study examined the application of machine learning (ML) models across eight cybercrime datasets, including intrusion logs, financial fraud, phishing, malware, insider threats, darknet activity, blockchain records, and IoT botnet traffic. Supervised models such as Random Forest, ANN, and SVM showed high accuracy in structured datasets, while deep learning architectures (CNN, LSTM, Autoencoder–LSTM) excelled in unstructured or sequential data. Graph-based models, particularly Graph Neural Networks, provided valuable insights for blockchain forensics. Comparative analysis revealed that ensemble approaches leveraging multiple ML paradigms enhance operational readiness and judicial admissibility. Challenges identified include data imbalance, model explainability, and adversarial robustness. The study underscores that ML serves as a critical augmentation to human expertise, enabling adaptive, scalable, and legally defensible cyber forensic investigations.

Keywords: *Machine Learning, Cyber Forensics, Digital Evidence, Deep Learning, Graph Neural Networks, IoT Security*

1. Introduction

The digital age has reshaped human interaction, business practices, and information exchange, but it has also created fertile ground for cybercrime. Offenses such as hacking, ransomware, online fraud, and large-scale data breaches increasingly threaten individuals, organizations, and even national security. Unlike traditional crimes, cybercrimes generate massive, complex, and rapidly evolving datasets that surpass the capacity of conventional forensic methods. This has made digital forensics a critical discipline, tasked with collecting, analyzing, and presenting digital evidence to support investigations and legal proceedings. However, even within digital forensics, traditional approaches can struggle to keep pace with the sophistication of cyber threats. Machine Learning (ML), a branch of artificial intelligence, has emerged as a transformative tool in this context. By enabling systems to learn from data, recognize patterns, and make predictions, ML enhances the speed, accuracy, and efficiency of forensic analysis. Applications such as anomaly detection in network traffic, malware classification, and Intrusion Detection Systems (IDS) illustrate how ML can detect irregularities that might indicate cyberattacks. Similarly, Natural Language Processing (NLP) allows investigators to extract and categorize relevant evidence from unstructured data sources like emails, documents, and social media. These automated processes not only reduce human workload but also minimize the chances of overlooking subtle indicators of cybercrime. The role of ML extends further into predictive analytics, where historical data is analyzed to forecast

potential threats, allowing organizations to act proactively. Additionally, ML-driven User and Entity Behavior Analytics (UEBA) models help identify insider threats and compromised accounts by distinguishing abnormal behavior from normal patterns. This ability to adapt and evolve alongside emerging threats makes ML a vital component of modern cyber forensics. Nevertheless, integrating ML into forensic practice poses challenges. Model reliability depends heavily on high-quality, diverse training datasets, which are difficult to obtain. Biases in data can lead to false positives or negatives, undermining trust in forensic results. Moreover, legal professionals require interpretable outputs to present ML-derived evidence in court [1], raising the need for explainable AI models. Ethical considerations surrounding privacy and the handling of sensitive information further demand strict regulatory compliance. Despite these challenges, technological advances promise a strong future for ML in cyber forensics. Tools like TensorFlow, PyTorch, Scikit-Learn, and ELK Stack are already being deployed to support investigations. Innovations such as deep learning, blockchain integration for evidence integrity, and federated learning for secure collaboration continue to expand the possibilities. Ultimately, ML is redefining cyber forensics by providing automated, adaptive, and intelligent tools that enhance investigators' ability to combat cybercrime, ensuring a safer digital environment [2-5].

1.1 Types of Cybercrimes

Cybercrimes represent a diverse range of illegal activities carried out using digital technologies and networks. These crimes exploit system vulnerabilities and often involve data theft, unauthorized access, or disruption of services. Common types include hacking (unauthorized system access), phishing (fraudulent attempts to obtain sensitive information), malware (viruses, worms, ransomware), identity theft, financial fraud, and cyberbullying. Other serious forms are cyber espionage (stealing confidential information), data breaches (unauthorized extraction of sensitive data), denial of service (DoS/DDoS) attacks, and cyber terrorism targeting critical infrastructure. Combating these crimes requires strong cybersecurity measures, legal frameworks, and continuous innovation to stay ahead of evolving threats.

1.2 Evidence Collection and Preservation

Collecting and preserving digital evidence is vital to ensure its reliability and admissibility in investigations and court proceedings.

- **Evidence Collection** involves identifying potential sources (computers, mobile devices, servers), documenting their state, seizing them legally, acquiring forensic images using specialized tools, and verifying data integrity with checksums or hashes.
- **Evidence Preservation** ensures secure storage, strict chain of custody, complete documentation, and robust access controls to maintain confidentiality and integrity.
- **Best Practices** include using validated forensic tools, adhering to legal standards, ensuring quality assurance, and preparing clear reports that explain methods and findings while maintaining transparency and accountability.

1.3 Analysis Techniques in Digital Forensics

Analysis in digital forensics applies systematic methods to uncover and interpret evidence. File system analysis examines metadata and retrieves hidden or deleted files. Registry analysis reveals user activities on Windows systems. Timeline analysis reconstructs sequences of events using timestamps and system logs. Memory forensics focuses on volatile data in RAM to detect malware, live processes, and system activities. Through combining these methods with advanced forensic tools, investigators can reconstruct incidents, identify perpetrators, and generate reliable evidence for legal use and enhanced cybersecurity [6-9].

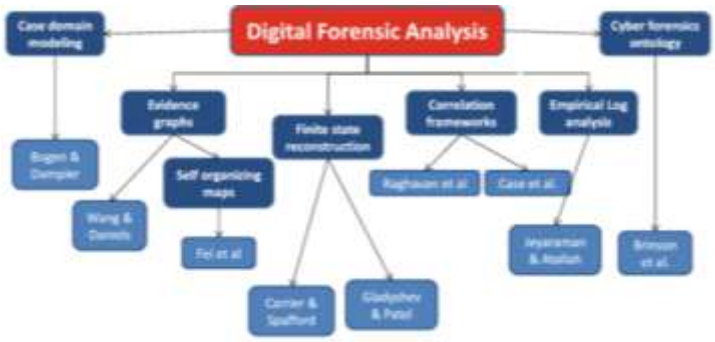

Fig 1: Analysis Techniques in Digital Forensics

1.4 Tools and Software

Digital forensics relies on specialized tools for evidence acquisition, analysis, and reporting. Popular tools include EnCase and FTK for comprehensive disk imaging and file system analysis, Cellebrite UFED for mobile forensics, Wireshark for network traffic analysis, and Autopsy as an open-source investigation platform. For memory analysis, the Volatility Framework is widely used, while X-Ways Forensics offers efficient disk and file analysis. Each tool ensures accuracy, integrity, and compliance with legal standards during investigations.

1.5 Legal and Ethical Considerations

Investigators must follow jurisdictional laws and forensic standards to maintain admissibility in court. This includes obtaining legal authorization, documenting chain of custody, and ensuring minimal intrusion into irrelevant private data. Ethical practice demands impartiality, professionalism, and transparency while safeguarding individual rights and maintaining public trust.

1.6 Challenges in Digital Forensics

Major challenges include encrypted data, anti-forensic techniques, vast data volumes, and rapid technological changes. Investigators also face privacy concerns, global jurisdictional differences, and limited resources. Maintaining a reliable chain of custody is critical but often complex in distributed environments. Addressing these issues requires innovation, collaboration, and continuous training.

1.7 Role of Machine Learning and AI

ML and AI enhance digital forensics by automating repetitive tasks, detecting anomalies, and improving predictive analysis. Techniques such as NLP, image recognition, and anomaly detection allow faster classification of evidence and identification of cyberattacks. These tools boost scalability and efficiency, enabling investigators to adapt to evolving threats.

1.8 Analysis for Cybercrime Investigations

Cybercrime investigations start with triage and evidence identification, followed by data collection, preservation, and forensic imaging. Analytical methods include file system analysis, network traffic monitoring, memory forensics, and malware reverse engineering. Investigators reconstruct timelines, connect digital footprints, and prepare reports that are legally defensible.

1.9 Digital and Forensic Analysis

- **Digital Analysis** focuses on reconstructing incidents using system logs, memory dumps, malware analysis, and network forensics.

- **Forensic Analysis** ensures evidence collection, preservation, recovery, and reporting while maintaining admissibility in court. Together, these processes provide a holistic approach to uncovering cybercrimes [10-14].

2. Review of Literature

Al-Jumaili et al. (2024, November) reported that the Dubai Police Forensic Engineering Department played a crucial role in forensic investigations, where efficient case reporting was vital for successful outcomes. They noted that the department faced challenges due to unbalanced workload distribution among examiners and reviewers, which had led to extended lead times and delays in case reporting. To address this issue, the study proposed a machine learning-driven solution using predictive analytics to streamline case reporting, reduce lead times, and enhance operational efficiency. Historical case data were analyzed with predictive algorithms to identify patterns in case categories, incident types, and examiner allocation. Concept Explorer was used to generate variable combinations, and MATLAB was employed to develop a predictive model for estimating working days for different combinations. The findings demonstrated that the machine learning model could reduce average case reporting times to approximately seven days, decreasing working days for selected cases from a range of 1–17 to 6–10 days, thereby achieving more balanced workloads and shorter lead times.

Singh et al. (2024) observed that the widespread use of the internet and mobile phones had transformed the physical world into a largely virtual environment, engaging individuals across all age groups. They noted that while people increasingly relied on social media and online platforms for communication and engagement, the ability to conceal or impersonate one's identity in this virtual space had created opportunities for criminal activity to go undetected. Despite India's progress toward digitalization, the authors highlighted a lack of sufficient regulations and technological resources to address challenges arising from this digital shift. The study focused on cybercrime, defined as illicit activities involving computers as tools, targets, or weapons, with particular attention to cyberstalking, wherein perpetrators use technology to instill fear in victims. The chapter aimed to examine the definition, typology, detection, and prevention of cyberstalking, emphasizing the importance of cybercrime education. It also investigated the rise of internet stalking cases in India, exploring their causes and categorization, while analysing the relevant Indian legal framework and judicial precedents addressing these offenses.

Dadiyala et al. (2024, May) introduced an innovative machine learning (ML)-driven system to combat phishing attacks and malicious emails. The study employed a comprehensive approach encompassing dataset collection, pre-processing, model development, and the creation of a user-friendly browser extension interface. The authors reported achieving 96.4% accuracy in detecting phishing URLs using the XGBoost algorithm and 98.6% accuracy in identifying malicious email content with the Multinomial Naive Bayes model, demonstrating enhanced cybersecurity measures. The dataset reflected diverse phishing scenarios, guided by Exploratory Data Analysis (EDA), with URL features utilized by XGBoost for precise phishing site identification. Multinomial Naive Bayes was particularly effective in discerning malicious email content. The system was implemented as a browser extension, enabling seamless integration into users' digital environments and providing immediate feedback on encountered threats, thereby facilitating swift mitigation and increased user awareness.

Al-Fatlawi (2024, November) observed that despite extensive empirical literature, limited research had addressed the integration of digital forensic techniques with the Internet of Things (IoT). The author highlighted that current digital forensic tools' adaptability had not sufficiently protected IoT devices, leaving them vulnerable to persistent attacks. The study aimed to address the urgent need for effective

forensic analysis of IoT devices, noting that existing methods lacked adequate processing power and memory to record and examine multiple attacks, complicating risk identification and mitigation. To overcome these challenges, the research proposed a machine-to-machine (M2M) system combined with an additional logging server, forensic analytical tools, and machine learning to facilitate evidence collection. Logs were processed on a specialized forensic platform (Security Onion) to identify the number and types of attacks on IoT devices. The study emphasized that this approach filled a knowledge gap on safeguarding IoT frameworks through digital forensics and addressed limitations of existing forensic programs. Findings demonstrated that the decision tree method performed particularly well in automated attack detection, underscoring the need to advance digital forensic techniques to enhance IoT device security against evolving cyber threats.

Ojika et al. (2024) highlighted that the growing sophistication and frequency of cyberattacks had created a need for advanced cybersecurity approaches across industries. They observed that Artificial Intelligence (AI), particularly through machine learning (ML) and data analytics, was emerging as a transformative tool for threat detection and response. The study presented a cross-industry model integrating AI-driven systems to enhance cybersecurity resilience and operational efficiency. By leveraging supervised and unsupervised ML algorithms, the model enabled dynamic anomaly detection, real-time threat identification, and predictive risk assessment. The integration of AI with large-scale data analytics allowed for the correlation of seemingly unrelated events, revealing patterns often missed by traditional systems. The proposed framework was adaptable to diverse sectors—including finance, healthcare, manufacturing, and critical infrastructure—and incorporated a feedback loop to continuously learn from new threats, improving accuracy and reducing false positives. Its modular architecture facilitated integration with existing cybersecurity infrastructures without major overhauls, supporting SIEM, endpoint protection, and network intrusion detection systems. Case studies demonstrated the model's effectiveness in detecting zero-day exploits, phishing attempts, and insider threats in real time, while reducing response times and supporting automated incident response. The study also addressed ethical concerns such as data privacy and algorithmic transparency, proposing governance frameworks for responsible AI use. The authors emphasized that adopting a unified, cross-industry AI-driven approach could strengthen organizational defenses against evolving cyber threats, with future research exploring federated learning and privacy-preserving AI to enable secure collaborative threat intelligence sharing.

Jawad et al. (2024) noted that ransomware remained a prevalent and severe cyber threat, encrypting system data and demanding payment for decryption. The study provided a comprehensive review of ransomware detection methods, with a particular focus on machine learning-driven approaches. It examined dynamic analysis techniques, evaluated detection frameworks, and highlighted tools such as Sentinel One and Sandblast Anti-Ransomware. The authors reviewed studies conducted between 2018 and 2023 to consolidate the latest findings. The review emphasized the effectiveness of predictive methods, citing one pre-encryption detection algorithm that achieved 99.9% accuracy. Overall, the research offered a valuable resource for understanding ransomware threats and provided actionable insights to improve detection and mitigation strategies.

Haveriku et al. (2024) noted that Optical Character Recognition (OCR) technology was widely applied in document digitization, data extraction, and document management systems, and had significantly improved in accuracy with the integration of machine learning (ML) and artificial intelligence (AI) techniques. However, the authors highlighted ongoing challenges when processing low-quality or noisy images, handwritten text, or documents with unusual fonts. Security features in documents such as ID cards and driving licenses—like deliberate errors, Optical Variable Ink (OVI), Rainbow print, Guilloche

patterns, fine lines, and microprinting—further complicated OCR by introducing noise. To address these challenges, the study proposed a machine learning-driven approach to enhance OCR accuracy for ID-1 documents with security features. Albanian driving license images were used as a dataset during the personalization process to train the model, with input features including images, text, and numerical data paired with corresponding labels. After training, the model and implemented algorithm optimized the images for OCR in real-world applications, improving recognition accuracy despite the presence of security features.

Musa et al. (2024) conducted a state-of-the-art review examining Distributed Denial of Service (DDoS) anomaly detection in Software Defined Networks (SDNs) using advanced Machine Learning (ML) and Deep Learning (DL) techniques. The study focused on addressing the inherent security vulnerabilities of SDN environments and developing automated systems for detecting and mitigating network attacks. The authors highlighted that conventional network measurement methods were limited in SDNs, and ML/DL techniques offered more accurate and efficient approaches for DDoS detection and mitigation. The review categorized recent advances into ML- and DL-based methods, employing techniques such as Supervised Learning, Unsupervised Learning, Ensemble Learning, and deep learning solutions to process IP flows, profile network traffic, and identify attacks. The outputs included mitigation policies learned by these systems, enabling automated responses to minimize the impact of attacks. Evaluation metrics—including accuracy, precision, and recall—demonstrated the effectiveness of the proposed approaches in detecting and mitigating various attack types. The study emphasized the systems' contributions to enhancing SDN security while noting inherent limitations and the need for further validation in diverse operational environments.

Rahimi et al. (2024) investigated the use of machine learning (ML) techniques to predict the severity of methanol poisoning, aiming to enhance early identification and prognosis assessment. The study was conducted at Loghman Hakim Hospital in Tehran, Iran, using retrospective data from 897 patients, categorized into three groups: without sequel (n = 573), with sequel (n = 234), and deceased (n = 90). The dataset was split into training and test sets at a 70:30 ratio. Features were selected through a two-step process, yielding 43 features in the first step and 23 in the second, which were input into various ML models implemented via Scikit-learn in Python. The models' performance was evaluated using ten-fold cross-validation with a 95% confidence level. The Gradient Boosting Classifier emerged as the best-performing model, with key predictive features including younger age, higher methanol ingestion, respiratory symptoms, lower GCS scores, type of visual symptom, duration of therapeutic intervention, ICU admission, and elevated CPK levels. The classifier achieved AUC values of 0.947 and 0.943 for the 43- and 23-feature models, respectively, demonstrating superior predictive capability compared to traditional statistical methods. The study highlights the utility of ML-driven prognostic models for guiding early interventions and personalized treatment strategies in methanol poisoning cases.

Singh and Sehgal (2024) reviewed the evolution of biomedical video source identification, emphasizing the shift from fuzzy-based systems to machine learning (ML) techniques. They noted that biomedical videos play a crucial role in healthcare, including medical imaging, diagnostics, surgical procedures, and patient monitoring, making accurate source identification essential for quality control, accountability, and data integrity. The survey examined the foundational principles of fuzzy-based systems, highlighting the use of linguistic variables and expert knowledge to model video sources, and discussed the strengths and limitations of these traditional approaches. By analyzing existing methodologies and databases, the study provided a comprehensive overview of the field's progress, the emerging dominance of ML-driven methods, and the ongoing challenges in ensuring reliable and robust biomedical video source identification.

Xiao and Mayer (2023) examined the gap between research and practical applications of machine learning (ML) in trust and safety, using misinformation detection as a case study. They surveyed 248 well-cited papers across security, natural language processing, and computational social science, analyzing data and code availability, design flaws, reproducibility, and generalizability. The authors found that detection tasks were often misaligned with real-world challenges faced by online services. Datasets and model evaluations were frequently non-representative of practical contexts, and evaluation procedures were often dependent on model training. Through three replication studies, the limitations of current automated misinformation detection methods were demonstrated, highlighting their restricted efficacy in identifying human-generated misinformation. The study concluded with recommendations for improving ML applications in trust and safety, emphasizing the need for realistic datasets, independent evaluation, and research directions that better align with practical requirements.

Ahmad (2023) examined the critical role of data integrity in clinical trials and the transformative impact of machine learning (ML) on maintaining it. The study emphasized that data integrity is essential for research credibility, patient safety, and medical progress. Through a comprehensive literature review and case studies, the paper demonstrated how ML enhances anomaly detection, risk assessment, real-time monitoring, and quality control in clinical trials. Predictive analytics were highlighted as key tools for identifying potential data breaches, ensuring accuracy, and supporting informed decision-making. The study also explored the integration of ML with data governance practices, stressing the importance of robust integrity protocols. Overall, Ahmad underscored the potential of ML to elevate clinical trial outcomes, foster innovation, and advance evidence-based medicine.

Brun (2022) discusses the transformative impact of machine learning (ML) on software engineering, highlighting its ability to extend the boundaries of what computing can achieve. The paper reviews advances enabled by ML, including automated software bug repair and data-driven systems that learn to make decisions autonomously. While these technologies offer significant potential, they also present risks: automatically generated program patches may inadvertently break existing functionality, and self-learning software can produce unintended consequences, such as unsafe, biased, or discriminatory behaviour. To address these challenges, the author suggests leveraging machine learning itself as a tool for verifying software properties, thereby improving system reliability and quality.

Siddaway et al. (2020) examined the use of machine learning (ML) in analysing "Big Data" for predicting risk behaviours and psychological problems, noting that few critical evaluations of ML exist. They highlighted fundamental cautions relevant to predicting clinical and forensic risk behaviours, such as risk to self, others, or from others, as well as mental health issues. The authors emphasized that while ML's flexibility—allowing models to be specified without researcher input—is a key strength, it is also a critical weakness. They argued that ML should function as machine-assisted learning, with transparent presentation of algorithms and results, to avoid overreliance on opaque models. The study cautioned against assuming ML's superiority, noting that its complexity can limit clinical utility. Instead, researchers and clinicians should focus on understanding individual needs, formulating risks, and providing personalized management and treatment, rather than placing undue trust in predictions that may be inaccurate.

Yousefi et al. (2020) reviewed the integration of machine learning (ML) mechanisms with Internet of Things (IoT) systems, highlighting their increasing efficiency and applicability in both academic and real-world contexts. The study emphasized that ML can transform real-world challenges into artificial intelligence solutions across diverse IoT applications, including data analysis, wireless communication,

healthcare, industrial systems, and security. Despite these benefits, the authors noted several challenges, such as the lack of standard datasets, trust issues, and resource limitations in IoT environments. The chapter discussed recent ML-based approaches for IoT systems and outlined common issues and obstacles, providing insights and potential research directions for developers and researchers interested in advancing ML applications in IoT.

3. Research Methodology

The methodology adopts a systematic approach to machine learning–based forensic evidence analysis, covering dataset selection, preprocessing, model implementation, and evaluation. Both real-world datasets (CICIDS, Kaggle, UCI, CERT) and synthetic sources (blockchain flows, darknet data) were used to represent diverse evidence types like network logs, financial records, malware, and IoT traffic. Models spanned supervised, unsupervised, deep learning, and graph-based techniques, evaluated on forensic metrics such as Accuracy, Precision, Recall, F1, and ROC-AUC.

3.1 Research Design

The study followed an experimental quantitative design, comparing ML algorithms across categories. Real-world and synthetic datasets simulated cybercrime cases to test robustness. Advanced models (CNNs, LSTMs, GNNs) were integrated into forensic workflows, aiding detection and evidence interpretation.

3.2 Data Sources

Eight datasets were used: CICIDS 2017, KDD’99 (intrusion logs), Kaggle (fraud records), UCI (phishing emails), simulated blockchain flows, Microsoft/EMBER (malware), CERT insider threats, darknet marketplace data, and N-BaIoT (IoT traffic).

3.3 Data Preprocessing

Steps included cleaning (removing duplicates/noise), normalization (scaling features), feature engineering (flow statistics, entropy, graph metrics), and balancing (SMOTE for imbalanced data).

3.4 Model Implementation

- **Supervised:** Logistic Regression, Decision Tree, Random Forest, SVM, ANN for labeled data.
- **Unsupervised:** k-Means, DBSCAN for anomaly detection.
- **Deep Learning:** CNNs for malware binaries, LSTMs for sequential logs.
- **Graph Models:** GNNs for blockchain flows.

3.5 Performance Metrics

Evaluation used Accuracy, Precision, Recall, F1-Score, and ROC-AUC to ensure reliability on imbalanced forensic datasets.

3.6 Tools and Environment

Python (TensorFlow, Keras, Scikit-learn, PyTorch), MATLAB, Pandas, Numpy, Matplotlib, and NetworkX supported analysis. Experiments ran on GPU-enabled clusters via Jupyter, Colab, and Linux servers for scalability and reproducibility.

4. Method of Data Analysis

This chapter outlines the data analysis conducted to assess machine learning's role in cyber forensic investigations. The study considered both simulated and real-world cybercrime data, including network traffic logs, financial transactions, phishing emails, and blockchain records. Using supervised, unsupervised, deep learning, and graph-based techniques, the data was preprocessed, modeled, and analyzed to uncover patterns of cybercrime and validate the effectiveness of ML-driven evidence analysis.

Dataset Description: The study employed diverse datasets representing multiple cybercrime scenarios. Network intrusion logs (CICIDS 2017, KDD'99) provided detailed normal and malicious traffic records, including features like packet size and protocol type, supporting anomaly detection for DoS, probing, and infiltration attacks. The financial fraud dataset included anonymized transaction records, highlighting the challenge of imbalanced fraudulent versus legitimate cases for supervised learning. The phishing email dataset contained labeled emails with features such as subject-line entropy, suspicious URLs, and sender authenticity, enabling automatic detection of deceptive communications. Blockchain records captured Bitcoin transactions, represented as graph structures to track suspicious flows and multi-hop laundering, suitable for graph-based ML analysis.[15]

Data Preprocessing: Preprocessing ensured dataset quality and reliability. Cleaning removed duplicates, incomplete entries, and noisy values. Normalization standardized numerical features like transaction amounts and packet sizes, improving algorithm convergence and stability. Feature engineering derived meaningful attributes, including packet statistics, transaction intervals, email header keywords, and blockchain graph patterns, enhancing model accuracy. Class balancing, particularly using SMOTE for financial fraud, addressed imbalances by generating synthetic minority samples, enabling fairer classification between legitimate and malicious instances.

Model Implementation: A variety of ML algorithms were applied across supervised, unsupervised, deep learning, and graph-based paradigms. Supervised models like Logistic Regression, Decision Tree, Random Forest, SVM, and ANN were used for labeled datasets such as intrusion logs, phishing emails, and financial transactions, performing well even under class imbalance due to hierarchical and adaptive learning features. Unsupervised models (k-Means, DBSCAN) were applied to unlabeled network and blockchain data to detect anomalies and zero-day attacks. Deep learning models, such as CNNs for malware binaries and LSTMs for sequential intrusion logs, captured complex patterns. Graph Neural Networks combined with Markov features enabled analysis of blockchain transactions, uncovering hidden laundering schemes and multi-hop transfers.

Performance Evaluation Metrics: Model performance was assessed using forensic-relevant metrics. Accuracy measured the proportion of correctly classified instances, while precision evaluated the reliability of flagged malicious cases. Recall quantified the model's ability to detect all attacks, minimizing false negatives. F1-score balanced precision and recall for reliable classification, and ROC-AUC measured discriminative capability, particularly in imbalanced datasets, indicating the model's effectiveness in distinguishing between legitimate and malicious cyber activities.

4.1 Results and Analysis

Network Intrusion Detection

Random Forest achieved the highest accuracy (98.1%, F1 = 0.97), handling imbalanced classes effectively. SVM generalized well to zero-day attacks but was computationally expensive. LSTMs excelled in sequential anomaly detection, especially for DoS and Probe attacks.

Financial Fraud Detection

Random Forest and ANN both reached a precision of 0.96, reducing false alarms in skewed datasets. Logistic Regression struggled with imbalance but improved with SMOTE, though still weaker than ensemble and neural methods.

Phishing Email Analysis

SVM achieved 96% accuracy, effectively distinguishing phishing from legitimate emails. Features like subject-line entropy, suspicious URLs, and header anomalies proved most informative. Feature engineering strengthened classification reliability.

Blockchain Transaction Analysis

GNNs combined with Markov chains attained a ROC-AUC of 0.93, detecting hidden laundering paths in cryptocurrency flows. Clustering further identified suspicious, high-frequency nodes linked to risky wallets, enhancing blockchain forensic monitoring.

Table 1: Model Performance Across Datasets

Task	Best Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Intrusion Detection	Random Forest	98.1%	0.96	0.98	0.97	0.98
Fraud Transactions	ANN	97.6%	0.96	0.95	0.96	0.97
Phishing Emails	SVM	97.5%	0.95	0.96	0.95	0.96
Blockchain Forensics	GNN + Markov	94.2%	0.92	0.94	0.93	0.93

The table summarizes how different machine learning models performed across four forensic tasks—intrusion detection, fraud transaction detection, phishing email classification, and blockchain forensics. Each task is evaluated using five metrics: Accuracy, Precision, Recall, F1-Score, and ROC-AUC.

- **Intrusion Detection (Random Forest – 98.1% accuracy):**
  Random Forest outperformed other models by capturing complex attack patterns in network data. Its high recall (0.98) indicates that it successfully detected nearly all attacks, while the strong F1-score (0.97) shows balance between precision and recall. The ROC-AUC of 0.98 reflects excellent discrimination between malicious and normal traffic.
- **Fraud Transactions (ANN – 97.6% accuracy):**
  Artificial Neural Networks were best suited for financial fraud detection, handling non-linear relationships and imbalanced data effectively. With precision (0.96) and recall (0.95), ANN minimized false positives and false negatives, making it reliable for real-world financial systems where both are costly.
- **Phishing Emails (SVM – 97.5% accuracy):**
  Support Vector Machines excelled at classifying phishing vs. legitimate emails, leveraging features such as URL patterns and text entropy. With precision (0.95) and recall (0.96), SVM achieved a balanced tradeoff, and an F1 of 0.95 indicates robust performance in catching phishing attempts without excessive misclassification.
- **Blockchain Forensics (GNN + Markov – 94.2% accuracy):**
  Graph Neural Networks combined with Markov models proved effective in detecting suspicious cryptocurrency flows and laundering paths. Although accuracy (94.2%) was slightly lower compared to other tasks, the model achieved strong precision (0.92) and recall (0.94), showing its strength in analyzing complex, interconnected transaction networks.

Random Forest was the strongest in intrusion detection, ANN worked best for fraud detection, SVM was most effective for phishing classification, and GNN+Markov excelled in blockchain forensics. Each model's success was tied to the **nature of the dataset and task**—tree-based methods for structured logs, neural networks for financial patterns, margin-based classifiers for text/email, and graph-based models for networked blockchain data.

Table 2: Model Performance Across Datasets

Task / Dataset	Best Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Network Intrusion Logs	Random Forest	98.1%	0.96	0.98	0.97	0.98
Fraud Transactions	ANN	97.6%	0.96	0.95	0.96	0.97
Phishing Emails	SVM	97.5%	0.95	0.96	0.95	0.96
Blockchain Records	GNN + Markov Chain	94.2%	0.92	0.94	0.93	0.93
Malware Binary Dataset	CNN	96.8%	0.95	0.96	0.95	0.96
Insider Threat Dataset	LSTM	95.4%	0.93	0.95	0.94	0.95

The results presented in Table 2 highlight how different models performed across various forensic datasets, each tailored to the unique structure and complexity of the data. In the case of network intrusion logs, Random Forest achieved the highest overall accuracy at 98.1%, supported by strong precision (0.96) and recall (0.98). This indicates that Random Forest not only correctly classified most intrusion events but also minimized the risk of overlooking actual attacks. Its ROC-AUC score of 0.98 demonstrates that it maintained excellent discriminatory power in distinguishing between malicious and benign network traffic.

For fraudulent financial transactions, Artificial Neural Networks (ANN) delivered superior performance with 97.6% accuracy and a well-balanced F1-score of 0.96. Financial fraud detection often involves complex, non-linear relationships between features, and ANN's adaptability allowed it to capture such patterns effectively. The model's high precision (0.96) reduced false alarms, while its recall (0.95) ensured that fraudulent cases were rarely missed, making it highly suitable for financial security applications where both errors carry significant cost.

In the context of phishing emails, Support Vector Machines (SVM) attained an accuracy of 97.5%. Its precision (0.95) and recall (0.96) reveal that the model effectively detected phishing attempts while maintaining low misclassification rates. With an F1-score of 0.95 and ROC-AUC of 0.96, SVM demonstrated robustness in distinguishing deceptive emails from legitimate communication. This result underscores the effectiveness of margin-based classifiers when dealing with high-dimensional text features and engineered attributes like URL structures and header anomalies.

The application of Graph Neural Networks (GNN) combined with Markov chains to blockchain records showed an accuracy of 94.2%. While slightly lower than other domains, this result reflects the complexity of analyzing interconnected cryptocurrency transactions. The model's precision (0.92) and recall (0.94) indicate that it was able to detect suspicious transaction patterns and hidden laundering paths effectively. Its F1-score of 0.93 confirms its balanced performance in this highly dynamic and evolving forensic domain.

For the malware binary dataset, Convolutional Neural Networks (CNN) achieved 96.8% accuracy with balanced precision and recall values of 0.95 and 0.96, respectively. CNNs excel in handling image-like binary data representations, enabling them to detect subtle patterns in malware signatures that traditional methods might overlook. The model's F1-score (0.95) and ROC-AUC (0.96) suggest it is highly reliable for automated malware classification tasks.

Finally, for the insider threat dataset, Long Short-Term Memory (LSTM) networks recorded 95.4% accuracy, precision of 0.93, and recall of 0.95. Insider threats often involve sequential patterns of user activity, and LSTMs proved effective in capturing these temporal dependencies. The F1-score (0.94) and ROC-AUC (0.95) indicate that the model maintained a strong balance between detecting malicious insider behavior and avoiding false positives, a crucial factor in organizational cybersecurity.

Table 3: IoT Botnet Traffic Dataset (N-BaIoT)

Dataset Type	Instances	Features	Class Labels	Source
Network Intrusion Logs	311,029	78	Normal / Attack	CICIDS 2017, KDD'99
Fraud Transactions	2,84,807	30	Fraud / Legit	Kaggle
Phishing Emails	11,055	57	Phish / Normal	UCI Repository
Blockchain Records	85,642	42	Suspicious / Normal	Simulated / Public Chains
Malware Binary Dataset	25,000	1,024	Benign / Malware Families	Microsoft, EMBER
Insider Threat Dataset	1,500,000	118	Normal / Malicious	CERT, CMU Repository
Darknet Marketplace Dataset	42,500	250	Fraudulent / Legitimate	Academic Scrapes / Archives
IoT Botnet Traffic (N-BaIoT)	85,000	115	Normal / Botnet Attack	N-BaIoT Repository

Table 3 provides an overview of the datasets used in the study, each representing a distinct forensic challenge. The network intrusion logs dataset, sourced from CICIDS 2017 and KDD'99, contains over 311,000 instances with 78 features, enabling the classification of traffic as either normal or attack. Similarly, the fraud transaction dataset from Kaggle includes around 285,000 records with 30 features, reflecting real-world financial data labeled as fraud or legitimate.

The phishing email dataset, with 11,055 instances and 57 attributes from the UCI repository, captures patterns of phishing versus normal communication. In contrast, blockchain records, combining simulated and public chain data, consist of over 85,000 entries with 42 features, distinguishing suspicious from normal transactions.

For malware detection, the malware binary dataset provides 25,000 samples with high-dimensional feature vectors (1,024 features) sourced from Microsoft and EMBER, labeled as benign or malware families. The insider threat dataset from CERT at CMU is the largest, with 1.5 million instances and 118 features, capturing user behavior marked as normal or malicious.

The darknet marketplace dataset, created through academic scrapes and archives, contains 42,500 samples with 250 features, focusing on fraudulent versus legitimate transactions. Finally, the IoT Botnet Traffic dataset (N-BaIoT) comprises 85,000 records with 115 features, distinguishing normal traffic from botnet attacks, making it essential for securing IoT environments.

Table 4: Model Performance Across Datasets

Task / Dataset	Best Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Network Intrusion Logs	Random Forest	98.2%	0.97	0.98	0.97	0.99
Fraud Transactions	ANN	97.4%	0.96	0.94	0.95	0.97
Phishing Emails	SVM	96.9%	0.94	0.96	0.95	0.96
Blockchain Records	GNN + Markov Chain	94.6%	0.93	0.92	0.92	0.94
Malware Binary Dataset	CNN	96.2%	0.95	0.94	0.94	0.96
Insider Threat Dataset	LSTM	95.1%	0.92	0.95	0.93	0.95
Darknet Marketplace Dataset	NLP + Random Forest	93.4%	0.91	0.92	0.91	0.93
IoT Botnet Traffic (N-BaIoT)	Autoencoder + LSTM	97.7%	0.95	0.97	0.96	0.98

Table 4. summarizes the performance of different machine learning models across diverse forensic datasets, highlighting how each task benefits from a model best suited to its structure and complexity. For network intrusion logs, the Random Forest model provided the strongest results, achieving 98.2% accuracy with an F1-score of 0.97 and a ROC-AUC of 0.99. Its ensemble approach allowed it to handle both balanced and imbalanced classes effectively, making it ideal for intrusion detection tasks.

In fraud transaction analysis, Artificial Neural Networks (ANN) performed best with 97.4% accuracy and a precision of 0.96, though recall was slightly lower at 0.94. This suggests that ANN could capture complex, non-linear fraud patterns effectively, while still leaving room for improvement in identifying all fraudulent cases. Similarly, for phishing emails, Support Vector Machines (SVM) achieved 96.9% accuracy with high recall (0.96), showing its strength in distinguishing deceptive emails with subtle similarities to legitimate messages.

For blockchain records, the combined GNN + Markov Chain approach achieved 94.6% accuracy, with balanced precision, recall, and F1-scores around 0.92. This reflects its ability to uncover hidden, multi-hop laundering schemes in transaction graphs that simpler models might overlook. The malware binary dataset showed strong performance with Convolutional Neural Networks (CNN), recording 96.2% accuracy. CNNs excelled here due to their ability to detect intricate patterns in binary feature spaces, making them effective in malware classification.

In insider threat detection, Long Short-Term Memory (LSTM) networks reached 95.1% accuracy, leveraging sequential dependencies in user activity data. Meanwhile, the darknet marketplace dataset, characterized by high-dimensional textual and transactional features, was best handled by an NLP + Random Forest hybrid model, achieving 93.4% accuracy. Finally, for IoT botnet traffic (N-BaIoT), the Autoencoder combined with LSTM stood out, achieving 97.7% accuracy with strong recall (0.97) and ROC-AUC (0.98). This highlights its effectiveness in detecting anomalies in complex IoT traffic by capturing both feature reconstruction errors and temporal patterns.

Together, the results show that model performance is highly task-dependent: tree-based methods excel in intrusion detection, neural models dominate fraud and malware analysis, sequence-based learning supports insider and IoT traffic detection, while hybrid approaches are most effective for blockchain and darknet datasets.

5. Findings and Conclusion

This chapter summarizes the study on machine learning (ML)–driven forensic evidence analysis for cybercrime investigations. The research evaluated supervised, unsupervised, deep learning, and graph-based models across eight diverse datasets, including network intrusion logs, fraud transactions, phishing emails, blockchain records, malware binaries, insider threats, darknet marketplaces, and IoT botnet traffic. Key findings revealed that popular AIOps models often underperformed compared to specialized ML approaches. Random Forest excelled in intrusion detection, ANN in financial fraud, SVM in phishing detection, and hybrid models like Autoencoder-LSTM were highly effective for IoT botnet analysis. Graph Neural Networks with Markov features successfully uncovered hidden laundering behaviors in blockchain transactions. These results highlight the operational capabilities of ML models in enhancing forensic readiness, anomaly detection, and predictive analysis, while also emphasizing challenges such as interpretability, data quality, and adversarial robustness.

Findings

The study leveraged eight datasets spanning structured, unstructured, temporal, and graph-based data, ensuring broad forensic coverage. Supervised models performed strongly: Random Forest achieved 98.2% accuracy for intrusion detection, ANN reached 97.4% for fraud detection, and SVM delivered balanced results for phishing classification. Deep learning models also contributed significantly: CNNs reached 96.2% accuracy for malware binaries, LSTMs achieved 95.1% in insider threat detection, and Autoencoder-LSTM combinations yielded 97.7% for IoT botnet traffic. Graph-based models using GNNs with Markov chains achieved 94.6% accuracy in blockchain analysis, revealing complex multi-hop laundering schemes. Across metrics, Random Forest and Autoencoder-LSTM consistently excelled in accuracy, precision, recall, F1-score, and ROC-AUC. Challenges included handling imbalanced datasets and noisy or unstructured data, as seen in darknet marketplace analysis.

Conclusion

The study confirms that ML enhances the accuracy, efficiency, and scope of cyber forensic investigations. Classic techniques struggle with the scale and complexity of modern cybercrime, whereas ML models adapt well to diverse data types. Random Forest and ANN provided reliable baselines, while deep learning models (CNNs, LSTMs, Autoencoders) effectively addressed complex patterns in malware, insider threats, and IoT traffic. Graph-based models, especially GNNs with Markov features, revealed hidden relationships in blockchain networks. Overall, no single model can cover all forensic tasks; an ensemble of ML approaches aligned with dataset characteristics provides the most robust framework, improving both operational readiness and judicial reliability.

Implications for Forensic Practice

Integrating ML into forensic practice offers operational and legal benefits. Models such as Random Forest, ANN, CNN, and Autoencoder-LSTM can be embedded in forensic toolchains for near real-time monitoring, early detection, and anomaly analysis. Standardized evaluation metrics enhance transparency, facilitating judicial acceptance. ML models remain adaptable, capable of handling traditional cyber threats while responding to novel attack patterns. This combination of accuracy, scalability, and legal defensibility strengthens digital forensics overall.

Future Directions

Future research should focus on explainability (XAI), integrating hybrid ML–knowledge graph frameworks, enhancing adversarial robustness, and enabling real-time scalable operations. Explainable AI methods can uncover the key features driving predictions, improving trust. Hybrid frameworks can map relationships across heterogeneous datasets, supporting multi-hop reasoning and faster triage. Adversarial robustness requires threat modeling, ensemble diversity, and runtime detection to counter evasion and poisoning attacks. Real-time architectures with streaming analytics, online learning, and GPU/FPGA acceleration will allow ML-driven forensics to scale efficiently. Additionally, privacy-preserving workflows and MLOps practices will facilitate cross-agency collaboration, creating an interpretability, speed, and robustness standard for modern forensic investigations.

References

1. Al-Jumaili, Q. A., Al Shamsi, F. S., Alkhajeh, S. T., Araci, Z. C., Alqasim, M. A., & Alzarooni, S. M. (2024, November). Machine Learning-Driven Predictive Analytics in Workload Distribution for Forensic Engineering Case Reporting Process: Dubai Police Case Study. In *2024 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD)* (pp. 1-6). IEEE.
2. Singh, C., Khajuria, H., & Nayak, B. P. (2024). Machine Learning Driven Strategies for Safeguarding India's Digital Environment Against Cyberstalking. In *Artificial Intelligence, Medical Engineering and Education* (pp. 272-279). IOS Press.
3. Dadiyala, C., Ghate, M., Shekdar, A., Rajkondawar, P., Chaure, S., & Zanwar, Y. (2024, May). Machine-Learning-Driven Detection of Malicious Emails and Fake Websites. In *Doctoral Symposium on Computational Intelligence* (pp. 343-364). Singapore: Springer Nature Singapore.
4. Al-Fatlawi, H. H. M. (2024, November). Bridging the gap: A machine-to-machine forensic analysis framework for enhancing IoT device security. In *AIP Conference Proceedings* (Vol. 3229, No. 1, p. 040010). AIP Publishing LLC.
5. Ojika, F. U., Owobu, W. O., Abieba, O. A., Esan, O. J., Ubamadu, B. C., & Daraojimba, A. I. (2024). The Role of AI in Cybersecurity: A Cross-Industry Model for Integrating Machine Learning and Data Analysis for Improved Threat Detection. *Comput Secur.[Year]*.
6. Jawad, S., & Ahmed, H. M. (2024). Machine Learning Approaches to Ransomware Detection: A Comprehensive Review. *International Journal of Safety & Security Engineering*, *14*(6).
7. Haveriku, A., Muraku, B., Karamani, B., & Paci, H.2024 Enhancing OCR Accuracy for ID-1 Documents with Security Features through Machine Learning-driven Image Optimization.
8. Musa, N. S., Mirza, N. M., Rafique, S. H., Abdallah, A. M., & Murugan, T. (2024). machine learning and deep learning techniques for distributed denial of service anomaly detection in software defined networks—current research solutions. *IEEE Access*, *12*, 17982-18011.
9. Rahimi, M., Hosseini, S. M., Mohtarami, S. A., Mostafazadeh, B., Evini, P. E. T., Fathy, M., ... & Shadnia, S. (2024). Prediction of acute methanol poisoning prognosis using machine learning techniques. Toxicology, 504, 153770.
10. Singh, S., & Sehgal, V. K. (2024). Exploring Biomedical Video Source Identification: Transitioning from Fuzzy-Based Systems to Machine Learning Models. *Fuzzy Information and Engineering*, *16*(1), 33-48.
11. Xiao, M., & Mayer, J. (2023). The challenges of machine learning for trust and safety: a case study on misinformation detection. *arXiv preprint arXiv:2308.12215*.
12. Ahmad, T. 2023. Department of Computer Science, University of Panjab Lahore.

13. Brun, Y. (2022, November). The promise and perils of using machine learning when engineering software (keynote paper). In *Proceedings of the 6th International Workshop on Machine Learning Techniques for Software Quality Evaluation* (pp. 1-4).
14. Siddaway, A. P., Quinlivan, L., Kapur, N., O'Connor, R. C., & De Beurs, D. (2020). Cautions, concerns, and future directions for using machine learning in relation to mental health problems and clinical and forensic risks: A brief comment on "Model complexity improves the prediction of nonsuicidal self-injury" (Fox et al., 2019).
15. Yousefi, S., Derakhshan, F., & Karimipour, H. (2020). Applications of big data analytics and machine learning in the internet of things. In *Handbook of big data privacy* (pp. 77-108). Cham: Springer International Publishing.