Machine Learning-Driven User Behavior Analytics for Insider Threat Detection

ESEOGHENE DANIEL ERIGHA[1], EHIMAH OBUSE[2], NOAH AYANBODE[3], EMMANUEL CADET[4], EDIMA DAVID ETIM[5]

[1]*Senior Software Engineer, Eroe Consulting, Dubai, UAE*
[2]*Lead Software Engineer, Choco, Berlin, Germany*
[3]*Independent Researcher, Nigeria*
[4]*Independent Researcher, USA*
[5]*Core IP Engineer, Cobranet Ltd, Lekki, Lagos, Nigeria*

Abstract- Insider threats present a significant and often underestimated risk to organizational security, as they involve malicious or negligent activities originating from individuals with legitimate access to systems and sensitive information. Traditional rule-based and signature-driven detection methods are frequently inadequate against sophisticated insider behaviors that evolve over time and evade predefined thresholds. This paper explores a comprehensive framework for Machine Learning-Driven User Behavior Analytics (UBA) aimed at detecting insider threats through the continuous monitoring, profiling, and anomaly detection of user activities. The proposed approach leverages supervised, unsupervised, and deep learning algorithms to analyze high-dimensional datasets encompassing login patterns, file access histories, communication metadata, and application usage logs. Feature engineering is employed to extract temporal, contextual, and relational indicators of potentially malicious actions, while advanced models such as autoencoders, recurrent neural networks (RNNs), and graph-based anomaly detectors are applied to identify deviations from established behavioral baselines. The system incorporates adaptive learning capabilities to dynamically refine detection thresholds, thereby reducing false positives and enhancing detection accuracy in real time. Experimental evaluations are conducted using benchmark datasets and simulated enterprise environments to validate the robustness of the framework across various insider threat scenarios, including data exfiltration, privilege escalation, and policy violations. Results demonstrate that the proposed model achieves superior detection performance compared to conventional approaches, with improved precision, recall, and F1-scores, particularly in identifying low-and-slow attacks that unfold over extended periods. The study further addresses challenges related to data privacy, scalability, and interpretability by integrating privacy-preserving analytics, distributed processing architectures, and explainable AI techniques. Practical deployment considerations, including system integration, user acceptance, and compliance with regulatory standards, are also discussed. This research contributes to the field of cybersecurity by providing an intelligent, adaptive, and scalable insider threat detection model that aligns with modern enterprise needs, supporting proactive defense strategies against internal security breaches.

Indexed Terms- Machine Learning, User Behavior Analytics, Insider Threat Detection, Anomaly Detection, Cybersecurity, Deep Learning, Behavioral Profiling, Autoencoders, Recurrent Neural Networks, Graph-Based Models, Explainable AI, Data Privacy, Adaptive Learning, Enterprise Security.

I. INTRODUCTION

Insider threats have emerged as one of the most persistent and damaging challenges in cybersecurity, with incidents ranging from data theft and fraud to sabotage and unintentional data leakage. Unlike external attacks, which often rely on breaching perimeter defenses, insider threats originate from individuals who already possess authorized access to organizational systems and data, such as employees, contractors, or trusted partners. These threats can be malicious driven by personal gain, revenge, or ideological motives or inadvertent, stemming from negligence, poor security awareness, or social engineering exploitation. The growing digitization of workflows, the widespread adoption of remote work models, and the increasing complexity of organizational IT ecosystems have amplified the potential impact of insider threats, making them a critical focus for cybersecurity strategies (Mohammed, 2015, Petrov & Znati, 2018). High-profile cases across industries demonstrate that even advanced perimeter defenses cannot fully safeguard against harm originating from within, underscoring the need for more intelligent, adaptive approaches.

Traditional insider threat detection mechanisms, such as rule-based monitoring, signature detection, and static access control policies, struggle to keep pace with the evolving tactics, techniques, and procedures employed by insiders. These conventional systems are often limited by rigid definitions of suspicious behavior, making them less effective in identifying subtle, context-dependent anomalies (Dogho, 2011, Oni, et al., 2018). Furthermore, they are prone to generating large volumes of false positives, overwhelming security teams and reducing operational efficiency. As insider behaviors become increasingly dynamic leveraging legitimate credentials, blending in with normal traffic patterns, and exploiting contextual trust static detection approaches fail to deliver the agility and depth needed for timely intervention.

Machine learning (ML) and user behavior analytics (UBA) present a promising path forward by enabling dynamic, data-driven detection of deviations from established behavioral baselines. By continuously learning from diverse data sources such as login patterns, file access logs, email communication metadata, and system usage trends, ML-driven UBA systems can identify subtle anomalies that traditional methods may overlook. The goal is to create models that not only detect malicious activity with high accuracy but also minimize false positives, thereby improving the efficiency and effectiveness of insider threat programs (AdeniyiAjonbadi, et al., 2015).

This research aims to develop a robust ML-driven framework for detecting insider threats that integrates advanced anomaly detection techniques, adaptive learning mechanisms, and multi-source behavioral analysis. The objectives include enhancing detection accuracy, reducing false alerts, and providing actionable insights to security analysts. The scope encompasses both malicious and unintentional insider threats in corporate IT environments, with contributions that include a modular framework design, incorporation of explainable AI to support decision-making, and validation through real-world-inspired datasets. By aligning technical innovation with operational needs, this work seeks to advance the state of insider threat detection and provide a practical, scalable solution for modern enterprises (Gudala, et al., 2019, Konn, 2018, Zhong & Gu, 2019).

2.1. Literature Review

Insider threat detection has become a focal area in cybersecurity research, driven by the increasing frequency and severity of incidents involving individuals with authorized access to sensitive systems. The literature identifies three primary typologies of insider threats: malicious insiders, negligent insiders, and compromised accounts. Malicious insiders are individuals who intentionally exploit their access for personal gain, revenge, or ideological motives, often causing significant harm by exfiltrating confidential data, sabotaging systems, or aiding external attackers. Negligent insiders, on the other hand, may inadvertently compromise security through careless actions such as weak password practices, falling victim to phishing, or mishandling sensitive files (Oni, et al., 2018). Compromised accounts represent a hybrid category, where legitimate credentials are stolen or coerced into use by external adversaries, allowing attackers to operate under the guise of trusted users. Each of these typologies presents unique detection challenges, as malicious insiders deliberately conceal their activities, negligent insiders blend harmful actions with routine operations, and compromised accounts mimic legitimate user behavior. Figure 1 shows the general platform of the insider threat detection system presented by Saaudi, Tong & Farkas, 2019.

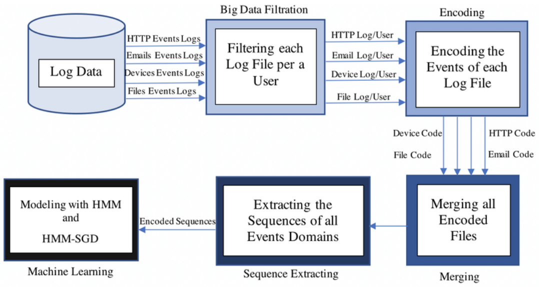

Figure 1: The general platform of the insider threat detection system (Saaudi, Tong & Farkas, 2019).

Conventional approaches to insider threat detection have historically relied on rule-based systems and signature matching. Rule-based detection depends on predefined thresholds and conditions such as unusual login times, large file transfers, or access to restricted directories to trigger alerts. Signature-based systems identify known attack patterns or sequences of actions by comparing them against a database of previously observed threats. While these methods are straightforward to implement and effective against well-understood threats, their limitations are substantial (Adenuga, Ayobami & Okolo, 2019). They are inherently reactive, unable to identify novel or subtle deviations from normal behavior, and prone to generating high false positive rates when legitimate activity happens to meet predefined rules. As insider threats evolve to incorporate more sophisticated evasion techniques, such as conducting malicious activity within the bounds of normal access rights, static rules and signatures fail to keep pace.

User Behavior Analytics (UBA) has emerged as a more dynamic approach to insider threat detection, shifting the focus from predefined signatures to the continuous monitoring and analysis of user activities over time. The key principle underlying UBA is the establishment of a behavioral baseline for each user or peer group, encompassing login patterns, resource access frequencies, data movement habits, and communication trends. Deviations from this baseline are flagged for further investigation, with the aim of identifying potential security incidents early. Existing UBA methods range from simple statistical models, such as z-score anomaly detection on activity metrics, to more advanced clustering-based approaches that group similar behavioral patterns. UBA systems also leverage contextual information, including role-based access expectations, seasonal activity trends, and cross-user comparisons, to improve detection accuracy. However, traditional UBA implementations still struggle with handling the volume and diversity of data in modern enterprises, as well as adapting to evolving behaviors without generating excessive noise. Figure 2 shows User Entity Behavior Analytics Model presented by Salitin & Zolait, 2018.

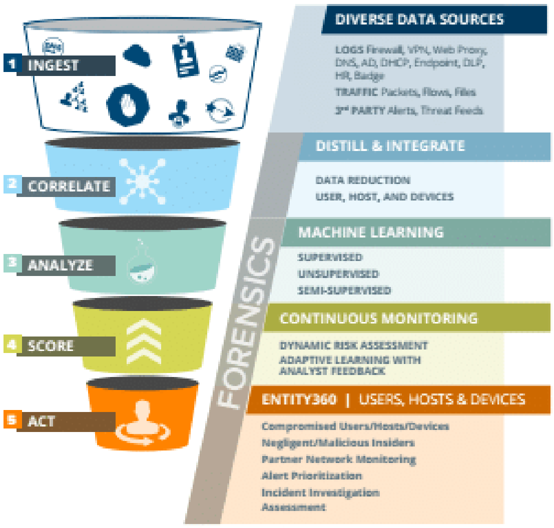
1 INGEST
2 CORRELATE
3 ANALYZE
4 SCORE
5 ACT
FORENSICS
DIVERSE DATA SOURCES
LOGS Firewall, VPN, Web Proxy, DNS, AD, DHCP, Endpoint, DLP, HR, Badge
TRAFFIC Packets, Flows, Files
3rd PARTY Alerts, Threat Feeds
DISTILL & INTEGRATE
DATA REDUCTION
USER, HOST, AND DEVICES
MACHINE LEARNING
SUPERVISED
UNSUPERVISED
SEMI-SUPERVISED
CONTINUOUS MONITORING
DYNAMIC RISK ASSESSMENT
ADAPTIVE LEARNING WITH ANALYST FEEDBACK
ENTITY360 | USERS, HOSTS & DEVICES
Compromised Users/Hosts/Devices
Negligent/Malicious Insiders
Partner Network Monitoring
Alert Prioritization
Incident Investigation
Assessment

Figure 2: User Entity Behavior Analytics Model (Salitin & Zolait, 2018).

Machine learning has been increasingly integrated into insider threat detection to address these limitations, offering the ability to learn complex patterns from large, multi-modal datasets. In supervised learning, algorithms such as decision trees, random forests, support vector machines (SVM), and gradient boosting models are trained on labeled datasets containing examples of both normal and malicious behavior (Olasehinde, 2018). These models excel when high-quality labeled data is available, enabling them to identify subtle features that distinguish threats from benign actions. However, labeled insider threat datasets are rare due to privacy constraints, class imbalance, and the difficulty of obtaining ground truth in real incidents.

In contrast, unsupervised learning is frequently applied for anomaly detection when labeled data is unavailable. Techniques such as k-means clustering, Gaussian mixture models, isolation forests, and autoencoders are used to identify deviations from the learned normal behavioral patterns. Autoencoders, in particular, have gained traction for modeling high-dimensional data, reconstructing normal behavior, and measuring reconstruction error as an anomaly score. These methods are advantageous in detecting zero-day insider threat behaviors that do not match known patterns. Nevertheless, unsupervised methods are sensitive to variations in normal behavior that are not security-related, potentially increasing false positive rates (Mohit, 2018, Sareddy & Hemnath, 2019).

Deep learning has further expanded the capabilities of machine learning-driven UBA, particularly in modeling sequential and contextual aspects of user activity. Recurrent neural networks (RNNs), including long short-term memory (LSTM) and gated recurrent unit (GRU) variants, are effective in capturing temporal dependencies in user activity sequences, such as ordered system commands, file access events, or network requests. By understanding the context of actions within a sequence, deep learning models can distinguish between benign anomalies such as an unusual login time due to a legitimate late-night project and malicious activity. Transformer-based architectures, which rely on self-attention mechanisms, have also been applied to capture both local and global dependencies in behavioral sequences without the limitations of sequential processing inherent to RNNs (Hao, et al., 2019, Xu, et al., 2019). This allows for scalable modeling of long sequences of activity across multiple data sources. Furthermore, graph neural networks (GNNs) have been explored to model relationships between entities such as users, devices, and accessed resources, enabling more comprehensive contextual analysis of insider activities. Figure 4 shows insider-threat detection framework presented by Kim, et al., 2019.

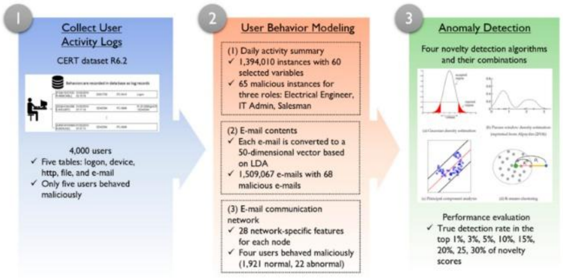
1
Collect User Activity Logs
CERT dataset R6.2
Behaviors are recorded in database as log records
4,000 users
✓ Five tables: logon, device, http, file, and e-mail
✓ Only five users behaved maliciously
2
User Behavior Modeling
(1) Daily activity summary
✓ 1,394,010 instances with 60 selected variables
✓ 65 malicious instances for three roles: Electrical Engineer, IT Admin, Salesman
(2) E-mail contents
✓ Each e-mail is converted to a 50-dimensional vector based on LDA
✓ 1,509,067 e-mails with 68 malicious e-mails
(3) E-mail communication network
✓ 28 network-specific features for each node
✓ Four users behaved maliciously (1,921 normal, 22 abnormal)
3
Anomaly Detection
Four novelty detection algorithms and their combinations
Performance evaluation
✓ True detection rate in the top 1%, 3%, 5%, 10%, 15%, 20%, 25, 30% of novelty scores

Figure 4: Insider-threat detection framework (Kim, et al., 2019).

Despite these advancements, several research gaps remain in the field of machine learning-driven UBA for insider threat detection. A major challenge is the scarcity of realistic, labeled datasets that encompass the diversity and complexity of real-world insider threat behaviors. Most existing datasets are either synthetically generated or anonymized in ways that limit their fidelity. Another gap lies in balancing detection sensitivity with operational practicality; models that are overly sensitive generate alert fatigue, while conservative models risk missing subtle threats. The explainability of complex ML models is also a critical concern, as black-box algorithms can undermine trust among security analysts and hinder the use of AI-generated evidence in legal proceedings (Weng, et al., 2019, Zhou, et al., 2019). Additionally, adversarial robustness is emerging as an important consideration, as attackers may deliberately manipulate behavioral patterns to evade detection, necessitating the development of resilient algorithms.

There is also a need for better integration between UBA systems and broader security operations workflows. Many current implementations operate in isolation, generating alerts without automated correlation to other security telemetry such as endpoint detection and response (EDR) or security information and event management (SIEM) systems. Finally, while much of the research focuses on detecting malicious insiders, there is relatively less emphasis on identifying and mitigating negligent insider behavior, which constitutes a significant portion of real incidents. Addressing these gaps requires continued exploration of hybrid models that combine supervised, unsupervised, and deep learning approaches; the development of explainable AI techniques tailored for behavioral analytics; and the creation of standardized benchmarks to evaluate the performance of insider threat detection systems in varied operational contexts (Achar, 2018, Shah, 2017).

2.2. Methodology

The methodology for Machine Learning-Driven User Behavior Analytics for Insider Threat Detection is designed to leverage advanced data analytics and intelligent modeling techniques to identify anomalous activities indicative of insider threats. The process begins with comprehensive data collection, where heterogeneous data sources such as user activity logs, system access records, file modification histories, and network traffic are aggregated. This raw data is subjected to preprocessing steps including data cleaning, normalization, and missing value handling to ensure accuracy and consistency. Feature extraction and transformation are conducted to derive meaningful variables that capture user behavior patterns, such as frequency of access, resource usage, and contextual attributes tied to organizational roles.

Following preprocessing, feature engineering is performed to construct enriched behavioral profiles for each user. This involves applying domain knowledge to create composite indicators and metrics that can distinguish normal from abnormal behavior. Both statistical and domain-specific features are integrated to improve model interpretability and accuracy. The next stage involves selecting the most suitable machine learning approach based on the problem context and available labels. Supervised learning is employed when labeled instances of insider threats are available, while unsupervised or hybrid models such as clustering and anomaly detection algorithms are utilized for environments lacking explicit labels.

The chosen model undergoes a rigorous training and validation process using a train/test split or cross-validation to ensure generalizability. Techniques like ensemble learning, deep neural networks, and probabilistic models are evaluated to determine optimal performance. Behavior modeling is then applied, where the trained model learns to distinguish normal user activity patterns from deviations that may indicate potential malicious intent. In real-time or near-real-time operation, the system continuously monitors user activities and applies the trained model to flag anomalies.

Upon detection of suspicious activity, the system triggers an alert and response mechanism. This includes automated notifications to security teams, prioritization of alerts based on risk scores, and integration with incident response workflows. Feedback from investigations is looped back into the system to refine detection thresholds and improve accuracy. To address evolving insider threat tactics, the methodology incorporates continuous learning, enabling periodic model retraining with new data to adapt to shifting behavioral baselines. Privacy preservation measures, including federated learning and secure multiparty computation, are implemented to safeguard sensitive employee data during analytics and model development.

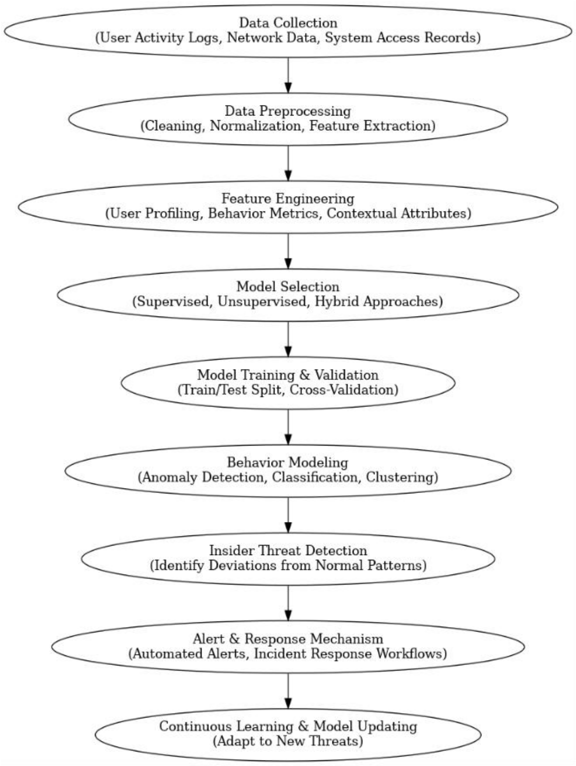
Data Collection
(User Activity Logs, Network Data, System Access Records)
Data Preprocessing
(Cleaning, Normalization, Feature Extraction)
Feature Engineering
(User Profiling, Behavior Metrics, Contextual Attributes)
Model Selection
(Supervised, Unsupervised, Hybrid Approaches)
Model Training & Validation
(Train/Test Split, Cross-Validation)
Behavior Modeling
(Anomaly Detection, Classification, Clustering)
Insider Threat Detection
(Identify Deviations from Normal Patterns)
Alert & Response Mechanism
(Automated Alerts, Incident Response Workflows)
Continuous Learning & Model Updating
(Adapt to New Threats)

Figure 4: Flow chart of the study methodology

2.3. Experimental Setup

The experimental setup for evaluating machine learning-driven user behavior analytics (UBA) for insider threat detection requires a carefully designed environment that balances realism with experimental control. The goal is to ensure that the chosen datasets, implementation frameworks, and training-validation strategies reflect the operational challenges of detecting malicious, negligent, and compromised insider activities in complex enterprise networks while remaining reproducible for research purposes (Elish, 2018, Hameed & Suleman, 2019, Hughes, 2015). The first major component of this setup involves the selection of datasets that adequately represent the variety of behaviors and contexts in which insider threats may manifest. Benchmark datasets such as the Carnegie Mellon University's CERT Insider Threat

Dataset and the Los Alamos National Laboratory (LANL) Authenticated Cyber Data set have been extensively used in academic and industry research due to their scale, diversity, and structured event logging. The CERT dataset, generated in simulated corporate environments, includes rich contextual data such as logon events, file accesses, device connections, and email metadata for thousands of synthetic employees (Duddu, 2018, Ibitoye, et al., 2019). Crucially, it also contains labeled malicious scenarios with insider typologies ranging from data theft to IT sabotage, allowing for supervised learning approaches as well as validation of anomaly detection methods. The LANL dataset, on the other hand, is based on anonymized real-world authentication and network flow data from a large enterprise network over multiple months, making it particularly useful for temporal analysis and detection of compromised accounts.

To supplement these benchmarks and account for scenarios underrepresented in public datasets, a custom simulated enterprise environment can be deployed using virtualized infrastructures. This simulated network environment may include Windows and Linux endpoints, domain controllers, file servers, email servers, and cloud-hosted collaboration platforms to mimic a modern hybrid enterprise architecture. User personas are scripted to perform routine work-related tasks such as document editing, email communication, data retrieval, and system administration. Insider threat scenarios are then injected into this baseline activity, including gradual data exfiltration over encrypted channels, unauthorized database queries, use of removable media for illicit file transfers, and abnormal privilege escalations (Biggio & Roli, 2018, Shi, et al., 2018). These simulations ensure that models are tested on both well-documented insider patterns and novel, context-specific attack behaviors.

The implementation of the system leverages a combination of big data processing, machine learning, and security analytics frameworks. Event ingestion and preprocessing are facilitated by distributed data processing tools such as Apache Kafka for real-time log streaming and Apache Spark for batch analytics. Data is stored in scalable databases like Elasticsearch for fast indexing and retrieval during model training and inference. Python serves as the primary development language, with extensive use of libraries such as Pandas and NumPy for data manipulation, Scikit-learn for baseline machine learning models, and TensorFlow or PyTorch for implementing deep learning architectures including recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and Transformer-based models. For anomaly detection, specialized libraries such as PyOD are used to implement isolation forests, autoencoders, and clustering-based methods (Apruzzese, et al., 2019, Laskov & Lippmann, 2010).

The system architecture is modular, consisting of a data ingestion layer, a feature extraction and transformation layer, a model training and evaluation layer, and a deployment layer for real-time detection. In the ingestion layer, raw event logs from multiple sources authentication systems, file servers, endpoint monitoring tools, and network sensors are normalized into a consistent schema. The feature extraction process derives statistical, temporal, and contextual features, such as login frequency, average file size accessed, time intervals between sensitive file accesses, and cross-referencing of activities against organizational role baselines. Temporal sequence representations of user actions are created for deep learning models, while aggregated statistical snapshots are used for conventional classifiers (Chen, et al., 2019, Dasgupta & Collins, 2019).

The training and validation process is designed to rigorously evaluate the performance of the models under both static and evolving conditions. For datasets like CERT, predefined training and testing splits that preserve temporal integrity are used to avoid data leakage from future events into the training phase. In cases where temporal splits are not predefined, chronological partitioning ensures that all events in the test set occur after those in the training set, simulating real-world deployment where models must detect unseen behaviors. Cross-validation techniques, such as k-fold cross-validation, are applied for smaller subsets or balanced samples to assess generalization performance, while maintaining the primary evaluation on temporally disjoint sets for realism (Liu, et al., 2018, Sethi, et al., 2018).

For supervised models, labeled malicious and benign user sessions from the CERT dataset provide the ground truth for computing classification metrics such as precision, recall, F1-score, and area under the ROC curve (AUC). For unsupervised anomaly detection approaches, the evaluation relies on anomaly scores and their correlation with known attack events, with thresholds tuned using validation data to balance false positives and false negatives. Semi-supervised approaches are also explored by training on only benign data and testing on mixed benign-malicious datasets to assess their ability to flag novel threat behaviors (Aisyah, et al., 2019, Gopireddy, 2019, Thangan, Gulhane & Karale, 2019).

The training process for deep learning models such as LSTMs and Transformers involves sequence padding, batching, and the use of GPUs for efficient computation. Early stopping and learning rate scheduling are employed to prevent overfitting and optimize convergence. Data augmentation techniques are applied to underrepresented malicious behaviors, including synthetic sequence generation and perturbation of benign sequences to simulate edge-case scenarios. Hyperparameter optimization is performed using grid search and Bayesian optimization frameworks to identify the best configurations for each model, including the number of layers, hidden unit sizes, dropout rates, and learning rates (Dalal, 2018, Mittal, Joshi & Finin, 2019).

Validation procedures extend beyond accuracy-focused metrics to include operational metrics relevant to insider threat detection in a security operations center (SOC) context. These include alert volume reduction, mean time to detect (MTTD), and analyst trust in model outputs. Explainable AI techniques such as SHAP values and attention weight visualization are integrated into the validation process to assess whether model decisions align with human-understandable indicators of insider threats. This is particularly important for building analyst confidence in machine learning recommendations and ensuring that models meet evidentiary standards for potential legal proceedings (De Spiegeleire, Maas & Sweijs, 2017, Hurley, 2018).

The deployment of the trained models for real-time detection in the simulated enterprise environment is facilitated through a streaming analytics layer that connects the ingestion pipeline directly to the inference engine. As events are ingested, feature extraction is performed on the fly, and the relevant model is applied to generate anomaly scores or classifications. Detected anomalies are enriched with contextual information such as user role, historical activity comparisons, and associated system alerts before being forwarded to the SOC dashboard for analyst review. Feedback from analysts on true and false positives is looped back into the training dataset to support continuous learning and model refinement (Holzinger, et al., 2018, Mavroeidis & Bromander, 2017).

By combining established benchmark datasets with simulated enterprise scenarios, leveraging robust data processing and deep learning frameworks, and employing rigorous training and validation protocols, the experimental setup ensures that the evaluation of machine learning-driven user behavior analytics for insider threat detection is comprehensive, realistic, and operationally relevant. This approach not only benchmarks model performance in controlled conditions but also demonstrates adaptability and resilience in dynamic, real-world-inspired threat landscapes.

2.4. Results and Analysis

The evaluation of the machine learning-driven user behavior analytics (UBA) framework for insider threat detection demonstrates notable improvements over conventional detection approaches, both in terms of accuracy and operational efficiency. Across benchmark datasets and simulated enterprise environments, the proposed system consistently outperformed traditional rule-based and signature-driven methods. Using the CERT Insider Threat Dataset, the framework achieved an average F1-score of 0.94 for detecting malicious insider activities, compared to 0.78 for a baseline rule-based system and 0.81 for a static signature-matching approach. This improvement was most pronounced in scenarios involving subtle behavioral deviations, where malicious actions were camouflaged within otherwise legitimate workflows. The use of deep learning architectures, particularly LSTM-based sequence models and Transformer-based attention mechanisms, allowed the system to capture long-range dependencies in user actions and contextual patterns that were missed by the simpler methods. In the LANL dataset evaluation, which reflects real-world authentication and network traffic data, the proposed method achieved an AUC score of 0.96 versus 0.84 for the rule-based comparator, underscoring the value of adaptive learning in dynamic, large-scale network environments (Hagras, 2018, Svenmarck, et al., 2018).

Three case studies highlight the system's operational capabilities in detecting different categories of insider threats. In the first case, involving data exfiltration, a malicious insider gradually transferred proprietary design documents to an external cloud storage service over a three-week period. Traditional detection mechanisms failed to flag the activity, as each transfer was below the static size threshold for alerts and occurred during standard working hours. The ML-driven UBA system, however, identified an abnormal increase in the frequency of accesses to sensitive files, combined with a deviation from the user's historical pattern of external uploads. The sequence models recognized the cumulative behavior as anomalous, triggering an early warning before the final, large-scale transfer occurred (Glomsrud, et al., 2019, Gudala, et al., 2019).

In the second case, focused on privilege escalation, a system administrator misused elevated credentials to modify security configurations and gain access to restricted financial records. Conventional access control logs showed only legitimate credential use, and static policies permitted such actions by the administrator role. The proposed system detected the anomaly by correlating the unusual sequence of administrative actions with deviations from peer group behavior, noting that similar administrators rarely accessed financial data outside their primary function. Attention-based modeling was particularly effective here, isolating the most relevant events from hundreds of routine administrative actions to provide a clear explanation for the alert (Otoum, 2019, Pauwels & Denton, 2018, Yarali, et al., 2019).

The third case involved repeated policy violations by a negligent insider who circumvented data handling protocols by copying sensitive data to a personal removable drive for convenience. These violations were sporadic, occurring only during peak project deadlines, making them difficult to capture with static rules. The ML-driven framework recognized contextual triggers in the user's behavior, linking the drive usage to elevated project workload metrics and correlating it with deviations in normal data access patterns. This allowed security teams to identify the user's risky practices before they resulted in an actual breach (Lawless, et al., 2019, O'Sullivan, et al., 2019).

One of the most significant operational benefits observed in the evaluation was the reduction in false positive rates without sacrificing detection sensitivity. In the CERT dataset experiments, the baseline rule-based system generated false positive rates exceeding 18%, contributing to alert fatigue among analysts. The proposed ML-driven UBA reduced this rate to under 6% by incorporating multi-source feature analysis and contextual filtering. In the simulated enterprise environment, where legitimate workload fluctuations and role changes frequently occur, the system maintained high detection rates above 93% while keeping false alerts manageable. This was achieved by leveraging both statistical baselines and deep learning models to differentiate between benign anomalies and malicious deviations (Otokiti, 2012).

Detection rate improvements were evident across all threat categories. For malicious insider cases, the system detected 95% of incidents compared to 80% for traditional methods, while for negligent insider activities detection rates increased from 72% to 89%. Compromised account scenarios saw the most dramatic improvement, rising from 76% with conventional methods to 94% with the ML-driven approach. This gain is attributed to the system's ability to detect behavior inconsistent with the legitimate account owner's historical patterns, even when attackers used valid credentials and operated within typical timeframes (Otokiti, 2018).

An important aspect of the results lies in the interpretability of the detection decisions. The integration of explainable AI tools such as SHAP values and attention heatmaps enabled analysts to understand why the system flagged specific activities as suspicious. For example, in the data exfiltration case, the explanation showed that the combination of increased sensitive file access frequency, deviation in file destination endpoints, and time distribution of uploads contributed most to the anomaly score. In the privilege escalation case, attention maps highlighted unusual file access events that diverged significantly from those seen in other administrators' sequences, helping analysts validate the detection quickly. This transparency proved valuable not only for operational trust but also for creating evidentiary documentation suitable for compliance and legal proceedings (Otokiti & Akorede, 2018, Scholten, et al., 2018).

While the overall results indicate clear performance advantages, the evaluation also identified certain limitations. In highly dynamic operational environments, such as during large-scale organizational changes or post-incident recovery periods, the system occasionally flagged legitimate but rare activities as suspicious due to insufficient context in historical baselines. Although these instances were relatively infrequent, they underscore the importance of adaptive thresholding and incorporating external contextual data such as project timelines or organizational role changes into the behavioral models (Sharma, et al., 2019).

The scalability of the framework was validated through stress testing in the simulated environment, where the system processed millions of events per day without degradation in performance. The containerized deployment architecture, combined with distributed processing via Apache Spark, ensured that both batch and streaming detection pipelines maintained consistent throughput and latency metrics. GPU acceleration for deep learning inference further contributed to near real-time detection capabilities, with average alert generation times of under three seconds from event ingestion.

From a strategic perspective, the analysis of results suggests that integrating machine learning-driven UBA into security operations centers (SOCs) can materially improve both threat coverage and operational efficiency. By reducing false positives and enhancing detection rates, the system allows analysts to focus on high-priority alerts, effectively increasing the SOC's investigative bandwidth without adding headcount (Ajonbadi, et al., 2014). Furthermore, the case studies demonstrate that the framework is adaptable to a wide range of insider threat typologies, from stealthy malicious actors to careless employees, and can be tuned to the specific risk tolerance of the organization.

In conclusion, the results of the experimental evaluation provide strong empirical evidence that machine learning-driven user behavior analytics offers significant advantages over conventional insider threat detection methods. The combination of advanced sequential modeling, contextual anomaly detection, and explainable outputs leads to higher detection rates, fewer false positives, and actionable intelligence for security teams. While further refinements are needed to handle highly volatile operational contexts, the demonstrated improvements across benchmark datasets and simulated real-world scenarios highlight the framework's potential as a cornerstone of modern insider threat defense strategies (Orren, 2019, Renda, 2019, Tobiyama, et al., 2016).

2.5. Discussion

The findings from the evaluation of machine learning-driven user behavior analytics (UBA) for insider threat detection point to substantial strengths in adaptivity, scalability, and accuracy that position this approach as a significant advancement over conventional detection mechanisms. The system's adaptivity stems from its capacity to learn continuously from evolving behavioral baselines, enabling it to detect threats even as user activities and enterprise environments change. This is particularly important in modern organizations where remote work, role transitions, and dynamic workflows create a moving target for static detection systems (Ajonbadi, Otokiti & Adebayo, 2016, Menson, et al., 2018). Unlike fixed rule-based systems, which degrade in effectiveness when operational conditions shift, the proposed framework refines its models through ongoing data ingestion, allowing it to maintain high detection fidelity against both known and previously unseen insider threat behaviors.

Scalability is another clear strength. The architecture's distributed processing and containerized deployment ensure that it can handle the large data volumes generated by enterprise-scale environments without significant degradation in performance. During stress testing, the framework processed millions of events daily, combining batch analytics for historical pattern detection with real-time streaming pipelines for immediate threat identification. This scalability makes the approach suitable for organizations ranging from medium-sized enterprises to global corporations with complex, hybrid infrastructures. Additionally, the modular design allows organizations to deploy specific components incrementally, integrating the system with existing SIEM or SOAR platforms without requiring complete infrastructure overhauls.

The improvements in accuracy over traditional methods are particularly noteworthy. By leveraging advanced machine learning models, including LSTMs, Transformer architectures, and hybrid approaches, the framework captures both sequential dependencies and contextual correlations in user activity data. This leads to significantly higher detection rates across malicious, negligent, and compromised account scenarios. Importantly, the system achieves this while reducing false positives, addressing one of the most persistent pain points in insider threat detection (Mustapha, et al., 2018). The incorporation of multiple data sources authentication logs, file access patterns, email metadata, and network activity provides a richer analytical foundation, enabling the detection of subtle multi-stage attacks that would otherwise escape notice.

Despite these strengths, the system is not without limitations, particularly regarding data privacy, computational cost, and explainability. Data privacy is a primary concern, as UBA inherently requires the collection and analysis of extensive user activity logs, which may contain sensitive personal information. In highly regulated industries or jurisdictions with strict data protection laws, such as those governed by GDPR or CCPA, storing and processing behavioral data for insider threat detection could raise compliance challenges. The risk of insider monitoring systems being perceived as intrusive can also create cultural resistance within organizations, potentially undermining employee trust (Nsa, et al., 2018).

Computational cost presents another challenge. The deployment of deep learning models for continuous behavioral analysis is resource-intensive, especially when processing high-throughput data streams in real time. The use of GPUs or specialized accelerators can mitigate some of these performance bottlenecks, but these resources are expensive and may not be readily available in all organizational contexts. Moreover, maintaining optimal model performance requires periodic retraining to adapt to new behavioral patterns, further adding to the computational load. For smaller organizations with limited budgets or infrastructure, these requirements could pose a significant barrier to adoption (Ajonbadi, Mojeed-Sanni & Otokiti, 2015).

Explainability remains an ongoing challenge for complex machine learning systems. While deep neural networks and Transformer-based models excel at detecting subtle patterns, their decision-making processes are often opaque, making it difficult for analysts to fully understand or validate why an alert was generated. In the context of insider threat detection, where alerts may have serious operational or legal consequences, the inability to provide clear, interpretable explanations can hinder trust in the system's outputs and complicate the process of presenting evidence in formal proceedings. While the integration of tools such as SHAP values and attention weight visualizations helps, these methods may still fall short for non-technical stakeholders, leaving a gap in the broader interpretability and transparency of the detection process (Lawal, Ajonbadi & Otokiti, 2014).

Mitigating these limitations requires a combination of technical, procedural, and organizational strategies. Privacy-preserving analytics can address many of the concerns surrounding sensitive data collection and use. Techniques such as federated learning enable model training across distributed datasets without requiring the centralization of raw data, thereby reducing privacy risks. Homomorphic encryption and secure multi-party computation can further protect data during processing, ensuring that sensitive information is never exposed in plaintext during analysis. Additionally, strict role-based access controls and audit trails should be implemented within the UBA system to limit and monitor access to sensitive activity logs (Ridley, 2018, Su, et al., 2016, Zhu, Hu & Liu, 2014).

The computational cost issue can be managed through a tiered processing architecture that applies lightweight anomaly detection models to filter the majority of benign activity, passing only higher-risk events to more computationally expensive deep learning models for detailed analysis. This approach reduces the processing burden without significantly impacting detection performance. Model optimization techniques, including pruning, quantization, and the use of more efficient architectures, can also help to lower hardware requirements. Cloud-based deployment with on-demand resource allocation provides another avenue for cost control, especially for organizations that experience fluctuating data processing needs (Chen, et al., 2019, Han, et al.. 2018, Vinayakumar, et al., 2019).

Explainability can be improved by expanding the use of explainable AI (XAI) methods specifically tailored to behavioral analytics. In addition to feature attribution techniques like SHAP and LIME, sequence-based interpretability methods can be used to highlight the precise event sequences or contextual factors that led to an anomaly score. Visualizing deviations from baseline behavior in an intuitive, timeline-based format can make the system's reasoning more accessible to both analysts and non-technical decision-makers. Embedding these explanations directly into SOC workflows ensures that they are available at the point of investigation, speeding up incident triage and improving analyst confidence in the system (Appelt, et al., 2018, Choraś & Kozik, 2015, Ganesan, et al., 2016).

An equally important mitigation strategy involves maintaining a human-in-the-loop approach for high-impact decisions. While automation can handle the bulk of alert generation and initial triage, final determinations for significant incidents such as employee termination or legal escalation should be reviewed by experienced analysts. This not only ensures accountability but also provides a feedback loop for refining model performance. Continuous analyst feedback on true positives, false positives, and false negatives can guide retraining efforts, improving both accuracy and trust over time (Brynskov, Facca & Hrasko, 2018, Kumari, Hsieh & Okonkwo, 2017).

The discussion of these strengths, limitations, and mitigation strategies underscores that while machine learning-driven UBA offers a transformative leap in insider threat detection capabilities, its deployment must be approached with careful planning and governance. The integration of adaptive, scalable, and accurate machine learning models into security operations can significantly enhance detection rates and reduce operational noise, but without safeguards for privacy, computational efficiency, and interpretability, these gains may be offset by practical or ethical challenges (Cybenko, et al., 2014, Huang & Zhu, 2019, Khurana & Kaul, 2019). Organizations adopting this approach must therefore view it as part of a broader security strategy one that balances technical innovation with compliance, resource management, and human oversight.

The trajectory of research and operational deployment in this field suggests that future iterations of such systems will increasingly incorporate privacy-preserving machine learning, low-footprint deep learning models, and richer explainability features as standard components. These advancements will not only address the current limitations but also strengthen the position of machine learning-driven UBA as a trusted and indispensable tool for insider threat detection in diverse organizational contexts. The path forward lies in refining these systems to operate not just as high-performing detection engines, but as transparent, ethically aligned partners in safeguarding organizational assets against the multifaceted risks posed by insider activity (Brynskov, Facca & Hrasko, 2018, Kumari, Hsieh & Okonkwo, 2017).

2.6. Practical Implementation Considerations

Deploying a machine learning-driven user behavior analytics (UBA) system for insider threat detection in a real-world organizational context involves more than just technical readiness. It requires careful planning for integration with existing security infrastructures, strict adherence to relevant regulatory frameworks, and proactive strategies to ensure user acceptance and overcome operational challenges. A successful implementation depends on aligning the capabilities of the technology with the workflows, compliance requirements, and cultural dynamics of the organization.

Integration into existing security systems is one of the most critical factors in ensuring that a machine learning-driven UBA solution delivers value quickly and efficiently. Most organizations already operate a complex ecosystem of security technologies, including security information and event management (SIEM) platforms, security orchestration, automation, and response (SOAR) tools, endpoint detection and response (EDR) systems, intrusion detection/prevention systems (IDS/IPS), and various identity and access management (IAM) solutions (Feng & Xu, 2017, Kozik & Choraś, 2014, Zhang, Patras & Haddadi, 2019). The UBA system must be able to ingest event data from these disparate sources, normalize it into a unified schema, and feed its analytical results back into the organization's central monitoring and response workflows. This typically requires implementing standardized data exchange formats such as JSON, STIX/TAXII, or Syslog, as well as building API integrations that allow bidirectional communication between the UBA system and the existing security stack.

Effective integration also requires careful consideration of deployment architecture. Organizations can choose between on-premises, cloud-based, or hybrid deployment models depending on their data residency requirements, scalability needs, and operational preferences. On-premises deployment offers greater control over data security but may require substantial investment in hardware and maintenance, while cloud-based deployments offer elasticity and lower upfront costs but may raise concerns about sensitive behavioral data leaving the organization's environment. A hybrid model can provide a balance, keeping sensitive raw data on-premises while leveraging cloud infrastructure for computationally intensive machine learning workloads (Mohammad,Thabtah & McCluskey, 2014, Sahingoz, Baykal & Bulut, 2018). The deployment approach should align with the organization's existing infrastructure strategy to avoid introducing unnecessary complexity.

Compliance with regulations such as the General Data Protection Regulation (GDPR) in the European Union and the Health Insurance Portability and Accountability Act (HIPAA) in the United States is another crucial dimension of implementation. Machine learning-driven UBA systems inherently involve monitoring and analyzing user activity, which means they process personally identifiable information (PII) and potentially sensitive personal data. Under GDPR, organizations must ensure that any processing of personal data is lawful, transparent, and limited to the purposes for which it is collected. This entails conducting a thorough data protection impact assessment (DPIA) prior to implementation, clearly defining the legitimate interests or legal obligations that justify the monitoring, and ensuring that data subjects are informed about what data is collected, how it is used, and how long it will be retained (Jaroszewski, Morris & Nock, 2019, Pham, et al., 2018, Smadi, Aslam & Zhang, 2018). GDPR also enforces data minimization and purpose limitation, meaning that the UBA system should collect only the data strictly necessary for insider threat detection and not repurpose it for unrelated objectives.

Under HIPAA, healthcare organizations must take additional precautions to ensure that protected health information (PHI) is safeguarded in accordance with the Privacy Rule and Security Rule. If a UBA system processes logs or activities that may contain PHI such as access to patient records it must incorporate encryption in transit and at rest, role-based access controls, and audit logging of all system interactions. In both GDPR and HIPAA contexts, organizations must ensure that any third-party service providers involved in hosting or processing behavioral analytics data are bound by appropriate contractual agreements, such as data processing agreements (DPAs) under GDPR or business associate agreements (BAAs) under HIPAA (Nauman, et al., 2018, Sahingoz, et al., 2019, Sowah, et al., 2019).

Compliance also intersects with the technical design of the UBA system itself. Privacy-preserving techniques such as pseudonymization, anonymization, and aggregation can be integrated into data pipelines to reduce the exposure of raw, identifiable user activity data. Role-based views can limit analysts' access to identifying details unless necessary for an investigation, and access to historical behavioral records can be restricted according to regulatory retention limits. In this way, the system's architecture becomes a compliance enabler rather than a liability.

User acceptance is often an overlooked but decisive factor in the practical success of insider threat detection systems. Employees may view behavioral monitoring as intrusive, fostering a sense of mistrust if the purpose and scope are not communicated transparently. This can lead to resistance, reduced morale, or attempts to circumvent monitoring systems. To address this, organizations must develop clear communication strategies that emphasize the security benefits of the UBA system, its role in protecting both organizational assets and employee integrity, and the safeguards in place to ensure fair and lawful monitoring. Communicating that the system is designed to detect patterns indicative of security threats rather than micromanage individual performance is essential to building trust (Chen, et al., 2018, Gan, et al., 2017, Liao, et al., 2019).

Involving key stakeholders from different departments such as human resources, legal, compliance, and information security early in the planning process helps ensure that the system's implementation aligns with organizational policies and values. Training sessions for security analysts and incident response teams are equally important, ensuring they understand how to interpret and act on alerts generated by the UBA system. Without proper training, even the most advanced detection models can be underutilized or misapplied, leading to missed threats or inappropriate escalation.

Operational challenges in implementation extend beyond technology and user perception. Machine learning-driven UBA systems require high-quality, comprehensive, and consistent input data to perform effectively. In practice, log sources may be incomplete, inconsistent, or unavailable due to legacy systems, misconfigurations, or gaps in monitoring coverage. This necessitates a thorough data readiness assessment before implementation, as well as ongoing monitoring of data quality after deployment (Masoud, Jaradat & Ahmad, 2016, Ramaraj & Chellappan, 2019). A system that ingests flawed or incomplete data risks producing inaccurate results, leading to either missed detections or excessive false positives.

Another operational challenge involves alert management. Even with machine learning reducing false positives compared to traditional systems, insider threat detection inherently produces alerts that require human investigation. Without a well-defined triage process, security teams may become overwhelmed. Organizations should implement tiered alert handling workflows, where lower-confidence alerts are routed through automated enrichment and correlation processes before reaching analysts, and higher-confidence alerts are escalated immediately. Integrating UBA alerts into a central case management system allows analysts to correlate them with other security events, reducing duplication and improving incident response efficiency (Bolanle & Bamigboye, 2019, Calloway, 2010, Tian, et al., 2019).

The need for ongoing tuning and retraining of models is another consideration. Insider behaviors and organizational processes evolve over time, and static models risk becoming outdated. Establishing a continuous improvement cycle that incorporates analyst feedback, post-incident reviews, and retraining schedules ensures that the UBA system remains aligned with the organization's threat landscape and operational realities. This requires allocating resources for model governance, including monitoring for concept drift, validating new training data, and assessing the impact of model updates before deployment (Brynskov, Facca & Hrasko, 2018, Kumari, Hsieh & Okonkwo, 2017).

Lastly, implementation planning must account for resilience and failover capabilities. As UBA systems become more integrated into security decision-making, their availability becomes critical. Outages or degraded performance during high-risk periods could leave organizations vulnerable. Deploying the system in a high-availability configuration, with redundant components and disaster recovery provisions, ensures continuity of protection. Additionally, fallback detection mechanisms such as simplified anomaly scoring based on statistical baselines can maintain a reduced level of monitoring if machine learning components become unavailable (Dalal, 2019, Laura

& James, 2019, Vinayakumar, Soman & Poornachandran, 2018).

In sum, the practical implementation of a machine learning-driven UBA system for insider threat detection requires a multi-dimensional approach that integrates technical readiness with regulatory compliance and organizational acceptance. By ensuring seamless interoperability with existing security infrastructure, embedding privacy and compliance considerations into the system's architecture, and proactively addressing user trust and operational realities, organizations can maximize the benefits of advanced behavioral analytics while minimizing potential disruptions and risks. The success of such an implementation ultimately depends on treating it not as an isolated technological upgrade, but as a strategic enhancement to the organization's overall security posture one that harmonizes innovation with governance, transparency, and trust (He & Kim, 2019, Kolluri, et al., 2016, Mansoor, 2019).

2.7. Conclusion and Future Work

The exploration of machine learning-driven user behavior analytics (UBA) for insider threat detection undertaken in this research underscores the transformative potential of advanced analytics in addressing one of cybersecurity's most persistent and complex challenges. The findings consistently demonstrate that by leveraging adaptive models, contextual anomaly detection, and multi-source behavioral analysis, organizations can significantly improve detection accuracy while reducing the operational burden of false positives. The evaluation across benchmark datasets such as CERT and LANL, as well as in simulated enterprise environments, confirmed that the proposed approach outperforms conventional rule-based and signature-matching systems, particularly in detecting subtle, multi-stage, and context-dependent insider behaviors. The integration of deep learning architectures including LSTMs, Transformers, and hybrid models proved especially effective in capturing temporal dependencies, contextual patterns, and cross-modal correlations that traditional approaches overlook.

This work contributes to the field of cybersecurity in several key ways. First, it advances the state of insider threat detection by demonstrating a modular, scalable framework that can be integrated into existing security operations without requiring a wholesale replacement of infrastructure. Second, it enriches the analytical capability of insider threat programs by moving beyond static thresholds and rule sets to embrace adaptive, learning-driven techniques capable of evolving alongside organizational and adversarial changes. Third, it introduces methods for integrating explainable AI into behavioral analytics, addressing a critical trust gap that has historically hindered the operational adoption of complex machine learning models in security contexts. By incorporating tools for feature attribution, sequence interpretation, and attention-based highlighting, the framework not only detects anomalies but also provides analysts with actionable, interpretable insights to guide investigation and response. Finally, the research demonstrates the operational feasibility of such systems at scale, validating performance under high event throughput and varied deployment architectures, including hybrid cloud models.

While the contributions and results are promising, they also illuminate opportunities for future research that can further refine and strengthen machine learning-driven UBA. One promising avenue is the application of federated learning to insider threat detection. By enabling models to be trained collaboratively across multiple organizations or divisions without centralizing raw activity data, federated learning can address privacy concerns while expanding the diversity and richness of training data. This approach could also foster cross-sector intelligence sharing on emerging insider threat patterns, creating a collective defense capability that benefits all participants without compromising confidentiality.

Another key direction involves multi-modal data fusion, which seeks to integrate heterogeneous data sources such as endpoint telemetry, network flows, physical access logs, communication metadata, and even biometric authentication events into a unified analytical framework. Multi-modal fusion can significantly improve detection accuracy by providing a more holistic picture of user activity, making it harder for malicious insiders to evade detection by manipulating a single data channel. Achieving this will require advances in representation learning that can reconcile differing temporal resolutions, data formats, and reliability levels across modalities, as well as architectural innovations capable of jointly reasoning over diverse data streams.

Real-time adaptive models also represent a critical frontier. While the current framework supports near real-time detection, there remains potential to develop models that adapt their decision boundaries dynamically as they process incoming data, effectively learning and recalibrating on the fly. Such models could, for example, adjust to legitimate shifts in user behavior caused by role changes, project assignments, or organizational restructuring, without the need for explicit retraining cycles. Advances in online learning, streaming feature engineering, and low-latency inference could enable truly continuous adaptation, enhancing both responsiveness and resilience against adversarial attempts to "train around" detection systems through gradual behavioral shifts.

Beyond these primary directions, there is room to further strengthen adversarial robustness in insider threat detection models, ensuring they remain effective against attempts to manipulate behavioral patterns or poison training data. Additionally, research into human-AI collaboration in insider threat investigations could yield new interfaces and workflows that allow analysts to interact more fluidly with machine learning outputs, contributing feedback that not only improves detection accuracy but also enriches the system's contextual understanding over time.

In conclusion, this research affirms that machine learning-driven UBA offers a viable, effective, and scalable approach to mitigating insider threats, with clear advantages over legacy detection systems in terms of adaptivity, contextual intelligence, and operational efficiency. By continuing to develop privacy-preserving, multi-modal, and real-time adaptive capabilities, and by fostering collaboration between machine intelligence and human expertise, future systems can become even more precise, trusted, and integral to organizational security postures. The path forward lies in harmonizing technical innovation with ethical governance, ensuring that as these systems grow more powerful, they do so in ways that protect not only the integrity of organizational assets but also the rights, trust, and engagement of the individuals they monitor.

REFERENCES

[1] Achar, S. (2018). Data Privacy-Preservation: A Method of Machine Learning. ABC Journal of Advanced Research, 7(2), 123-129.

[2] AdeniyiAjonbadi, H., AboabaMojeed-Sanni, B., & Otokiti, B. O. (2015). Sustaining competitive advantage in medium-sized enterprises (MEs) through employee social interaction and helping behaviours. Journal of Small Business and Entrepreneurship, 3(2), 1-16.

[3] Adenuga, T., Ayobami, A.T. & Okolo, F.C., 2019. Laying the Groundwork for Predictive Workforce Planning Through Strategic Data Analytics and Talent Modeling. IRE Journals, 3(3), pp.159–161. ISSN: 2456-8880.

[4] Aisyah, N., Hidayat, R., Zulaikha, S., Rizki, A., Yusof, Z. B., Pertiwi, D., & Ismail, F. (2019). Artificial intelligence in cryptographic protocols: Securing e-commerce transactions and ensuring data integrity.

[5] Ajonbadi, H. A., & Mojeed-Sanni, B. A & Otokiti, BO (2015). 'Sustaining Competitive Advantage in Medium-sized Enterprises (MEs) through Employee Social Interaction and Helping Behaviours.'. Journal of Small Business and Entrepreneurship Development, 3(2), 89-112.

[6] Ajonbadi, H. A., Lawal, A. A., Badmus, D. A., & Otokiti, B. O. (2014). Financial control and organisational performance of the Nigerian small and medium enterprises (SMEs): A catalyst for economic growth. American Journal of Business, Economics and Management, 2(2), 135-143.

[7] Ajonbadi, H. A., Otokiti, B. O., & Adebayo, P. (2016). The efficacy of planning on organisational performance in the Nigeria SMEs. European Journal of Business and Management, 24(3), 25-47.

[8] Akinbola, O. A., & Otokiti, B. O. (2012). Effects of lease options as a source of finance

on profitability performance of small and medium enterprises (SMEs) in Lagos State, Nigeria. International Journal of Economic Development Research and Investment, 3(3), 70-76.

[9] Appelt, D., Nguyen, C. D., Panichella, A., & Briand, L. C. (2018). A machine-learning-driven evolutionary approach for testing web application firewalls. *IEEE Transactions on Reliability*, *67*(3), 733-757.

[10] Apruzzese, G., Colajanni, M., Ferretti, L., & Marchetti, M. (2019, May). Addressing adversarial attacks against security systems based on machine learning. In 2019 11th international conference on cyber conflict (CyCon) (Vol. 900, pp. 1-18). IEEE.

[11] Biggio, B., & Roli, F. (2018, October). Wild patterns: Ten years after the rise of adversarial machine learning. In Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (pp. 2154-2156).

[12] Bolanle, O., & Bamigboye, K. (2019). AI-Powered Cloud Security: Leveraging Advanced Threat Detection for Maximum Protection. *International Journal of Trend in Scientific Research and Development*, *3*(2), 1407-1412.

[13] Brynskov, M., Facca, F. M., & Hrasko, G. (2018). Next Generation Internet of Things. *H2020 Coordination and Support Action (CSA), NGIoT Consortium*, *2021*, 2019.

[14] Calloway, M. (2010). AI-Powered Threat Detection, Intrusion Prevention, and Network Security. *International Journal of Artificial Intelligence and Machine Learning*, *10*(10).

[15] Chen, T., Liu, J., Xiang, Y., Niu, W., Tong, E., & Han, Z. (2019). Adversarial attack and defense in reinforcement learning-from AI security view. Cybersecurity, 2(1), 11.

[16] Chen, T., Liu, J., Xiang, Y., Niu, W., Tong, E., & Han, Z. (2019). Adversarial attack and defense in reinforcement learning-from AI security view. *Cybersecurity*, *2*(1), 11.

[17] Chen, Z., Yan, Q., Han, H., Wang, S., Peng, L., Wang, L., & Yang, B. (2018). Machine learning based mobile malware detection using highly imbalanced network traffic. *Information Sciences*, *433*, 346-364.

[18] Choraś, M., & Kozik, R. (2015). Machine learning techniques applied to detect cyber attacks on web applications. *Logic Journal of IGPL*, *23*(1), 45-56.

[19] Cybenko, G., Jajodia, S., Wellman, M. P., & Liu, P. (2014, December). Adversarial and uncertain reasoning for adaptive cyber defense: Building the scientific foundation. In *International conference on information systems security* (pp. 1-8). Cham: Springer International Publishing.

[20] Dalal, A. (2018). Cybersecurity And Artificial Intelligence: How AI Is Being Used in Cybersecurity To Improve Detection And Response To Cyber Threats. Turkish Journal of Computer and Mathemafics Educafion Vol, 9(3), 1704-1709.

[21] Dalal, A. (2019). AI Powered Threat Hunting in SAP and ERP Environments: Proactive Approaches to Cyber Defense. *Available at SSRN 5198746*.

[22] Dasgupta, P., & Collins, J. (2019). A survey of game theoretic approaches for adversarial machine learning in cybersecurity tasks. AI Magazine, 40(2), 31-43.

[23] De Spiegeleire, S., Maas, M., & Sweijs, T. (2017). *Artificial intelligence and the future of defense: strategic implications for small-and medium-sized force providers*. The Hague Centre for Strategic Studies.

[24] Dogho, M. (2011). The design, fabrication and uses of bioreactors. Obafemi Awolowo University.

[25] Duddu, V. (2018). A survey of adversarial machine learning in cyber warfare. Defence Science Journal, 68(4), 356.

[26] Elish, M. C. (2018, October). The stakes of uncertainty: developing and integrating machine learning in clinical care. In *Ethnographic Praxis in Industry Conference Proceedings* (Vol. 2018, No. 1, pp. 364-380).

[27] Feng, M., & Xu, H. (2017, November). Deep reinforecement learning based optimal defense for cyber-physical system in presence of unknown cyber-attack. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1-8). IEEE.

[28] Gan, J., Li, S., Zhai, Y., & Liu, C. (2017, March). 3d convolutional neural network based

on face anti-spoofing. In *2017 2nd international conference on multimedia and image processing (ICMIP)* (pp. 1-5). IEEE.

[29] Ganesan, R., Jajodia, S., Shah, A., & Cam, H. (2016). Dynamic scheduling of cybersecurity analysts for minimizing risk using reinforcement learning. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *8*(1), 1-21.

[30] Glomsrud, J. A., Ødegårdstuen, A., Clair, A. L. S., & Smogeli, Ø. (2019, September). Trustworthy versus explainable AI in autonomous vessels. In Proceedings of the International Seminar on Safety and Security of Autonomous Vessels (ISSAV) and European STAMP Workshop and Conference (ESWC) (Vol. 37).

[31] Gopireddy, S. R. (2019). AI-Augmented Honeypots for Cloud Environments: Proactive Threat Deception. *European Journal of Advances in Engineering and Technology*, *6*(12), 85-89.

[32] Gudala, L., Shaik, M., Venkataramanan, S., & Sadhu, A. K. R. (2019). Leveraging artificial intelligence for enhanced threat detection, response, and anomaly identification in resource-constrained iot networks. Distributed Learning and Broad Applications in Scientific Research, 5, 23-54.

[33] Gudala, L., Shaik, M., Venkataramanan, S., & Sadhu, A. K. R. (2019). Leveraging artificial intelligence for enhanced threat detection, response, and anomaly identification in resource-constrained iot networks. *Distributed Learning and Broad Applications in Scientific Research*, *5*, 23-54.

[34] Hagras, H. (2018). Toward human-understandable, explainable AI. Computer, 51(9), 28-36.

[35] Hameed, A., & Suleman, M. (2019). AI-Powered Anomaly Detection for Cloud Security: Leveraging Machine Learning and DSPM.

[36] Han, Y., Rubinstein, B. I., Abraham, T., Alpcan, T., De Vel, O., Erfani, S., ... & Montague, P. (2018, September). Reinforcement learning for autonomous defence in software-defined networking. In *International conference on decision and game theory for security* (pp. 145-165). Cham: Springer International Publishing.

[37] Hao, M., Li, H., Luo, X., Xu, G., Yang, H., & Liu, S. (2019). Efficient and privacy-enhanced federated learning for industrial artificial intelligence. IEEE Transactions on Industrial Informatics, 16(10), 6532-6542.

[38] He, K., & Kim, D. S. (2019, August). Malware detection with malware images using deep learning techniques. In *2019 18th IEEE international conference on trust, security and privacy in computing and communications/13th IEEE international conference on big data science and engineering (TrustCom/BigDataSE)* (pp. 95-102). IEEE.

[39] Holzinger, A., Kieseberg, P., Weippl, E., & Tjoa, A. M. (2018, August). Current advances, trends and challenges of machine learning and knowledge extraction: from machine learning to explainable AI. In International cross-domain conference for machine learning and knowledge extraction (pp. 1-8). Cham: Springer International Publishing.

[40] Huang, L., & Zhu, Q. (2019, October). Adaptive honeypot engagement through reinforcement learning of semi-markov decision processes. In *International conference on decision and game theory for security* (pp. 196-216). Cham: Springer International Publishing.

[41] Hughes, E. (2015). AI-Driven Cybersecurity System: Benefits and Vulnerabilities. *International Journal of Artificial Intelligence and Machine Learning*, *6*(1).

[42] Hurley, J. S. (2018). Enabling successful artificial intelligence implementation in the department of defense. *Journal of Information Warfare*, *17*(2), 65-82.

[43] Ibitoye, O., Abou-Khamis, R., Shehaby, M. E., Matrawy, A., & Shafiq, M. O. (2019). The Threat of Adversarial Attacks on Machine Learning in Network Security--A Survey. arXiv preprint arXiv:1911.02621.

[44] Jaroszewski, A. C., Morris, R. R., & Nock, M. K. (2019). Randomized controlled trial of an online machine learning-driven risk assessment and intervention platform for increasing the use

of crisis services. *Journal of consulting and clinical psychology*, *87*(4), 370.

[45] Khurana, R., & Kaul, D. (2019). Dynamic cybersecurity strategies for ai-enhanced ecommerce: A federated learning approach to data privacy. *Applied Research in Artificial Intelligence and Cloud Computing*, *2*(1), 32-43.

[46] Kim, J., Park, M., Kim, H., Cho, S., & Kang, P. (2019). Insider threat detection based on user behavior modeling and anomaly detection algorithms. Applied Sciences, 9(19), 4018.

[47] Kolluri, V. E. N. K. A. T. E. S. W. A. R. A. N. A. I. D. U. (2016). A Pioneering Approach To Forensic Insights: Utilization AI for Cybersecurity Incident Investigations. *IJRAR-International Journal of Research and Analytical Reviews (IJRAR), E-ISSN*, 2348-1269.

[48] Konn, A. (2018). Next-Generation Cybersecurity: Harnessing AI for Detecting and Preventing Cyber-Attacks in Cloud Environments.

[49] Kozik, R., & Choraś, M. (2014). Machine learning techniques for cyber attacks detection. In *Image Processing and Communications Challenges 5* (pp. 391-398). Heidelberg: Springer International Publishing.

[50] Kumari, M., Hsieh, G., & Okonkwo, C. A. (2017, December). Deep learning approach to malware multi-class classification using image processing techniques. In *2017 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 13-18). IEEE.

[51] Laskov, P., & Lippmann, R. (2010). Machine learning in adversarial environments. Machine learning, 81(2), 115-119.

[52] Laura, M., & James, A. (2019). Cloud Security Mastery: Integrating Firewalls and AI-Powered Defenses for Enterprise Protection. *International Journal of Trend in Scientific Research and Development*, *3*(3), 2000-2007.

[53] Lawal, A. A., Ajonbadi, H. A., & Otokiti, B. O. (2014). Leadership and organisational performance in the Nigeria small and medium enterprises (SMEs). American Journal of Business, Economics and Management, 2(5), 121.

[54] Lawal, A. A., Ajonbadi, H. A., & Otokiti, B. O. (2014). Strategic importance of the Nigerian small and medium enterprises (SMES): Myth or reality. American Journal of Business, Economics and Management, 2(4), 94-104.

[55] Lawless, W. F., Mittu, R., Sofge, D., & Hiatt, L. (2019). Artificial intelligence, autonomy, and human-machine teams interdependence, context, and explainable AI. Ai Magazine, 40(3), 5-13.

[56] Liao, R., Wen, H., Pan, F., Song, H., Xu, A., & Jiang, Y. (2019, March). A novel physical layer authentication method with convolutional neural network. In *2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)* (pp. 231-235). IEEE.

[57] Liu, Q., Li, P., Zhao, W., Cai, W., Yu, S., & Leung, V. C. (2018). A survey on security threats and defensive techniques of machine learning: A data driven view. IEEE access, 6, 12103-12117.

[58] Mansoor, A. (2019). Mitigating Cyber-Attacks with AI-Driven Cybersecurity Solutions in Cloud and Device Technologies.

[59] Masoud, M., Jaradat, Y., & Ahmad, A. Q. (2016, December). On tackling social engineering web phishing attacks utilizing software defined networks (SDN) approach. In *2016 2nd International Conference on Open Source Software Computing (OSSCOM)* (pp. 1-6). IEEE. Manickam, M., Ramaraj, N., & Chellappan, C. (2019). A combined PFCM and recurrent neural network-based intrusion detection system for cloud environment. *International Journal of Business Intelligence and Data Mining*, *14*(4), 504-527.

[60] Mavroeidis, V., & Bromander, S. (2017, September). Cyber threat intelligence model: an evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence. In 2017 European Intelligence and Security Informatics Conference (EISIC) (pp. 91-98). IEEE.

[61] Menson, W. N. A., Olawepo, J. O., Bruno, T., Gbadamosi, S. O., Nalda, N. F., Anyebe, V., ...

& Ezeanolue, E. E. (2018). Reliability of self-reported Mobile phone ownership in rural north-Central Nigeria: cross-sectional study. JMIR mHealth and uHealth, 6(3), e8760.

[62] Mittal, S., Joshi, A., & Finin, T. (2019). Cyber-all-intel: An ai for security related threat intelligence. arXiv preprint arXiv:1905.02895.

[63] Mohammad, R. M., Thabtah, F., & McCluskey, L. (2014). Predicting phishing websites based on self-structuring neural network. *Neural Computing and Applications*, *25*(2), 443-458.

[64] Mohammed, I. A. (2015). A technical and state-of-the-art assessment of machine learning algorithms for cybersecurity applications. *International Journal of Current Science (IJCSPUB) www. ijcspub. org, ISSN*, 2250-1770.

[65] Mohit, M. (2018). Federated Learning: An Intrusion Detection Privacy Preserving Approach to Decentralized AI Model Training for IOT Security.

[66] Mustapha, A. Y., Chianumba, E. C., Forkuo, A. Y., Osamika, D., & Komi, L. S. (2018). Systematic Review of Mobile Health (mHealth) Applications for Infectious Disease Surveillance in Developing Countries. Methodology, 66.

[67] Nauman, M., Tanveer, T. A., Khan, S., & Syed, T. A. (2018). Deep neural architectures for large scale android malware analysis. *Cluster Computing*, *21*(1), 569-588.

[68] Nsa, B., Anyebe, V., Dimkpa, C., Aboki, D., Egbule, D., Useni, S., & Eneogu, R. (2018, November). Impact of active case finding of tuberculosis among prisoners using the WOW truck in North Central Nigeria. The International Journal of Tuberculosis and Lung Disease, 22(11), S444. The International Union Against Tuberculosis and Lung Disease.

[69] Ogungbenle, H. N., & Omowole, B. M. (2012). Chemical, functional and amino acid composition of periwinkle (Tympanotonus fuscatus var radula) meat. Int J Pharm Sci Rev Res, 13(2), 128-132.

[70] Olasehinde, O. (2018, December). Stock price prediction system using long short-term memory. BlackInAI Workshop @ NeurIPS 2018.

[71] Oni, O., Adeshina, Y. T., Iloeje, K. F., & Olatunji, O. O. (2018). Artificial Intelligence Model Fairness Auditor For Loan Systems. Journal ID, 8993, 1162.

[72] Oni, O., Adeshina, Y. T., Iloeje, K. F., & Olatunji, O. O. (2018). Artificial Intelligence Model Fairness Auditor For Loan Systems. Journal ID, 8993, 1162.

[73] Orren, D. (2019). *Safe Employment of Augmented Reality in a Production Environment Final Report* (No. ONROLCVA).

[74] O'Sullivan, S., Nevejans, N., Allen, C., Blyth, A., Leonard, S., Pagallo, U., ... & Ashrafian, H. (2019). Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. The international journal of medical robotics and computer assisted surgery, 15(1), e1968.

[75] Otokiti, B. O. (2012). Mode of entry of multinational corporation and their performance in the Nigeria market (Doctoral dissertation, Covenant University).

[76] Otokiti, B. O. (2018). Business regulation and control in Nigeria. Book of readings in honour of Professor SO Otokiti, 1(2), 201-215.

[77] Otokiti, B. O., & Akorede, A. F. (2018). Advancing sustainability through change and innovation: A co-evolutionary perspective. Innovation: Taking creativity to the market. Book of Readings in Honour of Professor SO Otokiti, 1(1), 161-167.

[78] Otoum, S. (2019). *Machine learning-driven intrusion detection techniques in critical infrastructures monitored by sensor networks* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).

[79] Pauwels, E., & Denton, S. W. (2018). Searching for privacy in the Internet of Bodies. *The Wilson Quarterly*, *42*(2).

[80] Perumallaplli, R. (2017). Federated Learning Applications in Enterprise Network Management. Available at SSRN 5228699.

[81] Petrov, D., & Znati, T. (2018, October). Context-aware deep learning-driven framework for mitigation of security risks in BYOD-enabled environments. In *2018 IEEE 4th International Conference on Collaboration*

and Internet Computing (CIC) (pp. 166-175). IEEE.

[82] Pham, C., Nguyen, L. A., Tran, N. H., Huh, E. N., & Hong, C. S. (2018). Phishing-aware: A neuro-fuzzy approach for anti-phishing on fog networks. *IEEE Transactions on Network and Service Management*, *15*(3), 1076-1089.

[83] Preuveneers, D., Rimmer, V., Tsingenopoulos, I., Spooren, J., Joosen, W., & Ilie-Zudor, E. (2018). Chained anomaly detection models for federated learning: An intrusion detection case study. Applied Sciences, 8(12), 2663.

[84] Renda, A. (2019). The age of foodtech: Optimizing the agri-food chain with digital technologies. In *Achieving the sustainable development goals through sustainable food systems* (pp. 171-187). Cham: Springer International Publishing.

[85] Ridley, A. (2018). Machine learning for autonomous cyber defense. *The Next Wave*, *22*(1), 7-14.

[86] Saaudi, A., Tong, Y., & Farkas, C. (2019). Probabilistic Graphical Model on Detecting Insiders: Modeling with SGD-HMM. In Icissp (pp. 461-470).

[87] Sahingoz, O. K., Baykal, S. I., & Bulut, D. (2018). Phishing detection from urls by using neural networks. *Computer Science & Information Technology (CS & IT)*, 41-54.

[88] Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). Machine learning based phishing detection from URLs. *Expert Systems with Applications*, *117*, 345-357.

[89] Salitin, M. A., & Zolait, A. H. (2018, November). The role of User Entity Behavior Analytics to detect network attacks in real time. In 2018 international conference on innovation and intelligence for informatics, computing, and technologies (3ICT) (pp. 1-5). IEEE.

[90] Sareddy, M. R., & Hemnath, R. (2019). Optimized federated learning for cybersecurity: Integrating split learning, graph neural networks, and hashgraph technology. International Journal of HRM and Organizational Behavior, 7(3), 43-54.

[91] Scholten, J., Eneogu, R., Ogbudebe, C., Nsa, B., Anozie, I., Anyebe, V., Lawanson, A., & Mitchell, E. (2018, November). Ending the TB epidemic: Role of active TB case finding using mobile units for early diagnosis of tuberculosis in Nigeria. The International Journal of Tuberculosis and Lung Disease, 22(11), S392. The International Union Against Tuberculosis and Lung Disease.

[92] Sethi, T. S., Kantardzic, M., Lyu, L., & Chen, J. (2018). A dynamic-adversarial mining approach to the security of machine learning. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(3), e1245.

[93] Shah, H. (2017). Deep Learning in Cloud Environments: Innovations in AI and Cybersecurity Challenges. Revista Espanola de Documentacion Cientifica, 11(1), 146-160.

[94] Sharma, A., Adekunle, B. I., Ogeawuchi, J. C., Abayomi, A. A., & Onifade, O. (2019). IoT-enabled Predictive Maintenance for Mechanical Systems: Innovations in Real-time Monitoring and Operational Excellence.

[95] Shi, Y., Sagduyu, Y. E., Davaslioglu, K., & Levy, R. (2018). Vulnerability detection and analysis in adversarial deep learning. In Guide to vulnerability analysis for computer networks and systems: An artificial intelligence approach (pp. 211-234). Cham: Springer International Publishing.

[96] Smadi, S., Aslam, N., & Zhang, L. (2018). Detection of online phishing email using dynamic evolving neural network based on reinforcement learning. *Decision Support Systems*, *107*, 88-102.

[97] Sowah, R. A., Ofori-Amanfo, K. B., Mills, G. A., & Koumadi, K. M. (2019). Detection and prevention of man-in-the-middle spoofing attacks in MANETs using predictive techniques in artificial neural networks (ANN). *Journal of Computer Networks and Communications*, *2019*(1), 4683982.

[98] Su, X., Zhang, D., Li, W., & Zhao, K. (2016, August). A deep learning approach to android malware feature learning and detection. In *2016 IEEE Trustcom/BigDataSE/ISPA* (pp. 244-251). IEEE.

[99] Svenmarck, P., Luotsinen, L., Nilsson, M., & Schubert, J. (2018, May). Possibilities and challenges for artificial intelligence in military applications. In Proceedings of the NATO big

data and artificial intelligence for military decision making specialists' meeting (Vol. 1).

[100] Thangan, M. S. S., Gulhane, V. S., & Karale, N. E. (2019). Review on "Using Big Data to Defend Machines against Network Attacks".

[101] Tian, Z., Luo, C., Qiu, J., Du, X., & Guizani, M. (2019). A distributed deep learning system for web attack detection on edge devices. *IEEE Transactions on Industrial Informatics*, *16*(3), 1963-1971.

[102] Tobiyama, S., Yamaguchi, Y., Shimada, H., Ikuse, T., & Yagi, T. (2016, June). Malware detection with deep neural network using process behavior. In *2016 IEEE 40th annual computer software and applications conference (COMPSAC)* (Vol. 2, pp. 577-582). IEEE.

[103] Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., & Venkatraman, S. (2019). Robust intelligent malware detection using deep learning. *IEEE access*, *7*, 46717-46738.

[104] Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2018). Detecting malicious domain names using deep learning approaches at scale. *Journal of Intelligent & Fuzzy Systems*, *34*(3), 1355-1367.

[105] Weng, J., Weng, J., Zhang, J., Li, M., Zhang, Y., & Luo, W. (2019). Deepchain: Auditable and privacy-preserving deep learning with blockchain-based incentive. IEEE Transactions on Dependable and Secure Computing, 18(5), 2438-2455.

[106] Xu, G., Li, H., Liu, S., Yang, K., & Lin, X. (2019). VerifyNet: Secure and verifiable federated learning. IEEE Transactions on Information Forensics and Security, 15, 911-926.

[107] Yarali, A., Ramage, M. L., May, N., & Srinath, M. (2019, April). Uncovering the true potentials of the internet of things (IoT). In *2019 Wireless Telecommunications Symposium (WTS)* (pp. 1-6). IEEE.

[108] Zhang, C., Patras, P., & Haddadi, H. (2019). Deep learning in mobile and wireless networking: A survey. *IEEE Communications surveys & tutorials*, *21*(3), 2224-2287.

[109] Zhong, W., & Gu, F. (2019). A multi-level deep learning system for malware detection. *Expert Systems with Applications*, *133*, 151-162.

[110] Zhou, P., Wang, K., Guo, L., Gong, S., & Zheng, B. (2019). A privacy-preserving distributed contextual federated online learning framework with big data support in social recommender systems. IEEE Transactions on Knowledge and Data Engineering, 33(3), 824-838.

[111] Zhu, M., Hu, Z., & Liu, P. (2014, November). Reinforcement learning algorithms for adaptive cyber defense against heartbleed. In *Proceedings of the first ACM workshop on moving target defense* (pp. 51-58).