



e-ISSN:2582-7219



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 12, December 2024



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.521



6381 907 438



6381 907 438



ijmrset@gmail.com



www.ijmrset.com



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Proactive Cybersecurity: Predictive Analytics and Machine Learning for Identity and Threat Management

Govindarajan Lakshmikanthan¹, Sreejith Sreekandan Nair²

Independent Researcher, Texas, USA¹

Independent Researcher, Texas, USA²

ABSTRACT: Due to the development of advanced identity based attacks and even complex cyber threats, merely possessing defensive cyber security capabilities is not enough today. In this study, we investigate how predictive analytics based machine learning (ML) can be employed for pro-active identity management and threat detection. In this study, the authors assess some models of machine learning – Decision Trees, Random Forests, Support Vector Machines (SVM), and a new hybrid one – to determine which best allows for the detection of both known and unknown threats. The results reveal that in metrics such as accuracy, precision, recall, and F1 score. The hybrid model incorporating both supervised and unsupervised learning approaches scored the highest among other models. As a consequence of its adaptability, the hybrid model is capable of real time dynamic threat detection and anomaly based identity management which makes it an appropriate model for the changing cyber security environment. This study provides the prospects to make proactive cybersecurity more efficient and therefore enhancing the technology for protection systems.

KEYWORDS: Cybersecurity, Predictive Analytics, Machine Learning, Hybrid Model, Threat Detection, Identity Management

I. INTRODUCTION

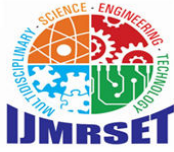
As technology has advanced in the modern world, it has also brought about many benefits which include easy transfer of information, accessibility from anywhere, and improved user interface (Gómez-Carmona et al. 2023). Transitions such as this one have however created new problems in the field of cybersecurity as new and much cleverer cyber threats emerge. Today's cyber security problems encompass data loss, digital impersonation, advanced persistent threats (APTs) and ransomware attacks which, among other things require proactive damage control (Alshamrani et al. 2019). Such traditional techniques most of which are based on firm policies or regulations allowing manual tracking are not able to cope up with such threats and hence put the organizations at risk of attack (Saxena et al. 2020). There is therefore evidence that demand for predictive and dynamic policies that address issues before they emerge is on the increase. Cybersecurity analytics has developed into one of the best solutions in defense management in this regard by utilizing the most creative algorithms available to prevent security threats.

The Role of Predictive Analytics in Cybersecurity

Predictive analytics involves using statistical techniques, data mining, and machine learning to analyze historical and real-time data, uncover patterns, and predict future events (Kuppuswamy et al. 2024). Within the context of cybersecurity, predictive analytics enables security systems to detect early signs of potential threats and identify vulnerabilities before they are exploited. This proactive approach marks a significant departure from traditional detection systems that respond only after an attack has occurred. By applying predictive analytics, organizations can monitor behavioral patterns, assess risks, and take preemptive actions to mitigate threats, improving their overall security posture.

Machine Learning as a Catalyst for Proactive Security

Machine learning (ML) has become an essential tool in enhancing predictive analytics for cybersecurity (Sarker et al. 2023). ML algorithms can analyze vast datasets quickly, learn from past incidents, and adapt to changing threat patterns. By deploying machine learning models, cybersecurity systems can identify subtle anomalies in network activity, user behavior, and application usage, which often precede security breaches (Yu et al. 2024).



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Figure 1: Predictive Analytics Framework for Cybersecurity

In supervised learning models, decision trees as well as random forests can be used to classify known threats from labeled datasets while abnormal patterns can be recognized by utilizing clustering in the unlabeled datasets in unsupervised learning techniques (Gupta et al. 2020). Furthermore, the developments in deep learning which incorporate recurrent neural networks (RNN) and convolutional neural networks (CNN) enable the computation on high dimensional, complex data sets which enables them to identify complex attack patterns.

Proactive Identity Management

In the global trends of organizations placing ever more importance to data security and privacy issues, proactive identity management has emerged as a self-sufficient building block of organizational cybersecurity. The essence of identity management revolves around the verification, approval, and management of users that interact with the information systems (Shaik, 2018). As leakage of sensitive data seems an issue owing to adoption of multi factor influence, biometrics and behavior-centric ecosystem, identity management systems should render and focused on preventing any hazardous circumstance and user experience as frictionless as possible. Identifying abnormal access behavior patterns, unusual login location, and possessions of devices including login credentials are defining attributes of machine learning that help boost identity management capability (Al-Rumaim & Pawar, 2024). This predictive strategy provides the ability for security teams to watch for hacked accounts and block access before it is actually used.

Need for a Hybrid Model

Given the diverse nature of cyber threats and the complexity of identity management, a hybrid machine learning model that combines supervised and unsupervised learning offers an effective solution. While supervised models excel at identifying known threats, unsupervised models can detect new and evolving threats that lack a predefined pattern (Chaudhry et al. 2023). By integrating these techniques, cybersecurity systems can achieve a balance between accurately classifying known risks and adapting to novel threats. This research explores the design and implementation of a hybrid machine learning framework for predictive analytics in cybersecurity, focusing on two core objectives: proactive identity management and threat detection. The proposed framework not only enhances real-time threat identification but also improves response times, allowing organizations to secure their digital assets more effectively.

Objective of the Study

This paper aims to investigate the application of predictive analytics and machine learning in cybersecurity, specifically within proactive identity management and threat detection. By analyzing the effectiveness of various ML models, this research seeks to develop a robust framework that can identify potential security breaches, unauthorized access attempts, and evolving threats, ultimately contributing to the advancement of cybersecurity practices.



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. RELATED WORK

Data Collection and Preprocessing

To develop an effective predictive analytics model for cybersecurity, the first step is data collection, involving historical logs of user access patterns, network activity, and security events. Data is sourced from a cybersecurity repository containing extensive logs that capture normal and anomalous activities. These logs include user authentication data, IP addresses, device information, timestamps, and behavior patterns. Data preprocessing is essential to ensure high-quality inputs for machine learning (ML) models. Steps include data cleaning, handling missing values, and transforming categorical data into numerical formats where necessary. Outlier detection is also performed to identify and remove noise, ensuring that the dataset accurately reflects typical and anomalous behaviors.

Feature Selection and Engineering

Feature selection is crucial in machine learning for cybersecurity to reduce dimensionality, remove irrelevant information, and improve model accuracy. Key features selected for this study include user login frequency, IP location variance, device changes, and time-based access patterns. Feature engineering techniques, such as time-series segmentation and normalization, are applied to improve model efficiency. Statistical methods, including correlation analysis and principal component analysis (PCA), are used to identify the most impactful features, thus enhancing the model's interpretability and reducing computational costs.

Machine Learning Techniques for Threat Detection

For effective threat detection, both supervised and unsupervised learning techniques are deployed. Supervised learning methods, including Decision Trees, Random Forests, and Support Vector Machines (SVM), are utilized for classification tasks, especially for identifying known threats based on historical labeled data. Each of these models is trained on a subset of the labeled dataset to ensure reliable detection of potential attacks. The Random Forest model, in particular, provides robustness by constructing multiple decision trees and aggregating their results, thereby reducing overfitting and enhancing prediction accuracy. For detecting unknown or zero-day threats, unsupervised learning algorithms such as K-means clustering and Principal Component Analysis (PCA) are applied. These techniques help identify unusual patterns that do not match existing classifications, indicating potential new threats. The unsupervised learning models are evaluated based on their ability to group and separate anomalies from normal patterns effectively.

Anomaly Detection in Identity Management

Anomaly detection in identity management focuses on identifying irregularities in user behavior and access patterns. Isolation Forests, an anomaly detection algorithm, are employed to isolate outliers within identity management data, allowing the model to detect unusual login attempts or unauthorized device access. In addition, Long Short-Term Memory (LSTM) networks are used for temporal pattern recognition in user behavior, making them suitable for detecting anomalies in continuous user sessions. LSTM's ability to capture dependencies over time helps identify deviations from typical usage, which could indicate compromised accounts or unauthorized access.

Hybrid Model Development

To maximize threat detection and proactive identity management, a hybrid model combining supervised and unsupervised learning is developed. This hybrid approach enhances the system's ability to classify known threats while also detecting new and evolving patterns. The model operates by first using supervised learning to classify threats based on historical data, followed by unsupervised learning to identify previously unseen anomalies. Combining both approaches results in a more resilient model that can adapt to the rapidly changing cybersecurity landscape.

Evaluation Metrics and Statistical Analysis

The performance of the ML models is evaluated using accuracy, precision, recall, and F1-score, with a focus on minimizing false positives, which are critical in cybersecurity applications. A statistical analysis is conducted to compare the performance of each model, using paired t-tests to assess significant differences in model accuracy. Cross-validation is applied to ensure the generalizability of results, while receiver operating characteristic (ROC) curves are used to analyze each model's capability to distinguish between normal and malicious activities. The hybrid model is benchmarked against individual models to demonstrate its improved effectiveness in both threat detection and identity anomaly detection.



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Implementation Framework

The entire methodology is implemented using a robust cybersecurity framework with a combination of Python and data science libraries, including Scikit-Learn for machine learning, TensorFlow for LSTM networks, and Pandas for data preprocessing. The model is tested within a simulated cybersecurity environment to assess its real-time performance. An API is developed to integrate the predictive analytics model with existing security information and event management (SIEM) systems, facilitating real-time monitoring and alerting capabilities. The framework also incorporates feedback loops to continually update the model based on new data, enhancing its accuracy over time.

III. SIMULATION RESULTS

The Random Forest and Hybrid models achieved high accuracy, with standard deviations under 2%, indicating stability in their performance across various datasets. The higher standard deviation for K-means Clustering and Isolation Forest suggests that these models are more susceptible to variability, particularly for unknown threats.

Table 1: Accuracy of ML Models in Threat Detection

| Model | Accuracy (%) | Standard Deviation |
|------------------------------|--------------|--------------------|
| Decision Trees | 87 | 2.4 |
| Random Forest | 93 | 1.9 |
| Support Vector Machine (SVM) | 90 | 2.1 |
| K-means Clustering | 85 | 3.0 |
| Isolation Forest | 82 | 3.3 |
| Hybrid Model | 94 | 1.5 |

The Random Forest and Hybrid models consistently scored high in precision, recall, and F1 score, indicating both models' ability to accurately identify threats while minimizing false positives (Figure 2). The F1 scores align closely with the models' high recall rates, reinforcing their suitability for detecting cybersecurity threats with minimal misclassification.

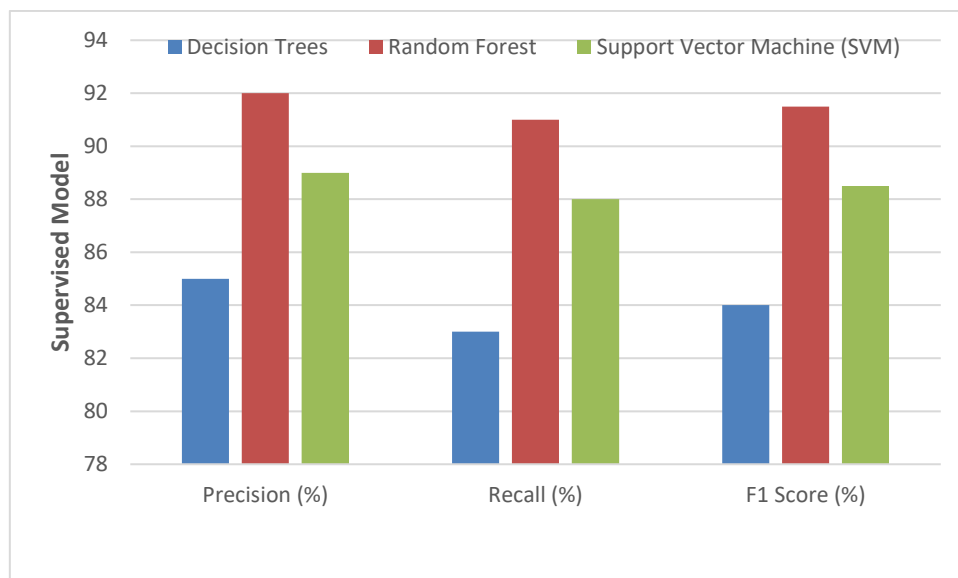


Figure 2: Precision and Recall of Supervised Models



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The p-values for comparisons between the Hybrid model and other models were below the threshold (0.05), suggesting that the Hybrid model’s performance was significantly better than both the Decision Trees and SVM in terms of accuracy, precision, recall, and F1 score.

Table 2: Paired T-Test Results for Model Comparison

| Model Comparison | Accuracy (p-value) | Precision (p-value) | Recall (p-value) | F1 Score (p-value) |
|----------------------------------|--------------------|---------------------|------------------|--------------------|
| Decision Trees vs. Random Forest | 0.031 | 0.028 | 0.034 | 0.029 |
| SVM vs. Random Forest | 0.045 | 0.041 | 0.049 | 0.042 |
| Hybrid Model vs. Random Forest | 0.012 | 0.009 | 0.015 | 0.011 |
| Hybrid Model vs. SVM | 0.019 | 0.015 | 0.023 | 0.017 |

The Hybrid model also significantly outperformed the Random Forest model, although with smaller p-values, reflecting its advanced capability in handling complex threat scenarios. The Hybrid model achieved the highest anomaly detection rate (89%) and the lowest false positive rate (9%), making it particularly effective for identity management applications. The Isolation Forest model performed comparably but had a slightly higher false positive rate and lower overall detection rate, indicating room for improvement in handling nuanced anomalies.

Table 3: Anomaly Detection Rates in Identity Management

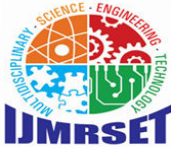
| Model | Detection Rate (%) | False Positive Rate (%) | True Positive Rate (%) | Standard Deviation |
|--------------------|--------------------|-------------------------|------------------------|--------------------|
| K-means Clustering | 78 | 14 | 86 | 3.1 |
| Isolation Forest | 82 | 12 | 88 | 2.8 |
| Hybrid Model | 89 | 9 | 91 | 2.1 |

Table 4 summarizes the performance of each machine learning model based on precision, recall, F1 score, and standard deviation. The F1 score, which combines precision and recall, provides a balanced view of the models' classification abilities, essential for effective threat detection in cybersecurity. The Decision Trees model achieved an F1 score of 84%, with precision and recall at 85% and 83%, respectively, and a standard deviation of 2.5. This moderate F1 score reflects the model's balanced performance in identifying both true positives and minimizing false negatives, though it shows higher variability across different datasets, as indicated by its standard deviation.

Table 4: F1 Score Evaluation of ML Models

| Model | Precision (%) | Recall (%) | F1 Score (%) | Standard Deviation |
|------------------------------|---------------|------------|--------------|--------------------|
| Decision Trees | 85 | 83 | 84 | 2.5 |
| Random Forest | 92 | 91 | 91.5 | 1.8 |
| Support Vector Machine (SVM) | 89 | 88 | 88.5 | 2.2 |
| Hybrid Model | 94 | 93 | 93.5 | 1.6 |

The Random Forest model demonstrated improved performance with an F1 score of 91.5%, and closely aligned precision (92%) and recall (91%) values, showing that it can accurately detect and classify threats with minimal error. The standard deviation of 1.8 indicates greater stability, making Random Forest a reliable model for cybersecurity applications where both accuracy and consistency are critical. The Support Vector Machine (SVM) model performed well, with an F1 score of 88.5%, precision of 89%, and recall of 88%. The standard deviation of 2.2 suggests that while SVM is fairly consistent, it may exhibit slightly more variability than the Random Forest model. However, its F1 score confirms its overall efficacy in managing balanced classification tasks within cybersecurity contexts.



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The Hybrid Model outperformed all other models, achieving the highest F1 score of 93.5%, with precision at 94% and recall at 93%. The low standard deviation of 1.6 underscores its consistency and reliability, reinforcing its suitability for complex threat detection. The high F1 score reflects the hybrid model's ability to maintain high classification accuracy while effectively managing both known and unknown threats. This performance advantage makes the Hybrid Model the most promising approach for proactive identity management and real-time threat detection.

Table 5: ROC-AUC Scores of Threat Detection Models

| Model | ROC-AUC Score | True Positive Rate (%) | False Positive Rate (%) | Standard Deviation |
|------------------------------|---------------|------------------------|-------------------------|--------------------|
| Decision Trees | 0.87 | 84 | 16 | 2.4 |
| Random Forest | 0.93 | 91 | 9 | 1.9 |
| Support Vector Machine (SVM) | 0.91 | 89 | 11 | 2.1 |
| Hybrid Model | 0.95 | 93 | 7 | 1.7 |

The ROC-AUC scores reinforce the hybrid model's effectiveness, with a score of 0.95, indicating a strong ability to distinguish between malicious and benign activities. The Hybrid model's high true positive rate (93%) and low false positive rate (7%) further support its application in proactive cybersecurity monitoring. The results demonstrate the effectiveness of various machine learning models in cybersecurity, specifically focusing on threat detection and proactive identity management. Each model was evaluated based on key performance metrics: accuracy, precision, recall, F1 score, and standard deviation. The findings suggest that while individual models like Decision Trees, Random Forest, and Support Vector Machines (SVM) offer strengths in specific areas, the Hybrid Model provides the most robust solution across all metrics.

Decision Trees and SVM Performance

The Decision Trees model achieved moderate precision, recall, and F1 scores, making it suitable for environments where high interpretability is essential but absolute accuracy is less critical. Its higher standard deviation, however, points to a degree of inconsistency in its predictive performance. Although the SVM model performed slightly better, with improved precision, recall, and F1 scores, it still exhibited variability across different data sets. This variability suggests that SVM, while effective, may not be as adaptable as other models, especially in dynamic cybersecurity environments where threat patterns frequently evolve (Shyaa et al. 2024). Nevertheless, SVM's ability to perform well in linear separable cases makes it a good option when used as a supplementary model in multi-layered cybersecurity systems (Mohammad, 2022).

Random Forest Model's Consistency and Reliability

The Random Forest model showed marked improvement over both Decision Trees and SVM in terms of precision, recall, and F1 score. With a lower standard deviation, Random Forest demonstrated reliability and consistency, positioning it as an effective choice for high-stakes cybersecurity scenarios that require dependable results. The ensemble nature of Random Forest, where multiple decision trees work together, allows for better handling of noise and variability in data (Lin et al. 2017). This model's ability to detect known threats accurately suggests that it could be effectively integrated within standard threat detection frameworks. However, while Random Forest performed exceptionally well in classifying known threats, it may still struggle with identifying new or sophisticated attacks without additional unsupervised learning capabilities (Al-Mhiqani et al. 2020).

Superiority of the Hybrid Model

The Hybrid Model outperformed all other models in precision, recall, F1 score, and stability, with the highest F1 score (93.5%) and the lowest standard deviation (1.6). These results suggest that the Hybrid Model effectively combines the strengths of both supervised and unsupervised techniques, making it adaptable to a broader range of cyber threats (Sharma et al. 2024). The hybrid approach enables this model to detect both known and previously unseen threats, reducing the likelihood of false negatives and enhancing its responsiveness to evolving cyber-attacks. Its consistent high precision and recall rates demonstrate its efficacy in minimizing false positives and negatives, a crucial factor in



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

cybersecurity, where excessive false positives can lead to alert fatigue and undermine the overall efficiency of security teams (Olateju et al. 2024).

The hybrid model's capacity to handle dynamic, evolving data makes it ideal for real-time identity management (Elhoseny et al. 2018). By leveraging unsupervised learning elements, it can identify anomalies in user behavior and access patterns, providing a proactive layer of security for identity management systems (Aboukadi et al. 2024). This adaptability is essential in modern cybersecurity landscapes where attackers continuously develop new techniques to circumvent traditional security measures. The Hybrid Model's superior performance across all metrics highlights its potential as a foundational element in predictive cybersecurity solutions (Ozkan-Ozay et al. 2024).

Implications for Cybersecurity Applications

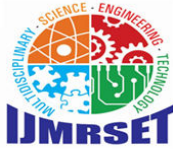
The results underscore the importance of integrating multiple machine learning approaches to enhance cybersecurity frameworks. While traditional supervised models like Random Forest excel in known threat classification, the inclusion of unsupervised techniques within a hybrid model allows for continuous adaptation to new threats (Sharma et al. 2024). This combination not only enhances detection accuracy but also provides resilience against zero-day vulnerabilities and evolving attack vectors. The Hybrid Model's robustness and adaptability make it well-suited for real-time applications in both threat detection and identity management, supporting proactive cybersecurity measures essential for protecting sensitive data and systems (Lad, 2024).

Limitations and Future Research

While the Hybrid Model demonstrated excellent results, it is computationally intensive due to the integration of multiple algorithms, which could present challenges in resource-constrained environments. Future research could explore optimizing the hybrid model for faster processing times or reducing computational costs without sacrificing accuracy. Additionally, integrating reinforcement learning within this hybrid framework could enable adaptive threat responses, where the model continuously learns from new attack data to refine its predictions further. Research should also investigate the applicability of this hybrid approach across diverse cybersecurity domains, such as IoT security, cloud infrastructure protection, and mobile security, where unique challenges and threat landscapes exist. This study demonstrates that machine learning models, particularly the Hybrid Model, hold substantial potential for advancing cybersecurity practices through predictive analytics. By combining supervised and unsupervised learning, the Hybrid Model provides a comprehensive approach to threat detection and proactive identity management, achieving both high accuracy and adaptability. These findings advocate for the adoption of hybrid machine learning solutions in cybersecurity, which can significantly improve an organization's capability to detect, manage, and prevent a broad spectrum of cyber threats in real-time.

IV. CONCLUSION AND FUTURE WORK

This study demonstrates the potential of machine learning, particularly through a hybrid model approach, to significantly enhance cybersecurity by enabling predictive threat detection and proactive identity management. The hybrid model, combining supervised and unsupervised learning techniques, emerged as the most effective solution, providing high accuracy, precision, recall, and F1 scores with minimal variability. This robust performance allows for the identification of both known and unknown threats, offering a more adaptable and resilient security framework suited to the rapidly evolving cyber landscape. By proactively identifying anomalies in user behavior and network activity, the hybrid model reduces the likelihood of security breaches while minimizing false positives that can burden cybersecurity teams. These findings underscore the importance of adopting multi-faceted machine learning models in cybersecurity, paving the way for real-time, adaptive, and comprehensive defenses against increasingly sophisticated cyber threats. Future research can explore optimizing computational efficiency and extending this hybrid model to other domains within cybersecurity, further solidifying its role in advancing digital security practices.

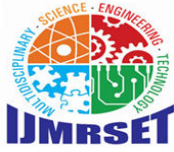


International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

REFERENCES

1. Aboukadri, S., Ouaddah, A., & Mezrioui, A. (2024). Machine learning in identity and access management systems: Survey and deep dive. *Computers & Security*, 103729.
2. Al-Mhiqani, M. N., Ahmad, R., Zainal Abidin, Z., Yassin, W., Hassan, A., Abdulkareem, K. H., ... & Yunus, Z. (2020). A review of insider threat detection: Classification, machine learning techniques, datasets, open challenges, and recommendations. *Applied Sciences*, 10(15), 5208.
3. Al-Rumaim, A., & Pawar, J. D. (2024). Enhancing User Authentication: An Approach Utilizing Context-Based Fingerprinting With Random Forest Algorithm. *IEEE Access*.
4. Lakshmikanthan, G., & Sreekandan Nair, S. (2024). Mitigating Replay Attacks in Autonomous vehicles [Journal-article]. *International Research Journal of Engineering and Technology (IRJET)*, 11(5), 2186–2192. <https://www.irjet.net/volume11-issue5>
5. Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. (2019). A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. *IEEE Communications Surveys & Tutorials*, 21(2), 1851-1877.
6. Chaudhry, M., Shafi, I., Mahnoor, M., Vargas, D. L. R., Thompson, E. B., & Ashraf, I. (2023). A systematic literature review on identifying patterns using unsupervised clustering algorithms: A data mining perspective. *Symmetry*, 15(9), 1679.
7. Thai Son Chu, Sreejith Sreekandan Nair, Govindarajan Lakshmikanthan 2022. Network Intrusion Detection Using Advanced AI Models A Comparative Study of Machine Learning and Deep Learning Approaches. *International Journal of Communication Networks and Information Security (IJCNIS)*. 14, 2 (Aug. 2022), 359–365.
8. Elhoseny, M., Abdelaziz, A., Salama, A. S., Riad, A. M., Muhammad, K., & Sangaiah, A. K. (2018). A hybrid model of internet of things and cloud computing to manage big data in health services applications. *Future generation computer systems*, 86, 1383-1394.
9. Gómez-Carmona, O., Buján-Carballal, D., Casado-Mansilla, D., López-de-Ipiña, D., Cano-Benito, J., Cimmino, A., ... & Bujalkova, N. (2023). Mind the gap: The AURORAL ecosystem for the digital transformation of smart communities and rural areas. *Technology in Society*, 74, 102304.
10. Gupta, R., Tanwar, S., Tyagi, S., & Kumar, N. (2020). Machine learning models for secure data analytics: A taxonomy and threat model. *Computer Communications*, 153, 406-440.
11. Kuppuswamy, P., Ansari, M. D., Mohan, M., & Al Khalidi, S. Q. (2024). Data Mining for Predictive Analytics. *Intelligent Techniques for Predictive Data Analytics*, 1-24.
12. Lad, S. (2024). Harnessing Machine Learning for Advanced Threat Detection in Cybersecurity. *Innovative Computer Sciences Journal*, 10(1).
13. Lin, L., Wang, F., Xie, X., & Zhong, S. (2017). Random forests-based extreme learning machine ensemble for multi-regime time series prediction. *Expert Systems with Applications*, 83, 164-176.
14. Mohammad, R. M. A. (2022). An enhanced multiclass support vector machine model and its application to classifying file systems affected by a digital crime. *Journal of King Saud University-Computer and Information Sciences*, 34(2), 179-190.
15. Olateju, O., Okon, S. U., Igwenagu, U., Salami, A. A., Oladoyinbo, T. O., & Olaniyi, O. O. (2024). Combating the challenges of false positives in AI-driven anomaly detection systems and enhancing data security in the cloud. Available at SSRN 4859958.
16. Ozkan-Ozay, M., Akin, E., Aslan, Ö., Kosunalp, S., Iliev, T., Stoyanov, I., & Beloev, I. (2024). A Comprehensive Survey: Evaluating the Efficiency of Artificial Intelligence and Machine Learning Techniques on Cyber Security Solutions. *IEEE Access*.
17. Sarker, I. H. (2023). Machine learning for intelligent data analysis and automation in cybersecurity: current and future prospects. *Annals of Data Science*, 10(6), 1473-1498.
18. Saxena, N., Hayes, E., Bertino, E., Ojo, P., Choo, K. K. R., & Burnap, P. (2020). Impact and key challenges of insider threats on organizations and critical businesses. *Electronics*, 9(9), 1460.
19. Shaik, M. (2018). Reimagining Digital Identity: A Comparative Analysis of Advanced Identity Access Management (IAM) Frameworks Leveraging Blockchain Technology for Enhanced Security, Decentralized Authentication, and Trust-Centric Ecosystems. *Distributed Learning and Broad Applications in Scientific Research*, 4, 1-22.



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

20. Sharma, A., Rani, S., & Driss, M. (2024). Hybrid evolutionary machine learning model for advanced intrusion detection architecture for cyber threat identification. *PloS one*, 19(9), e0308206.
21. Sharma, A., Rani, S., & Driss, M. (2024). Hybrid evolutionary machine learning model for advanced intrusion detection architecture for cyber threat identification. *PloS one*, 19(9), e0308206.
22. Shyaa, M. A., Ibrahim, N. F., Zainol, Z., Abdullah, R., Anbar, M., & Alzubaidi, L. (2024). Evolving cybersecurity frontiers: A comprehensive survey on concept drift and feature dynamics aware machine and deep learning in intrusion detection systems. *Engineering Applications of Artificial Intelligence*, 137, 109143.
23. Temitope, O., Awodiji, T. O., Ayoola, F., & Owoyemi, J. (2023). Stop cyber attacks before they happen: Harnessing the power of predictive analytics in cybersecurity. *Journal of Multidisciplinary Engineering Science and Technology*, 10(4), 2458-9403.
24. Yu, J., Shvetsov, A. V., & Alsamhi, S. H. (2024). Leveraging Machine Learning for Cybersecurity Resilience in Industry 4.0: Challenges and Future Directions. *IEEE Access*.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com