An Intelligent Machine Learning Framework for Cyber Attack Detection and Analysis

Ch. Satyanarayana Reddy[1], K. Pavani [2], N. Venkata Sravani[3]

#1. Assistant Professor in the Department of MCA, SRK Institute of Technology, Vijayawada.

#2. Assistant Professor & Head of Department of MCA, SRK Institute of Technology, Vijayawada.

#3. Student in the Department of MCA, SRK Institute of Technology, Vijayawada.

Abstract: Creating an automated and effective method for cyber-threat identification is one of the major issues in cybersecurity. This research introduces an artificial intelligence-based strategy that uses sophisticated artificial neural network techniques to identify cyber risks. The suggested system uses deep learning-based detection techniques to enhance cyber-threat identification and analysis while converting massive amounts of gathered security events into structured individual event profiles.In this work, several artificial neural network architectures, such as Fully Connected Neural Networks (FCNN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM), are integrated with event profiling techniques for data preprocessing to create an AI-SIEM (Artificial Intelligence–Security Information and Event Management) system. By differentiating between true-positive and false-positive security warnings, the suggested solution helps security analysts react to cyberthreats and malicious activity more swiftly and precisely.Large-scale security event logs are transformed into structured profiles via the event profiling technique so that deep learning models can interpret them effectively. While LSTM records long-term sequential behaviors and changing attack patterns within network operations, CNN is utilized to identify significant feature patterns from security event data. By discovering intricate connections between extracted characteristics, FCNN enhances classification and danger prediction tasks.Two benchmark datasets, NSLKDD and CICIDS2017, as well as two real-world security datasets gathered from operational environments, are used in experiments to assess the efficacy of the suggested AI-SIEM system. When compared to traditional machine learning methods, experimental results show that the suggested deep learning-based methodology offers superior cyber-threat detection capabilities, higher classification accuracy, and better discrimination between real threats and false warnings.

Index terms *- — Cybersecurity, Cyber-Threat Detection, Artificial Intelligence (AI), Security Information and Event Management (SIEM), AI-SIEM, Deep Learning, Artificial Neural Networks (ANN), Fully Connected Neural Network (FCNN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Event Profiling, Intrusion Detection, Security Event Analysis, Threat Classification, False Positive Reduction, NSL-KDD Dataset, CICIDS2017 Dataset, Network Security, Machine Learning, Security Analytics*

1. INTRODUCTION

Learning-based approaches for cyberattack detection have advanced considerably with the growth of Artificial Intelligence and machine learning technologies. Numerous studies have demonstrated promising outcomes in identifying malicious activities and protecting network infrastructures. However, defending IT systems against continuously evolving cyber threats and abnormal network behavior remains a major challenge. As cyberattacks become more sophisticated, organizations require reliable and intelligent security mechanisms capable of identifying both known and unknown threats effectively.

Traditionally, network intrusion detection mainly depends on two approaches: intrusion prevention systems (IPS) and Security Information and Event Management (SIEM) systems. IPS solutions commonly use signature-based techniques to inspect network traffic, protocols, and communication flows for suspicious activities. Once abnormal behavior is identified, security alerts are generated and forwarded to SIEM platforms for further analysis. SIEM systems collect, organize, and manage security logs and events generated from multiple network sources. Security analysts then investigate suspicious alerts by applying predefined rules, policies, thresholds, and event correlation methods. Despite these mechanisms, identifying intelligent cyberattacks accurately remains difficult because of the large volume of security data and high false alarm rates.

Recent developments in machine learning and deep learning have significantly improved automated intrusion detection systems. These learning-based approaches can identify hidden patterns and unusual behaviors within large-scale network data, enabling faster and more efficient cyber threat analysis. Machine learning solutions are particularly useful for detecting previously unknown attacks by learning behavioral patterns from historical security data. Compared with analyst-driven approaches that depend heavily on manually defined rules, machine learning techniques provide stronger adaptability for discovering new and evolving cyber threats.

However, existing learning-based intrusion detection methods still face several important limitations. One major challenge is the requirement for large-scale labeled datasets for training and evaluation. Obtaining accurately labeled security data is difficult and expensive, and many commercial SIEM systems do not maintain labeled datasets. Another limitation is the lack of generalized features within benchmark datasets, making it difficult to apply research models effectively in real-world environments. In addition, anomaly-based detection approaches often generate high false positive rates, increasing investigation workload and operational costs for security analysts. Attackers also continuously modify their behaviors to evade detection, reducing the long-term effectiveness of trained models. Furthermore, most security systems focus mainly on short-term event analysis and fail to capture long-range malicious behavioral patterns.

To overcome these challenges, the proposed AI-SIEM framework introduces a deep learning–based security analysis system capable of distinguishing genuine alerts from false alarms. The framework incorporates an event pattern extraction mechanism that correlates related security events and aggregates them into structured event profiles. This method enables efficient handling of large-scale security data while preserving important behavioral information. To reduce dimensional complexity and improve feature representation, the system applies the TF-IDF technique to transform collected security events into concise event profiles. Similarity calculations between TF-IDF event sets and predefined base points are then used to characterize normal and malicious activity patterns more effectively.

The generated event profiles are provided as input to multiple deep learning architectures including Fully Connected Neural Networks (FCNN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) models. These models improve intrusion classification capability while significantly reducing the number of false alerts presented to security analysts. The proposed framework was evaluated using benchmark datasets such as NSLKDD and CICIDS2017 along with real-world intrusion prevention system datasets. Experimental comparisons with traditional machine learning techniques including SVM, k-NN, RF, NB, and DT demonstrated improved performance in terms of accuracy, TPR, FPR, and F-measure.

2. LITERATURE SURVEY

a) Quantifying sensitivity and performance degradation of virtual machines using machine learning:

Virtualized data centers bring lot of benefits with respect to the reducing the high usage of physical hardware. But nowadays, as the usage of cloud infrastructures are rapidly increasing in all the fields to provide proper services on demand. In cloud data center, achieving efficient resource sharing between virtual machine and physical machines are very important. To achieve efficient resource sharing performance degradation of virtual machine and quantifying the sensitivity of virtual machine must be modeled, predicted correctly. In this work we use machine learning techniques like decision tree, K nearest neighbor and logistic regression to calculate the sensitivity of virtual machine. The dataset used for the experiment was collected using collected from open stack cloud environment. We execute two scenarios in this experiment to evaluate performance of the three mentioned classifiers based on precision, recall, sensitivity and specificity. We achieved good results using decision tree classifier with precision 88.8%, recall 80% and accuracy of 97.30%.

b) Efficient Outline Computation for Multi View Data Visualization on Big Data: Data representation in various perspectives with regard to visualization for managing massive amounts of data is a key component of big data analysis. In order to manage large amounts of data, the continuous parallel coordinate framework is an efficient data visualization tool that analyzes each attribute without updating or changing its values, without changing ongoing information structures, and presents data in a structural orientation based on attributes. Similarity Measure Centered with Multi Viewpoint (SMCMV) and related clustering algorithms have historically been used to show data in multi-attribute assessment based on multi-view data visualization procedures with various characteristics. Different sorts of characteristics are used to analyze and assess data based on various values in enormous amounts of multidimensional and large-scale data. Evaluating each characteristic independently is necessary for effective data processing in order to represent data in various factors with regard to the return of interest points in large-scale data. In order to assess data based on many criteria with regard to interest points from large amounts of data, we introduce and build a unique hybrid machine learning with sorting method in this research. The sorting algorithm consists of two fundamental processes in the data evolution process: the first is the evaluation of the sorted positional index, the second is the exploitation of the sorted positional index, and the third is the computational evaluation of selected and sequential data into a table. In order to compare the performance of current algorithms in terms of time, memory, and table index evaluation for sorted data, our developed method operates on real-world UCI repository data sets that are mostly utilized with sorting.

c) Advanced graphical-based security approach to handle hard AI problems based on visual security

The first consideration while analyzing human data from various web-based artificial intelligence (AI) applications is security. Using separate web apps to access data in different locations without protection is quite challenging. In order to utilize services safely in an external environment, a variety of security-related techniques were developed; nonetheless, they have some limits when it comes to safeguarding data from external attackers (hackers). In order to protect AI-related web-oriented applications from external attackers, we provide in this work an innovative and sophisticated security approach. In order to provide security services in our suggested method, we adhere to the fundamental characteristics of Captcha as a graphical password. We discuss pushing, pass-on, and guessing attacks in online applications using randomly selected Captcha passwords to access web services using Captcha graphical passwords. When compared to current security methods, our testing results demonstrate effective security relations in terms of Captcha creation, duration, and other factors found in online security apps.

d) Analysis of Different Pattern Evaluation Procedures for Big Data Visualization in Data Analysis:

Due to the rapidly increasing number and complexity of data, data visualization is the primary focus of big data analysis for processing and evaluating multivariate data. In general, data visualization may address three primary issues: 1. Evaluation of organized and unstructured patterns in large data research. 2. In data-indexed large data analysis, reduce the characteristics. 3. Rearranging properties in data

storage based on parallel indexes. Therefore, in this work, we examine several approaches for resolving the aforementioned three issues with the viability of each client need in big data analysis for visualization in real-time data stream extraction based on indexed data arrangement. In addition to evaluating quantitative exert review in real-time setups for processing data visualization, we have examined several prototypes in accessible parallel coordinates. In practice sessions, demonstrate various data visualization analysis findings for scientific and large-scale data generated by numerical simulation and analyzed in big data presentations.

e) Surveillance detection in high bandwidth environments:

The surveillance detection methods for enclave environments (ESD) and peering center environments (PSD) in System Detection are described in this study, and each method is assessed using data collected from two distinct network settings. PSD is assessed over 5 hours of tcpdump packet traces (110 million packets) collected from a peering center, whereas ESD is assessed over 74 hours of packet traces (344 million packets) from a big enclave. Although the systems may be operated in real-time, both surveillance detection modules were run offline over the audit data to provide surveillance detection warnings. Our findings demonstrate that both PSD and ESD may be adjusted to lower the number of warnings while reliably identifying large amounts of surveillance activity, including scattered and long-lived scans. Additionally, many of the behaviors found by ESD and PSD may be invisible to current IDS technologies.

3. METHODOLOGY

i) Proposed Work:

In this paper, the author presents an intelligent cyber-threat detection framework based on Artificial Intelligence–Security Information and Event Management (AI-SIEM) technology. The proposed system combines multiple deep learning algorithms such as Fully Connected Neural Networks (FCNN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) to improve the detection of cyber threats and malicious activities within enterprise networks. The framework operates using event profiling techniques, where security events and attack signatures are analyzed to identify abnormal behavior patterns and distinguish genuine threats from false alerts. The proposed AI-SIEM system processes large-scale security event data and applies deep learning methods to enhance automated intrusion detection and threat analysis. To evaluate the effectiveness of the proposed approach, the author compares its performance with several conventional machine learning algorithms including Support Vector Machines (SVM), Decision Tree, Random Forest, k-Nearest Neighbors (KNN), and Naive Bayes (Naïve Bayes). Experimental results indicate that deep learning–based methods achieve better detection accuracy and improved performance compared with traditional machine learning approaches. In this implementation, CNN and LSTM algorithms are utilized for cyber-threat detection and security event classification. CNN is applied to extract important features and spatial patterns from security event data, while LSTM is used to capture long-term sequential dependencies and evolving attack behaviors, thereby improving threat detection capability and reducing false-positive alerts.

ii) System Architecture:

The system architecture of the proposed model is designed as a two-phase deep learning framework for efficient cyber attack detection and attribution in ICS/IIoT environments. Initially, the input data is collected from the SWaT dataset and passed through a data preprocessing stage, where missing values are handled and normalization (Min-Max scaling) is applied. This ensures that the data is clean and suitable for model training. The preprocessed data is then fed into an autoencoder model, which performs deep feature extraction by learning compressed representations of the input data.

In the next stage, the extracted features are reduced using Principal Component Analysis (PCA) to eliminate redundant information and improve computational efficiency. These optimized features are then passed to a Decision Tree classifier, which performs the first phase of attack detection by classifying the data into normal or attack categories. In the second phase, the detected attack data is further

processed by a Deep Neural Network (DNN), which performs detailed classification and attribution of different attack types. Finally, the system outputs the predicted attack category along with performance evaluation metrics such as accuracy, precision, recall, and F1-score, ensuring a reliable and scalable cyber attack detection framework.

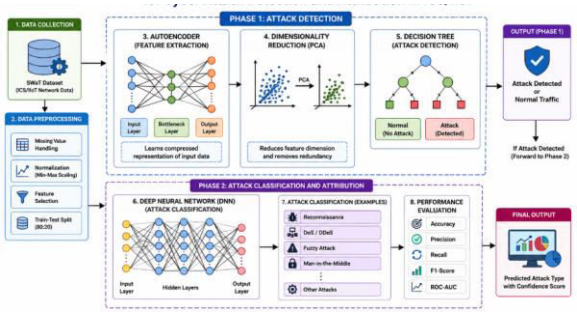


Fig1 proposed architecture

iii) Modules:

1. Dataset Upload Module

This module is responsible for uploading the SWaT dataset into the system. It reads the dataset and identifies different attributes and attack labels present in the data for further processing.

2. Data Preprocessing Module

In this module, missing values are handled and replaced appropriately. Min-Max normalization is applied to scale the data, and the dataset is split into training (80%) and testing (20%) sets to prepare it for model training.

3. Autoencoder Feature Extraction Module

This module uses an autoencoder deep learning model to extract meaningful and compressed features from the input data. It helps in handling imbalanced datasets and improves the quality of features.

4. PCA Dimensionality Reduction Module

The extracted features are passed through PCA to reduce dimensionality. This removes redundant data and improves computational efficiency while maintaining important information.

5. Decision Tree Detection Module

This module uses a Decision Tree classifier to detect whether the given input data represents normal activity or a cyber attack. It acts as the first phase of attack detection.

6. Deep Neural Network (DNN) Classification Module

In this module, a DNN model is trained using the detected attack data. It performs detailed classification and identifies the specific type of cyber attack.

7. Attack Prediction Module

This module takes new or unseen input data and predicts whether it is an attack or normal traffic. If it is an attack, the system also provides the attack category.

8. Performance Evaluation Module

This module evaluates the system performance using metrics such as accuracy, precision, recall, and F1-score. It also generates comparison graphs and tables for analysis.

iv) Algorithms:

1. K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a supervised machine learning algorithm used for classification and prediction tasks. The algorithm classifies a new data instance by identifying the K closest training samples based on a distance metric such as Euclidean distance. The class that appears most frequently among the nearest neighbors is assigned as the predicted class. In the proposed cyber-threat detection system, KNN is used to classify security events by comparing them with previously known attack patterns. KNN achieved an accuracy of approximately 71% during experimental evaluation.

2. Naïve Bayes (NB)

Naïve Bayes is a probabilistic classification algorithm based on Bayes' Theorem. It assumes that all features are independent of each other, which simplifies the classification process and reduces computational complexity. The algorithm calculates the probability of each class and assigns the input to the class with the highest probability. In this system, Naïve Bayes is used to classify security events into threat and non-threat categories. The model achieved an accuracy of approximately 67%.

3. Decision Tree (DT)

Decision Tree is a supervised learning algorithm that represents decisions using a tree-like structure consisting of nodes and branches. Each internal node represents a feature test, while leaf nodes represent classification outcomes. The algorithm recursively splits the dataset into smaller subsets until optimal decision boundaries are created. In the proposed framework, Decision Tree is used to classify cyber threats based on event profile characteristics. Experimental results showed an accuracy of approximately 53%.

4. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a powerful supervised learning algorithm that identifies the optimal hyperplane separating different classes within a high-dimensional feature space. The objective of SVM is to maximize the margin between classes, resulting in better generalization performance. In this cyber-threat detection system, SVM is used to distinguish malicious activities from normal events using extracted security features. The model achieved an accuracy of approximately 84%.

5. Random Forest (RF)

Random Forest is an ensemble learning algorithm that combines multiple decision trees to improve classification accuracy and reduce overfitting. Each tree is trained using randomly selected subsets of data and features, and the final prediction is obtained through majority voting. In the proposed system, Random Forest analyzes security event profiles and classifies them into different threat categories. The algorithm achieved an accuracy of approximately 78%, demonstrating reliable performance in cyber-threat detection.

6. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a specialized Recurrent Neural Network (RNN) designed to capture long-term dependencies and sequential patterns in data. LSTM uses memory cells and gating mechanisms to retain important information over extended periods. In the AI-SIEM framework, LSTM is used to analyze sequences of security events and identify evolving attack behaviors over time. The model effectively captures temporal relationships within network activities and achieved an accuracy of approximately 94%.

7. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is a deep learning architecture capable of automatically extracting important features from structured input data. Although traditionally used for image processing, CNN can also identify meaningful patterns within security event profiles. In the proposed system, CNN extracts high-level features from event vectors generated during preprocessing and performs threat classification with high precision. Among all evaluated algorithms, CNN achieved the highest performance with an accuracy of approximately 99%, making it the best-performing model for cyber-threat detection.

Performance Analysis

The accuracy comparison graph shows that deep learning models significantly outperform traditional machine learning algorithms. CNN achieved the highest accuracy of nearly 99%, followed by LSTM with approximately 94%. Traditional machine learning algorithms such as SVM, Random Forest, KNN, Naïve Bayes, and Decision Tree produced comparatively lower accuracy values. Therefore, CNN was selected as the primary model for deployment due to its superior classification performance and ability to identify complex cyber-threat patterns effectively.

4. EXPERIMENTAL RESULTS

The proposed two-phase ensemble deep learning framework was evaluated using the SWaT dataset, which contains both normal and attack scenarios in ICS environments. The dataset was preprocessed using Min-Max normalization and split into 80% training and 20% testing data. Performance evaluation was carried out using standard metrics such as accuracy, precision, recall, and F1-score. The first phase, which combines autoencoder-based feature extraction, PCA, and Decision Tree, effectively identified anomalous patterns, while the second phase using DNN provided accurate classification of attack types.

The experimental results demonstrate that the proposed model outperforms traditional machine learning approaches such as KNN, SVM, and Logistic Regression. The system achieved higher accuracy and better recall, indicating improved detection of both known and unknown attacks. Additionally, the false alarm rate was significantly reduced due to the two-stage architecture. The results confirm that the proposed framework is efficient, scalable, and suitable for real-time cyber attack detection and attribution in ICS/IIoT environments.

Accuracy: A test's accuracy is its capacity to distinguish healthy from ill cases. Find the percentage of instances with genuine positives and negatives to assess test accuracy.

Accuracy = TP + TN /(TP + TN + FP + FN)

$$Accuracy = \frac{(TN + TP)}{T}$$

Precision: Classification accuracy or positive cases constitute precision. The formula for accuracy is:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\Pr e\,cision = \frac{TP}{(TP + FP)}$$

Recall: A model's recall measures its ability to recognize all appropriate machine learning class instances. The ratio of accurately predicted positive observations to total positives indicates a model's class instance detection skill.

$$Recall = \frac{TP}{(FN + TP)}$$

mAP: Mean Average Precision ranks quality. It considers the number and order of relevant ideas. Calculating MAP at K uses the arithmetic mean of each user or query's Average Precision (AP).

$$\boldsymbol{mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k}$$

$$\boldsymbol{AP_k = \text{ the AP of class } k}$$

$$\boldsymbol{n = \text{ the number of classes}}$$

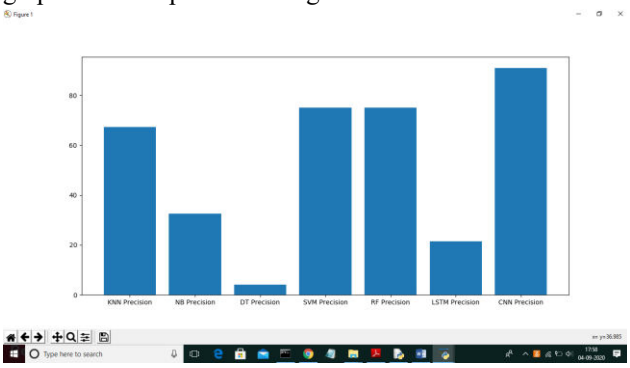
F1-Score: A high F1 score suggests an accurate machine learning model. Integrating recall and precision improves model correctness. Accuracy measures how often a model predicts a dataset correctly.

$$F1 = 2 \cdot \frac{(Recall \cdot \Pr e\,cision)}{(Recall + \Pr e\,cision)}$$

In the above graph, the X-axis represents the names of the implemented machine learning and deep learning algorithms, while the Y-axis represents the corresponding accuracy values achieved by each model during cyber-threat detection and classification. The graph provides a visual comparison of the performance of algorithms such as SVM, KNN, Random Forest, Naïve Bayes, Decision Tree, Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN).

From the graph, it can be observed that the deep learning models, particularly CNN and LSTM, achieve significantly higher accuracy compared with the conventional machine learning algorithms. Among all implemented models, CNN provides the best overall performance with the highest classification accuracy, demonstrating its effectiveness in extracting important feature patterns from security event data. LSTM also performs well by learning long-term sequential dependencies from the event profiles.

To further analyze model performance using additional evaluation metrics, the user can now click on the “Precision Comparison Graph” button to generate the corresponding precision comparison graph for all implemented algorithms.

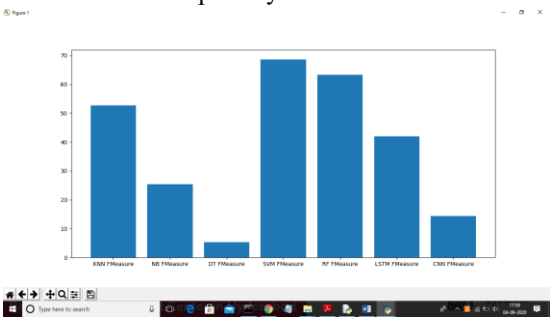


In the above graph, it can be observed that the Convolutional Neural Networks (CNN) model achieves the best precision performance compared with the other implemented machine learning and deep learning algorithms. The higher precision value indicates that CNN effectively minimizes false-positive alerts while accurately identifying malicious security events within the cyber-threat detection system.

To further evaluate the detection capability of all algorithms, the user can now click on the “Recall Comparison Graph” button. This module generates a graphical comparison of recall values for all implemented algorithms including SVM, KNN, Random Forest, Naïve Bayes, Decision Tree, Long Short-Term Memory (LSTM), and CNN. Recall measures the ability of each model to correctly identify actual cyber threats from the dataset and indicates how effectively the system detects malicious activities without missing attack events

In the above graph, the Long Short-Term Memory (LSTM) model demonstrates better recall performance compared with the other implemented algorithms. The higher recall value indicates that the LSTM model effectively identifies a larger number of actual cyber-threat events from the dataset while minimizing missed attack detections. This performance improvement is mainly due to the ability of LSTM networks to capture long-term sequential dependencies and evolving behavioral patterns within security event data.

To further analyze the overall classification effectiveness of all implemented algorithms, the user can now click on the “FMeasure Comparison Graph” button. This module generates a graphical comparison of F-Measure values for all machine learning and deep learning models including SVM, KNN, Random Forest, Naïve Bayes, Decision Tree, CNN, and LSTM. The F-Measure combines both precision and recall into a single evaluation metric, providing a balanced measure of classification performance and cyber-threat detection capability



From all the comparison graphs, it can be clearly observed that the deep learning models Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) perform significantly better than the conventional machine learning algorithms in cyber-threat detection and classification. The evaluation

metrics including accuracy, precision, recall, and F-Measure demonstrate that both CNN and LSTM achieve superior prediction performance while effectively identifying malicious security events from the dataset.

Among the implemented models, CNN provides the highest classification accuracy and precision, indicating its strong capability in extracting important feature patterns and minimizing false-positive alerts. On the other hand, LSTM achieves better recall performance by effectively capturing long-term sequential dependencies and evolving attack behaviors from historical security event data. The combined performance of CNN and LSTM confirms that deep learning–based approaches are more efficient and reliable for intelligent cyber-threat detection compared with traditional machine learning methods such as SVM, KNN, Random Forest, Naïve Bayes, and Decision Tree.

Overall, the experimental results demonstrate that the proposed AI-SIEM framework using CNN and LSTM provides improved intrusion detection capability, enhanced classification accuracy, reduced false alerts, and better adaptability for detecting continuously evolving cyber threats in large-scale network environments..

5. CONCLUSION

Using artificial neural networks and event profiling techniques, this study suggests an intelligent Artificial Intelligence–based Security Information and Event Management (AI-SIEM) system for sophisticated cyber-threat identification. The deployment of deep learning-based detection models to improve cyberattack recognition capabilities and the conversion of extensive security event data into condensed event profiles are the main contributions of the proposed work. By breaking down massive security datasets into structured profiles, the system enables effective handling and analysis of long-term security event histories.The suggested AI-SIEM architecture helps security analysts respond to urgent security warnings more effectively and efficiently by analyzing long-term behavioral trends across network events. Furthermore, the method reduces false-positive alarms, enabling the prompt identification of actual cyberthreats dispersed throughout large security.

Future Scope

Future research will focus on addressing the continuously evolving nature of cyberattacks by improving early threat prediction capability through multiple deep learning approaches. Special attention will be given to discovering long-term behavioral patterns within historical security event data to improve adaptive threat detection performance.

In addition, efforts will be made to improve the quality and precision of labeled datasets required for supervised learning models. Security Operations Center (SOC) analysts will contribute by manually labeling raw security events over extended periods to construct high-quality learning datasets. These improvements are expected to enhance the reliability, scalability, and prediction accuracy of future AI-SIEM systems for intelligent cyber-security applications.

REFERENCES

[1] S. Naseer, Y. Saleem, S. Khalid, M. K. Bashir, J. Han, M. M. Iqbal, and K. Han, ‘‘Enhanced network anomaly detection based on deep neural networks,’’ IEEE Access, vol. 6, pp. 48231–48246, 2018.

[2] B.-C. Zhang, G.-Y. Hu, Z.-J. Zhou, Y.-M. Zhang, P.-L. Qiao, and L.-L. Chang, ‘‘Network intrusion detection based on directed acyclic graph and belief rule base,’’ Electron. Telecommun. Res. Inst. J., vol. 39, no. 4, pp. 592–604, Aug. 2017.

[3] W. Wang, Y. Sheng, and J. Wang, ‘‘HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection,’’ IEEE Access, vol. 6, pp. 1792–1806, 2018.

[4] M. K. Hussein, N. Bin Zainal, and A. N. Jaber, ‘‘Data security analysis for DDoS defense of cloud based networks,’’ in Proc. IEEE Student Conf. Res. Develop. (SCOReD), Kuala Lumpur, Malaysia, Dec. 2015, pp. 305–310.

[5] S. S. Sekharan and K. Kandasamy, ‘‘Profiling SIEM tools and correlation engines for security analytics,’’ in Proc. Int. Conf. Wireless Commun.,

Signal Process. Netw. (WiSPNET), Mar. 2017, pp. 717–721.

[6] N. Hubballi and V. Suryanarayanan, ‘‘False alarm minimization techniques in signature-based intrusion detection systems: A survey,’’ Comput. Commun., vol. 49, p. 1Â17, Aug. 2014.

[7] A. Naser, M. A. Majid, M. F. Zolkipli, and S. Anwar, ‘‘Trusting cloud computing for personal files,’’ in Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC), Busan, South Korea, Oct. 2014, pp. 488–489.

[8] Y. Shen, E. Mariconti, P. A. Vervier, and G. Stringhini, ‘‘Tiresias: Predicting security events through deep learning,’’ in Proc. ACM CCS, Toronto, ON, Canada, Oct. 2018, pp. 592–605.

[9] K. Soska and N. Christin, ‘‘Automatically detecting vulnerable Websites before they turn malicious,’’ in Proc. USENIX Secur. Symp., San Diego, CA, USA, 2014, pp. 625–640.

[10] K. Veeramachaneni, I. Arnaldo, V. Korrapati, C. Bassias, and K. Li, ‘‘AI2 : Training a big data machine to defend,’’ in Proc. IEEE BigDataSecurity HPSC IDS, New York, NY, USA, Apr. 2016, pp. 49–54.

[11] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, ‘‘A detailed analysis of the KDD cup 99 data set,’’ in Proc. 2nd IEEE Symp. Comput. Intell. Secur. Defense Appl., Jul. 2009, pp. 53–58.

[12] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, ‘‘Toward generating a new intrusion detection dataset and intrusion traffic characterization,’’ in Proc. Int. Conf. Inf. Syst. Secur. Privacy (ICISSP), Jan. 2018, pp. 108–116.

[13] J. Song, H. Takakura, and Y. Okabe. (2006). Description of Kyoto University Benchmark Data. [Online]. Available: http://www.takakura.com/Kyoto_data/BenchmarkData-Description-v5.pdf

[14] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, ‘‘A deep learning approach to network intrusion detection,’’ IEEE Trans. Emerg. Topics Comput. Intell., vol. 2, no. 1, pp. 41–50, Feb. 2018.

[15] R. Vinayakumar, M. Alazab, K. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, ‘‘Deep learning approach for intelligent intrusion detection system,’’ IEEE Access, vol. 7, pp. 41525–41550, 2019.

Author Profiles

Mr. Ch. Satyanarayana Reddy Completed his MCA, He also a web developer and python developer, currently working has an Assistant Professor in the department of MCA at SRK Institute of Technology, Enikepadu, NTR District. His area of interest includes Artificial Intelligence and Machine Learning.

Ms. K. Pavani Working as Assistant & Head of Department of MCA, in SRK Institute of technology in Vijayawada. She done with MCA, M. Tech in Computer Science. Her area of interest includes Machine Learning with Python, AI and DBMS.

Ms. N. Venkata Sravani is MCA Student in the Department of Computer Applications at SRK Institute of Technology, Enikepadu, Vijayawada, NTR District. She has Completed Degree in B.Sc. (statistics) from Sri Harshini degree and PG college Ongole. Her area of interest are Deep Learning and Machine Learning.