Machine Learning for Human Intelligence
2023, Vol. 1, Issue 1

Research Article

A Multi-Model Machine Learning Framework for Robust Anomaly-Based Threat Detection

Sehrish Raza[1,⋆] and Aleena Muzammil[2]

[1,2]Institute of Computer Science, Women University, Multan, 60000, Pakistan.; Email: sehrish.raza@wum.edu.pk , aleena.muzammil@wum.edu.pk
⋆Corresponding author: Sehrish Raza (sehrish.raza@wum.edu.pk)

Article History

Academic Editor:
Dr. Muhammad Nabeel Asghar
Submitted: December 05, 2022
Revised: February 21, 2023
Accepted: March 1, 2023

Keywords:
Cybersecurity; AI-driven Solutions; Ensemble Learning; Security Frameworks; Digital Infrastructure

Abstract

Cybersecurity is essential in the contemporary, rapidly evolving digital environment. As AI-driven solutions become essential for protecting businesses, the increasing number and complexity of cyber threats often surpass traditional security measures, leading to considerable financial and reputational risks. This paper presents an enhanced cybersecurity framework utilizing an ensemble learning model that integrates machine learning and deep learning methods to tackle this difficulty. We assessed and contrasted various classifiers utilizing the HIKARI-2021 dataset from Kaggle, including Random Forest, Decision Tree, Gaussian Naive Bayes, K-Nearest Neighbors, Logistic Regression, Multi-Layer Perceptron, and Convolutional Neural Network. By merging these models using an ensemble technique, we harnessed their complementary capabilities, attaining a remarkable 96.32% accuracy—an appreciable enhancement above individual models. In addition to precision, the ensemble method improves adaptability, facilitating more dynamic and robust security frameworks. Our research underscores the effectiveness of ensemble learning in cybersecurity, illustrating its capacity to strengthen digital companies against emerging threats. This research provides practical solutions and facilitates further studies on AI-integrated cybersecurity, promoting innovation and a resilient global digital infrastructure.

1 Introduction

In the contemporary digital landscape, traditional security measures have failed to adapt to the advancements driven by modern technology and the ever developing nature of cyber threats [1]. Consequently, cybersecurity is a paramount concern for organizations today. The likelihood of cyberattacks escalates as an increasing number of corporations, governments, and individuals utilize digital devices and the internet, resulting in a greater volume of sensitive data stored online that attracts hackers and other cybercriminals. A disparity between emerging risks and unchanging security measures could jeopardize the company's financial stability and reputation [3]. Nonetheless, numerous studies have concentrated on cybersecurity tactics utilizing machine learning and deep learning methodologies [4]. These methods have demonstrated a degree of resemblance or efficacy in identifying and mitigating cyber risks [5]. The aim is to fortify digital enterprises against emerging cyber threats by providing dependable solutions, while improving predictive capabilities through classifiers and the creation of advanced, flexible security frameworks that ensure secure digital infrastructure via AI-driven methods.

Traditional security systems rely on established rules and predetermined patterns, rendering them ineffective against unrecognized or novel attacks. Artificial intelligence and machine learning possess the ability to learn and adapt in order to identify anomalous patterns and detect potential cyber-attacks. Nonetheless, employing a singular model occasionally fails to yield optimal results. Ensemble learning utilizes many models that amalgamate their strengths to enhance system security through protective solutions. Consequently, it is presently proposed that the ensemble learning method, which integrates many classifiers to achieve optimal performance by capitalizing on their individual strengths, is highly beneficial in enhancing cybersecurity. A cognitive cybersecurity technique employing an ensemble-meta-classifier that utilizes preprocessed security indications to develop a robust dataset [7]. The research was helpful in addressing data incompleteness and variety [8]. High-quality preprocessing introduces potential bias. The absence of rapid flexibility in model variance relative to the environment hinders model generalizability. Cyber dangers are ever developing; hence, models must be continuously updated to preserve their efficacy [9]. Cyberattacks are increasingly complex and more difficult to detect. Current methodologies can generate excessive false alarms or fail to detect novel threats. These limitations underscore the challenges and the imperative for continual enhancements of models employed in cybersecurity [11].

This paper identifies the shortcomings of current studies and presents an ensemble learning model that integrates deep learning and machine learning techniques. The models will incorporate implementations that utilize the optimal features of several classifiers to enhance predicted accuracy, while also developing sophisticated security frameworks that foster breakthroughs in cybersecurity. The project endeavors to develop an innovative cybersecurity system by integrating several machine learning and deep learning models to enhance the accuracy of cyber threat detection, tailored to effectively combat the evolving threat landscape.

The HIKARI-2021 dataset from Kaggle will be utilized, comprising cybersecurity incidents and associated attributes, including flow time, source and destination IP addresses, and labeled identifications for diverse categories of cyberattacks. A diverse array of classifiers was employed to train on this dataset, including random forest, decision tree, Gaussian Naive Bayes, K-Nearest Neighbors, logistic regression, multi-layer perceptron, and convolutional neural networks. The efficacy of each model was assessed utilizing established metrics such as recall, accuracy, and precision. This work employed an ensemble learning methodology after meticulously evaluating the advantages and disadvantages of the individual models. This methodology yielded a synthesis of predictions derived from the classifiers' strengths, enhancing the resilience of the authentication process and attaining an overall accuracy of 96.32%. This research seeks to address the limits of individual models and enhance their prediction capabilities. The contributions of this study are as follows:

1. Proposed a dependable and effective intrusion detection system utilizing ensemble learning techniques derived from machine learning and deep learning models.

2. Formulated a novel methodology by tackling class imbalance and preparing the HIKARI-2021 dataset. This entails transforming categorical variables into numerical format by label encoding and one-hot encoding methods to ensure data integrity and equilibrium, hence improving the precision of classification models.

3. Executed and assessed various ensemble models, including Random Forest, Decision Tree, Gaussian Naive Bayes, K-Nearest Neighbors, Logistic Regression, Multi-Layer Perceptron, and Convolutional Neural Network, illustrating the efficacy of ensemble methodologies in attaining elevated accuracy and resilience.

4. The performance assessment utilizes metrics such as recall, F1-score, accuracy, and precision, underscoring the effectiveness of the ensemble models in improving classification accuracy and dependability.

The remainder of the paper is structured as follows: Section 2 reviews the pertinent literature related to the proposed methodology; Section 3 delineates the methodology, encompassing the dataset and the applications of deep learning and machine learning models; Section 4 presents the results and discussion; and Section 5 concludes the study.

2 Related Work

Numerous studies have concentrated on enhancing fraud detection systems; nonetheless, they predominantly encounter three significant limitations: dependence on antiquated and restricted data sets, suboptimal detection rates accompanied by elevated false positives, and incapacity to adjust to advancing cyber threats. The suggested ensemble method directly tackles these deficiencies by integrating machine learning and deep learning techniques. This integrated approach not only addresses the inherent limits of singular models but also creates a more robust security framework against contemporary cybercrime.

The Ensemble Classifier Algorithm with a Stacking Process (ECASP) for bot detection illustrates that even sophisticated machine learning methodologies encounter inherent challenges. Although it attains 94.08% accuracy via optimal feature selection [9], its reliance on predetermined features renders it susceptible to evolving botnet architectures that necessitate ongoing updates. The XGBoost-based intrusion detection methodology attains 92.86% accuracy on the CICIDS2017 dataset; nevertheless, this achievement is hampered by the dataset's inadequate representation of developing real-world threats, hence diminishing its efficacy against innovative attack patterns. These examples demonstrate the challenges current solutions face regarding adaptability and generalization—limitations that our ensemble method explicitly mitigates through dynamic feature adaption and extensive threat coverage in the HIKARI-2021 dataset.

Additionally, the exploration of machine-learning-based detection of network intrusions utilizes three datasets: UNSW-NB15, NSL-KDD, and BoT-IoT [11]. Feature selection is conducted utilizing information gain and Pearson correlation, succeeded by classification employing diverse models. Stacked LightGBM and random forest attain the maximum predictive accuracy across all datasets; nevertheless, their reliance on certain specified datasets may not accurately represent developing zero-day assaults and growing cyber threats, thus undermining real-world applicability. A separate study examines the utilization of ensemble machine learning methods for network intrusion detection employing the NSL-KDD dataset [12]. It employed voting, bagging, and boosting techniques specifically utilized in the assessment of the voting classifier, random forest, and AdaBoost algorithms to enhance the efficacy of anomaly detection. The work is significantly constrained by its dependence on the NSL-KDD dataset, which may not sufficiently reflect contemporary cyber threats, hence limiting the generalizability of the suggested ensemble approaches in real-world attacks.

A method for detecting malicious URLs based on Cyber Threat Intelligence (CTI) via a two-stage ensemble learning approach was proposed [13]. To enhance detection accuracy, CTI features are derived from Google searches and data. Reclassification is performed with a Random Forest (RF) technique, succeeded by the application of a Multilayer Perceptron (MLP) for ultimate decision-making. The model demonstrates a 7.8% increase in overall accuracy and a 6.7% reduction in false positives compared to conventional URL-based detection methods. Nevertheless, it depends on external sources, such Google searches and Whois data, where delays in feature extraction may constrain real-time detection efficacy. This paper [14] introduces an ensemble learning strategy for detecting Cross-Site Scripting (XSS) threats through the integration of different Bayesian networks, using domain knowledge and threat information. A ranking methodology that evaluates nodes based on their influence on the output would enhance model interpretability for users. Direct comparisons demonstrate the model's efficacy on an actual dataset, particularly excelling in the identification of large-scale attacks. The technique relies on accurate domain expertise and threat intelligence, which constrains adaptability to entirely novel or developing assault patterns.

The HIKARI-2021 dataset serves as a superior baseline for network intrusion detection systems by rectifying significant deficiencies in current datasets. While NSL-KDD achieves an accuracy of 89% and CI-CIDS2017 reaches 92.86%, our model attains 96.32% on HIKARI, representing a 4-7% enhancement in the detection of contemporary assaults. In contrast to conventional datasets, HIKARI integrates both authentic and synthetic encrypted attack traffic, encompassing DDoS, port scanning, and malware, hence offering enhanced coverage compared to BoT-IoT's more limited IoT emphasis [15]. This distinctive composition directly addresses the shortcomings of prior methodologies by providing a robust framework for the development and evaluation of machine learning and deep learning models that can manage both established and novel threats. While conventional signature-based detection falters against innovative threats and anomaly detection methods are hindered by false positives, our ensemble strategy utilizes HIKARI's comprehensive coverage to address these enduring issues. By integrating various machine learning models (random forest, decision tree, Gaussian Naive Bayes) [18] and deep learning architectures (CNNs, MLPs), we illustrate how HIKARI facilitates superior intrusion detection relative to traditional datasets.

The efficacy of the IoT Intrusion Detection System is enhanced through the utilization of supervised machine learning and ensemble classifiers. Various models, including random forest, decision tree, logistic regression, and k-nearest neighbor, were trained on the TON-IoT dataset. Their integration employed both stacking and voting techniques. The ensemble classifiers surpassed the individual classifiers, achieving an accuracy exceeding 89.63%, hence enhancing IDS dependability and minimizing classification mistakes. Nevertheless, research concentrated solely on binary classification of normal versus abnormal traffic may hinder the detection and differentiation of specific attack types, hence constraining its application in the multifaceted real-world IoT threat landscape.

Recent studies [19] demonstrate both the potential and constraints of machine learning in cyber threat detection. Although models such as random forest, decision tree, support vector machine, and k-nearest neighbors demonstrate considerable classification capabilities, and deep learning techniques (CNNs, RNNs) excel in intricate pattern identification, substantial obstacles persist in attaining reliable real-world performance. Ensemble approaches such as bagging and boosting have enhanced accuracy through the amalgamation of classifiers; yet, they continue to encounter significant challenges in addressing complex attack patterns, especially when trained on traditional datasets. This elucidates the increasing focus on hyperparameter tuning and grid search [21] to enhance performance. The incorporation of hybrid datasets like as HIKARI-2021 [22] signifies a significant progression, as our research illustrates by merging these methodologies to address the adaptability constraints of previous ensemble techniques while preserving their accuracy advantages.

3 Methodology

A proposed method utilizes ensemble learning to enhance intrusion detection systems. The framework consists of multiple steps: initially, the HIKARI-2021 dataset was obtained from Kaggle. The preparation procedures are implemented to guarantee the quality and equilibrium of the dataset. The dataset was divided into two segments: 80% for training and 20% for testing. A variety of classifiers, including Random Forest, Decision Tree, Gaussian Naive Bayes, K-Nearest Neighbors, Logistic Regression, Multi-Layer Perceptron, and Convolutional Neural Network, were employed on the training data. The efficacy of individual classifiers was assessed and compared based on recall, F1-score, accuracy, and precision. An ensemble model is employed to integrate the capabilities of these individual classifiers and achieve optimal accuracy. Hyperparameter optimization was conducted to enhance the performance of each model. The suggested ensemble method resulted in significant enhancements in detection accuracy, thereby validating that the integration of ML and DL techniques offers a robust intrusion detection solution. This methodology addresses the limitations of singular models and enhances predictive capabilities, hence facilitating the development of a sophisticated security architecture and fortifying enterprises against continually evolving cyber threats.

3.1 Dataset HIKARI-2021

The HIKARI-2021 dataset offers an extensive array of choices for assessing a network intrusion detection system. The dataset comprises around 1.2 million records, with 80% allocated for training data (960,000 records) and the remaining 20% designated as testing data (240,000 records). It consists of both synthetic and authentic encrypted attack traffic produced by malware, port scanning, and DDoS assaults. The diversity of this dataset provides many attack vectors for the training and evaluation of deep learning and machine learning classification models. The integration of authentic and synthetic data facilitates an effective simulation of the real world through the HIKARI-2021 dataset, utilized by researchers for the development of various intrusion detection systems with efficacy and flexibility.

3.2 Data Preparation

A crucial element to consider in data analysis is pretreatment. It ensures the conversion of data from its unrefined condition to a format appropriate for deep learning and machine learning methodologies. Several essential stages in the preprocessing phase are incorporated to enhance model correctness.

3.2.1 Eliminating Null Values

Machine learning and deep learning techniques require categorical items, such as protocol types or service names, in network security datasets like the HIKARI-2021 dataset to be assigned a numerical representation. Categorical conversion is executed using various methods, including label encoding and one-hot encoding. One-hot encoding generates n binary columns, indicating presence or absence, where n is the number of categories, while label encoding allocates distinct numerical values to each category. The selection of the approach is contingent upon the algorithm utilized and the nature of the data. Therefore, categorical data must be appropriately managed to ensure the reliability and functionality of the models.

3.2.2 Managing Categorical Data

This research acquired the HIKARI-2021 dataset from Kaggle. This collection contains approximately 2 million records. Categorical variables underwent certain preprocessing transformations. Convert into numerical format using label encoding and one-hot encoding methods. The modification facilitated interoperability with machine learning and deep learning algorithms, resulting in rapid and precise training and testing of the models.

3.2.3 Label Encoding

Label encoding is a crucial first preprocessing step for transforming categorical data into an appropriate numerical representation for machine learning and deep learning models. In the HIKARI-2021 dataset, the detection labels 'normal' and 'attack' are converted to label 0 and label 1, respectively. This property ultimately enhances the performance and efficacy of the intrusion detection system by allowing the algorithms to understand and learn from data during model training.

3.2.4 Partitioning Data into Training and Testing Sets

A train-test split is important to prepare the HIKARI-2021 dataset for analysis and modeling activities in this work. The dataset is partitioned into two subsets: a training dataset and a testing dataset. Approximately 80% of the entries, totaling around 960,000, comprise the training dataset, while the remaining 240,000 serve as the testing dataset. Training data is subsequently input into machine learning and deep learning models to discern data patterns and attributes. Subsequently, the trained models are assessed with the testing data to ascertain their ability to generalize from unfamiliar data. This division is crucial for evaluating the models' efficacy in identifying intrusions and their associated accuracy and dependability in real-world scenarios. This train-test split aims to develop strong and reliable models for intrusion detection to address the predominantly evolving and varied cyber threats.

3.3 Machine Learning Models

Machine learning is a subset of artificial intelligence encompassing the creation of algorithms and methodologies that enable machines to learn from data and enhance their performance progressively. In contrast to conventional programming, users must offer explicit instructions; in contrast, machine learning algorithms identify patterns within data to anticipate outcomes, classify data, or enhance processes. The inherent characteristic of adaptive learning allows these machines to enhance with time, hence augmenting their accuracy and efficiency in diverse jobs [23].

Random Forest is an ensemble learning method employed to enhance classification accuracy by combining many decision trees. The final prediction is determined by the mean of the predictions from each individual tree in regression or the mode in classification [24].

$$Q_{left}(\theta) = \{(x, y)|x \leq t_m\} \tag{1}$$

$$Q_{right}(\theta) = Q_m - Q_{left}(\theta) \tag{2}$$

A decision tree is a model that resembles a tree and makes decisions based on feature values, partitioning data into subsets according to conditions on these features. Decision trees utilize Gini impurity or entropy for node splitting [25].

$$Gini = 1 - \sum_{i=1}^{c}(p_i)^2 \tag{3}$$

The Gaussian Naive Bayes classifier is based on Bayes' theorem, assuming feature independence. The subsequent formula is employed to ascertain the probability of a class based on features:

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x-\mu_y)^2}{2\sigma_y^2}\right) \tag{4}$$

KNN classifies a data item based on the predominant class among its k nearest neighbors in the feature space [24]. The Euclidean distance is commonly employed to ascertain the distance between data points.

$$d(p, q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \tag{5}$$

A logistic function is employed in logistic regression to denote the probability of a binary outcome [25]. The logistic function is computed as follows:

$$z = \sum_{i=1}^{n} w_i \cdot x_i + b \tag{6}$$

3.4 Deep Learning

The term "deep" denotes a category of machine learning that employs multi-layered neural networks. By emulating the human brain's information processing, these neural networks allow computers to learn and make decisions with minimal human intervention. Deep learning excels in complex tasks such as image and audio identification, natural language processing, and autonomous driving due to its ability to discern patterns within extensive datasets. It is a powerful artificial intelligence instrument due to its ability to learn from experience and data [26]. MLP is a feedforward artificial neural network with many layers of neurons equipped with activation function. CNNs are deep learning algorithms that excel in visual data analysis. Convolutional layers are employed to extract characteristics [27].

3.5 Ensemble Method

This research employs an ensemble technique to enhance classification accuracy beyond that of individual algorithms. Predictions from various foundational models were amalgamated, including K-Nearest Neighbors, Gaussian Naive Bayes, Random Forest, Decision Tree, Logistic Regression, Multi-layer Perceptron, and Convolutional Neural Networks. The research utilized various models to create a meta-model that optimally integrates their predictions, yielding a more effective penetration detection strategy than that employed by individual algorithms. This significantly enhanced overall performance while providing vital insights for improving network security applications.

3.6 Evaluation Metrics

The assessment of intrusion detection methods involves multiple criteria that assign numerical values to their effectiveness. These metrics encompass many aspects of class accuracy, robustness, and the algorithm's efficacy in identifying various cyber risks. A few common parameters are utilized in the assessment of intrusion detection [28].

The accuracy assesses the model's overall precision by dividing the count of correctly predicted instances by the total number of cases. FN represents False Negatives, FP denotes False Positives, TN indicates True Negatives, and TP signifies True Positives.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{7}$$

Precision indicates the proportion of true positive predictions relative to the total number of positive forecasts.

$$Precision = \frac{TP}{TP + FP} \times 100 \tag{8}$$

Recall assesses the model's ability to differentiate authentic positive instances from all other positive instances.

$$Recall = \frac{TP}{TP + FN} \times 100 \tag{9}$$

The F1-score is the harmonic mean of precision and recall, effectively balancing false positives and false negatives.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{10}$$

A confusion matrix displays the quantities of true positives, true negatives, false positives, and false negatives to evaluate the performance of a classification model.

4 Results and Discussion

This section delineates the results of employing various machine learning and deep learning algorithms on our dataset for intrusion detection, specifically Random Forest, Decision Tree, Gaussian Naive Bayes, K-Nearest Neighbor, Logistic Regression, Multi-Layer Perceptron, and Convolutional Neural Network. The efficacy of each algorithm was assessed utilizing critical performance metrics including accuracy, precision, recall, and F1 score. This research employed an ensemble strategy to enhance classification accuracy by utilizing the predictions from the base models as inputs for a meta-model. By using the distinct advantages of each algorithm, performance improved markedly, resulting in a more efficient and resilient intrusion detection system. The results demonstrate the extent to which ensemble learning may enhance security protocols.

4.1 Random Forest Performance Evaluation

The random forest classifier performed effectively on the intrusion detection dataset. The maximum accuracy achieved was 88.15%, indicating a substantial number of correctly identified occurrences. A precision score of 87.64% signifies that a substantial proportion of instances identified as positive were accurately classified. The recall is 88.15%, indicating the model's efficacy in identifying true positives with minimal false negatives. The F1 score is 87.89%, demonstrating a balanced precision and recall, which signifies the classifier's overall dependability and robustness.

4.2 Decision Tree Performance Evaluation

The overall accuracy was 87.82%, demonstrating that a significant number of instances were accurately identified. The model's precision is 87.55%. The 87.82% recall enables the identification of true positive instances. Correspondingly, an F1 score of 87.69% indicates that the classifier exhibited robustness, thereby achieving a satisfactory equilibrium between accuracy and recall.

4.3 Logistic Regression Performance Evaluation

The logistic regression model exhibited remarkable performance on our intrusion detection dataset, attaining an accuracy of 93.2%, so indicating a substantial number of occurrences were accurately identified. The precision of 86.96% indicates the accuracy of the model's predictions, but the recall of 93.25% emphasizes the model's effectiveness in identifying true positives. An F1 score of 89.99% indicates a balance between precision and recall, affirming the model's overall robustness and dependability.

4.4 Gaussian Naive Bayes Performance Evaluation

In comparison to the other models, the Gaussian Naive Bayes model exhibited commendable performance on our intrusion detection dataset, achieving an accuracy of 72.21%, a precision of 94.38%, a recall of 72.21%, and an F1 score of 79.13%. These data indicate that although the model demonstrates high accuracy in predicting good outcomes, its ability to identify actual positive outcomes is only moderate due to relatively lower memory. The balanced F1 score indicates a potential for improving detection capability by highlighting the trade-off between accuracy and recall.

4.5 Gradient Boosting Performance Evaluation

The Gradient Boosting Classifier demonstrated commendable performance on the intrusion detection dataset, with a minimum accuracy of 93.35%, precision of 91.14%, recall of 93.35%, and an F1 score of 90.97%. The data suggest that the model exhibited commendable performance, as evidenced by a high accuracy, indicating that the majority of cases were predicted accurately.

4.6 K-Nearest Neighbors Performance Evaluation

The K-Nearest Neighbors classifier shown enhanced performance in managing our intrusion detection dataset. The KNN classifier has an accuracy rate of 92.32%, signifying that most examples were accurately classified. The model's accuracy in predicting the positive class is 89.04%, and the recall for real positive occurrences is 92.32%. The model's robustness and dependability were underscored by an F1 score of 90.28%, indicating its balanced efficacy in precision and recall.

4.7 Multi-Layer Perceptron Performance Evaluation

The multi-layer perceptron classifier has strong performance on our intrusion detection dataset. This classifier accurately recognized an impressive 93.21% of cases, indicating that a significant proportion of instances was classified with precision. The precision of 89.57% indicates the model's effectiveness in making accurate positive classifications, while the recall of 93.21% measures its ability to identify actual positive events. The F1 score of 90.09% demonstrates a balance between precision and recall, indicating the model's overall reliability and robustness.

4.8 Convolutional Neural Network Performance Evaluation

The Convolutional Neural Network (CNN) shown exceptional performance on our intrusion detection dataset. The models have accurately categorized a substantial number of occurrences with an accuracy of 93.14%. An accuracy of 86.95% demonstrates the model's proficiency in predicting positive class instances within the training data, whereas a recall specificity of 93.14% reflects the model's effectiveness in identifying positive events. The F1 score of 89.94% signifies a balance between recall and precision, illustrating the model's robustness and stability.

4.9 Evaluation of Proposed Ensemble Models

Ensemble learning, utilizing several base models such as Random Forest, Decision Tree, Gaussian Naive Bayes, K-Nearest Neighbor, Logistic Regression, Multi-Layer Perceptron, and Convolutional Neural Networks, demonstrated effective performance with our dataset. The stacked model exhibited an accuracy of 96.33%, precision of 95.53%, and recall of 95.64%. This demonstrates the algorithm's capacity to identify genuine positive cases and provide very accurate positive predictions. Ensemble's notable achievement underscores its commitment to enhancing the intrusion detection system's efficacy and resilience through the integration of diverse algorithmic advantages, thus refining a more robust and precise detection mechanism.

4.10 Comparative Performance Analysis

This study compares the performance of several models, including the suggested ensemble model, Gradient Boosting, Logistic Regression, K-Nearest Neighbors, Multi-Layer Perceptron, Convolutional Neural Network, Random Forest, Decision Tree, and Gaussian Naive Bayes. The comparative evaluation criteria employed include accuracy, recall, F1-score, and precision.

Table 1 presents a comparative analysis of different machine learning algorithms utilizing four performance metrics: Accuracy, Precision, F1-Score, and Recall. The proposed LDX Model exhibits supremacy over all machines, achieving the highest accuracy of 96.33%, alongside balanced metrics of precision at 95.53%, recall at 96.33%, and an F1-score of 95.64%, hence possessing the most predictive potential.

Table 1: Comparison of Performance of the Proposed Model

Models	**Accuracy**	**Precision**	**F1 Score**	**Recall**
Proposed LDX Framework	96.33%	95.53%	95.64%	96.33%
Gradient Boosting	93.35%	91.14%	90.97%	93.35%
Logistic Regression	93.22%	86.98%	93.25%	89.99%
K-Nearest Neighbors	92.32%	89.04%	90.28%	92.32%
MLP	93.21%	89.57%	90.09%	93.21%
CNN	93.14%	89.04%	89.94%	93.14%
Random Forest	88.15%	87.64%	87.89%	88.15%
Decision Tree	87.82%	87.55%	87.82%	87.69%
Gaussian Naive Bayes	72.21%	94.38%	79.13%	72.21%

Among traditional models, Gradient Boosting demonstrates notable efficiency at 93.35%, closely succeeded by Logistic Regression, MLP, and CNN, all within the 93% range. KNN demonstrates a competitive performance of 92.32%, albeit with little worse precision. The performance of Random Forest and Decision Tree is suboptimal, with accuracies of 88.15% and 87.82%, indicating potential overfitting. Notably, Gaussian Naïve Bayes exhibits the lowest accuracy at 72.21%, while it achieves the best precision at 94.38%. This indicates a minimal occurrence of false positives, but it suffers from inadequate recall and overall performance. The results demonstrated the comparative superiority of ensemble and deep learning models over simpler techniques, consequently enhancing the robustness of the proposed LDX model in achieving optimal classification performance.

5 Conclusion and Future Work

The findings indicate that the efficacy of any intrusion detection system can be effectively enhanced by the utilization of ensemble models. Ensemble models, by integrating the strengths of diverse machine learning and deep learning techniques, provide enhanced accuracy and robustness in the detection of cyber threats. All algorithms were trained and assessed using our intrusion detection dataset, with assessment metrics including accuracy, precision, recall, and F1 score analyzed. Our findings indicate that the ensemble technique, which combines the strengths of various distinct algorithms, achieves optimal performance. The ensemble model achieved an accuracy of 96.33%, precision of 95.53%, recall of 96.33%, and F1 score of 95.64%. This indicates the efficacy of the ensemble strategy in markedly enhancing detection capabilities. Future endeavors will concentrate on: the integration of BERT and GPT models for the analysis of encrypted traffic, the assessment of framework adaptability via response time and false-positive rate benchmarks, and the evaluation against zero-day assaults in AWS/Azure cloud environments.

References

[1] Alotaibi, Y., & Ilyas, M. (2023). Ensemble learning framework for intrusion detection to augment the security of Internet of Things devices. *Sensors*, 23(12), 5568.

[2] Srinivasan, S., & Deepalakshmi, P. (2023). Augmenting cybersecurity by identifying botnets with ensemble classification-based machine learning. *Measurement: Sensors*, 25, 100624.

[3] Mokbal, F. M. M., et al. (2022). A proficient intrusion detection framework utilizing embedded feature selection and ensemble learning methodologies. *International Arab Journal of Information Technology*, 19(2), 237-248.

[4] Mishra, A. K., & Paliwal, S. (2023). Mitigating cyber dangers with the integration of feature selection and stacking ensemble learning: A perspective on LGBM and random forest for intrusion detection. *Cluster Computing*, 26(4), 2339-2350.

[5] Lower, N., & Zhan, F. (2020). An examination of ensemble techniques for cybersecurity. In *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 1001-1009). IEEE.

[6] Hossain, M. A., & Islam, M. S. (2023). Guaranteeing network security using a resilient intrusion detection solution utilizing ensemble-based machine learning. *Array*, 19, 100306.

[7] Jiang, Y., & Atif, Y. (2021). A selected ensemble model for cognitive cybersecurity analysis. *Journal of Network and Computer Applications*, 193, 103210.

[8] Akram, U., et al. (2023). IoTTPS: An Ensemble RKSVM Model-Based System for Internet of Things Threat Protection. *Sensors*, 23(14), 6379.

[9] Nadeem, M. W., et al. (2021). Fusion-Based Machine Learning Framework for Cardiac Disease Prognostication. *Computers, Materials & Continua*, 67(2).

[10] Rustam, F., et al. (2021). Classification of denial of service attacks with machine learning with several features. *Electronics*, 11(22), 3817.

[11] Majeed, R., et al. (2021). Advanced cyber-security system for IoT-enabled drones utilizing a voting classifier. *Electronics*, 10(23), 2926.

[12] Akram, U., et al. (2017). A comprehensive survey on Pi-Sigma neural network for time series prediction. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 9(3-3), 57-62.

[13] Fernandes, R., et al. (2021). The influence of identifiable features in machine learning classification algorithms utilizing the HIKARI-2021 dataset. IEEE.

[14] Sanober, I., & Mir, R. N. Feature Reduction and Dataset Balancing in Intrusion Detection Systems: A Comprehensive Evaluation on Hikari-2021 and Legacy Datasets. Available at SSRN 5009537.

[15] Parsazad, S., Saboori, E., & Allahyar, A. (2012, May). Fast feature reduction in intrusion detection datasets. In 2012 Proceedings of the 35th International Convention MIPRO (pp. 1023-1029). IEEE.

[16] Zhou, Y., & Wang, P. (2019). An ensemble learning methodology for detecting XSS attacks with domain knowledge and threat intelligence. *Computers & Security*, 82, 261-269.

[17] Rizwan, M., et al. (2022). Classification of depression from tweets utilizing compact deep transfer learning language models. *IEEE Access*, 10, 129176-129189.

[18] Ferriyan, A., et al. (2021). Creation of a network intrusion detection dataset utilizing authentic and encrypted synthetic attack traffic. *Applied Sciences*, 11(17), 7868.

[19] Mushtaq, M. F., et al. (2017). A novel cognitive architecture for humanoid robots. *International Journal of Advanced Computer Science and Applications*, 8(6), 64.

[20] Kwon, D., et al. (2023). Assessing imbalanced network data for attack detection. In *Proceedings of the 2023 Conference on Systems and Network Telemetry and Analytics* (pp. 23-26).

[21] El-Khatib, K. (2009). Impact of feature reduction on the efficiency of wireless intrusion detection systems. IEEE TRANSACTIONS on parallel and distributed systems, 21(8), 1143-1149.

[22] Vitorino, J., et al. (2023). A benchmark for adversarial robustness in enterprise network intrusion detection. In *International Symposium on Foundations and Practice of Security.*

[23] Akram, U., et al. (2019). An enhanced Pi-Sigma neural network with error feedback for predicting physical time series. *International Journal of Advanced Trends in Computer Science and Engineering*, 8(1), 1-7.

[24] Ferriyan, A., et al. (2021). Creation of a network intrusion detection dataset utilizing authentic and encrypted synthetic assault traffic. *Applied Sciences*, 11(17), 7868.

[25] Mushtaq, M. F., et al. (2017). A comprehensive survey of cryptographic encryption algorithms. *International Journal of Advanced Computer Science and Applications*, 8(11), 333-344.

[26] Mushtaq, M. F., et al. (2019). Key schedule algorithm utilizing three-dimensional hybrid cubes for block cipher. *International Journal of Advanced Computer Science and Applications*, 10(8).

[27] Sarwar, A., et al. (2023). Detection of IoT network attacks utilizing multi-novel features and the Extra Tree Random-Voting Ensemble Classifier (ER-VEC). *Journal of Ambient Intelligence and Humanized Computing*, 14(12).

[28] Lazzarini, R., Tianfield, H., & Charissis, V. (2023). Propose a stacking ensemble of deep learning models for the detection of intrusions in IoT systems. *Knowledge-Based Systems*, 279, 110941.