Cyberatttack Detection and Classification in IIoT systems using XGBoost and Gaussian Naïve Bayes: A Comparative Study

Mordi Alenazi
College of Computer and Information Sciences, Majmaah University, Majmaah, Saudi Arabia
441104425@s.mu.edu.sa (corresponding author)

Shailendra Mishra
College of Computer and Information Sciences, Majmaah University, Majmaah, Saudi Arabia
s.mishra@mu.edu.sa

Received: 27 April 2024 | Revised: 11 May 2024 | Accepted: 20 May 2024

 DOI: https://doi.org/10.48084/etasr.7664

ABSTRACT

The Industrial Internet of Things (IIoT) is experiencing rapid expansion, forming a vast network of interconnected devices, sensors, and machines that generate large volumes of data. In the context of Industry 5.0, ensuring the accuracy and reliability of this data is essential. This paper addresses the challenges of detecting and classifying cyberattacks within the IIoT by employing advanced analytical techniques. Specifically, we explore the application of Machine Learning (ML) algorithms, focusing on the comparison between the XGBoost and Naïve Bayes models. Our study uses the KDD-99 and NSL KDD datasets to evaluate the performance of these models in terms of accuracy, precision, recall, and F1 score. The results demonstrate that the XGBoost model significantly outperforms the Naïve Bayes model across all metrics, achieving an accuracy of 99%. This study contributes to the improvement of intrusion detection and classification of cyberattacks in IIoT environments.

Keywords-cyberattacks in IIoT; XGBoost; Naïve Bayes

I. INTRODUCTION

The progress of Industry 4.0 has enabled the integration of digital and physical technologies through IoT-based processes [1]. This integration is facilitated by a vast network of interconnected devices, machines, applications, and human engagement [2]. The Industrial Internet of Things (IIoT) has the potential to revolutionize the landscape for businesses and governments, heralding a new era of smart automation with the creation of smart factories, advances in smart healthcare systems, the development of smart homes and cities, and the enhancement of intelligent transportation systems. However, with the complex nature of IIoT comes a set of security challenges and privacy concerns that need to be addressed. The intricate Cyber-Physical Systems (CPSs) network faces cyber security risks as cyber-attackers evolve their methods and strategies, exploiting AI technologies to enhance their capabilities. Technological advancements demand continuous research to safeguard against the misuse of AI by cyber criminals [3]. Industrial installations often feature geographically dispersed infrastructures, increasing the potential for cyber adversaries to cause significant damage. These challenges can lead to threats potentially resulting in widespread operational failures with catastrophic consequences [4]. The rapid proliferation of internet-connected devices raises concerns about the resilience of industrial networks against increasingly sophisticated cyber threats. Existing cybersecurity measures in industrial settings fall short of addressing the magnitude and complexity of threats posed in the IIoT environment, necessitating the development of more robust and adaptable security frameworks to effectively safeguard industrial networks [5]. As IIoT systems become more autonomous and interconnected, they become more vulnerable to sophisticated cyberattacks, particularly with the advent of AI-powered cyberattacks that can bypass conventional security measures. This underscores the need for advanced, AI-based security solutions [6]. The use and evolution of Deep Learning (DL) methods has reduced web-based attacks, such as DDoS and SQL injection [7]. Machine Learning (ML) methods have emerged as highly effective tools in various fields, and they now present a promising avenue for enhancing cybersecurity in Industry 5.0. DL methods such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and transformer models stand out for their ability to autonomously learn complex patterns and representations directly from raw data [7]. This attribute is particularly valuable in cybersecurity, enabling DL models to uncover new and sophisticated attacks that traditional ML methods might miss [2]. This study aims to explore the utilization of ML algorithms in enhancing the security of IIoT systems. The objectives of this study are:

- To emphasize the efficiency and effectiveness of ML algorithms in identifying intrusion attacks within IIoT systems.
- To investigate the integration of ML algorithms, specifically XGBoost, into Intrusion Detection Systems (IDSs) for IIoT security enhancement.
- To examine existing research findings and current limitations regarding the use of DL and ML for intrusion detection in IIoT environments.
- To provide a comprehensive overview of the subject by detailing the discovered insights and limitations.

The novelty of this research lies in its detailed exploration of the XGBoost algorithm's capability to provide enhanced cyber threat detection within IIoT systems. Previous research has recognized the general effectiveness of XGBoost in various applications, however, our study goes a step further by providing robust empirical evidence of its superior performance in the nuanced context of IIoT. The proposed Extreme Gradient Boosting (XGBoost) model achieved an impressive 99.98% accuracy on the KDD-99 dataset and 99.97% on the NSL-KDD dataset surpassing previous benchmarks [4, 26]. By also testing Gaussian Naive Bayes (NB), this study identifies key limitations in accuracy and effectiveness across datasets, particularly for NSL-KDD. Such findings highlight areas where the XGBoost model excels, establishing a clearer path for future research.

II. RELATED WORK

Research papers focusing on cybersecurity in IIoT are critical to ground the research in the current state of knowledge, understand the complexity of the security landscape, identify critical vulnerabilities, and guide future research and practical applications.

Authors in [4] introduced a sophisticated detection and classification system designed to improve security within the Industrial Internet of Things (IIoT). This system employs Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), Gaussian NB, Bagging, XGBoost, and Adaboost and applies them to the KDD-99 dataset for the precise identification and categorization of Advanced Persistent Threats (APTs). Through a comparative analysis, the XGBoost classifier demonstrated notable performance with an accuracy of 98.7%, a precision of 96.1%, a recall rate of 97.5%, and an F1 score of 96.6%. Additionally, the Gaussian NB classifier also showed strong results with an accuracy of 97.2%, a precision of 96.3%, a recall of 97.4%, and an F1 score of 96.7%. Authors in [8] initially tackle current cybersecurity threats and evaluate how ML can address these issues. Authors in [9] conducted an extensive analysis on DL methods for detecting cybersecurity intrusions, focusing on federated learning. They tested the effectiveness of RNN, CNN, and Deep Neural Network (DNN) models using three innovative real IoT traffic datasets. Authors in [10] provide a comprehensive review of neural networks and DL techniques applied to cybersecurity, evaluating their effectiveness across various cybersecurity tasks. Authors in [11] offer a detailed overview of advanced DL solutions for detecting cyber-attacks, specifically in CPSs. They propose a six-step methodology for reviewing and applying DL techniques in CPS, covering scenario analysis, threat identification, problem formulation, model customization, data acquisition, and performance evaluation. The paper also addresses current challenges, outlines opportunities, and suggests directions for future research [11]. Authors in [12] examined the use of DL techniques for detecting cybersecurity breaches, emphasizing the importance of datasets in improving IDSs. The study showcased various neural network methods suited for cybersecurity tasks, such as DNNs, constrained Boltzmann machines, deep belief networks, CNNs, deep Boltzmann machines, and deep autoencoders.

Authors in [7] developed an innovative DL approach to detect web-based attacks in Industry 5.0 settings, utilizing CNNs, RNNs, and transformers. Their findings highlighted that the transformer-based model outperformed traditional ML and DL models in terms of accuracy, precision, and recall, thus significantly enhancing the protection of critical infrastructure and sensitive data against cyber threats [7]. Authors in [13] studied the effectiveness of XGBoost for IDSs in Wireless Sensor Networks (WSNs) facing imbalanced data and cyberattacks such as black hole, gray hole, flooding, and scheduling. The study compared the performance of XGBoost with DT and NB. The results indicate that while three categories of attacks showed moderate imbalance, the flood attack category was severely imbalanced. XGBoost outperformed the standard methods, showing the highest AUC values in all categories, demonstrating its superior detection ability in diverse attack scenarios. Authors in [14] conducted research using the CSE-CIC-IDS2018 dataset to evaluate the use of ML algorithms (DT, RF, NB, LR, Catboost, LightGBM, and XGBoost) for detecting malicious network traffic. The study focused on the impact of ensemble feature selection on the performance of the classifiers, particularly in terms of their ability to detect breakthroughs. In [15], the authors examined two cyber data sets, each containing five classes, using three different classification algorithms: RF, XGBoost, and the Keras Sequential model. The findings confirmed that these classifiers are effective in detecting cybersecurity threats. In [16], focus was given on the use of ML techniques for identifying cybercrime activities. The study delves into the mathematical foundation of the XGBoost algorithm and explains how this mathematical basis contributes to the efficient and quick detection of cybercrime activities. Authors in [17] evaluated the performance of various ML algorithms in detecting attacks on IoT networks using simulated attack data. The study found that the RF algorithm achieved a detection accuracy of 99.9%, demonstrating its effectiveness. Additionally, the research highlighted that blockchain technology could enhance security and privacy in IoT networks by establishing a tamper-proof decentralized communication system [17]. Authors in [18] propose enhancing IoT cybersecurity by introducing a combined model based on DL. The model for malware detection utilizes Deep CNNs (DCNNs). Additionally, TensorFlow Deep Neural Networks (TFDNNs) are introduced for the identification of software piracy threats based on source code plagiarism. In [19], IPIDS, a novel comprehensive IDS tailored for IoT applications, capable of accurately detecting

intrusions via MQTT, HTTP, and CoAP protocols, is presented. The findings demonstrate that IPIDS surpasses the performance of other models trained on the same datasets. Furthermore, the DT and Long Short-Term Memory (LSTM) models developed in this study reached an impressive accuracy of 99.9%.

Authors in [20] discussed the integration challenges of AI-based cybersecurity within the aviation industry, emphasizing the need to align proposed solutions with existing regulations. The study covers the necessity for design verification of AI/ML algorithms and the certification requirements for AI/ML solutions applicable to manufacturers and agencies in the sector. Authors in [21] introduced the Enhanced Naïve Bayes Posterior Probability (ENBPP) algorithm designed to improve accuracy and reduce processing time for cyber-attack predictions. This algorithm combines a modified NB posterior probability function with a revised risk assessment function, aiming to enhance threat prediction efficiency. The results showed a significant improvement compared to traditional methods, with prediction accuracy increasing to 92-96% and processing time decreasing to 0.028 s from 0.043 s. Authors in [22] assessed the security performance of a WSN using Gaussian NB, Multinomial NB, and Bernoulli NB. These were compared against three established algorithms: K-Nearest Neighbors (kNN), SVM, and Multilayer Perceptron (MLP), utilizing the Spearman correlation for univariate feature selection [22]. Authors in [23] examined the cybersecurity challenges Autonomous Transportation Systems (ATSs) face. They underlined the critical role of integrating AI to improve the efficiency of cyber defense tactics and ensure the security of ATS assets [23]. Authors in [26] employed a Hybrid Meta-heuristic approach that integrates an LSTM classification model for dimension selection. When applied to the KDD-99 dataset, the XGBoost classifier achieved an accuracy of 94.06%, a precision of 92.94%, a recall of 93.53%, and an F1 score of 94.50%. Additionally, using the NSL-KDD dataset, the XGBoost classification reported an accuracy of 95.50%, a precision of 92.00%, a recall of 98.00%, and an F1 score of 95.55%.

III. THE RESEARCH METHOD

Identifying and understanding cyberattack techniques are significant steps in cybersecurity research. These initiatives are critical because they help in the swift development and implementation of cybersecurity policies. The examination of historical cyberattack data informs researchers and practitioners about potential threats and vulnerabilities, enabling the creation of proactive strategies to enhance an organization's defense mechanisms. This systematic approach to data analysis and model application, as shown in Figure 1, is fundamental in strengthening cybersecurity measures.

A. Data Collection

Collecting data related to IIoT datasets is a key step to conducting analysis using datasets such as KDD Cup 1999 [24] and NSL-KDD [25]. These datasets encompass various types of cyber-attacks, including DoS, U2R, Probe, R2L, providing researchers with diverse examples of cyber threats to analyze and mitigate.

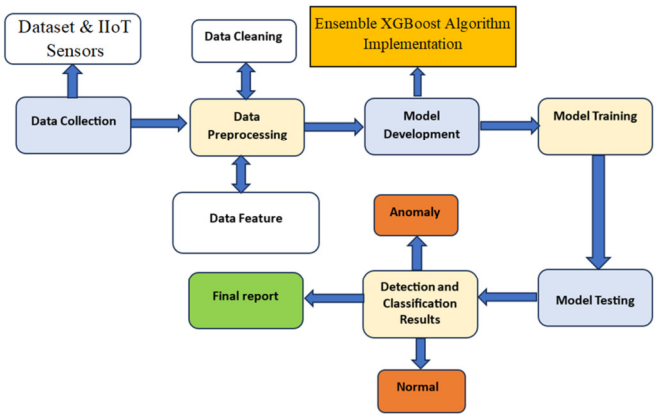

Fig. 1. Research methodology.

B. Data Preprocessing

Data preprocessing is crucial for effective model training, involving steps such as cleaning missing values and outliers, conducting feature engineering, normalizing data, and addressing class imbalances. These steps ensure that the dataset is well-organized and reliable for machine learning. In the context of the KDD99 and NSL-KDD datasets, this process involves partitioning the dataset into training (80%) and testing (20%) subsets, followed by normalization and one-hot encoding before moving on to model training and evaluation phases.

C. Ensemble XGBoost and Gaussian Naïve Bayes Implementation

Ensemble XGBoost is adept at handling complex data and significantly improves model accuracy. Meanwhile, the NB algorithm, respected for its simplicity and efficiency, introduces a complementary perspective to the modeling process. Algorithm 1 shows the steps taken for the implementation of the XGBoost model.

```
Algorithm 1: XGBoost Algorithm
Input: Dataset with features and target
labels.
Output: model_predictions,
model_performance
Def TrainAndEvaluateXGBoost(Dataset,
Features, Labels)
Reprocess Dataset to transform Features
and Labels
Training_set, testing_set <- split
preprocessed_data
Xgb_model <- initialize XGBoost
(parameters)
Train xgb_model on training_set
Best_params<-
tune_hyperparameters(xgb_model,
training_set)
Xgb_model <- reinitialize
XGBoost(best_params)
Model_predictions <-predict xgb_model
using testing_set
```

```
Model_performance<-evaluate
model_predictions against testingset
Labels
Return model_predictions,
model_performance
End def
```

D. Training and Testing

In this step, we train the cybersecurity models to recognize and predict cyber threats accurately. The models are taught with historical data, which include normal operations and past cyberattacks. This training allows the models to identify unusual patterns that could indicate a security risk. After training, we rigorously test the models to ensure they can detect actual threats without confusing them with normal behavior.

E. Cyber Attack Detection and Classification

After the model was tested, its predictions were used to identify anomalies within the data. Anomalies are unusual patterns or data points that stand out from the typical data. These could point to errors, unique outliers, or rare and significant events that need further investigation.

F. Final Report

The research culminates in a final report that outlines the methodology, evaluates the model's performance, and shares insights from the anomaly detection and classification process. It includes key performance metrics such as accuracy, precision, recall, and F1 score, along with any recommendations or conclusions drawn from the analysis.

G. The XGBoost Mathematical Model

1) Preprocessing and Feature Representation for XGBoost Implementation

- Feature Encoding: Encoding methods such as one-hot encoding are utilized to transform categorical attributes into numerical formats.
- Normalization: The scale of numerical attributes is adjusted to enhance learning efficiency.
- Label Encoding: Categorical target variables (such as attack types) are transformed into numerical codes.

2) Defining the Objective Function

In network intrusion detection scenarios, the objective function generally includes a classification loss function tailored to categorize network behaviors. For tasks distinguishing between normal and attack activities (binary classification), the binary logistic loss function is frequently selected. Conversely, when differentiating normal from various kinds of attacks (multi-class classification), a multi-class log loss function is utilized. For a set of n training data points and K trees, the objective function at iteration t is given by:

$$Obj^{(t)} = \sum_{i=1}^{n} l(y_i, y_i^{\wedge(t-1)} + f_t(x_i)) + \sum_{k=1}^{K} \Omega(f_k) \quad (1)$$

where l is the log loss function appropriate for binary or multi-class classification, y_i is the true label for instance i, $y_i^{\wedge(t-1)}$ is the predicted probability distribution over classes, for instance i from all trees up to the $(t\text{-}1)^{th}$, f_t is the new tree being added at iteration t, and Ω represents the regularization term.

3) Regularization

The regularization term in XGBoost modeling is:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \quad (2)$$

Ω aids in avoiding overfitting of the model, which is essential in cybersecurity contexts, given that False Positives (FP), i.e. identifying normal activities as attacks, and False Negatives (FN), i.e. overlooking actual attacks as normal can lead to significant consequences. In (2), T is the number of leaves in the tree, w_j is the score on the j leaf, and γ and λ are parameters that control the degree of regularization.

4) Optimization Process

The optimization process adheres to the standard approach utilized by XGBoost. Nonetheless, it is crucial to customize specific parameters, such as the learning rate and the trees' maximum depth, to suit the cybersecurity environment. This customization should take into account the intricacy of attack patterns and the need to strike a balance between minimizing FP and FN.

$$w_j^* = -\frac{G_j}{H_j+\lambda} \quad (3)$$

where G_j is the sum of gradients of the loss function for all instances in the j leaf and H_j is the sum of second-order derivatives (Hessians) of the loss function for all instances in the j leaf.

IV. IMPLEMENTATION

A computer with an AMD Ryzen 9 5900X processor and 32 GB of DDR4 RAM was used running Windows 10 and Python 3.9 as the primary programming language. Jupyter 3.2.0 served as the interactive development environment, providing an efficient workspace for experimentation. Key Python libraries played a crucial role. NumPy and Pandas were used for numerical computations and data manipulation, while Matplotlib and Seaborn enabled data visualization. The Scikit-learn library was used for ML tasks, including train_test_split for dividing data into training and testing sets. Metrics such as accuracy_score, precision_score, recall_score, f1_score, and confusion_matrix were essential for evaluating model performance. The Gaussian NB classifier (GaussianNB) and XGBoost classifier (XGBClassifier) were imported to build and evaluate classification models. Data handling began with loading the datasets into Pandas DataFrames, followed by dividing them into training and testing sets using the train_test_split. Both Gaussian NB and XGBoost classifiers were trained on the training dataset and evaluated on the testing dataset using various performance metrics. The results were visualized through Matplotlib and Seaborn, providing comprehensive insights into the models' accuracy, precision, recall, F1 score, and confusion matrices.

A. Evaluation Metrics

The evaluation regarded performance metrics accuracy, precision, recall, and F1 score. These metrics are derived from the counts of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

1) Accuracy

Accuracy reflects the model's overall effectiveness by determining the proportion of instances that have been correctly classified out of the total instances assessed.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

2) Precision

Precision measures the accuracy of the model in predicting positive labels.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

3) Recall

Recall measures the model's ability to correctly identify all actual positives.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

4) F1 score

F1 score is the harmonic mean of Precision and Recall offering a balanced metric between the two.

$$\text{F1 Score} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \quad (7)$$

B. Exploring Datasets

Tables I and II describe the considered datasets and the attack classes and types. Tables III and IV display the distribution and number of attack classes within the datasets.

TABLE I. DESCRIPTION OF KEY FEATURES AND SIZE OF KDD99 AND NSL-KDD

Feature	Description
Duration	Length (s) of the connection
Protocol type	Type of protocol used (TCP, UDP, etc.)
Service	Type of service requested or provided
Flag	Status of the connection (normal, error, etc.)
Failed logins	Number of failed login attempts
Other features	
Number of records	KDD99: 494,021 records
	NSL-KDD: 125,973 records

TABLE II. ATTACK CLASSES AND TYPES IN KDD99 AND NSL-KDD

Attack class	Attack Type
Normal	Data without intrusion.
DoS	back, land,Neptune, pod,smurf,teardrop, mailbomb,apache2,process table, udpstorm
U2R	buffer_overflow, loadmodule, perl, rootkit, httptunnel, ps, sqlattack ,xterm
Probe	ipsweep, nmap, portsweep, satan, mscan, saint
R2L	ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster, sendmail, named, snmpgetattack, snmpguess, xlock, xsnoop,worm

TABLE III. ATTACK TYPE DISTRIBUTION IN KDD99

Attack type	Count	Percentage
Normal	97278	19.69%
DoS	391458	79.23%
R2L	1126	0.22
U2R	52	0.01%
Probe	4107	0.83%

TABLE IV. ATTACK TYPE DISTRIBUTION IN NSL-KDD

Attack type	Count	Percentage
Normal	67342	53.45%
DoS	45927	36.45%
R2L	995	0.78%
U2R	52	0.04%
Probe	11656	9.25%

Figures 2 and 3 show the attack distribution over the considered datasets.

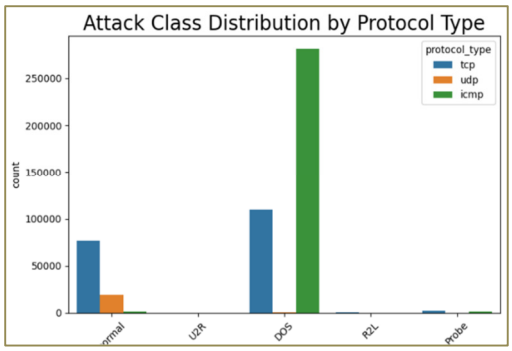

Fig. 2. KDD-99 attack protocol distribution.

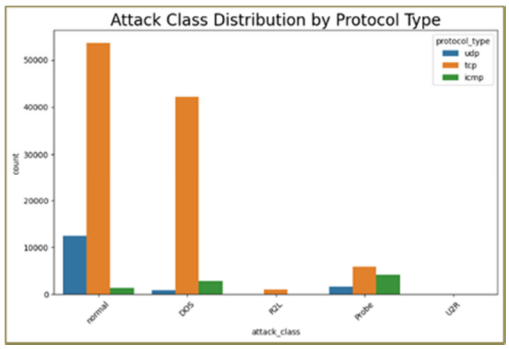

Fig. 3. NSL-KDD attack protocol distribution.

V. RESULTS AND DISCUSSION

A. Results

1) KDD-99

In the evaluation of the KDD-99 dataset (Table V and Figures 4 and 5), XGBoost showcased exceptional performance with high accuracy, precision, recall, and F1 score, indicating robustness in detecting attacks across various types. On the other hand, Gaussian NB exhibited lower accuracy and effectiveness, particularly in balancing precision and recall, leading to less reliable performance in distinguishing attacks from normal behavior. Figure 6 shows

the confusion matrix of the results for the XGBoost model. The model is very effective at classifying different types of network behavior with a high number of TP and a low number of FP and FN. The Gaussian NB classifier shows good accuracy in detecting DoS attacks and with the least possible errors in classification, as can be seen in Figure 7. But it faces challenges in accurately identifying Probe, R2L, and U2R attacks, as evidenced by high FN and FP.

TABLE V. PERFORMANCE OF XGBOOST AND NAIVE BAYES ON KDD-99 DATASET

	XGBoost	**NB**
Accuracy	99.98%	90.04%
Precision	98.21%	90.93%
Recall	99.73%	90.04%
F1 score	98.94%	89.43%

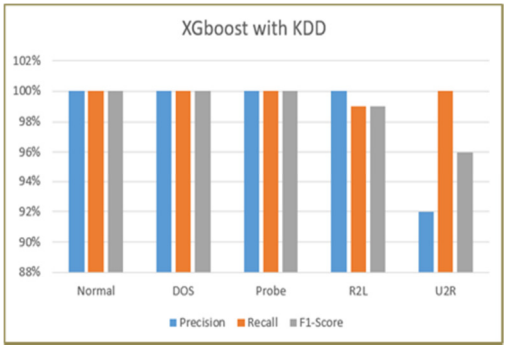

Fig. 4. Performance of XGBOOST on KDD-99 dataset.

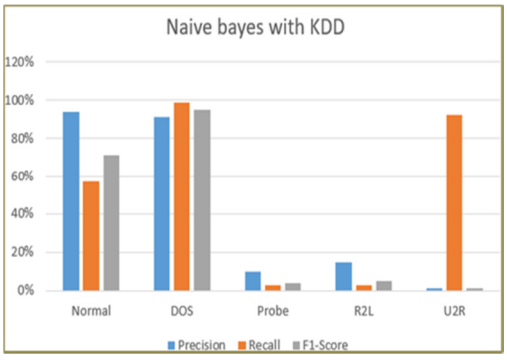

Fig. 5. Performance of NB on KDD-99 dataset.

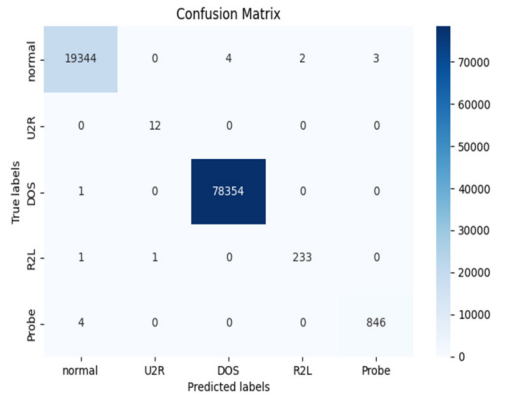

Fig. 6. Confusion matrix for XGBoost on KDD-99 dataset.

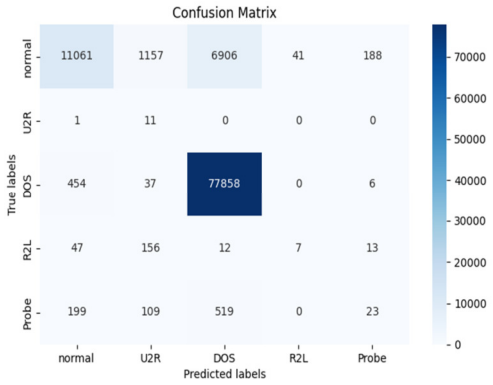

Fig. 7. Confusion matrix for NB on KDD-99 dataset.

2) NSL-KDD

In evaluating the models using the NSL-KDD dataset, Table VI and Figure 8 and 9 demonstrate that XGBoost maintained high performance. Conversely, Gaussian NB exhibited lower accuracy, precision, and recall. Figure 10 presents a confusion matrix for the XGBoost model, which effectively classifies Normal, DoS, and Probe attacks with high accuracy. On the other hand, as depicted in Figure 11, the confusion matrix of the NB model indicates proficient identification of Normal behavior and DoS attacks, but it reveals challenges in accurately classifying Probe, R2L, and U2R attacks, leading to a substantial number of FN.

TABLE VI. PERFORMANCE OF XGBOOST AND NAIVE BAYES ON NSL-KDD DATASET

	XGBoost	**NB**
Accuracy	99.97	39%
Precision	99.86	61%
Recall	98.45	39%
F1 score	99.12	27%

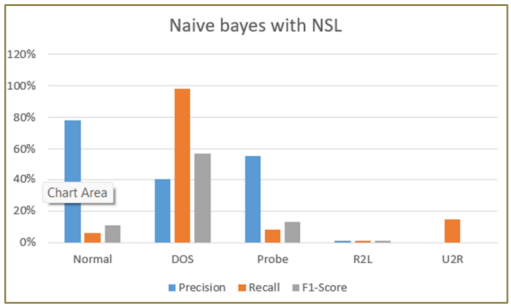

Fig. 8. Performance of NB on the NSL-KDD dataset.

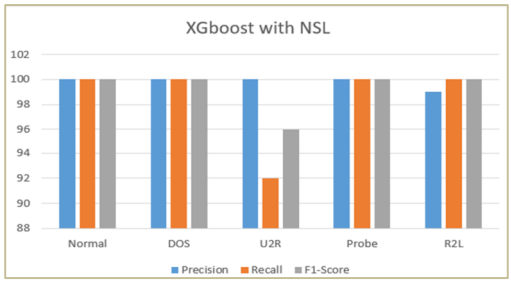

Fig. 9. Performance of XGBoost on NSL-KDD dataset.

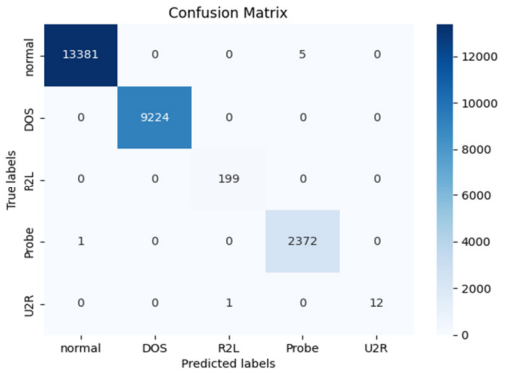

Fig. 10. Confusion matrix for XGBoost on the NSL-KDD dataset.

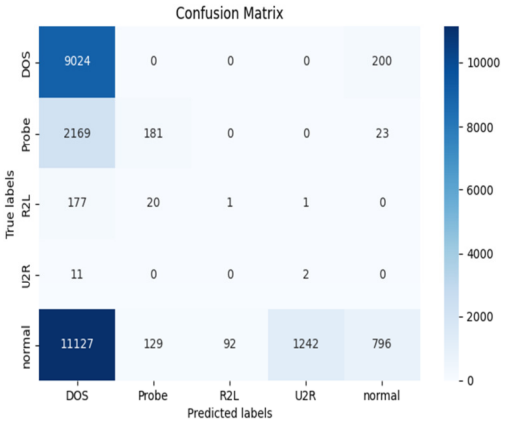

Fig. 11. Confusion matrix for NB on the NSL-KDD dataset

3) *Comparative Analysis*

Table VII and Figure 12 compare the performance metrics of the XGBoost and NB models across the KDD99 and NSL-KDD datasets, illustrating their accuracy, precision, recall, and F1 score.

TABLE VII. FINAL RESULT COMPARISON

Dataset	Model	Accuracy	Precision	Recall	F1 Score
KDD99	XGBoost	99.98	98.21	99.73	98.94
	NB	90.04	90.93	90.04	89.43
NSL-KDD	XGBoost	99.97	99.86	98.45	99.12
	NB	39	61	39	27

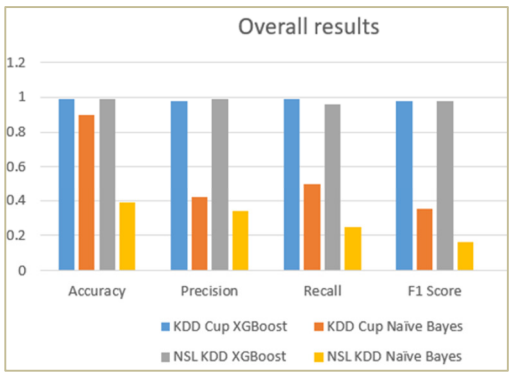

Fig. 12. Final result comparison between the XGBoost and NB models across the KDD99 and NSL-KDD datasets.

Table VIII states the comparative analysis of the proposed Model with recent relevant research. The proposed approach has undergone rigorous evaluation against the KDD-99 and NSL-KDD datasets. The results indicate a significant enhancement in detection capabilities, with our model achieving an outstanding performance on the KDD99 dataset and an exemplary performance on NSL-KDD. The performance metrics values not only underscore the precision of our model but also its ability to minimize FP and FN, a crucial aspect of IDS performance. When juxtaposed with the findings of [4, 26], our methodology not only surpasses their reported metrics, but also sets a new standard for accuracy and reliability in the field of cybersecurity, paving the way for more resilient and intelligent defense mechanisms against network threats. While [4, 26] laid the groundwork with their pioneering approaches, our research builds upon and optimizes these techniques. Our usage of XGBoost, coupled with extensive parameter tuning and feature selection methods, has evidently closed the gap in detection rates that previous models faced. Moreover, our model's ability to generalize across different datasets with minimal performance drop-off underscores its practicality for real-world applications.

TABLE VIII. PERFORMANCE COMPARISON WITH OTHER WORKS

Ref.	Model	Metric	KDD-99	NSL-KDD
[26]	XGBoost	Accuracy	94.06	95.50
		Precision	92.94	92.00
		Recall	93.53	98.00
		F1 score	94.50	95.55
[4]	Gaussian NB	Accuracy	97.2	-
		Precision	96.3	-
		Recall	97.4	-
		F1 score	96.7	-
	XGBoost	Accuracy	98.7	-
		Precision	96.1	-
		Recall	97.5	-
		F1 score	96.6	-
This study	XGBoost	Accuracy	99.98%	99.97
		Precision	98.21%	99.86
		Recall	99.73%	98.45
		F1 score	98.94%	99.12
	Gaussian NB	Accuracy	90.04%	39%
		Precision	90.93%	61%
		Recall	90.04%	39%
		F1 score	89.43%	27%

The results emphasize key trends and patterns in the data, beginning with the clear superiority of XGBoost over NB. XGBoost consistently achieved nearly perfect metric scores across both datasets, demonstrating its robustness and reliability in detecting diverse attack types. NB, on the other hand, experienced significant performance drops, particularly on the NSL-KDD dataset, underscoring its limitations in handling complex data patterns. In terms of effectiveness in detecting different attack types, XGBoost excelled across all classes, particularly in Normal, DoS, and Probe attacks. It maintained a high performance, even for R2L and U2R attacks, which have fewer instances. Conversely, NB struggled to distinguish between Normal and malicious traffic, leading to high FP for Normal and difficulty in classifying the more complex R2L and U2R attacks.

A significant unexpected finding was the comparatively poor performance of the Gaussian NB model, particularly on the NSL-KDD dataset. The model exhibited low accuracy (39%) and low F1 scores (27%), especially for certain attack categories like R2L and U2R. This result contrasted starkly with the XGBoost model, which achieved outstanding performance on both datasets, with accuracy nearing 100%. The poor performance of Gaussian NB may highlight the limitation of assuming feature independence, which often does not hold true for complex cybersecurity data. In real-world scenarios, feature interactions are crucial for accurately identifying nuanced attack patterns.

The impact of class imbalance was notable in the results, with XGBoost effectively handling this challenge due to its ensemble learning approach. However, NB struggled, revealing that simpler models may require additional preprocessing or oversampling techniques to enhance performance. Across both datasets, XGBoost consistently outperformed NB, suggesting that ensemble methods are generally better suited for managing the intricate patterns present in network traffic data. XGBoost's scalability was not affected by dataset size, demonstrating its ability to scale efficiently.

The strong performance of XGBoost suggests that ensemble methods can better adapt to changing data patterns, making them more suitable for modern cybersecurity challenges. It also underlines the need for tailored ML algorithms to suit the dynamic attack vectors present in IIoT networks.

B. Discussion

In this study, the XGBoost model consistently outperformed NB and previous models across all performance metrics. XGBoost achieved an impressive 99.98% accuracy on the KDD99 dataset and 99.97% on the NSL-KDD dataset, significantly surpassing the results attained in [4, 26]. These nearly flawless results show XGBoost's robust capability in distinguishing various attack types. In contrast, the NB model produced mixed results, performing well on KDD99 with a respectable 90.04% accuracy, but struggling with the more complex NSL-KDD dataset, only achieving 39% accuracy. This outcome aligns with literature findings that indicate that NB often faces challenges with imbalanced datasets and complex attack vectors due to its assumption of feature independence.

Studies like [26] emphasize the power of hybrid models for intrusion detection by blending DL with optimization algorithms like Artificial Raindrop and Harmony Search. Authors in [16] corroborate this by demonstrating the efficacy of classification algorithms like XGBoost when coupled with feature selection techniques such as ExtraTrees. This research further supports the use of hybrid approaches, which combine the strengths of multiple techniques to improve accuracy.

The XGBoost model's consistently high accuracy signifies its potential for real-world cybersecurity applications, as its precision and ability to detect various attack types ensure robust network security for IIoT and other similar environments. Although NB underperforms compared to XGBoost, its computational efficiency still makes it valuable in environments with limited computing resources.

However, the NSL-KDD dataset's inherent limitations in attack diversity and data imbalance negatively impacted NB's performance. Future research should focus on creating new datasets that better reflect the evolving cybersecurity landscape and provide comprehensive data for accurate model training. Despite XGBoost's high performance, it could remain susceptible to adversarial attacks, so future studies should explore hybrid models that combine ensemble methods with adversarial training.

Integrating explainable AI methods into models like XGBoost is crucial for practical applications, as this allows cybersecurity analysts to understand and trust the decision-making process. Our XGBoost model remains powerful due to its scalability, adaptability to different datasets, and ability to handle data imbalances, while the NB model is best suited for straightforward, resource-constrained datasets but struggles with complex cybersecurity data. Harnessing hybrid approaches that combine classifiers and optimization algorithms is promising, and future research should use different techniques to develop comprehensive cybersecurity solutions

VI. CONCLUSIONS AND FUTURE WORK

The study advances the knowledge in the Industrial Internet of Things (IIoT) domain by demonstrating the effectiveness of advanced machine learning algorithms, specifically XGBoost and Naive Bayes, in enhancing cybersecurity measures against an array of cyber threats. By overcoming the limitations inherent in traditional intrusion detection systems, our research highlights the superior accuracy and efficiency of these algorithms in detecting anomalies and security breaches within IIoT environments. The study's findings validate the robustness of the XGBoost algorithm in identifying and mitigating various cyber threats, outperforming conventional models in accuracy, precision, recall, and F1 score metrics. This advancement is crucial given the increasing complexity and volume of cyber-attacks in the IIoT domain. Moreover, the integration of machine learning techniques into cybersecurity frameworks as demonstrated in our research, offers a dynamic approach to protecting IIoT systems against evolving threats.

Future avenues of research could explore the fusion of different machine learning strategies to foster collaborative and distributed intrusion detection systems. This approach could further enhance real-time threat detection capabilities while maintaining data privacy and network efficiency in the IIoT context. Future studies ought to deliver a meticulous dissection of data preprocessing and feature selection, alongside the deployment of algorithms specifically crafted for IIoT systems. It is crucial to compile varied and exhaustive datasets that accurately delineate the subtleties of emergent cyber threats, particularly those involving advanced persistent threats, to ensure comprehensive coverage and robust defense mechanisms.

Our study not only contributes to the academic field but also serves as a valuable resource for industry practitioners aiming to bolster their cybersecurity defenses. our research thus

represents a significant step forward in securing IIoT networks, promoting safer and more reliable industrial operations.

REFERENCES

[1] S. Pal and Z. Jadidi, "Analysis of Security Issues and Countermeasures for the Industrial Internet of Things," *Applied Sciences*, vol. 11, no. 20, Jan. 2021, Art. no. 9393, https://doi.org/10.3390/app11209393.

[2] S. F. Tan and A. Samsudin, "Recent Technologies, Security Countermeasure and Ongoing Challenges of Industrial Internet of Things (IIoT): A Survey," *Sensors*, vol. 21, no. 19, Jan. 2021, Art. no. 6647, https://doi.org/10.3390/s21196647.

[3] A. J. G. de Azambuja, C. Plesker, K. Schützer, R. Anderl, B. Schleich, and V. R. Almeida, "Artificial Intelligence-Based Cyber Security in the Context of Industry 4.0—A Survey," *Electronics*, vol. 12, no. 8, Jan. 2023, Art. no. 1920, https://doi.org/10.3390/electronics12081920.

[4] S. H. Javed, M. B. Ahmad, M. Asif, S. H. Almotiri, K. Masood, and M. A. A. Ghamdi, "An Intelligent System to Detect Advanced Persistent Threats in Industrial Internet of Things (I-IoT)," *Electronics*, vol. 11, no. 5, Jan. 2022, Art. no. 742, https://doi.org/10.3390/electronics11050742.

[5] T. N. I. Alrumaih, M. J. F. Alenazi, N. A. AlSowaygh, A. A. Humayed, and I. A. Alablani, "Cyber resilience in industrial networks: A state of the art, challenges, and future directions," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 9, Oct. 2023, Art. no. 101781, https://doi.org/10.1016/j.jksuci.2023.101781.

[6] A.-A. Bouramdane, "Cyberattacks in Smart Grids: Challenges and Solving the Multi-Criteria Decision-Making for Cybersecurity Options, Including Ones That Incorporate Artificial Intelligence, Using an Analytical Hierarchy Process," *Journal of Cybersecurity and Privacy*, vol. 3, no. 4, pp. 662–705, Dec. 2023, https://doi.org/10.3390/jcp3040031.

[7] A. Salam, F. Ullah, F. Amin, and M. Abrar, "Deep Learning Techniques for Web-Based Attack Detection in Industry 5.0: A Novel Approach," *Technologies*, vol. 11, no. 4, Aug. 2023, Art. no. 107, https://doi.org/10.3390/technologies11040107.

[8] M. Ahsan, K. E. Nygard, R. Gomes, M. M. Chowdhury, N. Rifat, and J. F. Connolly, "Cybersecurity Threats and Their Mitigation Approaches Using Machine Learning—A Review," *Journal of Cybersecurity and Privacy*, vol. 2, no. 3, pp. 527–555, Sep. 2022, https://doi.org/10.3390/jcp2030027.

[9] M. A. Ferrag, O. Friha, L. Maglaras, H. Janicke, and L. Shu, "Federated Deep Learning for Cyber Security in the Internet of Things: Concepts, Applications, and Experimental Analysis," *IEEE Access*, vol. 9, pp. 138509–138542, 2021, https://doi.org/10.1109/ACCESS.2021.3118642.

[10] I. H. Sarker, "Deep Cybersecurity: A Comprehensive Overview from Neural Network and Deep Learning Perspective," *SN Computer Science*, vol. 2, no. 3, Mar. 2021, Art. no. 154, https://doi.org/10.1007/s42979-021-00535-6.

[11] J. Zhang, L. Pan, Q.-L. Han, C. Chen, S. Wen, and Y. Xiang, "Deep Learning Based Attack Detection for Cyber-Physical System Cybersecurity: A Survey," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 3, pp. 377–391, Mar. 2022, https://doi.org/10.1109/JAS.2021.1004261.

[12] M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," *Journal of Information Security and Applications*, vol. 50, Feb. 2020, Art. no. 102419, https://doi.org/10.1016/j.jisa.2019.102419.

[13] A. G. Putrada, N. Alamsyah, S. F. Pane, and M. N. Fauzan, "XGBoost for IDS on WSN Cyber Attacks with Imbalanced Data," in *International Symposium on Electronics and Smart Devices*, Bandung, Indonesia, Nov. 2022, pp. 1–7, https://doi.org/10.1109/ISESD56103.2022.9980630.

[14] J. L. Leevy, J. Hancock, R. Zuech, and T. M. Khoshgoftaar, "Detecting cybersecurity attacks across different network features and learners," *Journal of Big Data*, vol. 8, no. 1, Feb. 2021, Art. no. 38, https://doi.org/10.1186/s40537-021-00426-w.

[15] R. Alenezi and S. A. Ludwig, "Explainability of Cybersecurity Threats Data Using SHAP," in *Symposium Series on Computational Intelligence*, Orlando, FL, USA, Dec. 2021, pp. 1–10, https://doi.org/10.1109/SSCI50451.2021.9659888.

[16] G. Abdiyeva-Aliyeva, J. Aliyev, and U. Sadigov, "Application of classification algorithms of Machine learning in cybersecurity," *Procedia Computer Science*, vol. 215, pp. 909–919, Jan. 2022, https://doi.org/10.1016/j.procs.2022.12.093.

[17] N. A. Alsharif, S. Mishra, and M. Alshehri, "IDS in IoT using Machine Learning and Blockchain," *Engineering, Technology & Applied Science Research*, vol. 13, no. 4, pp. 11197–11203, Aug. 2023, https://doi.org/10.48084/etasr.5992.

[18] K. Aldriwish, "A Deep Learning Approach for Malware and Software Piracy Threat Detection," *Engineering, Technology & Applied Science Research*, vol. 11, no. 6, pp. 7757–7762, Dec. 2021, https://doi.org/10.48084/etasr.4412.

[19] R. Alsulami, B. Alqarni, R. Alshomrani, F. Mashat, and T. Gazdar, "IoT Protocol-Enabled IDS based on Machine Learning," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12373–12380, Dec. 2023, https://doi.org/10.48084/etasr.6421.

[20] A. B. Garcia, R. F. Babiceanu, and R. Seker, "Artificial Intelligence and Machine Learning Approaches For Aviation Cybersecurity: An Overview," in *Integrated Communications Navigation and Surveillance Conference*, Dulles, VA, USA, Apr. 2021, pp. 1–8, https://doi.org/10.1109/ICNS52807.2021.9441594.

[21] A. Sentuna, A. Alsadoon, P. W. C. Prasad, M. Saadeh, and O. H. Alsadoon, "A Novel Enhanced Naïve Bayes Posterior Probability (ENBPP) Using Machine Learning: Cyber Threat Analysis," *Neural Processing Letters*, vol. 53, no. 1, pp. 177–209, Feb. 2021, https://doi.org/10.1007/s11063-020-10381-x.

[22] S. Ismail and H. Reza, "Evaluation of Naïve Bayesian Algorithms for Cyber-Attacks Detection in Wireless Sensor Networks," in *IEEE World AI IoT Congress*, Seattle, WA, USA, Jun. 2022, pp. 283–289, https://doi.org/10.1109/AIIoT54504.2022.9817298.

[23] O. Illiashenko, V. Kharchenko, I. Babeshko, H. Fesenko, and F. Di Giandomenico, "Security-Informed Safety Analysis of Autonomous Transport Systems Considering AI-Powered Cyberattacks and Protection," *Entropy*, vol. 25, no. 8, Aug. 2023, Art. no. 1123, https://doi.org/10.3390/e25081123.

[24] "SIGKDD : KDD Cup 1999 : Computer network intrusion detection." https://www.kdd.org/kdd-cup/view/kdd-cup-1999/Dat.

[25] "NSL-KDD." https://www.kaggle.com/datasets/hassan06/nslkdd.

[26] M. G. Raj and S. K. Pani, "Intrusion Detection System using Long Short Term Memory Classification, Artificial Raindrop Algorithm and Harmony Search Algorithm," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 12, pp. 95–103, 2022, https://doi.org/10.14569/IJACSA.2022.0131214.