CONESCAPANHONDURAS2025paper15.pdf

Institute of Electrical and Electronics Engineers (IEEE)

Document Details

Submission ID

trn:oid:::14348:477909394

Submission Date

Aug 1, 2025, 1:12 PM CST

Download Date

Aug 12, 2025, 12:07 PM CST

File Name

CONESCAPANHONDURAS2025paper15.pdf

File Size

1.2 MB

5 Pages

4,039 Words

23,819 Characters

20% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Match Groups

59 Not Cited or Quoted 18%
Matches with neither in-text citation nor quotation marks

7 Missing Quotations 1%
Matches that are still very similar to source material

2 Missing Citation 1%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

16%	Internet sources
19%	Publications
0%	Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **59** Not Cited or Quoted 18%
  Matches with neither in-text citation nor quotation marks
- **7** Missing Quotations 1%
  Matches that are still very similar to source material
- **2** Missing Citation 1%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

16%	Internet sources
19%	Publications
0%	Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	**jurnal.polibatam.ac.id**	**1%**
2	Internet	**yamanashi.repo.nii.ac.jp**	**1%**
3	Internet	**ijeecs.iaescore.com**	**1%**
4	Internet	**www.astesj.com**	**1%**
5	Internet	**oa.upm.es**	**<1%**
6	Internet	**eprint.innovativepublication.org**	**<1%**
7	Internet	**kc.umn.ac.id**	**<1%**
8	Internet	**ebin.pub**	**<1%**
9	Internet	**aries.ucsd.edu**	**<1%**
10	Internet	**www.ijritcc.org**	**<1%**

11	Publication	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dhirendra Kumar Shukla. "Intelli...	<1%
12	Publication	Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical ...	<1%
13	Internet	etasr.com	<1%
14	Publication	H L Gururaj, Francesco Flammini, V Ravi Kumar, N S Prema. "Recent Trends in He...	<1%
15	Internet	aaltodoc.aalto.fi	<1%
16	Internet	publichealth.jmir.org	<1%
17	Internet	ceur-ws.org	<1%
18	Internet	assets-eu.researchsquare.com	<1%
19	Internet	www.mdpi.com	<1%
20	Internet	ses.library.usyd.edu.au	<1%
21	Internet	www.geeksforgeeks.org	<1%
22	Publication	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dhirendra Kumar Shukla. "Artific...	<1%
23	Publication	Teppei Honda, Kimika Ymamoto, Hiroshi Hasegawa. "TABIMAE Mind Search Syste...	<1%
24	Internet	kth.diva-portal.org	<1%

25	Publication	Ioannis Makris, Aikaterini Karampasi, Panagiotis Radoglou-Grammatikis, Nikolao...	<1%
26	Publication	Ibrahim, Juma. "An Architecture for Network Traffic Anomaly Detection System B...	<1%
27	Publication	Matúš Čavojský, Gabriel Bugár, Dušan Levický. "Comparative Analysis of Feed-For...	<1%
28	Publication	D. Jeya Mala, Anto Cordelia Tanislaus Antony Dhanapal, Saurav Sthapit, Anita Kha...	<1%
29	Publication	Jinyan Wang, Guangquan Xu, Wenqing Lei, Lixiao Gong, Xi Zheng, Shaoying Liu. "...	<1%
30	Publication	Md. Abdur Rahman, M. Shamim Hossain, M. Saiful Islam, Nabil A. Alrajeh, Ghulam...	<1%
31	Internet	aircconline.com	<1%
32	Publication	Ao Xiong, Meng Chen, Shaoyong Guo, Yongjie Li, Yujing Zhao, Qinghai Ou, Chuan ...	<1%
33	Publication	Hakan Can Altunay, Zafer Albayrak. "A hybrid CNN+LSTM-based intrusion detecti...	<1%
34	Publication	Idowu, Ifedotun Roseline. "Improved Meta-Heuristic Based RNS Techniques for In...	<1%
35	Publication	Singh, Gurvinder. "Real-Time Quantum Computing Anomaly Detection Model on ...	<1%
36	Publication	Yanru Chen, Shijia Liu, Zilin Wang, Dizhi Wu, Yang Li, Bin Xing, Bing Guo, Liangyin ...	<1%
37	Internet	dl.ifip.org	<1%
38	Internet	openaccess.city.ac.uk	<1%

39	Internet	openaccess.uoc.edu	<1%
40	Publication	MohammadHossien Alishahi, Paul Fortier, Ming Zeng, Quoc-Viet Pham, Xingwang...	<1%
41	Publication	Vu, Ly Thi. "Deep Neural Network for Anomaly Detection", University of Technolo...	<1%
42	Publication	Alaa S. Hrizat, Robert P. Post, Allison Goldberg. "Exploring the efficacy of an artifi...	<1%
43	Publication	M. Scholten. "The modular habitat model (MHM) for the ide, Leuciscus idus (L.) - a...	<1%
44	Publication	Mehdi Selem, Farah Jemili, Ouajdi Korbaa. "Deep Learning for Intrusion Detection...	<1%
45	Publication	Tibebu Bekele Shana, Neetu Kumari, Mayank Agarwal, Samrat Mondal, Upaka Ra...	<1%
46	Publication	Yassine Himeur, Abdullah Alsalemi, Faycal Bensaali, Abbes Amira. "Smart power c...	<1%

Intrusion Detection in IoT Networks: A Comparative Analysis of Machine Learning Models

***Abstract*—The growing adoption of IoT devices has increased the attack surface in networks, demanding efficient and adaptable intrusion detection systems. This study evaluates the viability of classical machine learning models Random Forest, LightGBM, and Naive Bayes to detect intrusions in IoT networks under computational resource constraints, using the UNSW-NB15 dataset, which contains more than two million records with 49 features. Preprocessing was applied, including cleaning, removal of irrelevant attributes, variance analysis, one-hot encoding, and normalization. Subsequently, the dataset was reduced through stratified sampling to 20,000 records for training and 10,000 for testing, maintaining class proportionality. The models were evaluated using precision, recall, and F1-score, resulting in Random Forest achieving the best performance (precision 0.813, recall 0.858, F1-score 0.814), highlighting its balance between accuracy and efficiency. LightGBM showed lower performance (precision 0.563, F1-score 0.367), attributed to hyperparameter sensitivity in a resource-limited environment, while Naive Bayes obtained intermediate results (precision 0.728, F1-score 0.623), proving useful in contexts where efficiency is a priority. The classical models Random Forest and Naive Bayes represent practical and effective alternatives for IoT environments with restricted resources, although the exploration of advanced techniques such as federated learning or lightweight neural networks is suggested for future research.**

Index Terms IoT network security, intrusion detection, machine learning, traffic type, UNSW-NB15.

I. Introduction

Intrusion detection in Internet of Things (IoT) networks is a priority topic in the field of cybersecurity due to the constant growth in the number of devices connected to the network. From smart appliances to industrial sensors, connectivity has become a fundamental part of many daily and business activities. However, this expansion has brought with it increased exposure to cyberattacks, as each new device represents a potential entry point for malicious actors [10].

Connected devices often have limitations in processing power, storage, and energy, which makes it difficult to implement advanced security systems. In addition, their geographical distribution and the diversity of manufacturers create heterogeneous environments where security standards vary considerably [6]. These characteristics make traditional protection methods not always suitable, requiring specific solutions that adapt to these conditions [7].

Detecting intrusions in these environments involves facing several technical challenges. First, network traffic in these systems can be highly variable, as different devices have different behavior patterns. Second, the volume of data generated is high, making real-time analysis difficult. And third, it is not feasible to apply techniques that require large computational resources directly on the devices, as this could affect their normal operation [2], [8].

Although in recent years deep learning-based solutions have been developed that offer high precision in attack detection, these require elevated computational capabilities for both training and execution, which limits their use in real devices. For example, Altunay and Albayrak developed a system that combines convolutional and long short-term memory networks, achieving more than 99% accuracy, but with a high computational cost [2]. Similarly, Fosić et al. achieved good results using Random Forest, XGBoost, and Support Vector Machines, but also in more controlled environments [1].

In search of more practical solutions for resource-constrained environments, other researchers have explored federated or distributed learning. These strategies allow processing to be carried out locally, avoiding the need to send all data to a central server and reducing bandwidth consumption, while improving privacy [8], [9]. However, these solutions still present significant technical challenges, such as model synchronization and adaptation to very diverse hardware.

In this work, a more accessible and straightforward alternative is proposed: to evaluate the performance of three classical supervised learning models—Random Forest, LightGBM, and Naive Bayes—on the UNSW-NB15 dataset. This dataset has been widely used in previous studies due to its variety of records that include both legitimate and malicious traffic [1]. The objective is to determine whether these models, known for their efficiency and low resource requirements, can achieve good performance in intrusion detection without the need for complex architectures.

Random Forest has proven to be robust against overfitting and effective with mixed data, both numerical and categorical. LightGBM, in turn, offers speed and accuracy in training thanks to its optimized structure, even with large volumes of data [1]. Naive Bayes is a simple and fast model that, despite its limitations in complex scenarios, can be useful when a lightweight solution is needed. Evaluating these models in a realistic environment provides a concrete view of the balance between accuracy and efficiency. The results obtained show that it is possible to achieve competitive performance levels with simple models, which opens the door to implementing practical security solutions on real devices, without requiring advanced resources. This approach can be key to improving security in smart homes, connected factories, and other Internet of Things applications, where technical limitations are a constant [10].

II. Materials and Methods

A. Data Acquisition

For the development of this study, the UNSW-NB15 dataset was used, generated by the Australian Centre for Cyber Security and widely employed in research on network intrusion detection. This dataset contains simulated network traffic records that include both benign behaviors and multiple types of attacks, such as exploits, denial of service, unauthorized access, port scanning, and malware, which allows for training and evaluating intrusion detection models in a controlled but representative environment. The dataset contains 49 features per record, combining numerical variables, categorical variables, and class labels. The full dataset contains more than two million records, organized in CSV files ready for analysis [5].

Given the focus of this work on scenarios with limited computational resources, a reduction in the size of the dataset was chosen. A stratified random sample of 20,000 records was selected for training and 10,000 records for testing, preserving the original proportion between normal and anomalous classes. This decision was based on the need to optimize memory usage and processing time without compromising the representativeness of the classes or the validity of the results, in line with common practices in similar research with limited resources [1], [2].

B. Data Preporcessing

Figure 1 illustrates the flowchart that summarizes the key stages of data preprocessing applied to the UNSW-NB15 dataset. This diagram highlights the sequential steps carried out, from the initial exploration and removal of irrelevant attributes to one-hot encoding and Min-Max normalization. This graphical representation clearly visualizes the dataset transformation process, ensuring its suitability for the supervised models used in the study.

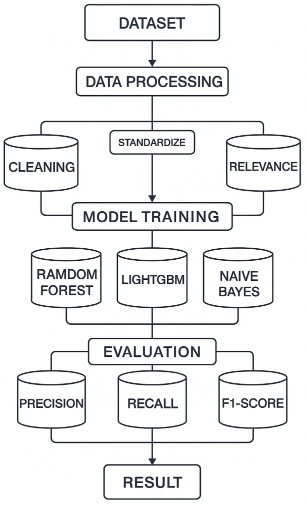

Fig. 1. Flowchart of intrusion detection using Random Forest, LightGBM, and Naive Bayes

Initially, the data were loaded and explored in a Python environment using the pandas, numpy, and matplotlib libraries. This stage made it possible to identify the presence of missing data, types of variables, and class distribution. It was confirmed that the dataset had a significant imbalance, with benign classes predominating, which may hinder the models' ability to correctly detect attack patterns, as noted in previous studies [1], [5].

During the attribute cleaning and selection phase, irrelevant columns such as unique identifiers and timestamps were removed, as they did not provide predictive value to the model. In addition, an exploratory analysis of variance and correlation between attributes was applied to discard redundant or low-relevance variables, following the methodological recommendations of similar studies [1], [4].

Subsequently, the encoding of categorical variables, such as protocol type or network service, was addressed using the one-hot encoding technique. This transformation generated binary variables that can be interpreted by supervised models, allowing categorical information to be preserved without introducing artificial order. This practice has been widely validated in classification applications with models such as Random Forest and Naive Bayes. Numerical variables, in turn, were normalized using Min-Max scaling, adjusting their values to the range (0, 1). This normalization was essential to ensure that no numerical feature disproportionately influenced the model, which is especially important in algorithms sensitive to data magnitude such as LightGBM [10].

Finally, the preprocessed dataset was divided into two subsets: one training set composed of 20,000 records and one testing set with 10,000 records. Both subsets were selected in a stratified manner to preserve the original proportion between benign and malicious classes. This strategy, in addition to reducing the computational load, allowed for an objective and reproducible evaluation of model performance under conditions that reflect real-world constraints in IoT environments [2].

C. Model Selection

The implementation of the selected machine learning models for the intrusion detection task was carried out. The choice of algorithms was based on their track record of good performance in network traffic classification problems and their low computational requirements, which is crucial in IoT network contexts with limited resources. After conducting preliminary tests with other classifiers such as logistic regression and Extra Trees, which showed lower performance or high training times, it was decided to work with three supervised models: Random Forest, LightGBM, and Naive Bayes. The Random Forest model was selected for its ability to handle both numerical and categorical variables, its resistance to overfitting, and its proven effectiveness in classification problems with heterogeneous and imbalanced data [1]. This model was implemented using scikit-learn, which included feature scaling, one-hot encoding, and hyperparameter tuning such as the number of trees, maximum depth, and class balancing. Its ensemble structure of multiple decision trees allows it to capture complex relationships 14 between variables without requiring high computational capacity. On the other hand, the LightGBM model was included, a decision tree-based algorithm with a highly optimized gradient boosting approach. LightGBM has proven 22 17 to be efficient in classifying large volumes of data, offering fast training times and relatively low memory consumption thanks to its histogram-based growth structure and leaf-wise 2 tree growth strategy [4].

However, its performance can be affected by the need 39 for careful hyperparameter configuration, which posed a challenge under the limitations of the working environment. The third model implemented was Naive Bayes, known for 12 its simplicity and speed. Despite assuming independence 18 between features a condition that is rarely met in real-world environments this model can provide acceptable results in traffic classification tasks, especially when a lightweight and fast solution is required. The Gaussian version of the classifier was used, which is appropriate when the numerical 45 variables exhibit an approximately normal distribution [10].

D. Comparasion with other models

23 In order to contextualize the performance of the models used in 16 this study, a comparison was made with previous works that have used both the UNSW-NB15 dataset and other datasets commonly applied in intrusion detection, such as NETFLOW. In addition, studies that worked with BoT-IoT and X-IIoTID, datasets specifically designed to represent malicious traffic scenarios in IoT environments, were considered as references, allowing for a broader framework of relative evaluation [1], [2].

18 This comparison includes various supervised machine learning approaches, such as Support Vector Machines, k-Nearest 33 Neighbors, as well as advanced deep learning methods, including hybrid architectures like convolutional networks 1 combined with LSTM [2], [12]. Additionally, boosting algorithms such as XGBoost are included, which have been widely used for their effectiveness in tasks involving tabular data classification and network traffic, especially when 37 significant computational resources are available [1], [4].

E. Computational Resources

3 During the training process, a personal computer with an 1 Intel Core i5-10300H processor at 2.50 GHz and 8 GB 12 of RAM was used, without access to graphical processing units (GPUs) or high-performance servers. This environment imposed clear limitations on the dataset size, training times, and the complexity of the models that could be efficiently implemented, which was decisive in the choice of algorithms and preprocessing techniques.

III. Results

A. Evaluation Metrics

To effectively evaluate the performance of classification models in detecting intrusions in IoT networks, standard metrics widely used in the fields of cybersecurity and machine learning were selected, namely: precision, recall, and F1-score. These metrics provide a comprehensive view of the model's ability to identify both normal events and attacks within network traffic. Precision (1) evaluates the proportion of instances correctly classified as malicious over the total number of positive predictions, which is crucial to avoid false alarms that could overload IoT systems. Recall (2), on the other hand, measures the model's ability to correctly detect malicious events present in the traffic, thereby minimizing undetected attacks and strengthening the overall security of the system. Finally, the F1-score (3) represents a harmonic mean between precision and recall, balancing both aspects to provide a more complete performance evaluation, especially relevant in scenarios with imbalanced classes where normal traffic outweighs malicious traffic. The inclusion of these metrics allows for a comparative analysis not only of the overall performance of the model but also of its effectiveness under realistic conditions, considering resource limitations and the need for fast and efficient processing in IoT environments.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{1}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{2}$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

B. Results

The results obtained with the Random Forest, LightGBM, and Naive Bayes models applied to the UNSW-NB15 dataset are summarized in Table I. The Random Forest model achieved the best performance, reaching a precision of 0.813, a recall of 0.858, and an F1-score of 0.814. These results reflect its ability to handle both numerical and categorical variables, as well as its robustness against class imbalances and data noise. Reducing the size of the training set was essential to control memory consumption and training time without significantly affecting model quality an important aspect for applications in resource-constrained environments with realtime requirements. On the other hand, LightGBM showed lower performance with a precision of 0.563, recall of 0.531, and F1-score of 0.367. Although known for its efficiency and speed, its performance here may have been affected by the dataset's complexity and the need for more detailed hyperparameter tuning. Additionally, its considerable resource consumption hindered its use under the study's constraints.

The Naive Bayes model achieved intermediate values with a precision of 0.728, recall of 0.722, and F1-score of 0.623. Its simplicity and fast training time are advantageous, although the assumption of feature independence does not always hold in network traffic. However, it remains a valid option when computational efficiency is a priority, maintaining acceptable performance for detection. Finally, XGBoost was discarded due to its high computational requirements, reaffirming the need to balance precision and efficiency in intrusion detection for IoT environments.

TABLE I
RESULTS OF THE CLASSIFICATION MODELS ON THE UNSW-NB15 DATASET

Model	Precision	Recall	F1-score
Random Forest	0.813	0.858	0.814
LightGBM	0.563	0.531	0.367
Naive Bayes	0.728	0.722	0.623

The graph presented in Figure 2 consists of two key parts for evaluating the performance of the binary classification model applied to intrusion detection. First, the ROC curve (Receiver Operating Characteristic), located on the left, shows the relationship between the true positive rate and the false positive rate. The shape of the curve, away from the diagonal and close to the upper left corner, indicates outstanding model performance, supported by an area under the curve (AUC) value of 0.942. Second, the Precision-Recall curve, on the right side of the image, illustrates the balance between precision and recall of the model, which is especially useful in contexts with imbalanced classes such as network traffic. The area under this curve (AP) is 0.967, confirming that the model maintains high precision even while maximizing the recall of positive instances. This evidence shows that the model achieves solid and reliable performance for critical tasks such as intrusion detection.

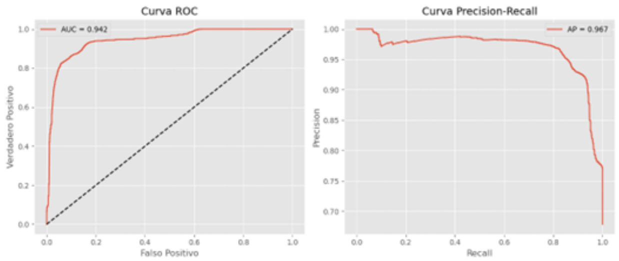

Fig. 2. Graph curve: ROC and Precision-Recall

C. Comparison with Other Studies

Table II presents a comparison of the results of this study with previous research that used different datasets, including NETFLOW and UNSW-NB15, as well as models like CNN+LSTM, XGBoost, SVM, and kNN. This comparison helps position the performance of our models, particularly Random Forest, within the broader context of the literature. While our results are lower than those achieved by hybrid models and more complex techniques, they stand out for offering a good balance between precision and efficiency under real-world resource constraints, which is crucial for applications in IoT environments [2].

When comparing the results obtained in this study with previous research, significant differences in model performance are observed across different datasets and configurations. Our Random Forest model achieved a precision of 0.813, recall of 0.858, and F1-score of 0.814, values that, although lower than those reported in other studies with different datasets, show solid performance under the specific conditions of limited resources and reduced training set size.

In the intrusion detection study on the NETFLOW dataset, the Random Forest model showed significantly higher results, with a precision of 97.68, recall of 98.47, and F1-score of 98.07 [1]. This suggests that the characteristics of the dataset and possibly the larger amount of available data positively influence the model's performance. Similarly, XGBoost and SVM also achieved high metrics on NETFLOW, surpassing the values we obtained with LightGBM (precision 0.563, recall 0.531, and F1-score 0.367) and Naive Bayes (precision 0.728, recall 0.722, and F1-score 0.623). Additionally, hybrid models combining autoencoders with LSTM have shown promising results in anomaly detection on time-series data [15].

Regarding the UNSW-NB15 dataset, another study implemented a CNN+LSTM hybrid model that achieved superior metrics (precision 99.1, recall 99.2, F1-score 99.1)[2]. Good results were also reported with SVM, Naive Bayes, and kNN, with precision ranging from 83.2 to 88.5 and F1-scores between 81.8 and 87.6, showing better performance than the LightGBM and Naive Bayes models trained in our study but in line with the performance of Random Forest. Federated learning approaches and anomaly detection based on LSTM and autoencoders, especially optimized for time-series data in IoT environments, have also been proposed [11].

These differences reflect how the type of model, dataset complexity, preprocessing, and available computational resources influence the results. Our work emphasizes a pragmatic balance between precision and efficiency, essential for applications in IoT environments with hardware constraints, a less critical situation in the compared studies where higher precision is achieved at the cost of greater resource consumption. Although models like CNN+LSTM and Random Forest over more complete and optimized datasets offer better metrics, our results show that with an efficiency-oriented approach adapted to real limitations, it is possible to achieve competitive performance in intrusion detection for IoT.

D. Contributions of the Study

This study provides practical evidence on the performance of classic machine learning models for intrusion detection in IoT networks, focusing on environments with computational resource limitations. It is confirmed that algorithms like Random Forest and Naive Bayes represent viable alternatives, as they offer a good balance between precision and efficiency,

TABLE II
COMPARISON OF RESULTS FROM PREVIOUS STUDIES ON INTRUSION DETECTION MODELS FOR NETWORKS

DATASET	Model	Precision	Recall	F1-Score
NETFLOW	RF	97.68	98.47	98.07
NETFLOW	XGB	96.84	97.12	96.98
NETFLOW	SVM	95.32	96.21	95.76
UNSW-NB15	CNN+LSTM	99.1	99.2	99.1
UNSW-NB15	SVM	88.5	86.7	87.6
UNSW-NB15	NB	83.2	80.5	81.8
UNSW-NB15	kNN	85.9	83.1	84.5

making them useful for implementations on devices with limited capacity, without significantly compromising detection quality. Additionally, it is observed that more complex models, such as XGBoost or deep learning-based techniques, although potentially more accurate under ideal conditions, require greater processing power and memory, which limits their practical application in IoT environments with restricted resources [3], [4].

In particular, XGBoost was discarded due to its high computational requirements, while LightGBM showed lower-than-expected performance, possibly related to the complexity of the reduced dataset and the need for more fine-tuned hyperparameter adjustments, which were difficult to optimize given the environment's limitations.

During the training process, additional challenges related to preprocessing arose, where the incorporation of encoders for categorical variables required the progressive inclusion of more interactions and feature combinations to improve data representation. This added complexity to the pipeline and demanded greater care in the selection and transformation of variables, especially under memory and computational time constraints.

Reducing the size of the dataset was a key strategy to control memory consumption and training time without significantly affecting model quality, although it represented a trade-off between data quantity and performance. The methodology applied, which includes dataset reduction and hyperparameter tuning, helps optimize model performance under these conditions, offering an approach that can be useful for future developments in this field [1], [5].

IV. CONCLUSIONS

Intrusion detection in IoT networks remains a significant challenge, particularly under constraints of limited computational resources. This study demonstrated that classical machine learning models such as Random Forest, LightGBM, and Naive Bayes can deliver acceptable performance in these environments, with Random Forest achieving the best results in terms of precision and recall. However, its performance did not surpass that of more advanced or hybrid models reported in previous research. LightGBM and Naive Bayes stood out for their speed and simplicity, making them suitable options when computational efficiency is prioritized, despite their lower precision. The XGBoost model, although known for its effectiveness, was not feasible in this study due to high resource demands and long training times, highlighting the practical limitations of using complex algorithms in resource-constrained IoT contexts.

In conclusion, while the evaluated models do not represent the highest achievable performance, they offer a solid and realistic foundation for building practical intrusion detection systems in IoT environments. Future research should explore the integration of more sophisticated or hybrid techniques to enhance detection capabilities without significantly compromising efficiency, ensuring adaptability to the constraints of real-world IoT deployments.

.

REFERENCES

[1] I. Fosić, D. Zagar, K. Grgić, and V. Križanović, "Anomaly detection in NetFlow network traffic using supervised machine learning algorithms," Computers, Materials Continua, vol. 70, no. 1, pp. 1015-1029, 2022.

[2] H. C. Altunay and Z. Albayrak, "A hybrid CNN+LSTM-based intrusion detection system for industrial IoT networks," IEEE Access, vol. 10, pp. 110819-110828, 2022.

[3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in *Magnetism*, vol. III, G. T. Rado and H. Suhl, eds. New York, NY, USA: Academic, 1963, pp. 271–350.

[4] T. T. Nguyen and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," IEEE Communications Surveys Tutorials, vol. 10, no. 4, pp. 56-76, 2008.

[5] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," 2015 Military Communications and Information Systems Conference (MilCIS), Canberra, Australia, 2015, pp. 1-6.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740–741, Aug. 1987.

[7] N. Mishra and S. Pandya, "Internet of Things applications, security challenges, attacks, intrusion detection, and future visions: A systematic review," *IEEE Access*, vol. 9, pp. 59353–59377, 2021, doi: 10.1109/ACCESS.2021.3075439.

[8] M. Le, T. Huynh-The, T. Do-Duy, T.-H. Vu, W.-J. Hwang, and Q.-V. Pham, "Aplicaciones del Aprendizaje Automático Distribuido para el Internet de las Cosas: una revisión exhaustiva," *arXiv preprint*, arXiv:2310.10549, 2023.

[9] P. García Santaclara, A. Fernández Vilas, and R. P. Díaz Redondo, "Prototipo de implementación de Aprendizaje Federado con dispositivos IoT," *arXiv preprint*, arXiv:2311.14401, 2023.

[10] M. R. Lyu, "Integración de Aprendizaje Automático en el Internet de las Cosas: desafíos y oportunidades", IEEE Internet of Things Journal, vol. 6, no. 3, pp. 456-465, junio 2019.

[11] Y. Liu *et al.*, "Deep anomaly detection for time-series data in industrial IoT: A communication-efficient on-device federated learning approach," *arXiv preprint*, arXiv:2007.09712, 2020.

[12] A. Shone, V. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 2, no. 1, pp. 41-50, Feb. 2018.

[13] W. Wang, M. Zhu, J. Wang, X. Zeng, and Z. Yang, "End-to-end encrypted traffic classification with one-dimensional convolution neural networks," in *Proc. 2017 IEEE Int. Conf. Intell. Secur. Informatics (ISI)*, Beijing, China, 2017, pp. 43–48.

[14] L. Liu, H. Zhu, W. Zhang, and Y. Liu, "Federated learning for privacy-preserving intrusion detection in industrial cyber-physical systems," *IEEE Trans. Ind. Informatics*, vol. 18, no. 3, pp. 1972–1983, Mar. 2022.

[15] Y. Wei *et al.*, "LSTM-Autoencoder based anomaly detection for indoor air quality time series data," *arXiv preprint*, arXiv:2204.06701, 2022.