INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

Optimizing Algorithms to integrate Advanced Machine Learning Methods for Predictive Maintenance in Industrial IoT

C. NAGA SWAROOPA
Research Scholar
Department of Computer Science and Engineering
J. S. UNIVERSITY, Shikohabad, UP

Dr. K. VIJAYA BHASKAR
Associate Professor
Department of Computer Science and Engineering
J. S. UNIVERSITY, Shikohabad, UP

ABSTRACT

Upkeep is performed before issues manifest through the use of a wide range of specialized machine learning approaches making it possible for systems or machines to predict and decrease a wide variety of machine failures. The term "predictive maintenance" (PdM) refers to a tool that is rapidly becoming an indispensable instrument for the purpose of improving the efficiency and dependability of industrial machinery while simultaneously enhancing the management of maintenance activities. Using predictive maintenance, which use machine learning algorithms to proactively identify and rectify potential equipment flaws, businesses have the potential to minimize unanticipated downtime, maintenance costs, and operational efficiency.

The Internet of Things (IoT) is a cutting-edge technology that enables commonplace objects to establish connections with one another and to create a global communication network via the sharing of data and the online replies to that data. Kevin Ashton was the first person to use the term "Internet of Things" in the year 1999. Things like self-driving cars, smart televisions, and RFID tags that are utilized in supply networks are all examples. Since its inception, the Internet of Things (IoT) has brought about world-wide transformations, and this trend is expected to continue in the years to come. The term "Internet of Things" refers to a network of networked computer devices that are able to detect, gather, store, and distribute information about their surrounding physical world.

Recently, research into deep learning (DL) techniques including Convolutional Neural Networks (CNN), Gated Recurrent Units (GRU), and Long Short-Term Memory (LSTM) has been concentrating on anomaly identification as a primary area of investigation. This thesis presents a unique hybrid DL-enabled strategy with the purpose of providing the essential security evaluations prior to successful attacks on infrastructure that is interconnected with the Internet of Things (IoT). XGBoost, Autoencoder, Convolutional Neural Networks (CNN), Gated Recurrent Units (GRU), Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN) are used in the construction, training, testing, and verification of a great number of hybrid models.

I. INTRODUCTION

In the twenty-first century, the Internet of Things (IoT) transformed into an indispensable portion of everyday lives, having wide range of applications ranging from health, home automation, smart metering, and smart cars to the Smart Factory. Tremendous advances in sub-evaluation, micro appliances manufacturing, and server connections had empowered innovative IoT products revolutionizing an extensive variety of communication-based operations.

1.1 Industrial Internet of Things (IIoT)

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

The IoT is frequently called a technical breakthrough capable of tackling the bulk of today's social challenges, such as smarter towns, traffic monitoring, pollutant management, and connected health, to name a very few. The IIoT is a subsection of IoT that comprises machine-to-machine (M2M) and industries technology breakthroughs with robotic techniques. Figure 1. depicts how IIoT evolved from IoT and Industry 4.0. IIoT opened doors for more efficient and sustainable management of manufacturing processes. Each of the available networks has a small coverage area and therefore requires small transmit power.

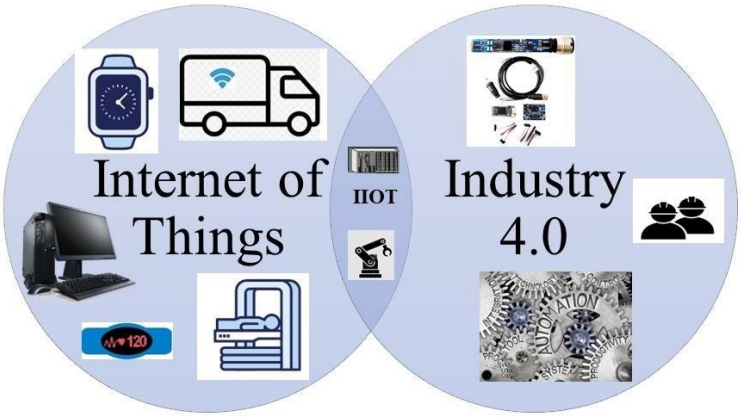


Figure 1. The Industrial Internet of Things

IoT technology helps IIoT to improve build operations and industrial processes. It is a method that captures, shares, and interprets information through a distributed system interconnected by communication networks, allowing for speedier decision-making. As instance, flaws inside the equipment of a commercial operation can be forecasted and rectified in a timely fashion before a section of it breaks or the entire equipment fails. IIoT helps to establish a connected home workplace by saving time, resources, and cost, and such small instruments are utilized through the conventional ICS.

IIoT is a term coined to define the amalgamation of smart electronic devices into industrial processes throughout the entire life cycle of a product (Zhang *et al.* 2017). IIoT offers industrial systems with connectivity and intelligence through sensors and actuators with pervasive computing and networking aptitudes. In smart industries, IIoT not only enables machine-to- machine communication but also provide interaction between human and smart machines. Hence, it is becoming a significant element of present and future industrial systems.

1.2 Intrusion Detection in IIoT Networks

The application of IoT in the industrial sector is highly critical when service interruption might be unbearable. Hence, protecting these IIoT systems against cyberattacks cannot be inflated. Therefore, the IDS model is a very significant security mechanism that keeps track of each data packet for abnormal behaviors.

In addition, the IDS can be categorized according to their placement tactics of the IIoT network where it is installed. In the centralized approach, IDS examines each data packet that goes in and out of the boundary router from the connected sensors (Kasinathan et al. 2013). However, this approach may not recognize internal attacks within an IIoT system (Wallgren et al. 2013). In the decentralized deployment approach, the IDSs are distributed across

ISSN: 2057-5688

the different devices in the IIoT network. Owing to the resource-constrained nature of IIoT networks, each system must be optimized autonomously, and the IDS must be lightweight (Mastorakis et al. 2020). The devices in an IIoT network that are configured to monitor attacks in adjacent expedients are called watchdogs (Gyamfi & Jurcut 2022). Whenever a watchdog identifies a cyberattack, it disseminates an alert to the other devices to defend them from the invader. The hybrid IDS deployment tactics assimilate the method of centralized and decentralized placement strategy.

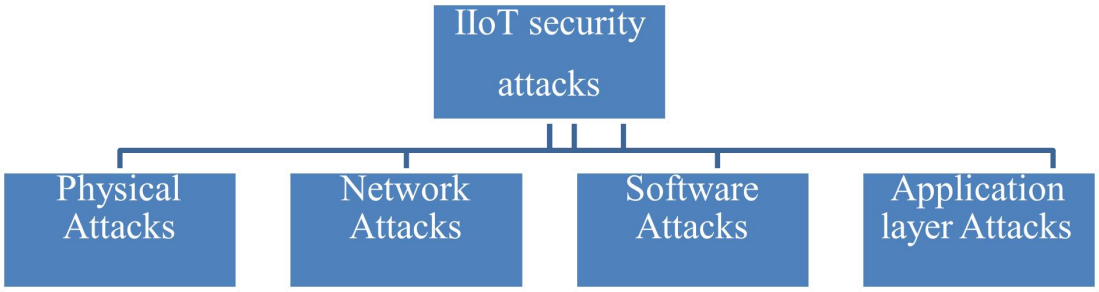
II. ATTACK VECTORS AGAINST IIOT AND IOT DEVICES

A digital attack is a malicious and purposeful attempt by a person or organization to breach another individual or organization's information structure. In most cases, the hacker tries to achieve an edge by breaching the victim's system (Sengupta et al., 2020). Figure 2 displays many kinds of cyber threat.

2.1 Iiot Security Model

IIoT security attacks are classified in four sections as illustrated in Figure 2: (a) Physical Attacks (b) Network Attacks (c) Software Attacks and (d) Application Layer Attacks.

Figure 2. IIoT security attacks associated with various layers Physical Attacks



A hacker might acquire unauthorized distant entry sites in physical assaults. As a result, equipment may fail at really weak permissible level as well as alarms may get failed for producing sound whenever normally ought. A further option is whether the attacker, after acquiring unauthorized access, changes all subscriber display parameters; such whenever an alert may go through, that operator is ignorant of that too. It might create considerable lag for normal reaction to something like an urgent situation, putting individuals around the facility in danger. PLCs and RTUs are examples of embedded elements, which execute programs. Firmware assaults, including such fault injection and backdoors, may make such units susceptible. Malicious hackers might exploit such firmware vulnerabilities for gaining entry into sensitive data and otherwise refuse assistance. One of the

for a device but also modify its data frame examples of physical attacks is reconnaissance threat. That is the early phase across any network assault. Attackers utilize scanning methods to analyze overall architecture of both the target system and locate gadgets and weaknesses. Reconnaissance attacks can be categorized as: Address scan, Function code scan, Device identification attack and Points scan. The address scan identifies ICS systems that are linked to a connection.

2.2 The proposed intrusion detection mechanism.

This part outlines the suggested invasion identification technique, which is built on ML processes. Our invasion identification mechanism's basic idea would be to teach ML algorithms to identify assaults in a local network. As a result, we may split our

ISSN: 2057-5688

evolution into two stages: (a) gather a relevant database to instruct and testing ML systems; (b) educate as well as evaluate ML prototypes on sample along with assessing its effectiveness in a home network.

➢ The dataset development

The database is the primary element utilized to educate and evaluate an ML technique since the ML methods understand from the information. As a result, selecting a research sample is a critical phase in the evolution of an ML program. In our situation, we are developing a ML system to identify assaults on IoT medical equipment on the local system. It involves using the database consisting network traffic traces, and so such database must contain both regular and targeted data.

➢ Numerical results and discussion

This part describes efficiency in ML algorithm produced in this study. Figure 3. depicts the accuracy measure for all ML systems employed in this study.

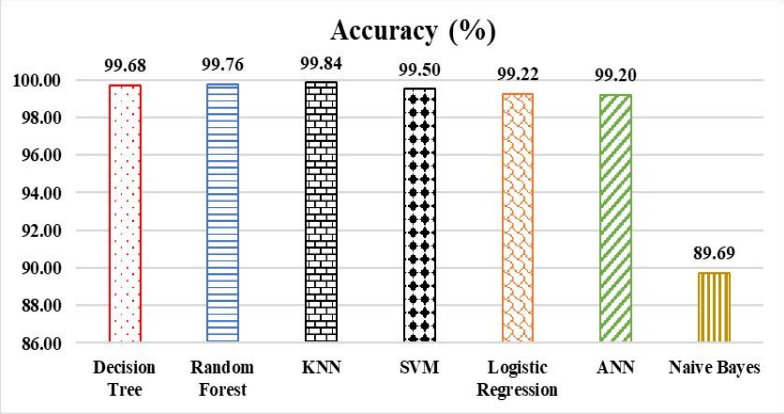


Figure 3. The accuracy Results

As seen in Fig. 3, the KNN method outperforms some other techniques. Nevertheless, the gap in accuracy between KNN and Decision Tree, Random Forest, SVM, Logistic Regression, and ANN methodologies seems small. A technique Naive Bayes performs very poorer. FAR measure is depicted in Figure 4.

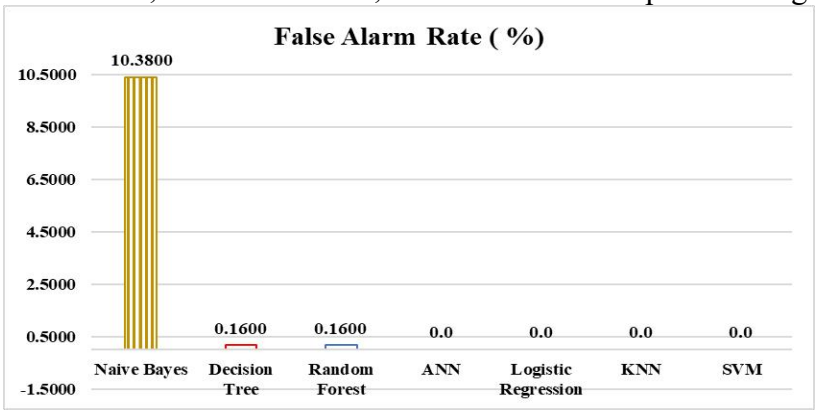


Figure 4. The False Alarm Rate Results

FAR estimates show steady activity, which the program incorrectly identified as unusual activity. As a result, these measurements show the proportion of false

ISSN: 2057-5688

alarms, so in this situation, the smaller the FAR number, the preferable. As demonstrated in Fig., the method Naive Bayes has the highest FAR value and the lowest productivity when this measure is considered. Figure 5. depicts the duration spent programming the ML algorithm, also known as preparation time.

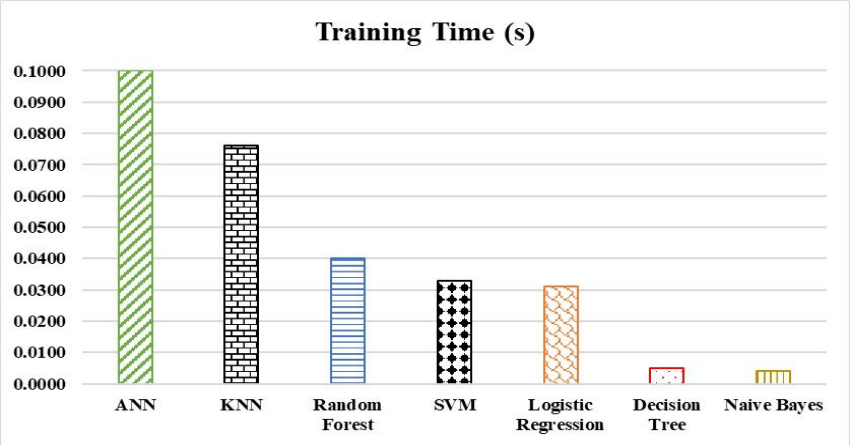


Figure 5. Training time

As shown in Figure , the ANN takes longer to build the machine than other techniques. When compared to the Logistical Regression method, the ANN algorithm is much more difficult to develop. Decision Tree and Naive Bayes, but at the other side, need lesser duration to build the algorithm.

III. A DEEP LEARNING-BASED PRIVACY-PRESERVING TECHNIQUE FOR IDS ON IIOT ENVIRONMENT

IDS framework is a retrofit approach for designing a protective shield against cyber threats. The basic idea of IDS models hinge on the behaviours of an assailant will diverge atypically from that of an authorized user and that abundant illicit actions are visible. To sneak into an IIoT network, assailants might intentionally exploit the weakness of the IIoT networks and create several threats, which cause revealing confidential private information, data alteration, or potential data loss. To mitigate adversaries or malicious nodes from distressing the normal performance of the system, some network security mechanisms are required to classify these adversaries in the network inevitably and enable the network to operate securely.

This work proposed a new DL-based privacy-preserving technique in an IIoT network to identify cyberattacks in the IIoT network. The proposed model includes (i) preprocessing methods such as data cleaning, normalization, and encoding to extract useful information from the dataset;(ii) ISSO-FS algorithm for optimal attribute selection method; (iii) SSAE-based classifier with BSA optimizer. A comprehensive set of experiments are conducted on NSL-KDD and CIC-DDoS dataset. The experimental results reveal the better performance of the BSA-IDS approach over the other IDS models regarding various performance measures. The working process of the proposed IDS approach is illustrated in Figure 6. The detailed working of these procedures is elucidated in the following sections.

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

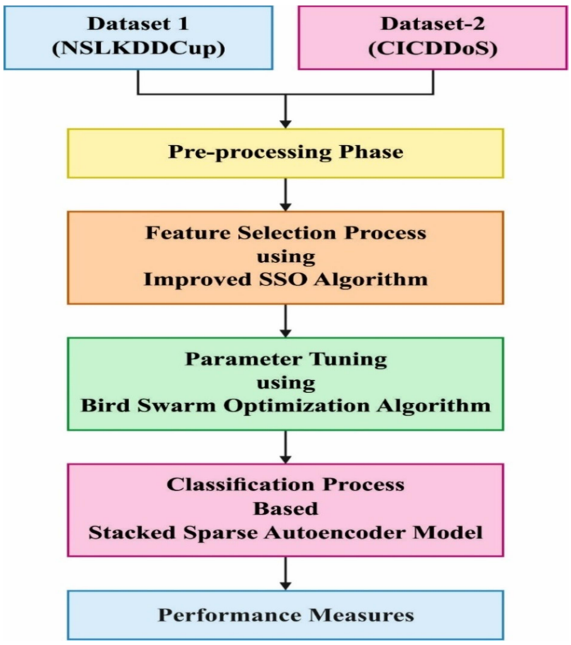


Figure 6. Process flow of the proposed BSA-IDS model

IV. A SECURE AND EFFICIENT IIOT ANOMALY DETECTION APPROACH USING A HYBRID DEEP LEARNING TECHNIQUE

The rapid growth of Internet of Things (IoT) technology has paved the way for the integration of smart devices in various domains, including industrial environments. The Industrial Internet of Things (IIoT) has revolutionized industries by enabling real-time monitoring, predictive maintenance, and data-driven decision-making. With the proliferation of IIoT devices, a massive amount of data is generated, providing valuable insights for optimizing industrial processes and improving productivity.

However, along with the benefits of IIoT, the increasing interconnectivity also introduces new challenges, particularly in terms of security and anomaly detection. Anomaly detection plays a crucial role in identifying abnormal behavior or events in IIoT systems, enabling timely response and mitigation of potential risks. Traditional anomaly detection approaches are often limited in their ability to handle the complexity and scale of IIoT data.

4.1 Anomaly

he cases that stand out as being different from all the rest is a typical requirement when studying real-world data sets. Such occurrences are referred to as anomalies, and the objective of anomaly detection (also known as outlier identification) is to identify all such occurrences in a data-driven-manner. An outlier is defined as an observation that deviates so significantly from other observations as to raise suspicion that it was generated by a different mechanism. Outliers can be caused by errors in the data, but sometimes they are indicative of a new, previously unknown, underlying process.

Deep neural networks have become increasingly common in the larger area of machine learning in recent years. From figure 10. we cobserver that Deep learning algorithms have demonstrated comparable performance to other machine learning techniques.

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

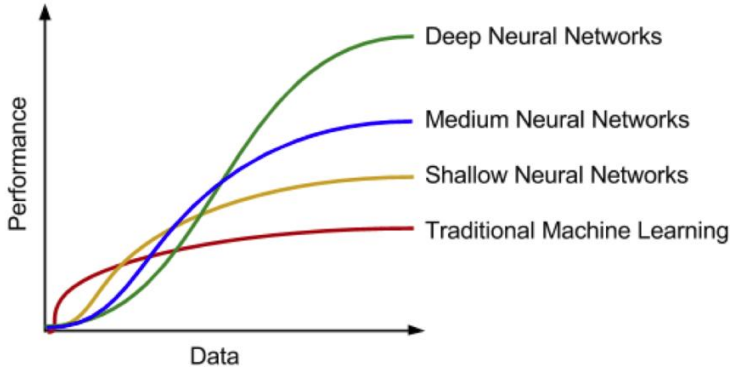


Figure 7. Performance Comparison of Deep Learning vs Traditional Algorithms.

Deep learning, a subset of machine learning, learns to represent the input as a deep hierarchy of concepts within layers of the neural network. This allows it to perform well and be flexible. Figure 11. represents how deep learning-based anomaly detection algorithms have been used for a variety of tasks in recent years. Research demonstrates that deep learning significantly outperforms conventional techniques. Over the years, researchers conduct several research works to address this challenge and develop effective anomaly detection models.

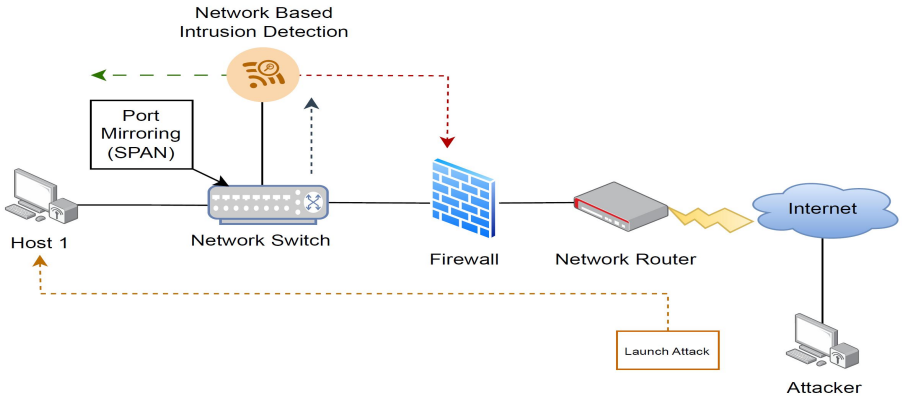


Figure 8. Cyber Network Intrusion Detection System

4.2 CNN+GRU Hybrid model:

The CNN+GRU model combines the strengths of Convolutional Neural Networks (CNNs) and Gated Recurrent Units (GRUs) to achieve high performance in anomaly detection. This architecture effectively captures both spatial and temporal information present in the data, making it well-suited for analyzing complex sequences such as those encountered in industrial IoT systems.

The CNN component of the model uses convolutional layers to extract spatial features from the input data. The convolutional layers apply filters to capture patterns and structures in the data, allowing the model to learn meaningful representations. By stacking multiple convolutional layers with increasing filter sizes and pooling layers, the CNN model can effectively capture hierarchical representations of the input.

The combination of the CNN and GRU models through the concatenation of their output layers allows for the fusion of both

ISSN: 2057-5688

spatial and temporal features. This fusion provides a comprehensive understanding of the data, enabling the model to make accurate predictions. By leveraging the complementary strengths of CNNs and GRUs, the CNN+GRU architecture achieves a balance between capturing local spatial features and modeling temporal dynamics.

The advantages of the CNN+GRU model lie in its ability to effectively capture complex patterns in industrial IoT data. The CNN component excels at extracting spatial features, detecting local anomalies, and recognizing spatial patterns. The GRU component, on the other hand, captures temporal dependencies, enabling the model to understand the temporal behavior of the data and detect anomalies that occur over time. By combining these two architectures, the CNN+GRU model can effectively identify anomalies that exhibit both spatial and temporal characteristics, providing a comprehensive solution for anomaly detection in industrial IoT systems.

Through extensive experimentation and evaluation, the CNN+GRU model has demonstrated its effectiveness in our work. It has achieved high accuracy, precision, recall, and F1 score in detecting anomalies in the industrial IoT dataset. The model's ability to capture intricate spatial and temporal patterns allows it to accurately identify anomalous instances, enabling proactive security measures in industrial IoT systems.

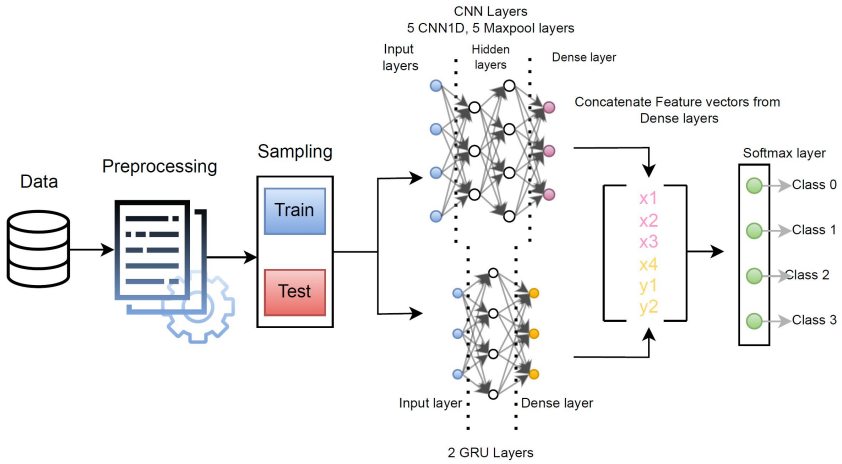


Figure 9. Hybrid CNN GRU mode

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

Psudo code:

Algorithm 2 CNN+GRU Model

```
1: Define the CNN model
2: cnn input _ ← Input(shape = input shape)
3: cnn layer _ ← Conv1D(64, 3, activation =′ relu′)(cnn _input)
4: cnn layer _ ← MaxPooling1D(2)(cnn layer)                    ▷ CNN Layers
5: Define the GRU model
6: gru input _ ← Input(shape = input shape)
7: gru layer _ ← GRU(units = 32, activation =′ tanh′)(gru input)   ▷ GRU layers
8: Concatenate the outputs from the CNN and GRU models
9: concat layer _ ← concatenate([cnn layer, gru layer])
10: Output layer for classifying anomalies
11: output layer _ ← Dense(num_classes, activation =′ softmax′)(concat layer)
12: Create the CNN+GRU model                                   ▷ CNN+GRU Model
13: model ← Model(inputs = [cnn input, gru input], outputs = output layer) _
```

V. FUSED IGAN-IDS MODEL USING AN EXTREMELY RANDOMIZED TREE WITH VI. IMPROVED CLASSIFICATION PERFORMANCE

Due to the uniqueness of components used in IIoT networks, the threats and attack patterns also differ from each other. Different effective IDSs have been extensively employed to protect these networks against these cyberattacks. Hence, the IIoT networks should be fortified with efficient processing elements to sense network intrusive events, categories of cyberattacks, and apprise their models routinely in real-time. However, most IDS systems struggle with a lack of datasets for training and testing which makes the implementation process challenging to identify the probable cyberattacks with expected accuracy. Even if there is a dataset, the number of records of each type of threat may not be sufficient for the detection system to get trained impeccably and identify the cyberattacks with higher accuracy. Moreover, imbalanced datasets are adding more challenges to the intrusion detection process.

6.1 IGAN in Intrusion Detection

The fused IGAN-IDS model identifies cyberattacks with higher classification accuracy. The proposed fused IGAN-IDS model contains five components: a database unit, sample synthesizer unit, IDS unit, controller unit, and loss calculation unit. Initially, the database unit gathers real-time attack samples from its data accumulator. The database unit may also acquire generated data from the producer of the sample synthesizer unit. All of these samples (i. e., original and generated data samples) are gathered unremittingly and simultaneously. After preprocessing these data, it is stored in the database with appropriate labels (i. e., normal/attack) to differentiate the data sources. The

ISSN: 2057-5688

generated samples are subdivided into undecided and fake samples. The samples labelled as fake are already certified and stored permanently in the dataset.

6.1.1 Performance Evaluation of Fused IGAN-IDS Model on NSL-KDD Dataset

To achieve more accurate results, the 10-fold CV method is used. Therefore, the entire dataset is split into ten parts. For each run, one part is used for testing, and the remaining parts are employed to train the algorithm. Now the average value of all 10 tests is considered for assessment..1.

From this table, it is observed that the fused IGAN-IDS model has achieved superior classification performance in terms of predictive accuracy of 99.7%, a sensitivity of 99.9%, specificity of 99.5%, precision of 99.5%, FPR of 4.2%, FNR of 0.2%, ρ-value of 5%, and JSS of 97.5%. Also, it is stimulating to perceive that the SD values obtained by the fused IGAN-IDS model are very small regarding accuracy (0.003), sensitivity (0.004), specificity (0.007), precision (0.005), FPR (0.006), FNR (0.001), ρ-value (0.003), and JSS (0.013). This reveals the reliability and robustness of the proposed model. Figures 6.2 and 6.3 show the superiority of the intended classifier with respect to evaluation metrics.

Table 2. Results of Fused IGAN-IDS on NSL-KDD for various folds

Fold	ACC	SEN	SPE	PRE	FPR	FNR	JSS	ρ-value
#1	0.997	0.999	0.994	0.992	0.034	0.003	0.982	0.010
#2	0.991	0.998	0.990	0.996	0.048	0.003	0.986	0.008
#3	0.999	0.997	0.996	0.997	0.042	0.003	0.960	0.002
#4	0.998	0.999	0.998	0.998	0.040	0.001	0.988	0.001
#5	0.994	0.999	0.998	0.999	0.042	0.003	0.980	0.003
#6	0.997	0.998	0.980	0.982	0.044	0.003	0.972	0.004
#7	0.999	0.999	0.997	0.999	0.031	0.002	0.989	0.002
#8	0.994	0.999	1.007	0.995	0.051	0.002	0.969	0.005
#9	0.998	0.999	0.998	0.996	0.042	0.003	0.970	0.006
#10	0.999	0.998	0.989	0.991	0.043	0.004	0.951	0.004
Mean	**0.997**	**0.999**	**0.995**	**0.995**	**0.042**	**0.002**	**0.975**	**0.005**
S. D	0.003	0.004	0.007	0.005	0.006	0.001	0.013	0.003

ISSN: 2057-5688

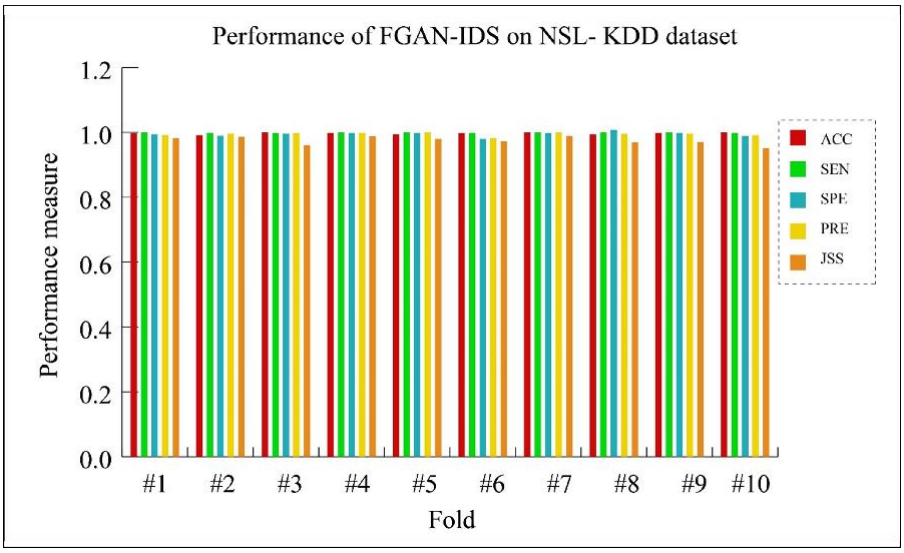


Figure 10. Results of IGAN-IDS model regarding ACC, SEN, SPE,PRE, and JSS

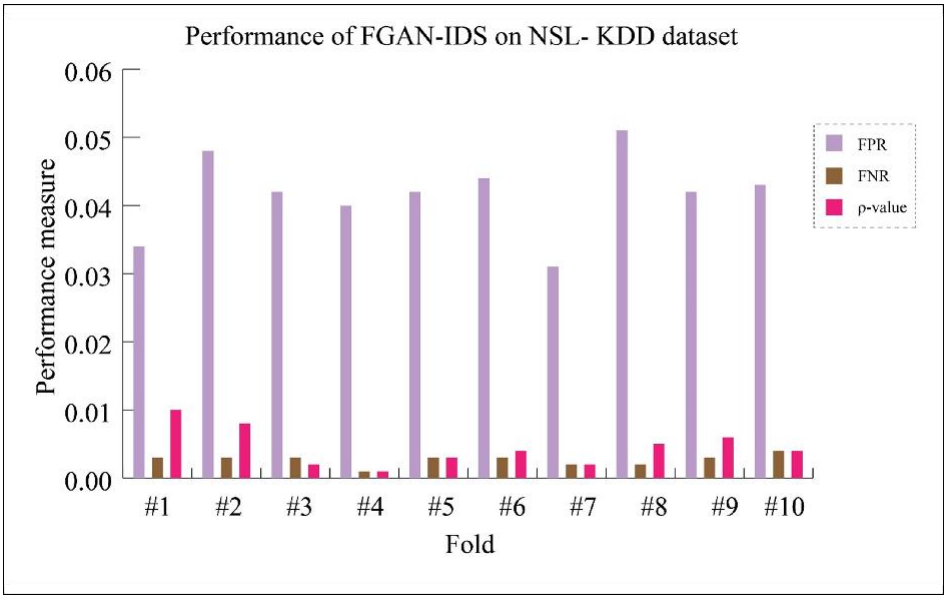


Figure 11. Results of IGAN-IDS on NSL-KDD regarding FPR, FNR,and ρ-value

To demonstrate the superiority of the fused IGAN-IDS model, the performance of this model is related to other IDS models.

VII. CONCLUSION

This study has developed and implemented a novel deep learning- based IDS model to achieve secure and privacy-preserving data transmission in an IIoT environment. This model includes preprocessing, ISSO-based feature selection, SSAE-based classification, and BSA-based parameter optimization. The design of the ISSO algorithm with β −HC concept and the BSA optimization helps to achieve an improved attack detection rate of the proposed IDS model. A comprehensive set of experiments are carried out on benchmark datasets such as NSL-KDD and CIC-DDoS. The results are analysed in terms of different aspects. The experimental results

ISSN: 2057-5688

emphasize the better performance of the BSA-IDS model over the other attack detection methods regarding different performance measures. An efficient feature selection algorithm, namely ERT-FS is fused with the IGAN model to select significant features from abnormal traffic. The network datasets NSL-KDD and CIC-DDoS are used for empirical analysis to demonstrate the performance of the proposed model regarding the accuracy, sensitivity, specificity, precision, ρ-value, JSS, FAR, ADR, and time consumption for training as well as testing for the classifier. In this thesis, we explored and evaluated various deep-learning models for anomaly detection in edge IIoT systems. Through extensive experimentation and analysis, we have gained valuable insights into the performance and capabilities of these models. The results demonstrate the effectiveness of different neural network architectures, including CNN, GRU, CNN+GRU, LSTM, XGBoost, and Autoencoder-based models in detecting anomalies in edge IIoT data.

REFERENCES

[1] X. Xie, C. Wang, S. Chen, G. Shi, and Z. Zhao, "Real-time illegal parking detection system based on deep learning," in Proceedings of the 2017 international conference on deep learning technologies, 2017, pp. 23–27.

[2] T. Schlegl, P. Seeb¨ock, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," in Information Processing in Medical Imaging: 25th International Conference, IPMI 2017, Boone, NC, USA, June 25-30, 2017, Proceedings. Springer, 2017, pp. 146–157.

[3] J. Xiong, S. Bharati, and P. Podder, "Machine and deep learning for iot security and privacy: Applications, challenges, and future directions," Security and Communication Networks, vol. 2022, p. 8951961, 2022. [Online]. Available: https://doi.org/10.1155/2022/8951961

[4] V. Chandola, A. Banerjee, and V. Kumar, "Outlier detection: A survey," ACM Computing Surveys, vol. 14, p. 15, 2007.

[5] D. M. Hawkins, Identification of outliers. Springer, 1980, vol. 11.

[6] Y. Himeur, K. Ghanem, A. Alsalemi, F. Bensaali, and A. Amira, "Artificial intelligence based anomaly detection of energy consumption

[7] in buildings: A review, current trends and new perspectives," Applied Energy, vol. 287, p. 116601, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261921001409

[8] N. Alexopoulos, E. Vasilomanolakis, N. R. Iv´ank´o, and M. Mu¨hlh¨auser, "Towards blockchain-based collaborative intrusion detection systems," in Critical Information Infrastructures Security, G. D'Agostino and A. Scala, Eds. Cham: Springer International Publishing, 2018, pp. 107–118.

[9] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," IEEE Transactions on Industrial Electronics, vol. 65, no. 7, pp. 5990–5998, 2018.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," nature, vol. 521, no. 7553, pp. 436–444, 2015.

[11] S. Tofigh, M. O. Ahmad, and M. Swamy, "A low-complexity modified thinet algorithm for pruning convolutional neural networks," IEEE Signal Processing Letters, vol. 29, pp. 1012–1016, 2022.