Detecting IOT Attacks Using Artificial intelligence

MSc Research Project

Cyber Security

Hemanth Murukutla

Student ID: X21219214

School of Computing

National College of Ireland

Supervisor: Mr. Michael Prior

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: ………Hemanth Murukutla…………………………………………………

Student ID: …………X21219214…………………………………………………………
……

Programme: ……M.Sc Cyber Security……… Year: ……2023….

Module: ………M,Sc Research Project ……………………………………………….

Supervisor: ………Michael Prior……………………………………………………….
Submission Due Date: …………14/08/2023………………………………………………………….

Project Title: Detecting IOT Attacks Using AI.

Word Count: …………………….. Page Count……………19………………………….

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ……………Hemanth……………………………………………….

Date: ……………14/08/2023……………………………………………….

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□

You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Abstract

The rapid expansion and significant significance of IIoT networks have led to a concerning escalation in cyber vulnerabilities, thereby demanding the creation of more advanced detection methodologies. Traditional cybersecurity techniques have proven ineffective in protecting these complex systems, emphasizing the necessity for novel and advanced approaches. The objective of the research is to assess the efficiency of machine learning models, with a specific focus of the Variational Autoencoder (VAE), Long Short-term Memory Model (LSTM), Recurrent Neural Network (RNN) and the Gated Recurrent Unit (GRU), in recognizing cyber threats within IIoT networks. The Edge IoT dataset, which is an extensive set of network logs gathered from several common IIoT settings, was used as the basis for this investigation. This set of data was utilized during the training, testing, and evaluation of the various machine learning models.. The research study validates the usefulness of the GRU model, exhibiting an amazing detection accuracy rate of 97%. These findings demonstrate the great potential of applying the GRU model for detecting cyber risks in IIoT networks. This research contributes to the larger mission of improving cybersecurity by fortifying our interconnected industrial systems against cyber-attacks.

1 Introduction

With the increasing adoption of digital technologies, the IIoT has emerged as a critical component in industrial area's digital transformation (Yazdinejad *et. al.*, 2023). It has transformed many old industrial processes, providing significant efficiency and productivity gains through supply chain automation, predictive maintenance, and real-time monitoring. These developments, however, introduce substantial vulnerabilities (Malik *et. al.*, 2021).

The IIoT's confluence of operational and information technology (IT) has increased the attack surface, making it a tempting target for cybercriminals (Yazdinejad *et. al.*, 2023). Given the crucial nature of these networks and the potentially catastrophic repercussions of successful attacks cyber threat detection in IIoT networks has become increasingly important (Yazdinejad *et. al.*, 2023).

The detection of cyber threats in a network comprises identifying, analyzing, and responding to potential security concerns. Beyond data security, it ensures the ongoing operation of industrial systems that govern vital infrastructures such as energy grids, water supply systems, manufacturing plants, and transportation networks in the context of IIoT.

Due to the complexity and variety of these systems, as well as the special characteristics of industrial operations, detecting cyber threats in IIoT networks presents significant problems. Due to changes in architecture, performance requirements, and the nature of threats, traditional security methods built for IT networks frequently fall short when applied to IIoT (Jahromi, Karimipour, and Dehghantanha, 2023). As a result, novel approaches and solutions are required to address the special security requirements of the IIoT.

Machine learning algorithms such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Logistic Regression, Artificial Neural Networks (ANN), and Feedforward Neural Networks (FNN) have been used extensively in these studies (Arora, Kaur, and Teixeira, 2022). However, the main focus of this research has been on detecting and preventing cyber threats in specific scenarios such as network traffic or specific application settings.

The purpose of the research is to examine the detectability of cyber threats within the domain of IIoT framework. To achieve this, a Machine Learning-based IIoT Network Security Framework is proposed that integrates data from two sources and applies advanced deep learning models such as Variational AutoEncoder, Long Short Term Memoy (LSTM), Recurrent Neural Network (RNN), and GRU. The framework enables the extraction and analysis of crucial features from network traffic data, enabling the classification of diverse cyber threats.

The proposed framework aims to identify the cyber-attacks based on the pattern in the data stream of IIoT systems. Understanding these variables is crucial in mitigating a cyber-attack, and enhancing overall resilience. By identifying patterns in the data, the proposed model will aid in early detection and prevention strategies. Proactively addressing a cyber threat at an early stage allows for effective measures to prevent threats and improve security. This study will give vital insights on the rising worry of cyberattacks, which will be beneficial to people and businesses all across the globe.

Research Question

"How do different Recurrent Neural Network architectures perform in detecting cybersecurity threats in an IIoT network."

The research is divided into several sections. In the second part, "Related Work," we discuss the ML techniques that are used to identify cyber-attacks as well as the research that has been done in this area in the past. Section 3 provides a full description of the ML-based Network Security Framework. In Section 4, the plan of the suggested framework is discussed, followed by a step-by-step implementation guide in Section 5. Section 6 evaluates the models implemented in the study. Section 7 presents a comprehensive argument of the testing and analysis, while Section 8 decides the research then discusses upcoming work in this area.

2 Related Work

This Part of study discusses the critical analysis of research conducted in cyber security for IIoT devices. Papers from 2019 to 2023 have been reviewed in detail to understand the different methodologies adopted in the detection of cyber threats in IIoT environment. In depth study of these research will help to understand their findings and also the limitations associated with them.

With a growing reliance on technology and an enhance in creation of IoT devices, the need of robust cybersecurity measures, such as advanced intrusion detection systems (IDS), cannot be stressed. Detecting novel digital threats, on the other hand, remains a challenge, necessitating the creation of complex frameworks to improve IDS efficiency.

A deep learning strategy based on Long Short-Term Networks (LSTMs) was developed for IoT cyberattack detection in a paper by Iwendi et al. (2021). This strategy outperformed many current models in terms of accuracy, F1 scores, and recall, with values of 99.09%, 99.46%, and 99.5%, respectively. However, the deep learning model's limitations were not thoroughly explained, leaving questions about its applicability in different contexts or with different datasets.

Hanif et al. (2019) proposed an alternate strategy for threat detection based on an Artificial Neural Network (ANN). In repeated 10-fold cross-validation, this system overcame the authentication difficulties unique to IoT, with an average precision of 84% and a false positive rate of less than 8%. While promising, this strategy may be hampered by the ANN's three-layer architecture, potentially oversimplifying the threat landscape's complexity.

Ali et al. (2020) performed a systematic assessment of AI and Machine Learning applications in cybersecurity threat detection, with a focus on classifiers. The most popular classifiers were identified as Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), and ANN. The study, while offering a taxonomy for comprehending these technologies, falls short of giving a full comparison of their efficacy.

The proliferation of IoT systems and smart gadgets has piqued network attackers' interest, resulting in the emergence of botnets targeted at acquiring control of these systems. To tackle this, powerful machine learning and deep learning solutions must be combined with suitable feature engineering to predict and fight against network vulnerabilities.

Panda et al. (2021) addressed this issue in a recent study by identifying cyberattacks using the UNSW-NB15 dataset, which was specifically intended for IoT-Botnet analysis. To construct a representative dataset, they used scatter search-based feature engineering tools and K-Medoid sampling. Three advanced machine learning algorithms were used in their method: A2DE, JChaid*, and HGC. In terms of detection rate, precision, recall, F1-score, and computational efficiency, the scatter search-based DMLP classifier outperformed the others. However, it is critical to explore how this approach performs on different types of data and in different attack scenarios.

In another study, Podder et al. (2021) thoroughly examined the use of deep learning technology in cybersecurity. They examined several deep learning methodologies for cybersecurity, differentiated between shallow and deep learning, and assessed the efficacy of deep learning methods in combating cyberattacks. Their findings showed that the Restricted Boltzmann Machine (RBM) algorithm performed well on customised datasets, while the Long Short-Term Memory (LSTM) strategy performed well on the KDD Cup 99 dataset. The study, however, did not investigate the performance of these approaches in real-time, dynamic threat settings.

As the subject of cybersecurity progresses, it becomes evident that machine learning and deep learning techniques have enormous potential for identifying and mitigating cyberattacks on IoT networks. However, more study is needed to evaluate these methods across various attack scenarios and data kinds, as well as to assess their usefulness in real-time circumstances.

AI integration in cybersecurity has become critical for enterprises to battle rising cyber threats and maintain data confidentiality. These threats, which are motivated by a variety of factors such as political rivalry, profit-driven techniques, information theft, and extreme group goals, frequently have malevolent intent. Tao et al. (2021) present a comprehensive overview of AI-based cybersecurity research in their study, laying the groundwork for future advances in the subject.

IoT devices have become great targets for cybercriminals, especially nation-state-sponsored attackers, as they continue to play an important role in our linked society. Saharkhizan et al. (2020) propose a deep learning method for detecting cyberattacks to address the issues of safeguarding IoT systems. Long Short-Term Memory (LSTM) modules are integrated into a collection of detectors, and a decision tree is used to combine these modules to generate an aggregated output. Real-world tests on Modbus network traffic data show an astounding detection accuracy rate of more than 99%.

Pacheco et al. (2020) highlight the possibility of varied devices and systems sharing resources to build advanced information services in the context of IoT. The growing attack surface, on the other hand, creates substantial security challenges, particularly for IoT components such as gateways (Fog Nodes). To solve this, the authors suggest an artificial neural network-based adaptive intrusion detection system (IDS). Despite the complexity of the adaptive system, their approach successfully describes the usual behavior of fog nodes and effectively detects anomalies from multiple sources, attaining a high detection rate and low false alarm rates.

Ghillani (2022) emphasizes the importance of deep learning approaches developed from artificial neural networks (ANNs) in boosting cyber risk analytics and organizational resilience in their study. They investigate several approaches for addressing various cybersecurity concerns, such as multilayer perceptrons, convolutional neural networks, and recurrent neural networks. The sensitivity to feature scaling and the computational cost of solving complicated security models, on the other hand, remain problems for these networks.

Given the surge in botnet-based attacks, Soe et al. (2020) propose a machine learning-based system for identifying botnet attacks in IoT devices. Using three separate machine learning algorithms, their solution employs a sequential detection architecture and an effective feature selection methodology to achieve high-performance lightweight detection with around 99% accuracy for botnet attacks. The machine learning-based botnet attack detection system proposed in this study achieved a high detection rate, but it's unclear how it performs against unknown, newly emerging botnet threats.

This literature review explores the role of Artificial Intelligence (AI), particularly deep learning and machine learning techniques, in cybersecurity. The studies presented provide a comprehensive overview of various AI-driven models and strategies for combating cyber threats and ensuring organizational resilience.

3 Research Methodology

This section discusses the research framework adapted for the study undertaken. In the view of building a machine learning-based intrusion detection system, data is of prominence. The data that is selected, is discussed in great depth in this section along with the source of collected data, the structure in which the data is available along with the exploratory study of the data.

3.1 Data Collection

The dataset that has been used for the study is acquired from the Kaggle repository[1]. This dataset is available in CSV file format and is made available on Kaggle by the author(s) themselves. This dataset is available in two separate files, these files collectively contain 237701 samples corresponding to different attack types. Following is the list of attacks present in the dataset.

1. DDoS_UDP – DDoS attack on UDP protocol
2. DDoS_ICMP – DdoS attack on ICMP protocol
3. DoS slowloris
4. SQL Injection
5. DDoS_HTTP – DDoS attack on HTTP protocol
6. DDoS_TCP – DDoS attack on TCP protocol
7. Vulnerability Scanner
8. Password
9. Uploading
10. Port Scanning
11. XSS
12. Fingerprinting
13. MITM
14. Backdoor
15. Ransomware

Along with being a large dataset, the dataset is created with 61 different attributes associated with a network data packet. The different number of attributes present in the dataset are listed in table 3.1 below.

frame.time	ip.src_host	ip.dst_host	arp.dst.proto_ipv4	arp.opcode	arp.hw.size	arp.src.proto_ipv4	icmp.checksum
icmp.seq_le	icmp.transmit_timestamp	icmp.unused	http.file_data	http.content_length	http.request.uri.query	http.request.method	http.referer
http.request.full_uri	http.request.version	http.response	http.tls_port	tcp.ack	tcp.ack_raw	tcp.checksum	tcp.connection.fin
tcp.connection.rst	tcp.connection.syn	tcp.connection.synack	tcp.dstport	tcp.flags	tcp.flags.ack	tcp.len	tcp.options
tcp.payload	tcp.seq	tcp.srcport	udp.port	udp.stream	udp.time_delta	dns.qry.name	dns.qry.name.len
dns.qry.qu	dns.qry.type	dns.retransmission	dns.retransmit_request	dns.retransmit_request_in	mqtt.conack.flags	mqtt.conflag.cleansess	mqtt.conflags
mqtt.hdrflags	mqtt.len	mqtt.msg_decoded_as	mqtt.msg	mqtt.msgtype	mqtt.proto_len	mqtt.protoname	mqtt.topic
mqtt.topic_len	mqtt.ver	mbtcp.len	mbtcp.trans_id	mbtcp.unit_id			

Table 3.1: Dataset Attributes

This list includes various aspects of network data, such as IP addresses, HTTP requests, TCP and UDP connection details, DNS queries, and MQTT (a lightweight messaging protocol for small sensors and mobile devices) attributes, among others.

This dataset hence is an extensive collection of data relating to the cyber security domain IIoT.

[1] https://www.kaggle.com/datasets/mohamedamineferrag/edgeiiotset-cyber-security-dataset-of-iot-iiot

3.2 Exploratory Data Analysis (EDA)

This step of the methodology deals with gaining significant insights into the dataset (Morgenthaler, 2009). It is an essential step in any data analysis study as it helps to identify null values present in the dataset, it also helps to identify the data types that are present in the dataset ranging from string values, numerical values, the data types of attack labels, etc. The EDA part of the study is done using the Pandas library available for Python. The info() function enlists the attributes in the dataset along with corresponding datatypes. Isnull() function of the dataframe object of Pandas, identifies the missing values in the dataset. And if there are any null values in the dataset, the rows corresponding to the nulls are dropped using the dropna() function.

3.3 Data Pre-processing

The knowledge acquired from the EDA part helps to identify that certain pre-processing steps are necessary to be undertaken to process the data further. These steps include label consolidation, class balancing, label encoding, and feature selection.

Label Consolidation: Label consolidation is an optional step in a data analysis methodology. This step of the methodology is required when the number of unique labels in the dataset is large. As in the selected dataset, the number of unique labels is **15**. This becomes a multi-label classification problem. In such a problem, the models that are used, require a very high level of hyper tuning which requires computational resources and is not always feasible. For a multi-label classification problem, it is seen that the models tend to perform poorly. Hence, to avoid this degradation of performance, and also, we are just looking for the presence of a threat in the data, we are opting for the label consolidation step.

In this step, we are replacing, the attack labels such as ddos_tcp, ddos_udp, vulnerability_scanner, backdoor etc. into three broad categories viz. normal, ddos, and web attack. We will train our models to classify the samples into these three classes only.

Label Encoding: This step in the methodology involves the conversion of categorical labels into numerical form which can be interpreted by machine learning models. To encode the labels, the study makes use of the LabelEncoder module of the sklearn library available for Python. This module first lists the labels in ascending alphabetical order and is given a number starting from 0. With respect to the study, the module will label, ddos as 0, normal as 1, and web attack as 2.

Class Balancing: Class imbalance is a problem in machine learning modeling. This is because the presence of a class imbalance in the dataset makes the model biased towards the class in the majority (Sisodia, Reddy, and Bhandari, 2017).. Also, the accuracy of the model cannot be determined efficiently as if the data contains 97% percent of samples belonging to one class, the model will have an accuracy of 97% regardless of the fact that it cannot classify the samples in a minority. To avoid such a scenario, it is necessary to balance out the class samples in the dataset. As the dataset at hand is large and contains more than 230000 samples, the classes can be balanced by selecting 20000 samples from each of the categories viz. normal, DDoS, and web attack. This is exactly what is done in the presented methodology. This has another benefit. As the number of samples under study is reduced to 60000, the computational requirements for the models to train on the data are very low and they can achieve convergence faster.

4 Design

This section explains the design of the framework in detail. Once, the pre-processed data is obtained, the next important step of modeling can be performed. As the study mainly focuses on the implementation of the cyber attack detection system, the design section of the paper deals with the selection criterion for the models.

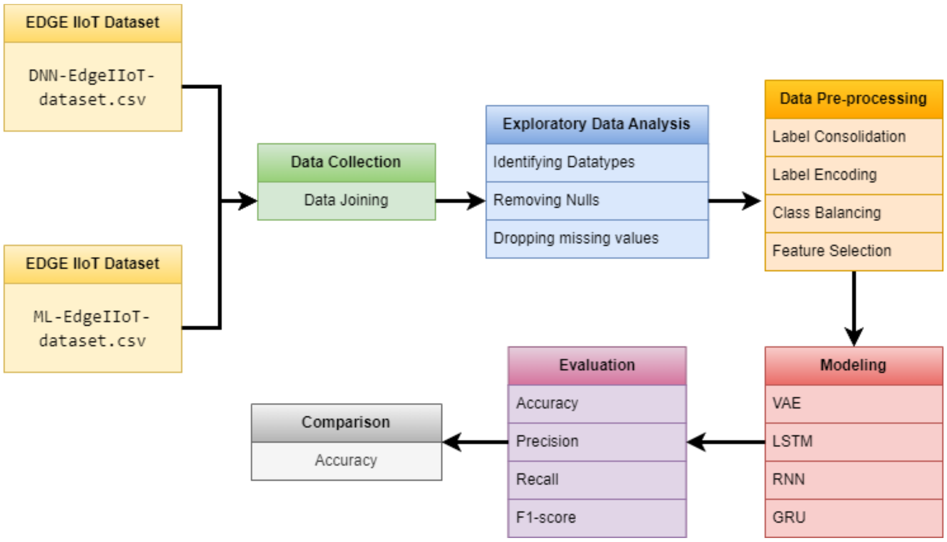

Figure 4.1: Methodology Process Flow

4.1 Feature Selection

Once, the pre-processed data is obtained, the next step in the data analysis stems from the fact that the efficiency of machine learning models depends heavily upon the dimensionality of the data. The dimensionality of the data corresponds to the number of attributes that are present. For the machine learning models to perform optimally, it is required that the number of attributes should be small enough to improve the model performance and large enough to include the most relevant of attributes.

One way to achieve a reliable feature selection is by using SelectKBest module of Sklearn library. SelectKBest method lists K features based on their relevance (Ayyanar *et. al.*, 2022). The relevance of the features is calculated using Chi2 statistic. In the context of feature selection, the chi-square test is used to test the independence between each feature and the target (Bisong and Bisong, 2019). The perception behind this is that if a feature and the target are independent, then the feature would not be useful for predicting the target and could be removed. Conversely, if they're not independent, the feature might be essential in prediction and should be kept.

For the presented study, we have chosen the value of K to be 25. This is solely based on intuition as the number is high enough to include the most relevant features. This step in the process gives us data of 25 dimensions.

4.2 Machine Learning Models

This study makes use of four different machine-learning models for detecting cyber attack from a data packet. These models are variational autoencoder, RNN, LSTM, and GRU. These models are chosen for the study based on following reasons.

Variational Autoencoder RNN (VAE): This model learns the underlying representations of the input features, which is useful for dealing with complicated, high-dimensional, and structured data, such as network data. It also works well when the class labels are imbalanced or there is noise in the data (Kingma and Welling, 2019).

Long Short-Term Memory (LSTM): This model is a type of recurrent neural network (RNN) that is capable of learning long-term dependencies in the data (Jozefowicz, Zaremba, and Sutskever, 2015). It makes decisions by considering the current input, and also what it has learned from the inputs it received previously. This feature makes LSTM models extremely powerful in tasks where the prediction at the current step depends not just on the current input, but on a series of previous inputs (Jozefowicz, Zaremba, and Sutskever, 2015).. Considering that network security involves analyzing sequences of packets to identify patterns, LSTM is a suitable choice.

Recurrent Neural Network (RNN): Just like the LSTM, the RNN also has the ability to use its internal state (memory) to process sequences of inputs, which makes it ideal for analyzing sequences of network traffic data. However, standard RNNs suffer from vanishing gradient problem, limiting their ability to learn long-range dependencies (Jozefowicz, Zaremba, and Sutskever, 2015).. But they are relatively simpler and faster to train than LSTMs or GRUs.

Gated Recurrent Unit (GRU): This is a type of recurrent neural network that is similar to LSTM but has fewer gates and thus fewer parameters. It combines the forget and input gates into a single "update gate" (Jozefowicz, Zaremba, and Sutskever, 2015).. It also merges the cell state and hidden state. The resulting model is simpler than standard LSTM models and has been growing in popularity due to its similar performance and faster training times (Jozefowicz, Zaremba, and Sutskever, 2015)..

5 Implementation

Our Machine Learning-based IIoT Network Security Framework is developed leveraging the power of Python Programming Language, utilizing Jupyter Notebook as our chosen Integrated Development Environment (IDE, Python 3.8.5). Key Python libraries were employed to this end, including but not limited to, Pandas for data manipulation, Numpy for numerical operations, TensorFlow for deep learning applications, and Scikit-learn for machine learning tasks.

The CTDS harnesses two datasets obtained from cyber-attack logs and network activity, which were later merged. The data was then subjected to a cleaning process, in which we dealt with missing values and performed class balancing for better model training.

Further, a crucial step in our pipeline was the feature selection process, where we used chi-squared statistics to select the 25 most relevant features. This helped us reduce the complexity of our models without compromising the model performance significantly.

We split our processed dataset in an 95:05 ratio for training and testing, ensuring a good balance between learning from the data and validating the findings, using Scikit-learn's train_test_split function with a random state set at 1234 for reproducibility of results.

As for the modeling aspect, we experimented with four deep learning models, specifically the Variational RNN, LSTM, a simple RNN, and a GRU model, all of which were implemented using the TensorFlow library. Each model was trained on the dataset and their performance was evaluated based on their prediction accuracy. The different models allowed us to observe the varied performance and choose the most effective model for our cybersecurity threat detection application.

The hyperparameters for the models are selected through the trial-and-error method, wherein the combination of hyperparameters giving the highest accuracy is chosen for model construction. Table 4.1 below shows the hyperparameter values for the models implemented in the study.

Model	Layers	Layer Configurations (Neurons)
Variational Autoencoder	3 encoder and 3 output and decoder layers	Encoder layers: 512, 128, 64 Dropout rate = 0.2 Decoder layers: 64,4
LSTM	4 layers (1 LSTM and 3 Dense)	LSTM:512 Dense:128,32,4
RNN	3 layers (1 SimpleRNN and 2 Dense)	SimpleRNN:32 Dense:10,4
GRU	3 layers (1 GRU and 2 Dense)	GRU:32 Dense:10,4

Table 5.1: Machine Learning model parameters

6 Evaluation

In this section, we will explore the evaluations of the models implemented as part of this study. We will thoroughly analyze and critique the performance of different models chosen for the framework. The main goal of this research is to effectively detect potential cybersecurity threats. To achieve this, we have tested various machine learning models and evaluated their predictive abilities using our consolidated dataset.

The dataset used in this study contains network activity records. We have carefully selected key features to form the basis for evaluating these models. Our objective is to determine the effectiveness of these models in predicting cyber threats, which can significantly improve existing cybersecurity measures.

The following sections will present a chronological account of the implementation of four deep learning models: Variational AutoEncoder, LSTM, RNN, and GRU. The evaluation will not only assess their individual performance but also make comparative observations to determine their relative effectiveness in this specific application scenario.

6.1 Evaluation of Variational Autoencoder

Table 6.1 below depicts the values for the metrics obtained for the Variational Autoencoder model for the classification of the data packets as Normal, DDoS, or Web Attack. From the table, it can be observed that the variational autoencoder model performed well in the detection in terms of metric values.

Metric	Value (%)
Accuracy	90.67

Table 6.1: Evaluation metrics for Variational Autoencoder

Figure 6.1 below depicts the classification report for the variational autoencoder model. 0, 1, 2 labels in the report correspond to DDoS, Normal, and Web Attack respectively. From the figure, it can be observed that the model has been successful in identifying the normal traffic well.

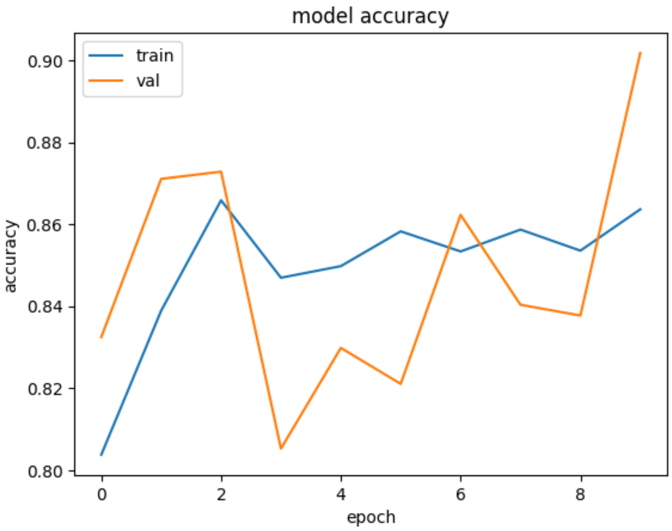

Figure 6.1: Training performance for variation autoencoder

6.2 Evaluation of LSTM

Table 6.2 below depicts the values for the metrics obtained for the LSTM model for the classification of the data packets. From the table, it can be observed that the LSTM model performed very poorly in the detection in terms of metric values.

Metric	Value (%)
Accuracy	32.10

Table 6.2: Evaluation metrics for LSTM

The performance of the model for training and validation set of data has been depicted in Figure 6.2 below. From the figure, it can be seen that the model has under-fitted.

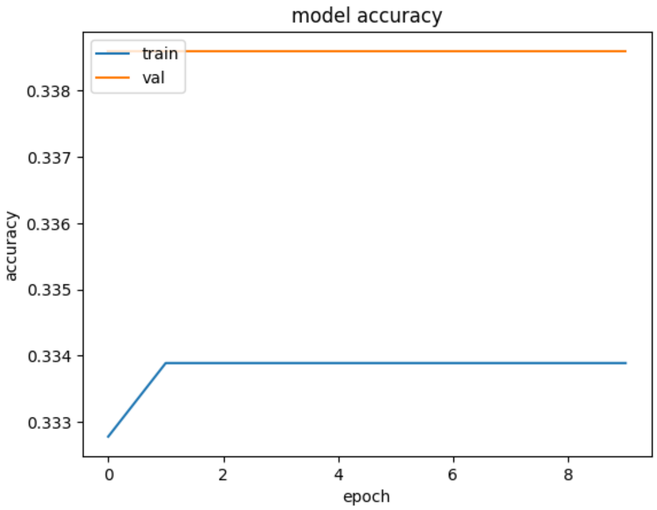

Figure 6.2: Training performance of the LSTM model

6.3 Evaluation of RNN

Table 6.3 below depicts the values for the metrics obtained for the RNN model for the classification of the data packets. From the table, it can be observed that the RNN model performed mediocrely in the detection.

Metric	Value (%)
Accuracy	59.83

Table 6.3: Evaluation metrics for RNN

The performance of the model for training and validation set of data has been depicted in Figure 6.3 below. From the figure, it can be seen that the model has fitted well.

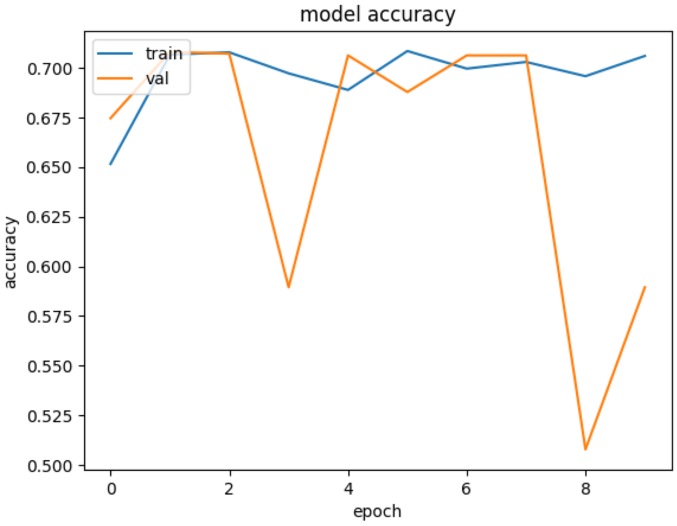

Figure 6.3: Training performance of the RNN model

6.4 Evaluation of GRU

Table 6.4 below depicts the values for the metrics obtained for the GRU model for the classification of the data packets. From the table, it can be observed that the GRU model performed well in the detection in terms of metric values.

Metric	Value (%)
Accuracy	97.33

Table 6.4: Evaluation metrics for GRU

The training performance of the GRU model has been depicted in Figure 6.4 below. From the figure, it can be seen that the model has fitted well and there is no evidence for over or under-fitting.

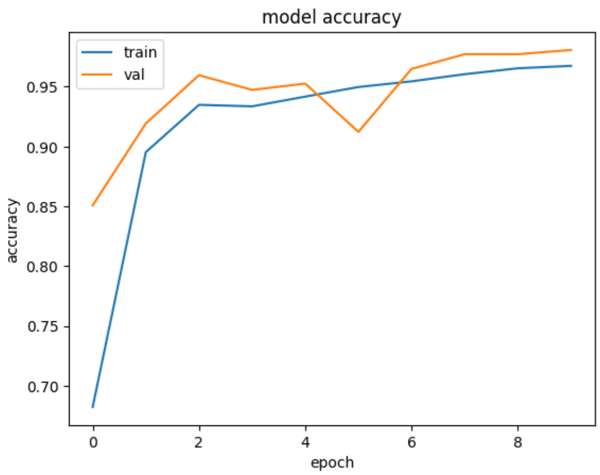

Figure 6.4: Training performance for the GRU model

6.5 Comparative Analysis

Table 6.5 below shows the comparison of the models implemented.

Model	Accuracy (%)
VAE	90.67
LSTM	32.10
RNN	59.83
GRU	**97.33**

Table 6.5: Comparison of the model performances

Figure 6.9 below shows a graphical representation of the results obtained from the evaluation of the models.

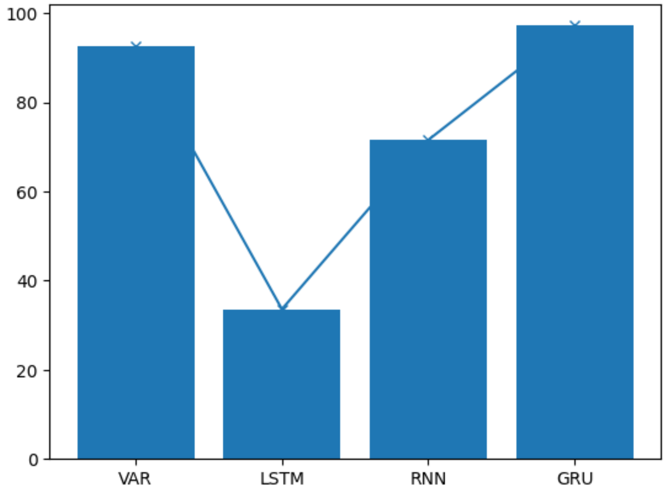

Figure 6.5: Graphical comparison of the model performances

7 Discussion

The following discussion presents an in-depth examination of the performance outcomes of four separate machine learning models: Variational Autoencoder (VAE), Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Gated Recurrent Unit (GRU). These models were tested for their ability to detect cyber threats. Several criteria, including precision, recall, F1-score, and accuracy, were used to assess their performance. The study was carried by utilizing a test dataset of 1500 cases.

The Variational Autoencoder produced promising results, with an overall accuracy of 90.67%. These results can be explained by the model's ability to correctly classify positive cases, but with some difficulty in finding all relevant examples. This observation highlights a common difficulty with Variational Autoencoders, which value precision above recall. It is worth mentioning that this constraint may be alleviated through improved feature engineering or model tuning.

In stark contrast, the LSTM model really underperformed. The significantly poor overall accuracy of 32.10, is far from ideal. Several variables, including LSTM's sensitivity to parameter settings, overfitting, and even the intrinsic difficulty associated with handling extended sequences within the data, may contribute to this performance disparity. LSTM models can struggle to grasp the dependencies found in such data sets.

The RNN model performed mediocrely, with an overall accuracy of 59.83%. The poor results indicate that the model failed to identify a large number of relevant cases, possibly due to RNNs' failure to adequately capture long-term dependencies in the data. This observation highlights the importance of using more complicated or optimized RNN designs to improve performance.

Among the tested models, the GRU model stood out as the best performer, with an amazing overall accuracy of 97.33%. For all classes, the model demonstrates it's ability to properly detect true positives while minimizing both types of errors. This is due to the use of gating mechanisms within the GRU design, which allows for the capture of both short-term and long-term dependencies within the data, resulting in higher overall performance.

8 Conclusions and Future Work

In this study, four machine learning models, namely Variational Autoencoder (VAE), Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Gated Recurrent Unit (GRU), were thoroughly evaluated in the context of cyber threat detection in Industrial Internet of Things (IIoT) systems. IIoT systems are vital in many businesses, making their protection critical. An extensive IIoT dataset was used to assess the effectiveness of these models, offering a realistic and complex testing ground.

The evaluation results gave intriguing insights into the performance of each model. The VAE model had good precision but low recall, indicating that it had trouble finding all positive instances in the dataset. The LSTM model demonstrated great precision for one class but fell short in overall accuracy, implying that generalizing its prediction capability across all classes will be difficult. The RNN model performed poorly, indicating possible limits in dealing with the complexity of the IIoT dataset. The GRU model, on the other hand, was the most successful, with excellent precision, recall, and an impressive overall accuracy of 97.33%. This highlights the GRU model's ability to learn from temporal dependencies in the data.

These models' various levels of effectiveness in dealing with the IIoT dataset provide useful insights into the applicability and usability of machine learning in IIoT cybersecurity. The GRU model's outstanding performance, in particular, highlights the effectiveness of machine learning models in boosting cybersecurity measures in industrial systems.

Moving forward, some areas merit further investigation and development. To begin, fine-tuning the models with alternative parameter settings and advanced feature extraction approaches to improve their performance on the IIoT dataset could be considered. This would entail tweaking each model's hyperparameters and investigating feature engineering strategies particular to the IIoT domain.

Second, resolving the recall deficiencies reported in the Variational Autoencoder and RNN models may entail investigating new model topologies or sophisticated feature engineering methodologies. Incorporating attention mechanisms, for example, or domain-specific information, could potentially improve the recall performance of these models. Furthermore, the performance of the LSTM model could be enhanced by altering sequence lengths or using strategies to prevent overfitting, such as regularization approaches.

Third, the GRU model's strong performance offers the possibility to investigating more elaborate or hybrid model designs. Researchers could investigate merging advantageous components of the previous models with the GRU model to develop unique architectures capable of delivering even better outcomes in terms of accuracy, precision, and recall.

Finally, evaluating the models on larger or more diversified IIoT datasets could improve their resilience and reliability in protecting against a broader range of cyber threats. This would entail gathering data from a broader range of IIoT systems across numerous businesses and scenarios.

This research is significant not only for its immediate findings, but also for its implications for future work in the field. The findings of this study provide encouragement for more research and development efforts aimed at improving IIoT cybersecurity through the use of advanced machine learning models. Researchers can contribute to the development of more effective cybersecurity solutions for IIoT systems by addressing the limits of existing models and investigating new techniques.

References

Ali, R., Ali, A., Iqbal, F., Khattak, A.M. and Aleem, S., 2020. A systematic review of artificial intelligence and machine learning techniques for cyber security. In Big Data and Security: First International Conference, ICBDS 2019, Nanjing, China, December 20–22, 2019, Revised Selected Papers 1 (pp. 584-593). Springer Singapore.

Arora, P., Kaur, B. and Teixeira, M.A., 2022. Security in industrial control systems using machine learning algorithms: An overview. *ICT Analysis and Applications*, pp.359-368.

Awotunde, J.B. and Misra, S., 2022. Feature extraction and artificial intelligence-based intrusion detection model for a secure Internet of things networks. In Illumination of artificial intelligence in cybersecurity and forensics (pp. 21-44). Cham: Springer International Publishing.

Ayyanar, M., Jeganathan, S., Parthasarathy, S., Jayaraman, V. and Lakshminarayanan, A.R., 2022, April. Predicting the Cardiac Diseases using SelectKBest Method Equipped Light Gradient

Boosting Machine. In *2022 6th International conference on trends in electronics and informatics (ICOEI)* (pp. 117-122). IEEE.

Bisong, E. and Bisong, E., 2019. More supervised machine learning techniques with scikit-learn. Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners, pp.287-308.

Calderon, R., 2019. The benefits of artificial intelligence in cybersecurity.

Christawan, E., Ariadi, S., Thalib, P., Astika, D. and Suyanto, B., 2023. Enhancement of Polri's Role in Dealing with Disinformation and Radicalism Extremism Terrorism and Separatism Propaganda in Cyberspace.

Das, M.B., Yuko, A., Chapman, T.B. and Jain, V., 2022. Silver Hues.

Deshmukh, A., Sreenath, N., Tyagi, AK and Jathar, S., 2022, January. Internet of Things based smart environment: threat analysis, open issues, and a way forward to future. In 2022 International Conference on Computer Communication and Informatics (ICI) (pp. 1-6). IEEE.

Dixit, S., Bohre, K., Singh, Y., Himeur, Y., Mansoor, W., Atalla, S. and Srinivasan, K., 2023. A Comprehensive Review of AI-enabled Models for Parkinson's disease diagnosis. Electronics, 12(4), p.783.

Elsisi, M., Tran, M.Q., Mahmoud, K., Mansour, D.E.A., Lehtonen, M. and Darwish, M.M., 2021. Towards secured online monitoring for digitalized GIS against cyber-attacks based on IoT and machine learning. Ieee Access, 9, pp.78415-78427.

Ghillani, D., 2022. Deep learning and artificial intelligence framework to improve cyber security. Authorea Preprints.

Hanif, S., Ilyas, T. and Zeeshan, M., 2019, October. Intrusion detection in IoT using artificial neural networks on UNSW-15 dataset. In 2019 IEEE 16th international conference on smart cities: improving quality of Life using ICT & IoT and AI (HONET-ICT) (pp. 152-156). IEEE.

Islam, U., Muhammad, A., Mansoor, R., Hossain, M.S., Ahmad, I., Eldin, E.T., Khan, J.A., Rehman, A.U. and Shafiq, M., 2022. Detection of distributed denial of service (DDoS) attacks in IOT-based monitoring systems of the banking sector using machine learning models. Sustainability, 14(14), p.8374.

Iwendi, C., Rehman, S.U., Javed, A.R., Khan, S. and Srivastava, G., 2021. Sustainable security for the Internet of Things using artificial intelligence architectures. ACM Transactions on Internet Technology (TOIT), 21(3), pp.1-22.

Jahromi, A.N., Karimipour, H. and Dehghantanha, A., 2023. An ensemble deep federated learning cyber-threat hunting model for Industrial Internet of Things. *Computer Communications*, *198*, pp.108-116.

Jozefowicz, R., Zaremba, W. and Sutskever, I., 2015, June. An empirical exploration of recurrent network architectures. In *International conference on machine learning* (pp. 2342-2350). PMLR.

Kingma, D.P. and Welling, M., 2019. An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, *12*(4), pp.307-392.

Krishnan, S., Neyaz, A. and Liu, Q., 2021. IoT network attack detection using supervised machine learning.

Larriva-Novo, X., Villagrá, V.A., Vega-Barbas, M., Rivera, D. and Sanz Rodrigo, M., 2021. An IoT-focused intrusion detection system approach based on preprocessing characterization for cybersecurity datasets. Sensors, 21(2), p.656.

Malik, P.K., Sharma, R., Singh, R., Gehlot, A., Satapathy, S.C., Alnumay, W.S., Pelusi, D., Ghosh, U. and Nayak, J., 2021. Industrial Internet of Things and its applications in industry 4.0: State of the art. *Computer Communications*, *166*, pp.125-139.

Morgenthaler, S., 2009. Exploratory data analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*(1), pp.33-44.

Pacheco, J., Benitez, V.H., Felix-Herran, L.C. and Satam, P., 2020. Artificial neural networks-based intrusion detection system for Internet of Things fog nodes. IEEE Access, 8, pp.73907-73918.

Panda, M., Abd Allah, A.M. and Hassanien, A.E., 2021. Developing an efficient feature engineering and machine learning model for detecting IoT-botnet cyber attacks. IEEE Access, 9, pp.91038-91052.

Podder, P., Bharati, S., Mondal, M., Paul, P.K. and Kose, U., 2021. Artificial neural network for cybersecurity: A comprehensive review. arXiv preprint arXiv:2107.01185.

Robinson, N., Tidd, B., Campbell, D., Kulić, D. and Corke, P., 2023. Robotic vision for human-robot interaction and collaboration: A survey and systematic review. ACM Transactions on Human-Robot Interaction, 12(1), pp.1-66.

Saharkhizan, M., Azmoodeh, A., Dehghantanha, A., Choo, K.K.R. and Parizi, R.M., 2020. An ensemble of deep recurrent neural networks for detecting IoT cyber attacks using network traffic. IEEE Internet of Things Journal, 7(9), pp.8852-8859.

Sharma, P., Jain, S., Gupta, S. and Chamola, V., 2021. Role of machine learning and deep learning in securing 5G-driven industrial IoT applications. Ad Hoc Networks, 123, p.102685.

Sisodia, D.S., Reddy, N.K. and Bhandari, S., 2017, September. Performance evaluation of class balancing techniques for credit card fraud detection. In *2017 IEEE International Conference on power, control, signals and instrumentation engineering (ICPCSI)* (pp. 2747-2752). IEEE.

Soe, Y.N., Feng, Y., Santosa, P.I., Hartanto, R. and Sakurai, K., 2020. Machine learning-based IoT-botnet attack detection with sequential architecture. Sensors, 20(16), p.4372.

Tao, F., Akhtar, M.S. and Jiayuan, Z., 2021. The future of artificial intelligence in cybersecurity: A comprehensive survey. EAI Endorsed Transactions on Creative Technologies, 8(28), pp.e3-e3.

Yazdinejad, A., Kazemi, M., Parizi, R.M., Dehghantanha, A. and Karimipour, H., 2023. An ensemble deep learning model for cyber threat hunting in industrial internet of things. *Digital Communications and Networks*, *9*(1), pp.101-110.

Zeadally, S., Adi, E., Baig, Z. and Khan, I.A., 2020. Harnessing artificial intelligence capabilities to improve cybersecurity. Ieee Access, 8, pp.23817-23837.

Video link –

https://studentncirl-my.sharepoint.com/:v:/g/personal/x21219214_student_ncirl_ie/EZaRCzat5AtApcbCMlQQNbIB2iLg4KhvyRzxFQQ6UQ844g?e=eAbOR3&nav=eyJyZWZlcnJhbEluZm8iOnsicmVmZXJyYWxBcHAiOiJTdHJlYW1XZWJBcHAiLCJyZWZlcnJhbFZpZXciOiJTaGFyZURpYWxvZyIsInJlZmVycmFsQXBwUGxhdGZvcm0iOiJXZWIiLCJyZWZlcnJhbE1vZGUiOiJ2aWV3In19