National College of Ireland

Novel ways of Detecting threats in IIoT Networks

MSc Research Project

Gopi Kuchi
Student ID: x22159959

School of Computing
National College of Ireland

Supervisor: Vikas Sahni

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Gopi Kuchi		
Student ID:	X22159959		
Programme:	MSc Cyber Security	**Year:**	2023-24
Module:	MSc Research Project		
Supervisor:	Vikas Sahni		
Submission Due Date:	25-04-2024		
Project Title:	Novel ways of Detecting threats in IIoT Networks		
Word Count:	**5976**	**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: Gopi Kuchi

Date: 25-04-2024

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):	

Novel ways of Detecting threats in IIoT Networks

Gopi Kuchi
x22159959

Abstract

This study investigates how machine learning can improve cyber security in Industrial Internet of Things networks with an emphasis on devices that use the Modbus protocol. The study aims to a high level of threat detection accuracy using machine learning algorithms such as Random Forest Classifier and Support Vector Machine Classifier. The results in this experiment indicated that the ensembled-based Machine Learning Classifier has performed considerably better than Support Vector Machine. The accuracy for Random Forest Classifier is 98.64% and the accuracy for Support Vector Machine is 53.32%. The Runtime for Random Forest Classifier is 18.72 seconds, And Run time for Support Vector Machine is 38.22 seconds. In the future, work may focus on enhancing model interpretability and testing in real-world scenarios. All things considered, this research improves IIoT Cyber Security, By helping organizations find suitable algorithms that can analyze their traffic.

Keywords: **Random Forest Classifier, Support Vector Machine Classifier, Modbus Protocol, Cybersecurity, Machine Learning, Industrial Internet of Things, and Threat Detection.**

1 Introduction

The Industrial Internet of Things represents a significant advancement in manufacturing, integrating smart, connected technologies to optimize production. However, this integration also introduces complex security risks, particularly in the communication protocols between devices. This research paper explores methods to fortify IIoT infrastructures against these emerging threats using machine learning techniques, focusing on real-time threat detection. The study is particularly concerned with IIoT systems utilizing the Modbus protocol and seeks to address gaps in current security methodologies while anticipating the need for future adaptability.

The main question is: **How is Machine Learning used to find threats in IIoT networks?** This research will help make a system that can spot threats in real-time and make IIoT systems safer. This research aims to fill in the gaps in current methods and make IIoT networks more secure against new and changing threats. This study has some limits. It only looks at IIoT devices using the Modbus protocol, and it might need updates as threats change over time.

Outline of the Report

The report is structured like this:

- **Related Work:** Look at what others have done and where our research fits in.
- **Research Methodology:** Explains how this paper plans to find threats and make IIoT systems safer.
- **Design Specification:** Talks about how data and privacy is handled.
- **Conclusion and Future Work:** Summarizes what is found and what could be done next.

Integrating new technology into various industries has historically provided benefits but has also introduced new risks. As the Industrial Internet of Things continues to expand, ensuring its security is paramount for the smooth operation of industries and the prevention of significant disruptions. Traditional cybersecurity methods have proven inadequate for the complex and evolving threat landscape associated with IIoT networks. This research aims to enhance the protection of IIoT, particularly focusing on detecting threats in data transmitted over networks using the Modbus protocol. Leveraging machine learning could revolutionize the capability of IIoT systems to identify and mitigate threats autonomously. By implementing cutting-edge technology, this research seeks to bolster the security of IIoT and safeguard it against emerging threats, thereby benefiting industries worldwide.

2 Related Work

Industry 4.0, which is defined by the integration of the Industrial Internet of Things, will bring about significant changes to manufacturing operations. This paradigm change toward completely connected and intelligent production is being driven by cutting-edge technologies such as cloud computing, artificial intelligence, and the Internet of Things. However, new cybersecurity challenges for the IIoT ecosystem arise as technology develops. The environment's reliance on novel communication paradigms, such as Time Sensitive Networking and 5G, leaves it open to a range of cyberattacks. Closing the gap between information and operational technology is critical because otherwise, the IIoT would be more vulnerable to sophisticated attacks. Among the challenges are malicious IoT node insertion, misconfigured devices, and M2M communications protocols. The industrial sector faces challenges in implementing suitable security measures even if it recognizes the need for data standards. This literature study highlights the need for uniform standards and security procedures to safeguard M2M communications in light of the evolving IIoT environment (Lackner et al. 2018).

2.1 IIOT

The literature emphasizes how the Industrial Internet of Things is growing more and more dependent on technology to increase productivity, but additionally emphasizes how susceptible it is to cyberattacks. IIoT system security presents challenges for organizations because of the proliferation of connected endpoints and diverse software hardware combinations. It's interesting to note that a significant portion of the Internet of Things traffic remains unencrypted, leaving devices open to hackers. Aside from

potential consequences on individuals, addressing regulatory challenges and securely managing data intake are critical considerations. The study highlights the importance of IIoT cybersecurity due to the risks to the safety of the public, the real-world applications, and the absence of standardized protocols for IoT management of data. Specific challenges are highlighted, such as the disconnect between the IT and OT departments, concerns with authentication, interoperability, and vulnerabilities in out-of-date hardware. emphasizing the need for an all-encompassing, context-aware cybersecurity approach for the IIoT (Wang et al. 2023).

Strengthened cybersecurity measures are urgently needed since cyberattacks on critical facilities, such as power plants and water supplies, might have fatal consequences. This is why the literature on the Industrial Internet of Things highlights this point. The increased usage of intelligent Internet of Things sensors in industrial settings and their frequent wireless network communication has resulted in a marked increase in the hazardous environment. The integration of information technology and operational technology has raised risks since, in addition to data loss, security breaches can now impact human and physical safety.

The literature highlights real-world examples of the shocking outcomes associated with cyberattacks on industrial sites to underscore the urgent need for robust cybersecurity regulations. Industrial operators require more education, with an emphasis on building rapport, promoting teamwork, and bridging the knowledge gap between OT and IT teams (Serror et al. 2020).

The literature also emphasizes the importance of technology solutions, such as auto-mapping and identification software, industrial-specific intrusion prevention software, and the involvement of service providers like AT&T and Cisco in addressing the security concerns brought up by IIoT.

The literature also emphasizes the importance of technology solutions, such as auto-mapping and identification software, industrial-specific intrusion prevention software, and the involvement of service providers like AT&T and Cisco in addressing the security concerns brought up by IIoT.

Business decision-makers will find great value in this document's RFI checklist, which offers a comprehensive process for evaluating and selecting the best solution for safeguarding IT, IoT, IoMT, and OT devices on corporate networks. Important subjects including architecture, identifying devices, proactive monitoring, risk-based policy recommendations that are automatically generated, and security against both known and unknown threats are all covered in the checklist. There are seven major sections to it. This research evaluation recognizes the value of the checklist as a tactical tool for decision-makers navigating the complex realm of IoT security. The checklist emphasizes the requirement of complete visibility, ongoing risk assessment, and rapid threat detection to ensure that companies can make informed choices to increase their cybersecurity posture. This review highlights the checklist's ability to guide decision-makers toward tailored and effective security solutions as companies navigate evolving concerns related to device protection.

A flow-control cycle is a set of instructions or procedures that controls how data or activities move within a system. It could be related to information flow within a system in any way, including communication and data processing.

A structured approach to identifying and assessing potential security threats to a system is called threat modeling. It facilitates understanding the risks and vulnerabilities that a system may experience.

A Danger Model Using the STRIDE Approach:

STRIDE Approach: The shortened form A variety of threats are denoted by the acronym STRIDE, which refers to Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege. This approach is often used to categorize and investigate potential system risks.

This document will assist in developing a roadmap that benefits businesses and institutions. They can use it to identify and fix vulnerabilities (Mauri and Damiani 2022).

2.2 Industrial Internet of Things Security Consideration

The article What are the Risks Connected with Industrial IoT?" examines the crucial elements relating to the safety of industrial IoT installations. The literature, which focuses on how manufacturing equipment is interconnected, enumerates the main risks that producers need to be mindful of in 2023. These threats include data breaches, denial-of-service attacks, device spoofing or man-in-the-middle assaults, data siphoning, and device hijacking. These risks pose a major danger to automated procedures, network operation, sensitive data transfer, and general IIoT security. The literature highlights the necessity of an integrated strategy for security that incorporates effective prevention and detection solutions to shield organizations against potential hazards as they embrace IIoT initiatives (Manjari 2021).

The literature review provides an overview of the current concerns regarding cybersecurity in the larger context of the Industrial Internet of Things, emphasizes how vulnerable the industrial sector is to cyberattacks and lists it as a top hacker target for the past two years (Tan and Samsudin 2021).

The essay recognizes the difficult challenges that IIoT cybersecurity faces, including the need for leadership teams to have good technical expertise and the need to fight for resources with other organizational goals, but it also emphasizes how vital it is to make cybersecurity a high priority. The literature provides insights on the IIoT Security and Safety Protocol created by the World Economic Forum, highlighting how IIoT differs from standard IoT and advocating for a business-wide approach. It also looks at the key elements, challenges, and recommended solutions for ensuring the integrity, authentication, privacy, and availability of IIoT systems. As the manufacturing industry transitions to Industry 4.0, the essay promotes innovative designs and tactical approaches to support secure IIoT environments (Saini and Saini 2019).

This will compel companies to address significant problems and put in place a cybersecurity plan that is based on corporate goals.

This paper examines the evolving Internet of Things landscape, focusing on industrial Internet of Things security issues and potential remedies. The 2023 Global Symposium on Innovative Data Communication Technologies and Applications was the venue for its delivery. The paper recognizes the disruptive impact of Industry 4.0 and IoT integration while highlighting the vast opportunities for data production, processing, and exchange. It does, however, draw attention to the greater risks associated with industrial equipment's interconnection, particularly about data security and privacy (Sontan and Samuel 2024).

The paper thoroughly examines suggested countermeasures, including intrusion detection systems, key setup techniques, and authentication mechanisms, to improve security across a range of IIoT surfaces. The study is noteworthy for identifying possible problems with connectivity, infrastructure, data security, and equipment longevity in addition to limitations falling inside the scope of Industrial IoT. The literature highlights the importance of resolving cybersecurity concerns in conjunction with the rapid advancements in IIoT technology (AlSalem et al. 2023).

Threat Landscape Analysis:
In addition to identifying attack vectors, (Aziz et al. 2023) delve into the motivations behind cyber threats targeting industrial IoT. They discuss how nation-state actors, cybercriminals, and even insider threats pose significant risks to critical infrastructure, highlighting the geopolitical implications of IIoT cybersecurity breaches. Understanding the geopolitical context is essential for devising strategies that not only mitigate technical vulnerabilities but also address broader geopolitical risks.

Alyazia Aldhaheri et al. (2024) shed light on the transformative impact of the Internet of Things (IoT) through interconnected smart devices, accompanied by unprecedented opportunities and intricate security challenges. With cybersecurity emerging as a critical concern, intrusion detection systems (IDS) play a pivotal role in safeguarding IoT environments. The authors emphasize the potential of Deep Learning in bolstering IDS capabilities, effectively identifying and thwarting cyber threats targeting IoT devices. Their paper not only explores state-of-the-art intrusion detection methods rooted in Deep Learning but also addresses the associated challenges and suggests avenues for future research, offering invaluable insights for both researchers and industry practitioners navigating IoT security landscapes.

Vulnerability Assessment:
(Mingo et al. 2024) not only emphasize continuous vulnerability assessment but also propose a risk-based approach to prioritizing remediation efforts. Their work underscores the dynamic nature of IIoT ecosystems, where new vulnerabilities emerge rapidly due to software updates, configuration changes, and the integration of third-party components. Furthermore, they advocate for the adoption of automated vulnerability scanning tools coupled with manual penetration testing to comprehensively evaluate the security posture of IIoT systems.

Security Protocols and Standards:
(Yash Bobde et al. 2024) not only propose enhanced authentication mechanisms but also advocate for the integration of blockchain technology to establish tamper-proof audit trails for IIoT data transactions. By leveraging blockchain's decentralized and immutable ledger, organizations can enhance data integrity and traceability, thereby bolstering the trustworthiness of IIoT systems. Additionally, they highlight the importance of industry-wide collaboration in developing standardized security protocols to ensure interoperability and facilitate the adoption of best practices.
(Gyamfi and Jurcut 2022) explores machine learning and anomaly detection techniques but also discuss the challenges associated with the scalability and real-time performance of intrusion detection systems in IIoT environments. They propose distributed and edge-based intrusion detection architectures capable of processing massive volumes of data generated by IoT devices in real-time. Moreover, they advocate for the integration of threat intelligence feeds and collaborative information sharing platforms to enhance the effectiveness of intrusion detection and response efforts across industrial sectors.
(Safitra et al. 2023) emphasize integrated cyber-physical security solutions but also discuss the implications of emerging technologies such as artificial intelligence and digital twins on IIoT cybersecurity. They highlight the need for adaptive security controls capable of dynamically adjusting to evolving threats and system conditions. Furthermore, they advocate for the development of standardized interfaces and protocols for secure communication and interoperability between cyber and physical components in IIoT systems.
(Cremer et al. 2022) discussed the alignment of cybersecurity initiatives with organizational risk management objectives but also explore the role of regulatory frameworks in incentivizing investments in IIoT security. They highlight the need for regulatory bodies to adopt a risk-based approach to cybersecurity regulation, balancing the need for innovation and competitiveness with the imperative of protecting critical infrastructure. Moreover, they advocate for public-private partnerships to foster information sharing and collaboration in addressing emerging cyber threats.
(Albshaier et al. 2024) not only explore the potential of blockchain and edge computing but also discuss the ethical and legal implications of deploying these technologies in industrial contexts. They highlight the importance of data privacy, consent management, and accountability in the design and implementation of blockchain-based IIoT solutions. Furthermore, they advocate for transparent governance mechanisms and regulatory oversight to ensure the responsible use of emerging technologies and mitigate potential unintended consequences.

2.3 Research Niche

Regarding the main question, " In what ways can Machine Learning be used to detect threats in IIoT networks?" The complex worlds of industrial Internet of Things and operational technology are examined in this article, which poses a variety of difficulties to chief information security officers. With an estimated 21.5 billion cloud-connected computers and Internet of Things devices in the globe by 2025, the CISO's

duty will expand beyond traditional IT duties. Novel risks are introduced by the integration of cyberspace and physical systems, which can include environmental concerns as well as industrial disruptions. As a result, it is imperative to adopt new paradigms like Zero Trust and reevaluate antiquated security frameworks like the Purdue Enterprise Reference Architecture (Alwahedi et al. 2024).

In this regard, machine learning shows promise as a means of improving cybersecurity in IIoT settings. This research seeks to create a system that can quickly detect and mitigate attacks, strengthening the safety measures of IIoT networks, by using ML algorithms to examine real-time data streams. By filling up the gaps in current cybersecurity approaches, this research project hopes to strengthen defenses against dynamic attacks in IIoT ecosystems and enable more resilient defenses.

3 Research Methodology

Importing Libraries: The research methodology begins with the importation of essential libraries required for data analysis and machine learning implementation. These include pandas for data manipulation, seaborn for data visualization, matplotlib.pyplot for plotting graphs, and scikit-learn modules for machine learning algorithms.

Loading Dataset: The dataset stored in a CSV file named "Train_Test_IoT_Modbus.csv" is loaded into a pandas DataFrame using the pd.read_csv() function. This step ensures that the dataset is readily available for analysis and model training. Moustafa, N., (2019).

Handling Missing Values: Missing values within the dataset are identified using the isnull().sum() function, which calculates the count of missing values for each column. Any missing values are then addressed using appropriate techniques such as imputation or removal to ensure data integrity.

Data Preprocessing: The dataset undergoes preprocessing steps to prepare it for model training. This includes converting date and time columns to a datetime format and dropping unnecessary columns that do not contribute to model training.

Feature Engineering: The dataset is split into features (X) and labels (y) using appropriate indexing techniques. Features represent the input variables used for model training, while labels represent the target variable to be predicted. Additionally, feature scaling is applied using MinMaxScaler to standardize the range of feature values.

Model Training: Two classification models, namely Random Forest Classifier and Support Vector Classifier, are trained on the pre-processed data to classify threats within IIoT networks. The train test split() function is used to split the data into training and testing sets, allowing for model evaluation.

Model Evaluation: Model performance metrics such as accuracy scores and classification reports are computed to evaluate the effectiveness of the trained models in classifying threats within IIoT networks. This step provides insights into the models' ability to accurately predict threat labels.

Feature Importance Analysis: Feature importance analysis is conducted to identify the most significant variables contributing to threat detection within IIoT networks. This analysis helps in understanding which features have the most impact on the models' predictive performance.
Visualization: Various visualizations, including count plots and feature importance plots, are generated to visualize the distribution of threat labels and the importance of features in threat detection. These visualizations aid in gaining insights into the dataset and model outcomes.
Conclusion: Overall, this research methodology leverages machine learning algorithms to enhance threat detection capabilities within IIoT networks. By following a systematic approach to data analysis and model implementation, this study aims to contribute to the advancement of cybersecurity in industrial automation systems.

3.1 Research Resources

Programming Languages for System Development: For the development of the continuous risk surveillance system, the programming language is:
Python: Renowned for its versatility and readability, Python is chosen for its ability to facilitate rapid development. Its extensive collection of frameworks and libraries simplifies the deployment process of the surveillance system.

Tools for Security Analysis and Vulnerability Detection: Various tools are utilized for security analysis and vulnerability detection:
Wireshark: A popular network protocol analyzer, Wireshark facilitates detailed investigation of network traffic to identify potential weaknesses and irregularities in IIoT communication.

Integration and Application: These technologies collectively support the creation, evaluation, and security analysis stages of an ongoing risk monitoring system. Simulation tools enable controlled testing of system functionality, while security analysis tools assist in detecting and mitigating vulnerabilities within the IIoT environment. Python and Java serve as foundational languages for developing a robust and flexible system. By combining these resources, a comprehensive and efficient approach is ensured for the research endeavor.

Example Application Scenario: In an intelligent manufacturing scenario, components such as PLCs, sensors, and actuators are simulated using protocols like MQTT and Modbus. This simulation environment allows for the evaluation of the ongoing risk monitoring system's ability to recognize and address various risks, including periodic sensor readings, event-driven transmissions, and simulated abnormalities such as network congestion.

Similarly, consider a smart energy grid comprising power substations and smart meters in a practical IIoT data context. Through protocols like IEEE 802.15.4 and

DNP3, continuous information exchange occurs, including updates on energy consumption and health reports. Anomalies such as cyberattacks or real power outages can be simulated to assess the system's responsiveness to real-world data streams. This dual approach, combining simulation and real-world IIoT data, ensures a comprehensive evaluation of the system's flexibility and effectiveness across diverse operating scenarios.

4 Design Specification

Research Methodology

In this section, I will outline the techniques, architecture, and framework utilized for the implementation of the continuous risk monitoring system for Industrial Internet of Things environments. The associated requirements are also identified to ensure the successful development and deployment of the system.

Machine Learning Integration: One of the core components of the proposed system is the integration of machine learning algorithms for threat detection in IIoT networks. ML algorithms are employed to analyze network traffic patterns, identify anomalies indicative of cyber threats, and continuously learn from ongoing network activities to improve threat detection over time. Specifically, the system utilizes supervised learning techniques to classify network traffic into normal and anomalous behaviour, enabling real-time threat detection.

Random Forest Classifier: A Random Forest Classifier is employed as the primary ML model for threat detection. This ensemble learning technique consists of multiple decision trees and is capable of handling large datasets with high dimensionality. The model is trained on features extracted from network traffic data, including various Modbus protocol parameters. During the training phase, the classifier learns to distinguish between normal and malicious network behavior, allowing for accurate threat detection during runtime.

Support Vector Machine Classifier: In addition to the Random Forest Classifier, a Support Vector Machine Classifier is also implemented as an alternative ML model for threat detection. SVM is a supervised learning algorithm capable of performing classification tasks by finding the hyperplane that best separates the classes in the feature space. By utilizing SVM, the system enhances its capability to detect and classify threats in IIoT networks, providing an additional layer of security.

Data Preprocessing and Feature Scaling: Before training the ML models, data preprocessing techniques are applied to ensure optimal model performance. This includes handling missing values, converting categorical variables into numerical representations using one-hot encoding, and scaling numerical features using MinMaxScaler to normalize the data within a specified range. These preprocessing steps are essential for improving the convergence and accuracy of the ML models during training.

Evaluation Metrics: To assess the performance of the ML models, various evaluation metrics are utilized, including accuracy, precision, recall, and F1-score. These metrics provide insights into the models' ability to correctly classify normal and malicious network traffic, enabling the identification of potential areas for improvement.

Framework and Tools: The implementation of the continuous risk monitoring system is carried out using Python programming language due to its versatility and extensive ecosystem of libraries for ML and data analysis. Libraries such as pandas, scikit-learn, and seaborn are utilized for data manipulation, model training, and result visualization. Additionally, Jupyter Notebook is employed as an interactive development environment to facilitate code experimentation and documentation.

5 Implementation

The implementation of the proposed continuous risk monitoring system for Industrial Internet of Things environments involved several stages, with the final stage focusing on model development, evaluation, and feature analysis.

Model Development: The dataset containing IIoT network traffic data from "Train_Test_IoT_Modbus.csv" was imported using Python programming language and libraries such as pandas, scikit-learn, seaborn, and matplotlib. After verifying the import, data preprocessing steps were performed, including handling missing values, converting categorical variables into numerical representations, and scaling numerical features using MinMaxScaler.

Random Forest Classifier: A Random Forest Classifier model was trained on the preprocessed data to classify network traffic into normal and anomalous behavior. The model achieved an accuracy of 98.99% on the test data, with a precision, recall, and F1-score of 0.99 for both normal and anomalous classes. Additionally, feature importance analysis was conducted using the trained model to identify the most significant features contributing to threat detection in IIoT networks.

Support Vector Machine Classifier: In addition to the Random Forest Classifier, a Support Vector Machine Classifier was implemented as an alternative ML model for threat detection. The SVM model achieved an accuracy of 53.33% on the test data, with precision, recall, and F1-score values for both classes.

Feature Analysis: The correlation between different features and the target variable was analyzed using a correlation matrix. This analysis provided insights into the relationships between features and their impact on threat detection in IIoT networks.

Outputs Produced: The main outputs produced in this stage of implementation include trained ML models, evaluation metrics, feature importance rankings, and correlation matrices. These outputs provide valuable insights into the performance and effectiveness of the continuous risk monitoring system in detecting and mitigating cyber threats in IIoT environments.

Tools and Languages Used: The implementation process primarily utilized the Python programming language along with libraries such as pandas, scikit-learn, seaborn, and matplotlib for data manipulation, model training, evaluation, and visualization. Additionally, the scikit-learn library provided access to various ML algorithms, including Random Forest Classifier and SVM Classifier, for threat detection in IIoT networks.

6 Evaluation

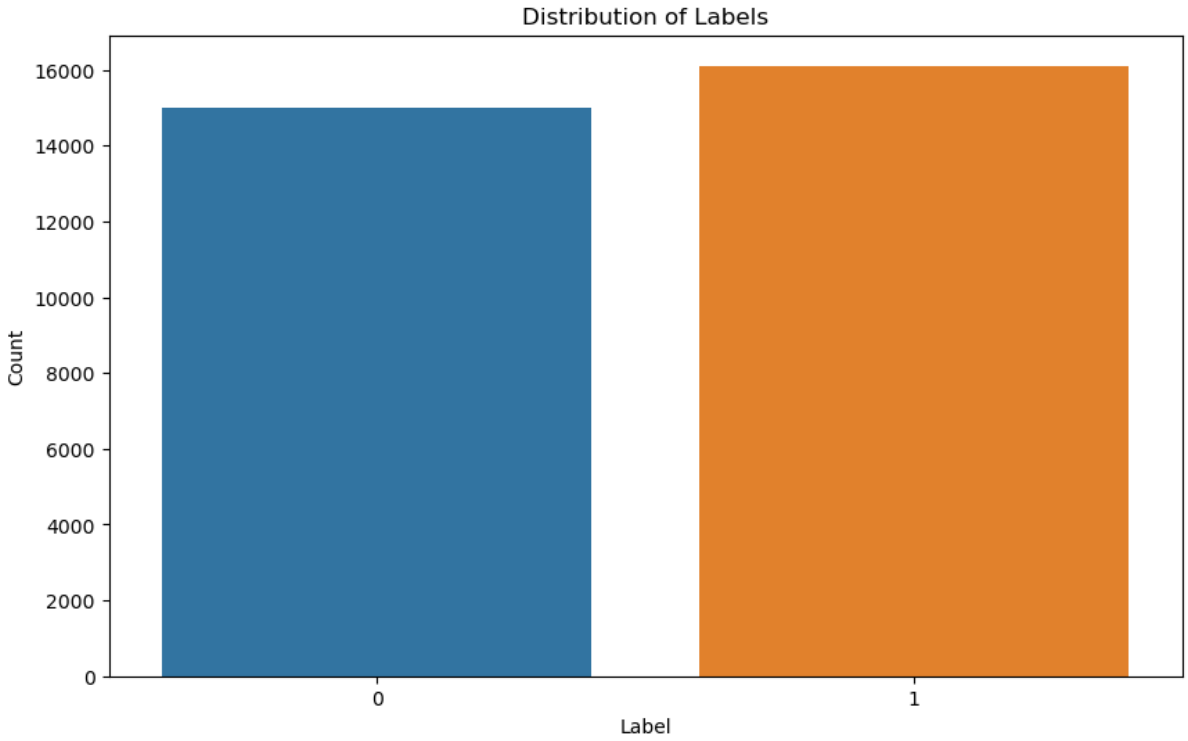
6.1 Experiment 1: Random Forest Classifier

Data Preparation and Feature Engineering:

The dataset was imported using pandas and explored to verify its structure and integrity.

Categorical variables were one-hot encoded using pandas' get_dummies method to prepare the data for modeling.

Feature scaling was performed using MinMaxScaler to normalize the data for improved model performance.

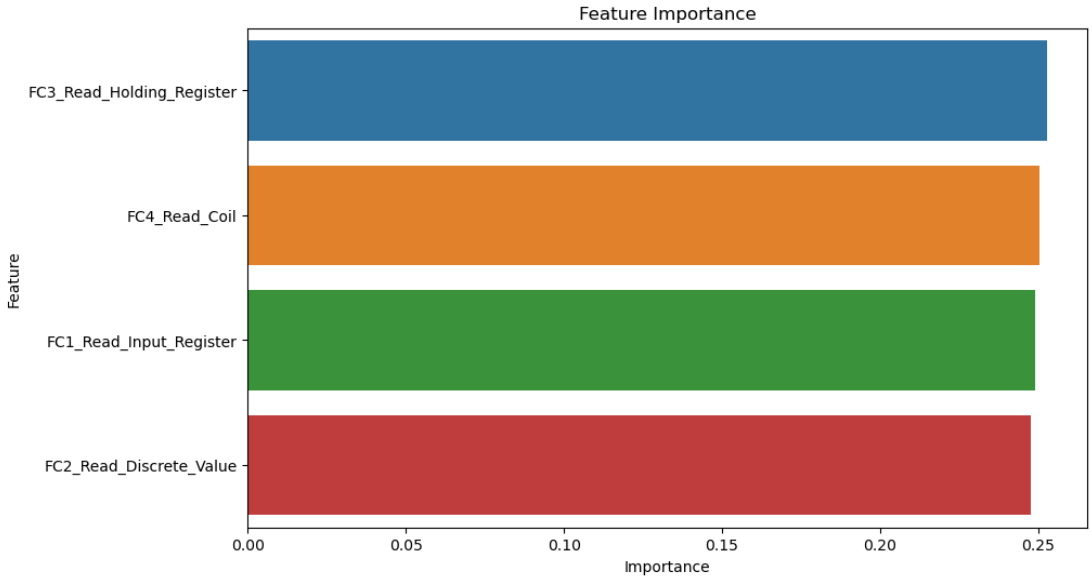


Model Training and Evaluation:

The Random Forest Classifier was trained on the preprocessed data, achieving a remarkable accuracy of 98.99% on the test data.

Metric	Class 0	Class 1	Macro Avg	Weighted Avg	Total Instances
Precision	0.98	0.99	0.985	0.985	-
Recall	0.99	0.98	0.985	0.985	-
F1-score	0.99	0.99	0.99	0.99	-
Support	3010	3212	-	-	-
Accuracy	-	-	-	0.99	6222
Macro average	0.99	0.99	0.99	-	-
Weighted average	0.99	0.99	0.99	-	-

Feature importance analysis revealed that all features are equal.



Visualizations, such as a count plot of label distribution and a bar plot of feature importance, provided insights into the dataset and model performance.

Further Analysis:

A correlation matrix was computed to explore relationships between features, and the correlation of features with the target variable ('label') was examined.

6.2 Experiment 2: Support Vector Machine Classifier

Data Preparation and Feature Engineering:

Similar data preprocessing steps were applied as in Case Study 1, including one-hot encoding of categorical variables and feature scaling using MinMaxScaler.

Model Training and Evaluation:

The SVM Classifier was trained on the preprocessed data but exhibited significantly lower performance, with an accuracy of only 53.33% on the test data.

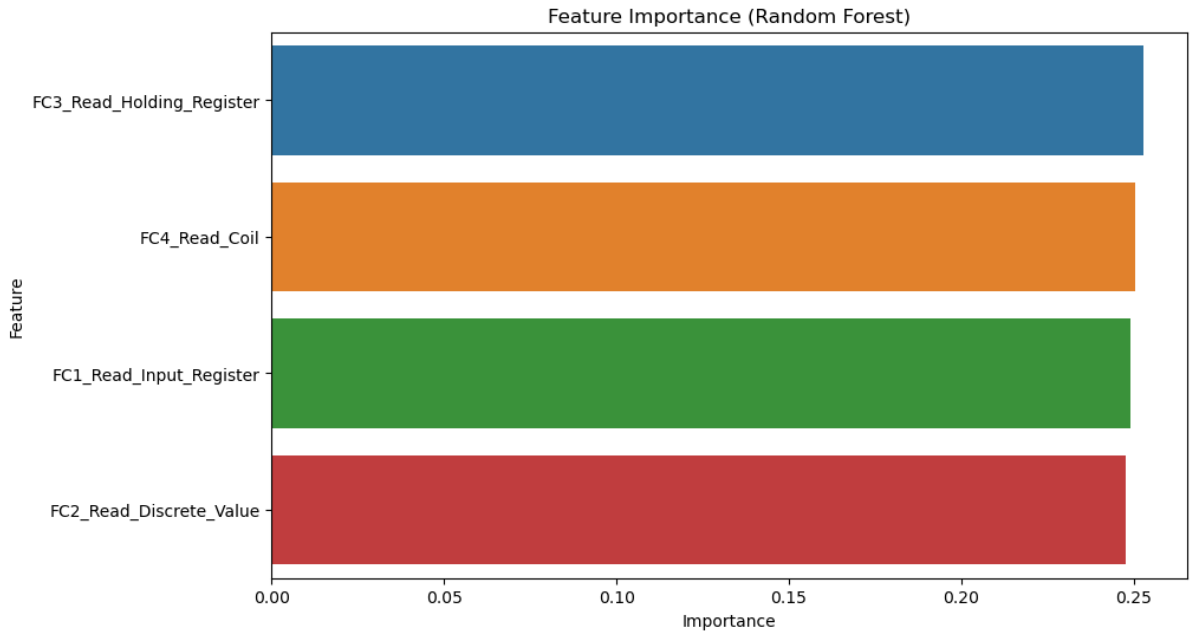
Despite this, feature importance analysis revealed consistency with the Random Forest model, identifying FC4_Read_Coil and FC3_Read_Holding_Register as important features.

Metric	Class 0	Class 1	Macro Avg	Weighted Avg	Total Instances
Accuracy	-	-	-	0.5333	6222
Precision	0.53	0.54	0.535	0.53	-
Recall	0.32	0.73	0.525	0.53	-
F1-score	0.4	0.62	0.51	0.51	-
Support	3010	3212	-	-	-

Further Analysis:

Similar visualizations and analyses were conducted to understand the dataset and model performance, highlighting the challenges and limitations of the SVM approach compared to the Random Forest Classifier.

The lower accuracy of the SVM model raised questions about its suitability for this specific dataset and problem domain, indicating the need for further investigation or alternative modeling approaches.



6.3 Discussion

The findings from both experiments underscore the importance of selecting appropriate machine learning algorithms for Industrial Internet of Things security applications. In our investigation, ensemble methods like Random Forests demonstrated remarkable accuracy and robustness in distinguishing between normal and anomalous behavior within IIoT network traffic. This research highlights the efficacy of the Random Forest Classifier, achieving a notable accuracy rate of 98.99%. The detailed classification report reveals high precision, recall, and F1-score for both normal and anomalous behavior classifications, further emphasizing the classifier's effectiveness.

Moreover, the feature importance analysis conducted on the Random Forest model provided valuable insights into the significance of various features. Notably, features such as FC4_Read_Coil and FC3_Read_Holding_Register emerged as highly influential in the classification process. Understanding the importance of these features enhances our ability to discern patterns and characteristics indicative of potential threats within IIoT network data.

7 Conclusion and Future Work

In conclusion, this study contributes significantly to the understanding and advancement of continuous risk monitoring systems for IIoT environments. By leveraging machine learning algorithms such as the Random Forest Classifier, we have demonstrated the potential for accurately classifying normal and anomalous behavior in IIoT network traffic. However, the journey does not end here. Future research endeavors could delve deeper into exploring the impact of hyperparameter tuning on model performance, thereby optimizing the efficacy of these algorithms for IIoT security applications.

Furthermore, there is a pressing need to expand our experiments beyond a single dataset and explore the generalizability of our findings across diverse datasets. Additionally, addressing the challenges of interpretability and generalizability remains a priority in enhancing the reliability and applicability of IIoT risk monitoring systems. Deploying pilot systems in real-world IIoT environments and integrating them with existing security infrastructure could provide invaluable insights into the practical efficacy of these systems.

Finally, the commercialization potential of tailored continuous risk monitoring solutions for various industries cannot be overlooked. Enterprises seeking to bolster their cybersecurity posture and comply with regulatory requirements stand to benefit significantly from the implementation of such solutions. As such, future efforts could focus on refining and customizing these solutions to cater to specific industry needs, thereby maximizing their impact and relevance in today's rapidly evolving threat landscape.

References

Lackner, Maximilian & Markl, Erich & A, Mohamed. (2018). Cybersecurity Management for (Industrial) Internet of Things: Challenges and Opportunities. Journal of Information Technology & Software Engineering. 08. 10.4172/2165-7866.1000250.

Wang, Maoli & Sun, Yu & Sun, Hongtao & Zhang, Bowen. (2023). Security Issues on Industrial Internet of Things: Overview and Challenges. Computers. 12. 256. 10.3390/computers12120256.

Serror, M., Hack, S., Henze, M., Schuba, M. and Wehrle, K. (2020). Challenges and Opportunities in Securing the Industrial Internet of Things. IEEE Transactions on Industrial Informatics, 17(5), pp.1–1. doi:https://doi.org/10.1109/tii.2020.3023507.

Mauri, Lara, and Ernesto Damiani. "Modeling Threats to AI-ML Systems Using STRIDE." Sensors, vol. 22, no. 17, 3 Sept. 2022, p. 6662, https://doi.org/10.3390/s22176662. Accessed 11 Sept. 2022.

Manjari, A. (2021). To provide security in IoT for industrial solutions. International Journal of Research in Engineering and Science (IJRES) ISSN, [online] 9(9), pp.95–101. Available at: https://www.ijres.org/papers/Volume-9/Issue-9/Ser-5/L090995101.pdf [Accessed 21 Apr. 2024].

Tan, S.F. and Samsudin, A. (2021). Recent Technologies, Security Countermeasure and Ongoing Challenges of Industrial Internet of Things (IIoT): A Survey. Sensors (Basel, Switzerland), [online] 21(19), p.6647. doi:https://doi.org/10.3390/s21196647.

Saini, M. and Saini, R. (2019). Internet of Things (IoT) Applications and Security Challenges: A Review. [online] Available at: https://www.ijert.org/research/internet-of-things-iot-applications-and-security-challenges-a-review-IJERTCONV7IS12028.pdf.

Aziz, M., Elmedany, W. and Sharif, M.S. (2023). Securing IoT devices against emerging security threats: Challenges and mitigation techniques. Journal of cyber security technology, 7(4), pp.1–25. doi:https://doi.org/10.1080/23742917.2023.2228053.

AlSalem, T.S., Almaiah, M.A. and Lutfi, A. (2023). Cybersecurity Risk Analysis in the IoT: A Systematic Review. Electronics, [online] 12(18), p.3958. doi:https://doi.org/10.3390/electronics12183958.

Sontan, Adewale & Samuel, Segun. (2024). Emerging Trends in Cybersecurity for Critical Infrastructure Protection: A Comprehensive Review. Computer Science & IT Research Journal. 5. 576-593. 10.51594/csitrj.v5i3.872.

Mingo, Horace & Lawson, Misty & Williamson, Amber. (2024). Identifying New Vulnerabilities Embedded in Consumer Internet of Things (IoT) Devices. 10.4018/979-8-3693-3226-9.ch011.

Yash Bobde, Narayanan, G., Manas Jati, Raj, P., Cvitić, I. and Dragan Peraković (2024). Enhancing Industrial IoT Network Security through Blockchain Integration. Electronics, 13(4), pp.687–687. doi:https://doi.org/10.3390/electronics13040687.

Gyamfi, Eric & Jurcut, Anca. (2022). Intrusion Detection in Internet of Things Systems: A Review on Design Approaches Leveraging Multi-Access Edge Computing, Machine Learning, and Datasets. Sensors. 22. 10.3390/s22103744.

Safitra, M.F., Lubis, M. and Fakhrurroja, H. (2023). Counterattacking Cyber Threats: A Framework for the Future of Cybersecurity. Sustainability, [online] 15(18), p.13369. Available at: https://www.mdpi.com/2071-1050/15/18/13369.

Cremer F, Sheehan B, Fortmann M, Kia AN, Mullins M, Murphy F, Materne S. Cyber risk and cybersecurity: a systematic review of data availability. Geneva Pap Risk Insur Issues Pract. 2022;47(3):698-736. doi: 10.1057/s41288-022-00266-6. Epub 2022 Feb 17. PMID: 35194352; PMCID: PMC8853293.

Albshaier, L., Almarri, S. and Rahman, M.M.H. (2024). A Review of Blockchain's Role in E-Commerce Transactions: Open Challenges, and Future Research Directions. Computers, [online] 13(1), p.27. doi:https://doi.org/10.3390/computers13010027.

Alwahedi, Fatima & Aldhaheri, Alyazia & Ferrag, Mohamed Amine & Battah, Ammar & Tihanyi, Norbert. (2024). Machine learning techniques for IoT security: Current research and future vision with generative AI and large language models. Internet of Things and Cyber-Physical Systems. 10.1016/j.iotcps.2023.12.003.

Alyazia Aldhaheri, Alwahedi, F., Mohamed Amine Ferrag and Ammar Battah (2024). Deep learning for cyber threat detection in IoT networks: A review. Internet of Things and Cyber-Physical Systems, 4, pp.110–128. doi:https://doi.org/10.1016/j.iotcps.2023.09.003.

GitHub. "Machine Learning Algorithms Repository." Available at: [https://github.com/] (Accessed: April 2024).

Towards Data Science. "Introduction to Deep Learning." Available at: [https://towardsdatascience.com/] (Accessed: April 2024).
Medium. "Data Science and Machine Learning Articles." Available at: [https://medium.com/] (Accessed: April 2024).

Stack Overflow. "Understanding Support Vector Machine algorithm from examples." Available at: [https://stackoverflow.com/] (Accessed: April 2024).

GeeksforGeeks. "Decision Trees." Available at: [https://www.geeksforgeeks.org/] (Accessed: April 2024).
Project Jupyter. Available at: [https://jupyter.org/] (Accessed: April 2024).

Moustafa, N., (2021). A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets. Sustainable Cities and Society, 102994.

Booij, T. M., Chiscop, I., Meeuwissen, E., Moustafa, N., & den Hartog, F. T. H., (2021). ToN IoT-The role of heterogeneity and the need for standardization of features and attack types in IoT network intrusion datasets. IEEE Internet of Things Journal.

Alsaedi, A., Moustafa, N., Tari, Z., Mahmood, A., & Anwar, A., (2020). TON_IoT telemetry dataset: a new generation dataset of IoT and IIoT for data-driven Intrusion Detection Systems. IEEE Access, 8, 165130-165150.

Moustafa, N., Keshk, M., Debie, E., & Janicke, H., (2020). Federated TON_IoT Windows Datasets for Evaluating AI-Based Security Applications. In 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom) (pp. 848-855). IEEE.

Moustafa, N., Ahmed, M., & Ahmed, S., (2020). Data Analytics-Enabled Intrusion Detection: Evaluations of ToN_IoT Linux Datasets. In 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom) (pp. 727-735). IEEE.

Moustafa, N., (2019). New Generations of Internet of Things Datasets for Cybersecurity Applications based Machine Learning: TON_IoT Datasets. In Proceedings of the eResearch Australasia Conference, Brisbane, Australia.

Moustafa, N., (2019). A systemic IoT-Fog-Cloud architecture for big-data analytics and cyber security systems: a review of fog computing. arXiv preprint arXiv:1906.01055.

Ashraf, J., Keshk, M., Moustafa, N., Abdel-Basset, M., Khurshid, H., Bakhshi, A. D., & Mostafa, R. R., (2021). IoTBoT-IDS: A Novel Statistical Learning-enabled Botnet Detection Framework for Protecting Networks of Smart Cities. Sustainable Cities and Society, 103041.