INSTITUTE OF AERONAUTICAL ENGINEERING
DUNDIGAL, HYDERABAD.

TWO WEEKS SUMMER INTERNSHIP

ON

Modelling, Detecting, and Mitigating Threats Against Industrial Healthcare Systems: A Combined Software Defined Networking and Reinforcement Learning Approach

BY

NANDI VARDHAN REDDY KUMMETHA

20951A1248

INFORMATION TECHNOLOGY

UNDER THE GUIDANCE

OF

MR.SHIVA

ABSTRACT

The emergence of the Internet of Medical Things (IoMT) has brought about significant advancements in the healthcare industry, offering benefits like remote medical assistance, real-time monitoring, and enhanced control. However, alongside these valuable healthcare services, the IoMT has also introduced notable concerns regarding cybersecurity and privacy. This article focuses on the IEC 60 870-5-104 protocol, a widely used protocol in industrial healthcare systems. Our aim is to investigate and evaluate the seriousness of cyberattacks targeting the IEC 60 870-5-104 protocol by employing a quantitative threat model that incorporates Attack Defense Trees and the Common Vulnerability Scoring System v3.1.

Additionally, we propose an intrusion detection and prevention system (IDPS) specifically designed to detect and mitigate IEC 60 870-5-104 cyberattacks. This IDPS makes effective use of machine learning (ML) and software-defined networking (SDN) technologies. ML techniques are employed to identify IEC 60 870-5-104 cyberattacks by analyzing two key elements: 1) network flow statistics based on the Transmission Control Protocol/Internet Protocol (TCP/IP) and 2) payload flow statistics related to the IEC 60 870-5-104 protocol. By leveraging these technologies, our proposed IDPS can automatically discern and counteract potential cyber threats associated with the IEC 60 870-5-104 protocol.

KEYWORDS:

Internet of Medical Things (IoMT), cybersecurity, privacy, IEC 60 870-5-104 protocol, quantitative threat model, Attack Defence Trees, Common Vulnerability Scoring System (CVSS), intrusion detection and prevention system (IDPS), machine learning (ML), software-defined networking (SDN), network flow statistics, payload flow statistics, cyberattacks, healthcare systems.

TABLE OF CONTENTS

1 INTRODUCTION

The rise of interconnected systems and digital technologies in industrial healthcare systems has revolutionized the healthcare ecosystem, enabling improved patient care, remote monitoring, and efficient healthcare delivery. However, this digital transformation has also brought forth unprecedented cybersecurity challenges, exposing industrial healthcare systems to various threats. In this article, we delve into the critical issue of modelling, detecting, and mitigating threats against industrial healthcare systems. Specifically, we propose a combined approach that harnesses the power of Software Defined Networking (SDN) and Reinforcement Learning (RL) techniques to enhance the security of these systems.

1.1 Evolution of Industrial Healthcare Systems:

We begin by examining the evolution of industrial healthcare systems, highlighting the integration of advanced technologies and the benefits they bring. We explore the adoption of interconnected devices, the Internet of Things (IoT), and the emergence of the Industrial Internet of Things (IIoT) in healthcare settings. This sets the stage for understanding the security challenges that arise due to the increased connectivity and interdependencies within these systems.

1.2 Cybersecurity Threats in Industrial Healthcare Systems:

Next, we analyse the diverse range of cybersecurity threats that industrial healthcare systems face. We discuss the potential risks associated with unauthorized access, data breaches, malware attacks, and the manipulation of medical devices and data. By outlining the severity and implications of these threats, we emphasize the urgent need for robust security measures to safeguard patient privacy, system integrity, and overall public health.

1.3 Software Defined Networking (SDN) in Industrial Healthcare Systems:

To address the unique security challenges of industrial healthcare systems, we introduce SDN as a promising solution. We delve into the fundamental concepts of SDN, including its centralized control, programmability, and network virtualization capabilities. We highlight the advantages of using SDN for enhancing security in industrial healthcare systems, such as fine-grained access control, real-time monitoring, and rapid response to emerging threats.

1.4 Reinforcement Learning (RL) for Threat Detection and Mitigation:

Building upon the foundation of SDN, we explore the integration of RL techniques to improve threat detection and mitigation in industrial healthcare systems. We discuss how RL algorithms can learn from interactions with the system's environment to make intelligent decisions and dynamically adapt security policies. We examine the potential applications of RL in anomaly detection, intrusion prevention, and response automation, emphasizing its ability to handle complex and evolving threats.

1.5 Combined SDN and RL Approach for Industrial Healthcare Systems:

In this section, we propose a novel combined approach that leverages the benefits of SDN and RL in synergy. We outline how SDN provides the infrastructure and control framework for implementing RL-based security mechanisms. We discuss the integration of RL agents within the SDN architecture, enabling autonomous threat detection, decision-making, and adaptive mitigation strategies. We highlight the advantages of this approach, such as agility, scalability, and the ability to learn and adapt to emerging threats.

The increasing reliance on interconnected systems in industrial healthcare settings necessitates robust security measures to protect against cybersecurity threats. By combining SDN and RL techniques, we can enhance threat modelling, detection, and mitigation in industrial healthcare systems. This integrated approach offers a promising solution to tackle the evolving nature of threats and ensure the safety, privacy, and integrity of patients and healthcare infrastructure.

SYSTEM PROPOSAL

2.1 Existing System

The existing system on modelling, detecting, and mitigating threats against industrial healthcare systems combines the power of Software Defined Networking (SDN) and Reinforcement Learning (RL) to enhance the security of these systems. This approach builds upon previous research and advancements in both SDN and RL to develop a comprehensive solution for addressing cybersecurity challenges in industrial healthcare environments.

1.SDN-Based Security Framework: The existing system incorporates a security framework based on SDN principles. SDN allows for centralized control and management of the network infrastructure, enabling efficient security policy enforcement and real-time monitoring. It provides the foundation for implementing security measures such as access control, traffic filtering, and threat detection.

2. Reinforcement Learning for Threat Detection:

Reinforcement Learning techniques are integrated into the system to enhance threat detection capabilities. RL algorithms learn from interactions with the environment, enabling autonomous decision-making and the ability to adapt to evolving threats. The RL agents are trained on large datasets of network traffic and security events to recognize patterns and anomalies associated with potential threats.

3. Anomaly Detection and Intrusion Prevention:

The existing system focuses on anomaly detection as a crucial aspect of threat detection. Through the analysis of network traffic patterns and behaviour, the system can identify deviations from normal behaviour that may indicate a security threat. Once an anomaly is detected, the RL agents take proactive measures to prevent the intrusion by dynamically adjusting security policies, isolating affected devices, or triggering alarms for further investigation.

4. Response Automation and Mitigation Strategies:

Upon threat detection, the system employs RL-based decision-making to determine appropriate mitigation strategies. RL agents leverage their learned knowledge and predefined security policies to make informed decisions on how to respond to specific threats. This includes dynamically updating firewall rules, redirecting network traffic, or initiating countermeasures to neutralize the threat.

5. Continuous Learning and Adaptation:

The existing system incorporates a feedback loop for continuous learning and adaptation. RL agents continuously monitor the effectiveness of their actions and adjust their decision-making strategies based on feedback and evolving threat landscapes. This enables the system to improve its threat detection and mitigation capabilities over time.

The existing system combines SDN and RL to provide a comprehensive approach to modelling, detecting, and mitigating threats against industrial healthcare systems. By leveraging the benefits of SDN's centralized control and programmability and RL's adaptive decision-making, the system offers an advanced and proactive security solution for safeguarding the integrity and privacy of industrial healthcare environments.

2.1.1 Disadvantages

The existing system lacks implementation details regarding SDN-based mitigation, including problem formulation and the methodology employed. Furthermore, there is a need for more comprehensive methods for testing and training on large datasets.

SDN-BASED MITIGATION-PROBLEM FORMULATION AND METHODOLOGY: The current system falls short in terms of describing how SDN is utilized for mitigating threats in industrial healthcare systems. It does not provide a clear problem formulation, outlining the specific challenges and vulnerabilities present in these environments. Additionally, the methodology for implementing SDN-based mitigation techniques is not adequately explained.

INSUFFICIENT TESTING AND TRAINING ON LARGE DATASETS: The existing system also lacks sufficient information on the testing and training methods used, particularly concerning large datasets. Training machine learning algorithms on extensive and diverse datasets is crucial for their effectiveness in threat detection and classification. However, the article does not elaborate on the size, diversity, or sources of the datasets used for training. Furthermore, the testing methodology, including the evaluation of system performance, robustness, and scalability, is not sufficiently detailed.

The current system needs to provide a more thorough explanation of how SDN is applied for threat mitigation in industrial healthcare systems, including problem formulation and a detailed methodology. Additionally, there is a need to expand on the testing and training methods, particularly concerning large datasets, to ensure the system's effectiveness and reliability in real-world scenarios.

2.2 Proposed System

The proposed threat modelling approach, known as IEC 60 870-5-104, combines two important concepts: Attack-Defence Trees (ADT) and the Common Vulnerability Scoring System (CVSS). This approach helps in understanding cyberattack paths and assessing their associated risks.

In the context of ADT, we have two types of nodes: attacking nodes and defending nodes. Attacking nodes represent the goals and actions that a cyber attacker may use to compromise the security of a target system. Defending nodes, on the other hand, represent the defences that can be employed by the defender to address or mitigate a cyberattack.

Each node can have multiple child nodes of the same type, allowing for further refinement into specific subgoals and actions. If a node doesn't have any child nodes of the same type, it signifies a basic action. Additionally, a node can have child nodes of the opposite type, which defines a countermeasure.

Refinements in the ADT can be categorized into two types: conjunctive and disjunctive. In conjunctive refinement, the goal of a refined node is achieved only if all of its children also accomplish their goals. This type of refinement uses an AND operator. On the other hand, disjunctive refinement uses an OR operator, meaning the goal of a refined node is achieved if at least one of its children accomplishes its goal. In parallel, CVSS is a framework used for assessing vulnerabilities or attacks. It assigns a severity score between 0 and 10 to quantify the level of risk associated with each vulnerability.

The proposed IEC 60 870-5-104 threat modelling approach combines ADT and CVSS. ADT focuses on modelling the cyberattack paths using attacking and defending nodes, with refinements indicating specific subgoals and actions. CVSS, on the other hand, provides a scoring system to quantify the severity of vulnerabilities or attacks. Together, these components enhance our understanding of cyber threats and their potential impacts.

2.2.1 Advantages

The implemented system consists of two key components:

1)Detection performance

2)Mitigation performance, which handle operations on datasets effectively.

To enhance dataset prediction, a notification and response module (NRM) was developed as part of the system.

2.3LITERATURE SURVEY

Numerous research papers have explored cybersecurity challenges in the healthcare sector. Yaqoob et al. investigated the vulnerabilities present in smart medical devices and propose suitable countermeasures. Chenthara et al. discussed the cybersecurity and privacy challenges faced by e-health solutions in cloud-computing environments. Wolker-Roberts et al. also addressed countermeasures against internal threats in healthcare critical infrastructures.

Vijayakumar et al. present an anonymous authentication framework for wireless body area networks, while Sun et al. conducted an extensive survey on security and privacy issues in the Internet of Medical Things (IoMT). Furthermore, the subsequent discussion focuses on related works concerning three main areas:

1) IEC 60 870-5-104 threat modelling,

2) intrusion detection against IEC 60 870-5-104, and

3) cyberattack mitigation and prevention through Software-Defined Networking (SDN).

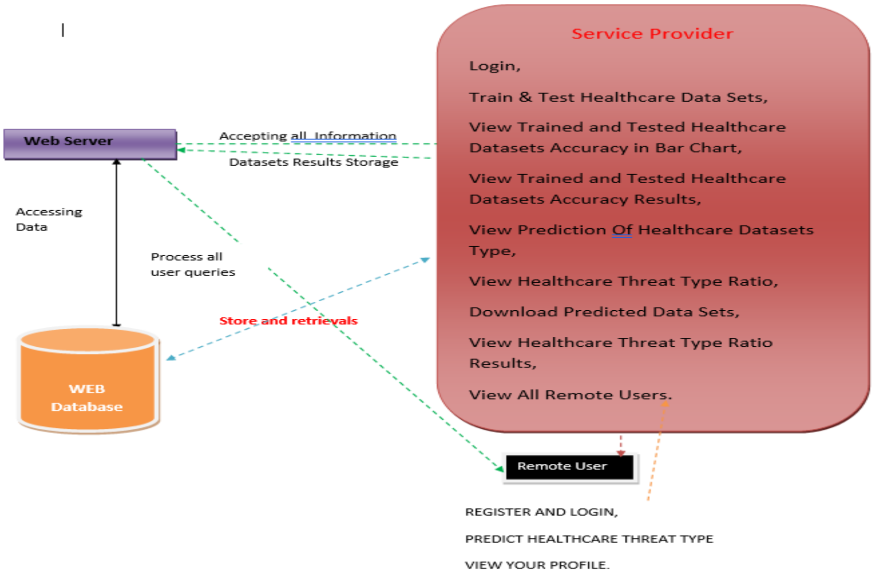
In a study, the authors perform an abstract threat analysis of IEC 60 870-5-104 industrial systems using coloured Petri nets (CPN). They categorize cyberattacks into two types: physical attacks and cyberattacks. Physical attacks involve activities carried out by an attacker with physical access to the target system, while cyberattacks exploit vulnerabilities in IEC 60 870-5-104. The latter category encompasses unauthorized access, man-in-the-middle (MITM) attacks, Denial-of-Service (DoS), and traffic analysis. Each cyberattack type is associated with CPN transitions, and their risks are quantified using the AlienVault OSSIM risk model.

Hodo et al. employed various machine learning (ML) algorithms to detect cyberattacks on an emulated industrial environment using the IEC 60 870-5-104 protocol. They evaluate the performance of ML classifiers such as Random Forest, OneR, J48, IBk, and Naive Bayes, using a dataset comprising replay attacks, DoS attacks, and address resolution protocol spoofing attacks. The evaluation indicates that J48 achieves the highest classification performance.

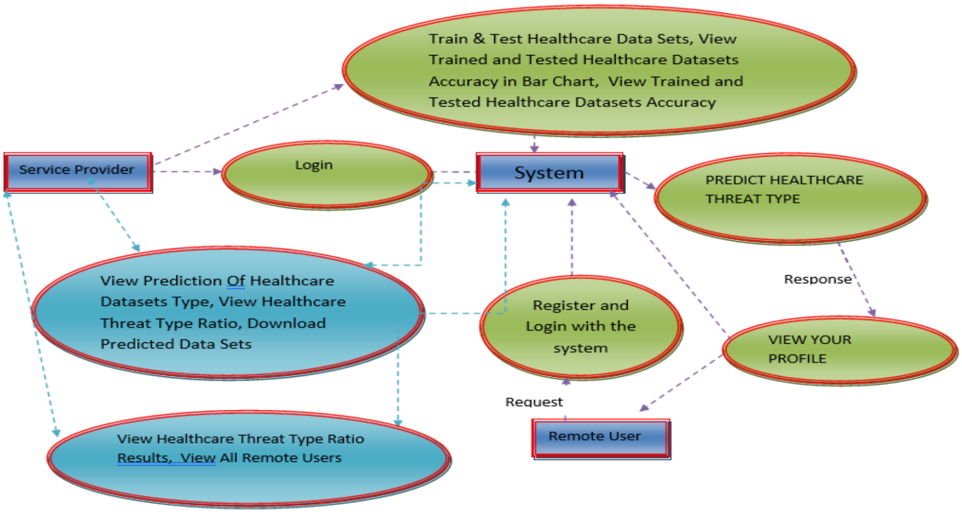
Yang et al. developed Snort-compliant signature and specification rules for detecting IEC 60 870-5-104-related cyberattacks. Signature rules define patterns associated with malicious behaviour, while specification rules define normal behaviour. Additionally, in the same authors propose a specification-based Intrusion Detection System (IDS) that recognizes anomalies in IEC 60 870-5-104. The IDS utilizes a detection state machine based on finite state machines, and experimental results demonstrate its effectiveness.

3 SYSTEM DIAGRAMS

3.1 ARCHITECTURE DIAGRAM

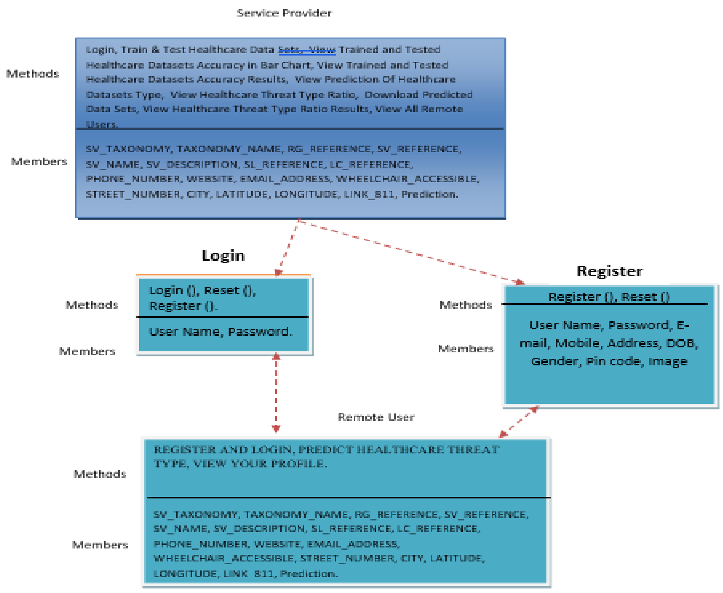

Service Provider
Login,
Train & Test Healthcare Data Sets,
View Trained and Tested Healthcare Datasets Accuracy in Bar Chart,
View Trained and Tested Healthcare Datasets Accuracy Results,
View Prediction Of Healthcare Datasets Type,
View Healthcare Threat Type Ratio,
Download Predicted Data Sets,
View Healthcare Threat Type Ratio Results,
View All Remote Users.
Web Server
Accepting all Information
Datasets Results Storage
Accessing Data
Process all user queries
Store and retrievals
WEB Database
Remote User
REGISTER AND LOGIN,
PREDICT HEALTHCARE THREAT TYPE
VIEW YOUR PROFILE.

3.2FLOW DIAGRAM

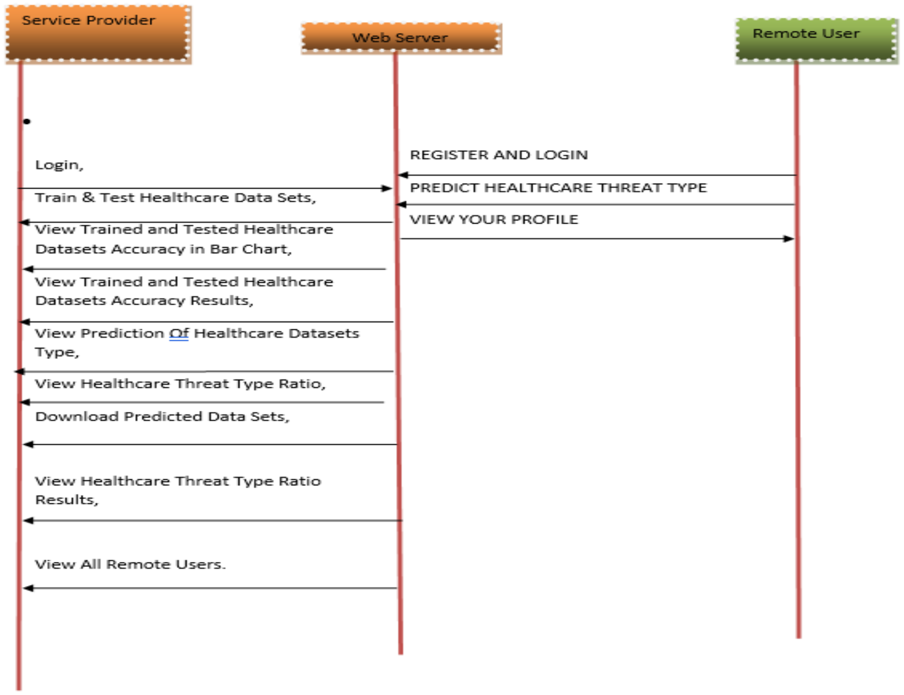

Train & Test Healthcare Data Sets, View Trained and Tested Healthcare Datasets Accuracy in Bar Chart, View Trained and Tested Healthcare Datasets Accuracy
Service Provider
Login
System
PREDICT HEALTHCARE THREAT TYPE
View Prediction Of Healthcare Datasets Type, View Healthcare Threat Type Ratio, Download Predicted Data Sets
Register and Login with the system
Response
VIEW YOUR PROFILE
Request
Remote User
View Healthcare Threat Type Ratio Results, View All Remote Users

3.3 UML Diagram

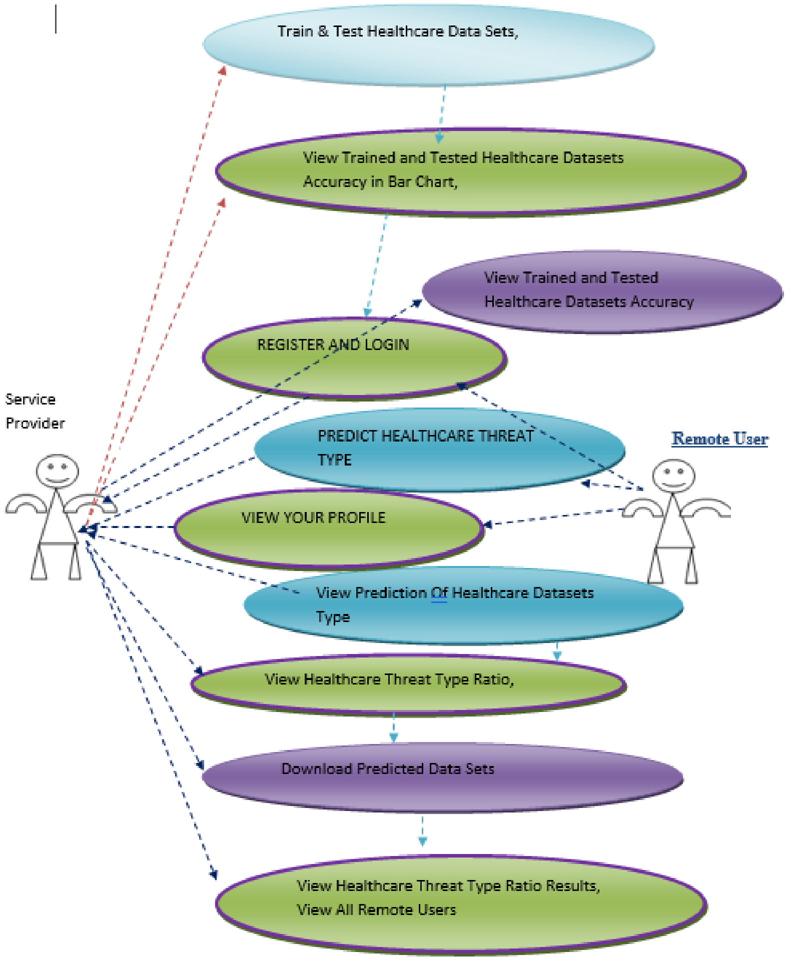
CLASS DIAGRAM


Service Provider
Methods
Login, Train & Test Healthcare Data Sets, View Trained and Tested Healthcare Datasets Accuracy in Bar Chart, View Trained and Tested Healthcare Datasets Accuracy Results, View Prediction Of Healthcare Datasets Type, View Healthcare Threat Type Ratio, Download Predicted Data Sets, View Healthcare Threat Type Ratio Results, View All Remote Users.
Members
SV_TAXONOMY, TAXONOMY_NAME, RG_REFERENCE, SV_REFERENCE, SV_NAME, SV_DESCRIPTION, SL_REFERENCE, LC_REFERENCE, PHONE_NUMBER, WEBSITE, EMAIL_ADDRESS, WHEELCHAIR_ACCESSIBLE, STREET_NUMBER, CITY, LATITUDE, LONGITUDE, LINK_811, Prediction.
Login
Methods
Login (), Reset (), Register ().
Members
User Name, Password.
Register
Methods
Register (), Reset ()
Members
User Name, Password, E-mail, Mobile, Address, DOB, Gender, Pin code, Image
Remote User
Methods
REGISTER AND LOGIN, PREDICT HEALTHCARE THREAT TYPE, VIEW YOUR PROFILE.
Members
SV_TAXONOMY, TAXONOMY_NAME, RG_REFERENCE, SV_REFERENCE, SV_NAME, SV_DESCRIPTION, SL_REFERENCE, LC_REFERENCE, PHONE_NUMBER, WEBSITE, EMAIL_ADDRESS, WHEELCHAIR_ACCESSIBLE, STREET_NUMBER, CITY, LATITUDE, LONGITUDE, LINK_811, Prediction.

SEQUENCE DIAGRAM


Service Provider
Web Server
Remote User
Login,
Train & Test Healthcare Data Sets,
View Trained and Tested Healthcare Datasets Accuracy in Bar Chart,
View Trained and Tested Healthcare Datasets Accuracy Results,
View Prediction Of Healthcare Datasets Type,
View Healthcare Threat Type Ratio,
Download Predicted Data Sets,
View Healthcare Threat Type Ratio Results,
View All Remote Users.
REGISTER AND LOGIN
PREDICT HEALTHCARE THREAT TYPE
VIEW YOUR PROFILE

USE CASE DIAGRAM


Train & Test Healthcare Data Sets,
View Trained and Tested Healthcare Datasets Accuracy in Bar Chart,
View Trained and Tested Healthcare Datasets Accuracy
REGISTER AND LOGIN
Service Provider
PREDICT HEALTHCARE THREAT TYPE
Remote User
VIEW YOUR PROFILE
View Prediction Of Healthcare Datasets Type
View Healthcare Threat Type Ratio,
Download Predicted Data Sets
View Healthcare Threat Type Ratio Results, View All Remote Users

4 IMPLEMENTATION

4.1ALGORITHMS

Decision tree classifiers

The decision tree classifier is a robust machine learning algorithm that is applied to both classification and regression tasks. It constructs a predictive model represented as a tree-like structure, with nodes representing features or attributes, branches representing possible feature values, and leaf nodes indicating the predicted class labels or outcomes.

To build the decision tree, the algorithm selects the most informative feature at each internal node to effectively split the data. This selection is based on criteria such as information gain, Gini impurity, or entropy, aiming to maximize the distinction between different classes.

When classifying new instances, the decision tree starts from the root node and follows a path down to a leaf node, considering the feature values encountered along the way. At each internal node, a decision is made based on the feature value, guiding the traversal through the branches until reaching a leaf node that provides the final predicted class label or outcome.

One of the key advantages of decision tree classifiers is their interpretability. The structure of the tree closely resembles human decision-making, making it easier to understand and explain the reasoning behind the predictions. Moreover, decision trees have the capability to handle both categorical and numerical features, enabling them to capture complex relationships within the data.

However, decision trees are susceptible to overfitting, which occurs when the model becomes overly complex and fails to generalize well to unseen data. To address this, techniques like pruning (removing unnecessary branches), limiting the depth of the tree, and setting a minimum number of samples at each node can be employed to prevent overfitting and enhance the model's performance on new data.

In summary, decision tree classifiers are versatile and intuitive algorithms that employ a tree-like structure to make predictions based on the features of the input data. They offer interpretability, can handle various types of data, and find applications in diverse domains for classification tasks..

Gradient boosting

Gradient boosting is a widely used machine learning approach employed in tasks such as regression and classification. It creates a prediction model by combining multiple weak prediction models, usually in the form of decision trees. When decision trees serve as the weak learners, the resulting algorithm is called gradient-boosted trees, which often outperform methods like random forest. The construction of a gradient-boosted trees model follows a step-by-step process, similar to other boosting techniques. At each step, the model focuses on minimizing a loss function that can be differentiated. By iteratively adding new weak learners to the ensemble, the model aims to gradually reduce the overall prediction error.

An advantage of gradient boosting lies in its flexibility to optimize various differentiable loss functions. This adaptability enables the model to cater to different problem domains and fine-tune the learning process according to specific objectives. Through the optimization of the chosen loss function, the gradient-boosted trees model strives to find the optimal combination of weak learners that collectively produce the most accurate predictions.

In comparison to ensemble methods like random forest, gradient-boosted trees often exhibit superior performance. This is because the iterative nature of gradient boosting allows the model to focus on challenging instances and assign higher weights to misclassified samples, thereby enhancing the learning process.

To summarize, gradient boosting is a powerful machine learning technique that constructs a prediction model by combining weak learners, typically decision trees. Through an iterative optimization process, the model achieves improved performance and demonstrates the ability to handle complex tasks in regression and classification.

K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a classification algorithm renowned for its simplicity and effectiveness. It classifies data points by comparing their similarity to other instances in the training dataset. KNN is considered non-parametric and lazy because it doesn't make assumptions about the data distribution and postpones learning until a test example is presented.

When a new data point needs to be classified, KNN identifies its k-nearest neighbors in the feature space. The feature space comprises variables that categorize the data, including both numerical and categorical attributes. By examining the k-closest neighbors, KNN determines the majority class label and assigns it to the new data point, basing predictions on the consensus of its neighbors.

The learning process of KNN is instance-based, relying on the specific instances available in the training dataset. This characteristic contributes to its lazy nature, as KNN doesn't create a generalized model during training. Instead, it dynamically searches for the nearest instances to the input vector during testing or prediction. However, it's important to note that finding the nearest neighbors can be computationally demanding, particularly with large training datasets or complex feature spaces.

To summarize, K-Nearest Neighbors (KNN) is a powerful classification algorithm that classifies data points by comparing their similarity to neighboring instances. Its simplicity and ability to handle different types of variables make it an intuitive approach for classification tasks.

Logistic regression Classifiers

Logistic regression analysis is a statistical method used to examine the relationship between a categorical dependent variable and a set of independent variables. When the dependent variable has two categories, it is referred to as binary logistic regression, while multinomial logistic regression is used when the dependent variable has three or more categories.

Compared to discriminant analysis, logistic regression is often considered more versatile and suitable for modelling various scenarios. One advantage is that logistic regression does not assume a normal distribution for the independent variables, unlike discriminant analysis.

This program calculates binary logistic regression and multinomial logistic regression models using both numeric and categorical independent variables. It provides essential information such as the regression equation, measures of goodness of fit, odds ratios, confidence intervals, likelihood, and deviance. Additionally, the program conducts a comprehensive analysis of residuals, including diagnostic reports and plots. It also offers the option to perform variable subset selection, aiming to identify the best regression model with the fewest independent variables. Furthermore, the program enables the generation of confidence intervals for predicted values and presents ROC curves, aiding in determining the optimal classification cut off point. It also facilitates result validation by automatically classifying rows that were not used during the analysis.

In summary, logistic regression analysis is a valuable tool for examining the relationship between categorical variables. The program described provides a range of features and statistical outputs to support the analysis and interpretation of logistic regression models.

Naïve Bayes

The Naive Bayes approach is a supervised learning method that makes a simplistic assumption: the presence or absence of one feature in a class is independent of the presence or absence of any other feature. Despite this assumption, Naive Bayes has demonstrated robustness and efficiency, comparable to other supervised learning techniques. One explanation for its performance is related to representation bias. Naive Bayes is considered a linear classifier, similar to linear discriminant analysis, logistic regression, and linear SVM. The difference lies in how the classifier's parameters are estimated, which influences the learning bias.

While Naive Bayes is popular in research, it is not widely used by practitioners seeking practical results. Researchers appreciate its ease of implementation, simple parameter estimation, fast learning even with large databases, and reasonably good accuracy compared to other methods. However, end-users often struggle to interpret and deploy the model, failing to recognize the benefits of this technique.

To address these challenges, a new presentation of Naive Bayes results is proposed, aiming to improve interpretability and deployment. The tutorial first explains the theoretical aspects of the Naive Bayes classifier, followed by its implementation on a dataset using Tanagra. The obtained results, including the model parameters, are compared to other linear approaches such as logistic regression, linear discriminant analysis, and linear SVM, revealing high consistency. This consistency contributes to the method's strong performance. In the second part of the tutorial, various tools like Weka, R, Knime, Orange, and RapidMiner are employed on the same dataset, with a focus on understanding the obtained results.

Random Forest

Random forests, also referred to as random decision forests, are a popular ensemble learning method employed for tasks like classification and regression. The underlying principle of random forests involves creating multiple decision trees during the training phase. In classification scenarios, the final prediction is determined by aggregating the majority vote from the individual trees, while in regression tasks, the average prediction of the trees is considered.

One significant advantage of random forests is their ability to address the overfitting problem commonly encountered with single decision trees. By constructing a diverse set of trees and combining their outputs, random forests enhance the reliability and robustness of the predictions. Although they generally outperform individual decision trees, their accuracy may be slightly lower compared to gradient boosted trees. However, the performance of random forests can be influenced by specific characteristics of the dataset being analysed.

The concept of random decision forests was initially introduced by Tin Kam Ho in 1995, where the random subspace method was employed to implement the "stochastic discrimination" approach proposed by Eugene Kleinberg. Leo Breiman and Adele Cutler further extended the algorithm by incorporating the "bagging" technique and introducing random feature selection. In fact, they even registered the term "Random Forests" as a trademark in 2006.

Random forests are often considered as "blackbox" models in practical business applications due to their ability to generate reliable predictions across diverse datasets with minimal configuration requirements. Their versatility and effectiveness in handling various tasks have led to their widespread adoption in the field of machine learning.

SVM

In classification tasks, discriminant machine learning techniques focus on finding a discriminant function that can accurately predict labels for new instances. Unlike generative approaches, which involve calculating conditional probability distributions, discriminant classifiers directly assign data points to specific classes without explicitly modelling the underlying probability distributions. Although they may have less modelling power compared to generative approaches, discriminant methods offer advantages such as reduced computational requirements and the ability to handle high-dimensional feature spaces, particularly when only posterior probabilities are of interest. From a geometric standpoint, training a classifier involves finding a multidimensional surface that effectively separates different classes in the feature space.

Support Vector Machines (SVM) are a specific type of discriminant technique. Unlike other classification algorithms like genetic algorithms (GAs) or perceptrons, SVM solves the convex optimization problem analytically, resulting in a consistent optimal hyperplane parameter. In contrast, solutions obtained by GAs or perceptrons can vary significantly depending on the initialization and termination criteria. By using a specific kernel to transform data from the input space to a higher-dimensional feature space, SVM training yields unique model parameters for a given training set. On the other hand, perceptrons and GA classifiers may produce different models each time training is initiated. The primary objective of GAs and perceptrons is to minimize training errors, which often leads to multiple hyperplanes meeting the error criteria.

4.2 MODULES

1)Service Provider Module

2) View and Authorize Users

3) Remote User

Service Provider Module:

The Service Provider module requires a valid username and password for logging in. Once successfully logged in, the user can perform various operations, including:

1. Login: The user can log in using their credentials.
2. Train & Test Healthcare Data Sets: This functionality allows the user to train and test healthcare datasets using specific algorithms or models.
3. View Trained and Tested Healthcare Dataset Accuracy in Bar Chart: The user can visualize the accuracy of the trained and tested healthcare datasets through a bar chart representation.
4. View Trained and Tested Healthcare Dataset Accuracy Results: The user can view detailed results of the accuracy achieved by training and testing healthcare datasets.
5. View Prediction of Healthcare Dataset Type: The user can observe the predicted type or category of healthcare datasets based on the trained models.
6. View Healthcare Threat Type Ratio: This feature provides insights into the ratio or proportion of different healthcare threat types within the datasets.
7. Download Predicted Data Sets: The user can download the datasets with predictions made by the trained models.
8. View Healthcare Threat Type Ratio Results: Detailed results and analysis of the healthcare threat type ratio can be viewed using this functionality.
9. View All Remote Users: The user can access information about all the registered remote users of the system.

View and Authorize Users Module:

This module is specifically designed for administrators, allowing them to perform the following tasks:

1. View Users: The administrator can see a list of all registered users. This list includes details such as the user's name, email address, and physical address.

2. Authorize Users: The administrator has the authority to authorize or grant permission to registered users for accessing certain features or functionalities of the system.

Remote User Module:

In this module, multiple users can register and access the system. Users need to register with their relevant details, which are then stored in the system's database. Once registration is successful, the user can log in using their authorized username and password. The module offers the following operations:

1. Register and Login: Users can register by providing necessary information and subsequently log in using their registered credentials.

2. Predict Healthcare Threat Type: After logging in, users have the ability to predict the type of healthcare threat based on the available data and models.

3. View Your Profile: Users can access and view their own profile information, which includes personal details and any relevant preferences they may have specified.

The Service Provider module enables authorized users to perform various operations related to healthcare datasets, while the View and Authorize Users module empowers administrators to manage user registrations and authorizations. The Remote User module allows registered users to utilize the system's functionalities, including healthcare threat prediction and profile viewing.

SYSTEM REQUIREMENTS

H/W System Configuration:-

➢ Processor	-	Pentium –IV
➢ RAM	-	4 GB (min)
➢ Hard Disk	-	20 GB
➢ Key Board	-	Standard Windows Keyboard
➢ Mouse	-	Two or Three Button Mouse
➢ Monitor	-	SVGA

SOFTWARE REQUIREMENTS:

➢ Operating system	:	Windows 7 Ultimate.
➢ Coding Language	:	Python.
➢ Front-End	:	Python.
➢ Back-End	:	Django-ORM
➢ Designing	:	Html, css, javascript.
➢ Data Base	:	MySQL (WAMP Server)

CONCLUSION AND FUTURE ENHANCEMENT

6.1 Conclusion

As the healthcare industry undergoes digital transformation and integrates the Internet of Medical Things (IoMT), there is a growing concern about the security vulnerabilities of legacy healthcare systems. In this article, the focus was on examining the IEC 60 870-5-104 protocol commonly used in industrial systems within the healthcare sector. The objective was to develop a quantitative threat model to assess the potential severity of cyber attacks targeting the specific commands of the IEC 60 870-5-104 protocol.

To address these security concerns, an Intrusion Detection and Prevention System (IDPS) was proposed that combines Machine Learning (ML) techniques with Software-Defined Networking (SDN). The IDPS utilizes a CART classifier, which leverages TCP/IP network flow statistics and IEC 60 870-5-104 payload flow statistics to detect potential cyber attacks. Additionally, SDN is employed for mitigation purposes and is treated as a Multi-Armed Bandit (MAB) problem solved using the Thompson Sampling (TS) method.

The evaluation of the proposed IDPS demonstrated its effectiveness in detecting and mitigating IEC 60 870-5-104 cyberattacks. Moving forward, the researchers have outlined their future plans to enhance the IDPS by extending its capabilities to detect multi-step cyberattacks involving the IEC 60 870-5-104 protocol as well as other industrial and IoMT protocols commonly used in the healthcare sector, such as Modbus, MQTT, and EtherCAT. This enhancement will involve the adoption of ML-based association rules techniques.

In summary, the article addresses the security challenges posed by the adoption of the IEC 60 870-5-104 protocol in healthcare systems. It proposes an IDPS that combines ML and SDN, demonstrates its effectiveness, and outlines future plans to expand its capabilities to address a wider range of cyberattacks targeting various protocols utilized in the healthcare industry.

6.2 Future Enhancement

In the context of modelling, detecting, and mitigating threats against industrial healthcare systems, future enhancements can be made to further improve the security measures. One potential approach involves combining Software-Defined Networking (SDN) with Reinforcement Learning (RL) techniques.

1. Modelling: Future enhancements can focus on refining the models used for threat detection and mitigation. This includes incorporating more sophisticated algorithms and techniques from the field of RL to capture complex patterns and behaviour of attackers in industrial healthcare systems. The models can be trained on extensive datasets that encompass a wide range of attack scenarios, enabling more accurate identification and prediction of threats.

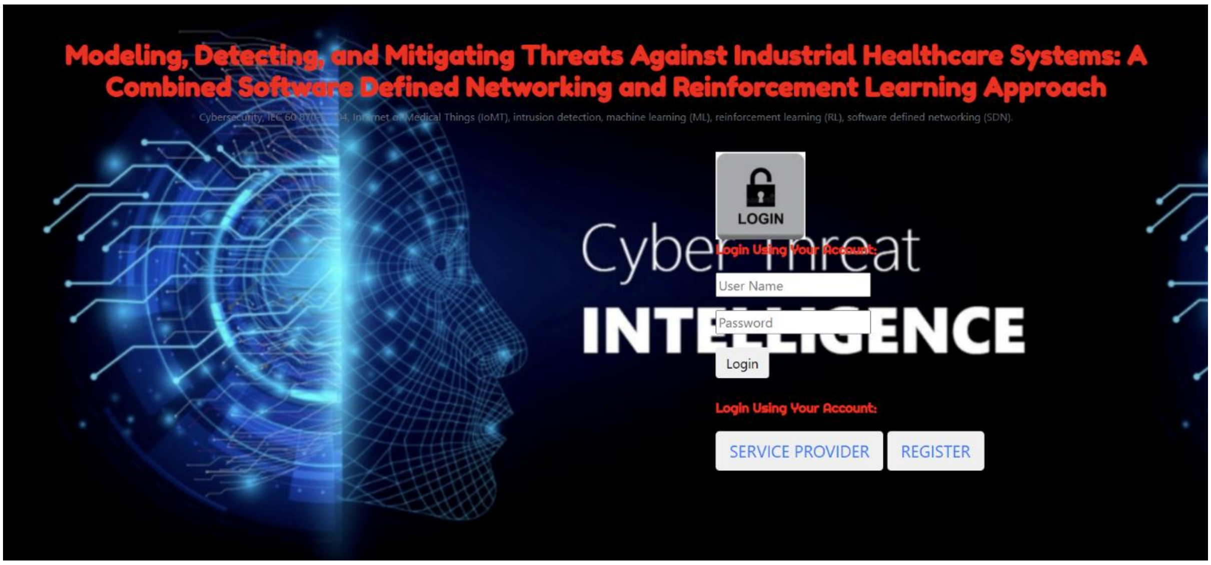
2. Detecting: To enhance threat detection capabilities, RL algorithms can be applied to learn and adapt to evolving attack strategies. Reinforcement Learning agents can continuously monitor network traffic and system behaviour, learning from past experiences and adjusting their detection mechanisms accordingly. This dynamic approach allows for real-time detection of new and emerging threats in industrial healthcare systems.

3. Mitigating: Reinforcement Learning can also play a role in developing effective mitigation strategies. RL agents can be trained to autonomously respond to detected threats by dynamically adjusting network configurations and implementing proactive security measures. By learning from past experiences, these agents can optimize response actions, minimize the impact of attacks, and strengthen the overall resilience of industrial healthcare systems.

4. Integration of SDN and RL: The combination of SDN and RL can further enhance the effectiveness of threat mitigation. SDN provides the flexibility and agility to dynamically control network behaviour, while RL algorithms enable intelligent decision-making in response to detected threats. Integrating these two approaches can result in a powerful framework that adapts to changing threat landscapes and effectively defends against sophisticated attacks in industrial healthcare systems.

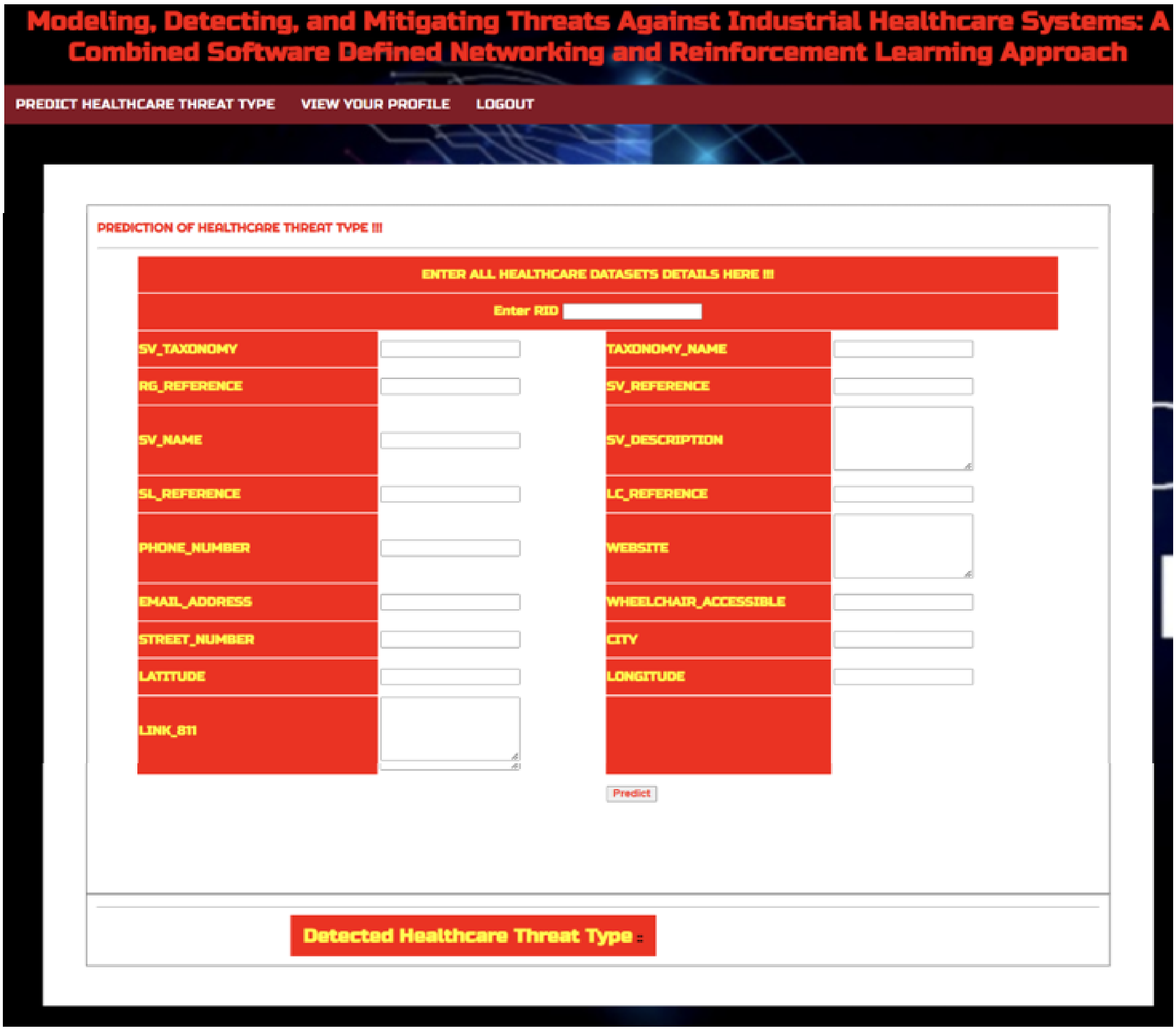
By leveraging the capabilities of SDN and RL together, future enhancements can enable more proactive and intelligent defence mechanisms. These approaches have the potential to provide robust protection against evolving threats, minimize the impact of attacks, and ensure the security and integrity of industrial healthcare systems.

7 SAMPLE CODINGS AND SAMPLE SCREEN SHOTS



Modeling, Detecting, and Mitigating Threats Against Industrial Healthcare Systems: A Combined Software Defined Networking and Reinforcement Learning Approach

Cybersecurity, IEC 60 870-5-104, Internet of Medical Things (IoMT), intrusion detection, machine learning (ML), reinforcement learning (RL), software defined networking (SDN).



Modeling, Detecting, and Mitigating Threats Against Industrial Healthcare Systems: A Combined Software Defined Networking and Reinforcement Learning Approach

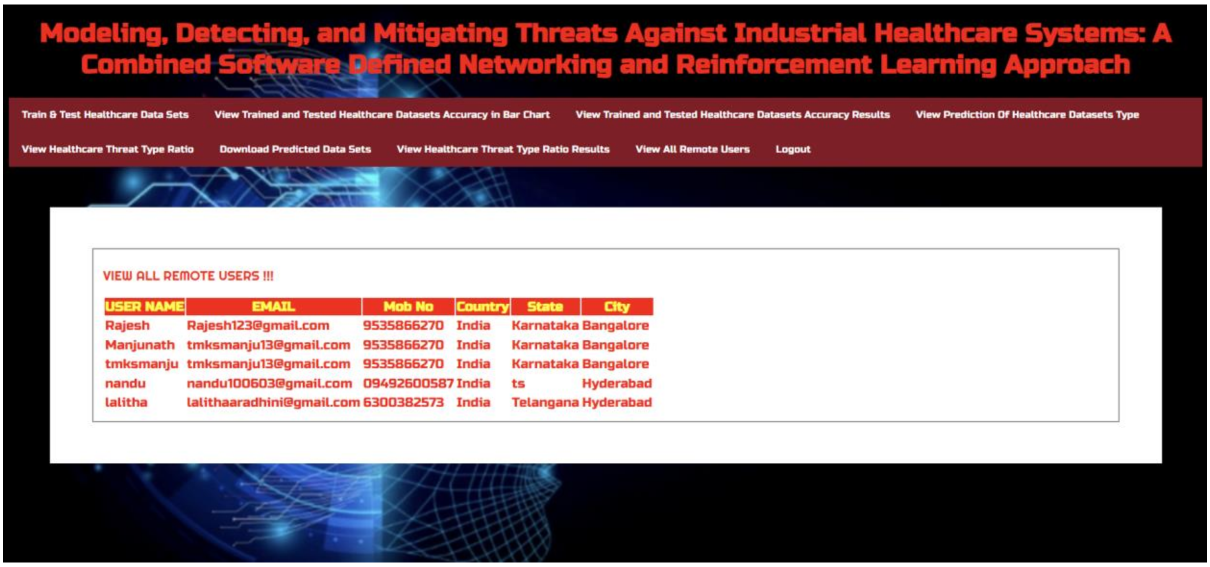
PREDICTION OF HEALTHCARE THREAT TYPE !!!

ENTER ALL HEALTHCARE DATASETS DETAILS HERE !!!

Enter RID

SV_TAXONOMY		TAXONOMY_NAME	
RG_REFERENCE		SV_REFERENCE	
SV_NAME		SV_DESCRIPTION	
SL_REFERENCE		LC_REFERENCE	
PHONE_NUMBER		WEBSITE	
EMAIL_ADDRESS		WHEELCHAIR_ACCESSIBLE	
STREET_NUMBER		CITY	
LATITUDE		LONGITUDE	
LINK_811			

Detected Healthcare Threat Type :

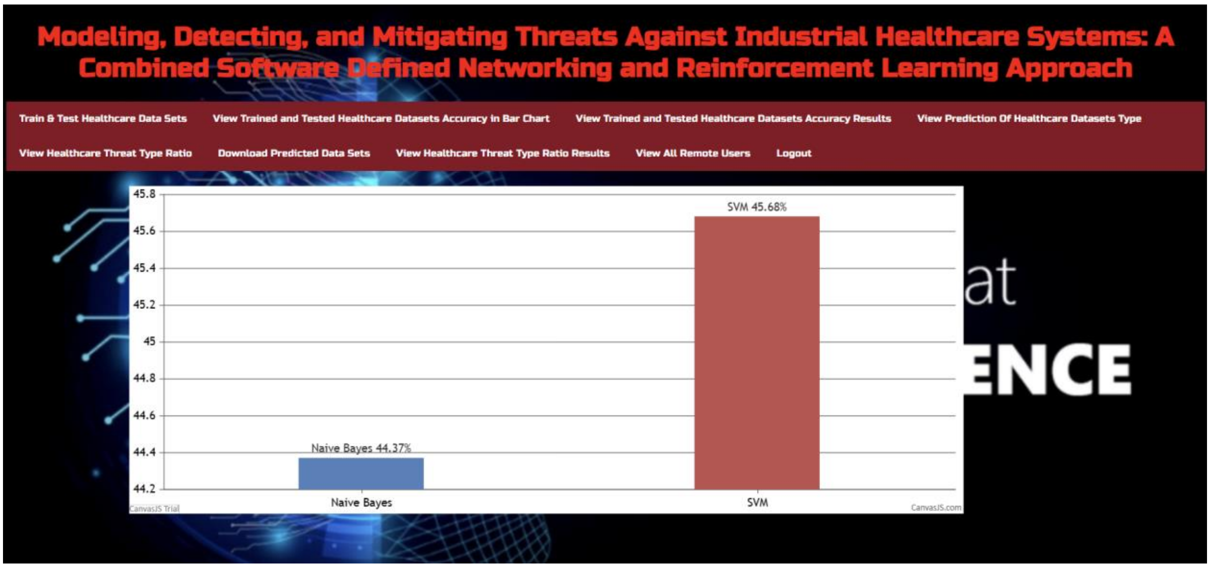


Modeling, Detecting, and Mitigating Threats Against Industrial Healthcare Systems: A Combined Software Defined Networking and Reinforcement Learning Approach

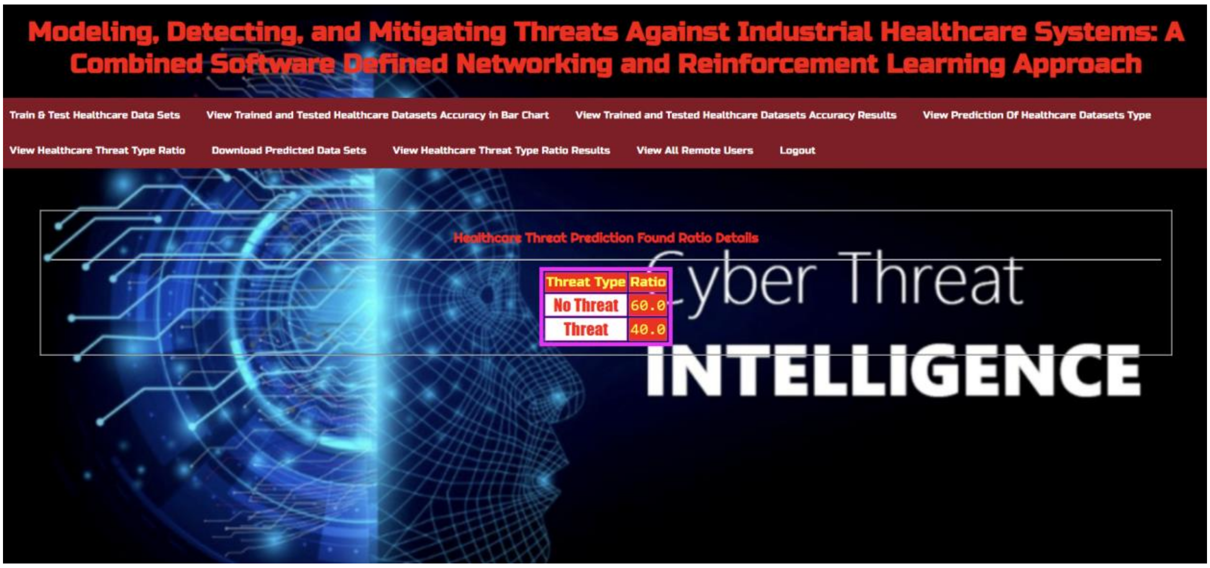
Train & Test Healthcare Data Sets
View Trained and Tested Healthcare Datasets Accuracy in Bar Chart
View Trained and Tested Healthcare Datasets Accuracy Results
View Prediction Of Healthcare Datasets Type
View Healthcare Threat Type Ratio
Download Predicted Data Sets
View Healthcare Threat Type Ratio Results
View All Remote Users
Logout

VIEW ALL REMOTE USERS !!!

USER NAME	EMAIL	Mob No	Country	State	City
Rajesh	Rajesh123@gmail.com	9535866270	India	Karnataka	Bangalore
Manjunath	tmksmanju13@gmail.com	9535866270	India	Karnataka	Bangalore
tmksmanju	tmksmanju13@gmail.com	9535866270	India	Karnataka	Bangalore
nandu	nandu100603@gmail.com	09492600587	India	ts	Hyderabad
lalitha	lalithaaradhini@gmail.com	6300382573	India	Telangana	Hyderabad



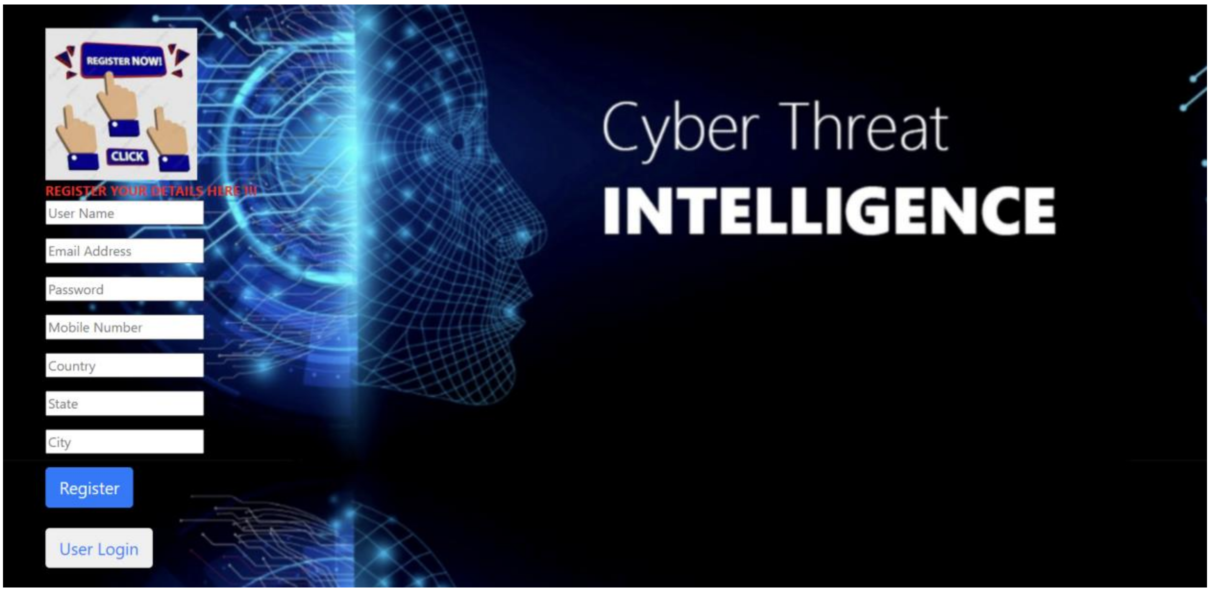
Modeling, Detecting, and Mitigating Threats Against Industrial Healthcare Systems: A Combined Software Defined Networking and Reinforcement Learning Approach


Modeling, Detecting, and Mitigating Threats Against Industrial Healthcare Systems: A Combined Software Defined Networking and Reinforcement Learning Approach

Train & Test Healthcare Data Sets
View Trained and Tested Healthcare Datasets Accuracy in Bar Chart
View Trained and Tested Healthcare Datasets Accuracy Results
View Prediction Of Healthcare Datasets Type
View Healthcare Threat Type Ratio
Download Predicted Data Sets
View Healthcare Threat Type Ratio Results
View All Remote Users
Logout
Cyber Threat
INTELLIGENCE

Healthcare Threat Prediction Found Ratio Details

Threat Type	Ratio
No Threat	60.0
Threat	40.0



REGISTER YOUR DETAILS HERE !!!

CODE:

```
#!/usr/bin/env python

"""Django's command-line utility for administrative tasks."""

import os

import sys

def main():

    """Run administrative tasks."""

    os.environ.setdefault('DJANGO_SETTINGS_MODULE', 
'modeling_detecting_and_mitigating_threats.settings')

    try:

        from django.core.management import execute_from_command_line

    except ImportError as exc:

        raise ImportError(

            "Couldn't import Django. Are you sure it's installed and "

            "available on your PYTHONPATH environment variable? Did you "

            "forget to activate a virtual environment?"

        ) from exc

    execute_from_command_line(sys.argv)



if __name__ == '__main__':

    main()
```

8 REFERENCE

[1] T. Yaqoob, H. Abbas, and M. Atiquzzaman, “Security vulnerabilities, attacks, countermeasures, and regulations of networkedmedical devices—A review,” *IEEE Commun. Surv. Tut.*, vol. 21, no. 4, pp. 3723–3768, Oct./Dec. 2019.

[2] M.Conti, D. Donadel, and F. Turrin, “A survey on industrial control system testbeds and datasets for security research,” 2021, *arXiv:2102.05631*.

[3] E. Hodo, S. Grebeniuk, H. Ruotsalainen, and P. Tavolato, “Anomaly detection for simulated IEC-60870-5-104 traffic,” in *Proc. 12th Int. Conf. Availability, Rel. Secur.*, 2017, pp. 1–7.

[4] Y. Yang, K. McLaughlin, T. Littler, S. Sezer, B. Pranggono, and H. Wang, “Intrusion detection system for IEC 60870-5-104 based SCADA networks,” in *Proc. IEEE Power Energy Soc. Gen.Meeting*, 2013, pp. 1–5.

[5] P.Radoglou-Grammatikis, P. Sarigiannidis, I. Giannoulakis,E.Kafetzakis, andE. Panaousis, “Attacking IEC-60870-5-104SCADAsystems,” in *Proc. IEEE World Congr. Serv. (SERVICES)*, 2019, pp. 41–46.

[6] P. R. Grammatikis, P. Sarigiannidis, A. Sarigiannidis, D. Margounakis, A. Tsiakalos, and G. Efstathopoulos, “An anomaly detection mechanism for IEC 60870-5-104,” in *Proc. 9th Int. Conf. Modern Circuits Syst. Technol.*, 2020, pp. 1–4.

[7] Y. Yang, K. McLaughlin, S. Sezer, Y. Yuan, and W. Huang, “Stateful intrusion detection for IEC 60870-5-104 SCADA security,” in *Proc. IEEE PES Gen. Meeting Conf. Expo.*, 2014, pp. 1–5.

[8] P. Radoglou-Grammatikis *et al.*, “Spear SIEM: A security information and event management system for the smart grid,” *Comput. Netw.*, vol. 193, 2021, Art. no. 108008.

[9] S. Chenthara, K. Ahmed, H. Wang, and F. Whittaker, “Security and privacy-preserving challenges of e-health solutions in cloud computing,” *IEEE Access*, vol. 7, pp. 74361–74382, 2019.

[10] S. Walker-Roberts, M. Hammoudeh, and A. Dehghantanha, “A systematic review of the availability and efficacy of countermeasures to internal threats in healthcare critical infrastructure,” *IEEE Access*, vol. 6,

pp. 25167–25177, 2018.
[11] P. Vijayakumar,M. S. Obaidat, M. Azees, S. H. Islam, and N. Kumar, “Efficient and secure anonymous authentication with location privacy for IoTbased wbans,” *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2603–2611, Apr. 2020.
[12] Y. Sun, F. P.-W. Lo, and B. Lo, “Security and privacy for the internet of medical things enabled healthcare systems:Asurvey,” *IEEE Access*, vol. 7, pp. 183339–183355, 2019.
[13] S. Meng *et al.*, “Security-aware dynamic scheduling for real-time optimization in cloud-based industrial applications,” *IEEE Trans. Ind. Informat.*, vol. 17, no. 6, pp. 4219–4228, Jun. 2021.
[14] H. Lin, “SDN-based in-network honeypot: Preemptively disrupt and mislead attacks in IoT networks,” 2019, *arXiv:1905.13254*.
[15] T. Xing, Z. Xiong, D. Huang, andD. Medhi, “SDNIPS: Enabling softwaredefined networking based intrusion prevention system in clouds,” in *Proc. 10th Int. Conf. Netw. Serv. Manage. Workshop*, 2014, pp. 308–311.
[16] B. Kordy, S. Mauw, S. Radomirovi´c, and P. Schweitzer, “Attack-defense trees,” *J. Log. Comput.*, vol. 24, no. 1, pp. 55–87, 2014. 17] P. Johnson, R. Lagerström, M. Ekstedt, and U. Franke, “Can the common vulnerability scoring system be trusted? A Bayesian analysis,” *IEEE Trans. Dependable Secure Comput.*, vol. 15, no. 6, pp. 1002–1015, Nov./Dec. 2018.
[18] A. Gatouillat, Y. Badr, B. Massot, and E. Sejdi´c, “Internet of medical things: A review of recent contributions dealing with cyber-physical systems in medicine,” *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3810–3822, Oct. 2018.
[19] M. H. Rehmani, F. Akhtar,A.Davy, andB. Jennings, “Achieving resilience in SDN-based smart grid: A multi-armed bandit approach,” in *Proc. 4th IEEE Conf. Netw. Softwarization Workshops*, 2018, pp. 366–371.
[20] D. Russo, B. Van Roy, A. Kazerouni, I. Osband, and Z. Wen, “A tutorial on thompson sampling,” 2017, *arXiv:1707.02038*.
[21] P. Radoglou-Grammatikis, P. Sarigiannidis, G. Efstathopoulos, and E. Panaousis, “Aries: A novel multivariate intrusion detection system for smart grid,” *Sensors*, vol. 20, no. 18, 2020, Art. no. 5305.