WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Smart Manufacturing Attack Detection and Classification with Advanced Fully Convolutional Neural Networks

Mr. B. Srikanth Reddy [a], Bethapudi Priyanka [b*] and Pallapothula Praneetha [b]

[a] *Assistant Professor, Vijayawada, 520001, India*
[b] *Student, Vijayawada, 520001, India*

A B S T R A C T

Operational integrity and data security are seriously threatened by the spread of Industrial Internet of Things (IoT) devices in smart manufacturing systems, which have created vulnerabilities open to cyberattacks. Because IoT settings are dynamic, traditional approaches for identifying and categorizing these assaults frequently lack scalability and flexibility. In order to improve attack detection and classification, this research presents a unique method based on a fully convolutional neural network (FCNN) architecture, enhanced by decision tree and logistic regression techniques. The suggested approach combines the robustness and interpretability provided by decision tree and logistic regression models with the benefits of FCNNs in capturing spatial relationships in input data. These methods work together to improve the system's ability to recognize and classify different kinds of assaults that are directed at IoT devices in smart manufacturing settings. In order to enable proactive mitigation of possible risks, the FCNN model effectively extracts features from log data streams. These characteristics are then fed into logistic regression and decision tree classifiers to produce accurate predictions in real-time.

Keywords: Industrial Internet of Things (IoT), Smart manufacturing, cyber-attacks, attack detection, attack classification, fully convolutional neural network (FCNN), logistic regression, decision tree, anomaly detection, data security.

Introduction

With previously unheard-of levels of automation, efficiency, and connection, the quick integration of Industrial Internet of Things (IoT) devices into smart manufacturing systems has completely changed industrial operations [1]. But with new weaknesses brought about by technological progress, essential infrastructure is now vulnerable to cyberattacks that might have disastrous effects. Strong cybersecurity measures are essential for securing assets, maintaining operational continuity, and protecting sensitive data as companies depend more and more on networked IoT devices to maximize services. Cybersecurity professionals have a special challenge when operational technology (OT) and information technology (IT) merge in smart industrial settings. IoT ecosystems, in contrast to standard IT networks, comprise a wide range of embedded devices, sensors, and actuators found in industrial machinery and production lines [2]. While improving real-time monitoring and management, these networked gadgets also provide a large attack surface that may be exploited by malevolent parties. As a result, protecting IoT infrastructures from cyberattacks becomes an urgent societal necessity that is essential to preserving the robustness and integrity of contemporary industrial systems.

The growing frequency and sophistication of cyberattacks on industrial facilities highlight the seriousness of the cybersecurity threat facing smart manufacturing systems. Recent research indicates that there has been a notable increase in cyber events inside the industrial sector, with manufacturing businesses reporting a notable rise in both the quantity and intensity of assaults [3]. These assaults, which pose serious hazards to worker safety, intellectual property protection, and production continuity, vary from ransomware infections and data breaches to operational interruptions and sabotage efforts. The necessity of proactive attack detection and categorization on IoT devices is becoming more and more obvious in the face of these growing cyberthreats. Because smart industrial settings can have complicated network topologies and dynamic attack vectors, conventional cybersecurity solutions frequently fall short in these situations. Moreover, the sheer amount and speed at which IoT devices create data [4] need sophisticated analysis methods that can distinguish unusual activity that may be a sign of a security breach. Therefore, there is an urgent need for novel approaches that make use of cutting-edge technology to strengthen industrial systems' defenses against cyberattacks.

In order to comprehend the scope of the cybersecurity problem that smart manufacturing systems face, it is essential to look at current industry data and patterns. In 2022 and 2023, there was a 20% surge in cyber events [5] for industrial firms, mostly targeting the manufacturing sector, as per research published by a prominent cybersecurity firm. During the same period, there was a notable 60% increase in ransomware assaults directed on manufacturing plants, highlighting the susceptibility of industrial networks to disruptive malware. In addition, the average cost of a cyber event for a manufacturing organization increased to more than $6 million in 2023, including costs associated with reputational harm, operational losses, and remediation [6]. These

concerning figures highlight the critical need for creative cybersecurity solutions made to address the particular difficulties presented by IoT-centric smart manufacturing.

Literature Survey

Shahin et al. proposed that cyberattacks have increased as a result of the industrial sector's growing usage of IoT devices. Deep learning algorithm-based intrusion detection systems have been created to safeguard these gadgets [7]. Utilizing long short-term memory (LSTM) architecture on three industrial IoT datasets (BoT-IoT, UNSW-NB15, and TON-IoT), this research proposes two classifications and detection approaches. Analysis of these models' performance demonstrates that state-of-the-art performance in cybersecurity threat detection may be attained by augmenting the LSTM with convolutional neural networks (CNN) and fully convolutional neural networks (FCN). Shahin et al. with the expanding use of IoT devices in the industrial sector has led to a rise in cyberattacks [8]. To protect these devices, intrusion detection systems based on deep learning algorithms have been developed. This study suggests two classification and detection strategies using long short-term memory (LSTM) architecture on three industrial IoT datasets (BoT-IoT, UNSW-NB15, and TONIoT). The performance analysis of these models shows that convolutional neural networks (CNN) and fully convolutional neural networks (FCN) may be added to the LSTM to achieve state-of-the-art performance in cybersecurity threat identification.

Tuptuk et al. an important development in the manufacturing sector is the emergence of smart manufacturing systems, which are designed to react instantly to supply chain dynamics and consumer requests. This falls under Industry 4.0, which combines digital surroundings with cloud computing, machine learning, data analytics, and Internet of Things technologies [9]. The combination of these technologies does, however, bring with it both new difficulties and advantages. Complex smart manufacturing technology integration opens the door to assaults from industrial espionage and sabotage, which might have disastrous national-wide impacts as well as economic losses, production losses, injuries, and fatalities. The security of current industrial and manufacturing systems, susceptibilities to prospective cyberattacks in the future, knowledge of these threats, and readiness for such problems are all covered in this article. Hichem et al. suggests a tiered architecture for Industrial Internet-ofThings (IIoT) applications in smart manufacturing that combines Blockchain technology (BCT) with Machine Learning (ML). Sensing, network/protocol, transport, application, and advanced services are the five layers that make up the architecture [10]. While ML uses classifiers to identify attacks such as DoS, DDoS, injection, MitM, brute force, cross-site scripting, and scanning assaults, BCT collects sensor access control data. Four metrics are used to evaluate the architecture: accuracy, precision, sensitivity, and Matthews Correlation Coefficient. It is compared to other comparable architectures in the literature. Within an enhanced framework for sensor access management in IoT networks, the suggested architecture considerably minimizes the amount of threats and assaults, including injection, DDoS, brute force, and XSS. To confirm the efficacy of the architecture, more research is required

Methodology

For this investigation, network traffic data had to be gathered from a variety of sources. A number of features are included in the dataset: the destination port, the flow duration, the total number of forward and backward packets, the length of forward and backward packets, the maximum and minimum lengths of forward packets, the mean and standard deviation of forward packet lengths, and more. The data underwent a number of preparation stages to confirm its acceptability before it could be used for modeling. Numerical characteristics were adjusted to make them comparable, and missing values were handled. Furthermore, the 'Label' column and other category information were encoded. For the 'Label' column, label encoding was used to transform category data into numbers. To make model training easier, a distinct integer value was given to each category. Decision trees and Logistic Regression were the two machine learning models used in this investigation. A wellliked supervised learning technique for regression and classification applications is the decision tree. Based on the values of the input characteristics, the choice tree algorithm divides the data into subgroups. It asks the data a series of queries in order to make judgments. On the other hand, a statistical technique called logistic regression is employed to examine datasets where an outcome is determined by a number of distinct variables. It works especially well for jobs involving binary categorization. The likelihood that a specific observation falls into a specific category is predicted by the model. The result of interest in this study was the classification of network traffic as "Normal" or "Anomaly."

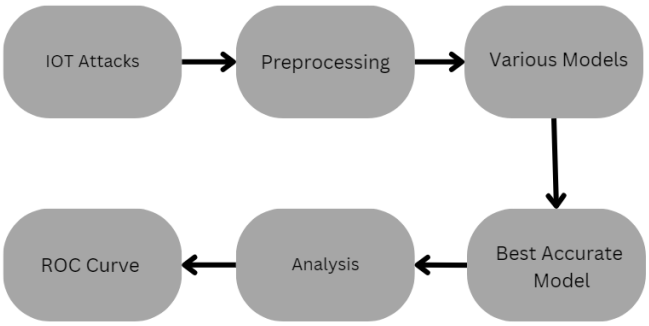

Fig. 1 - Working Methodology

During the data preparation phase, categorical feature encoding, numerical feature scaling, and missing value handling were all done. Label encoding was used for categorical features such as 'Label'. By giving each category a distinct integer value, this procedure transforms categorical information into numerical data. For example, the labels 'Normal' and 'Anomaly' may have corresponding encodes of 0 and 1. The data was prepared for model training after preprocessing. The two artificial intelligence models used were Logistic Regression and Decision Trees. Using the preprocessed data, these models were trained to identify whether the network activity was "normal" or "anomaly." After each model was trained, its performance was assessed using suitable assessment measures, including F1-score, accuracy, precision, and recall. These measurements provide information on how successfully each model identified network traffic. In order to determine how well each model classified network traffic as "Normal" or "Anomaly," the model results were examined. Conclusions about each model's appropriateness for the task of classifying network traffic were made in light of this investigation. The objectives of this technique were to gather and preprocess network traffic data, train predictive models, assess the models' performance, and make inferences from the outcomes. The final objective was to efficiently use machine learning models and the gathered data to categorize network traffic as "Anomaly" or "Normal."

4. Models

4.1 Decision Tree

A strong and well-liked supervised learning technique for applications involving regression and classification is decision trees. Based on the values of the input characteristics, the choice tree algorithm divides the data into subgroups. It asks the data a series of questions in order to make judgments. A "test" on an attribute is represented by every internal node in the tree, the result of the test is represented by each branch, and the class label is represented by each leaf node. Decision trees' interpretability is one of its main benefits. Users may comprehend the model's decision-making process by visualizing decision trees with ease. Because of their openness, decision trees are especially helpful in areas like this one, where it's critical to understand why a specific categorization.

Decision trees require very minimal data preparation and are capable of handling data that is numerical as well as categorical. Decision trees, however, are vulnerable to overfitting, which implies that noise in the data might be captured by them instead of the underlying correlations. Overfitting can be minimized by employing strategies like pruning, establishing a limit depth for the tree, or requiring a minimum quantity of data at a leaf node. 'Normal' or 'Anomaly' traffic was classified using decision trees trained on processed network traffic information in this study. Following training, the decision tree model's performance was assessed using suitable assessment measures, including F1-score, accuracy, precision, and recall. These metrics gave information about the decision tree model's classification performance for network traffic.

4.2 Logistic Regression

A statistical technique called logistic regression is applied to datasets where an outcome is determined by a number of distinct variables. For jobs involving binary classification, where the desired output might have two alternative values, it is especially well-suited. The model in logistic regression forecasts the likelihood that an observation falls into a certain category. Logical regression predicts the likelihood that an observation will fall into one of two categories, as opposed to linear regression, which makes predictions about a continuous result. Predicted values are mapped to probabilities using the logistic function, commonly referred to as the sigmoid function. Because logistic regression guarantees that the probabilities that are expected fall between 0 and 1, it is suitable for jobs involving binary categorization.

The ease of use and interpretability of logistic regression is one of its main benefits. The model's coefficients show how strongly and in which direction each of the independent variables and the outcome's log-odds are correlated. This facilitates comprehension of how each variable affects the expected likelihood of the result. Unlike more intricate models like decision trees, logistic regression is less prone to overfitting and can handle both numeric and categorical independent variables. Nonetheless, logistic regression makes the assumption that there is a linear relationship between the variables that are not dependent and the outcome's log-odds. If this supposition is incorrect, the model could not function correctly. Processed data on network traffic was used to train logistic regression to determine if the traffic was "normal" or "anomalous." Following training, the F1-score, accuracy, precision, and recall were among the suitable assessment measures used to assess the logistic regression model's performance. The effectiveness of the logistic regression model's classification of network traffic was revealed by these measures.

5. Results

Following the collection and preparation of the network traffic data, the data was visualized, and the effectiveness of the machine learning models—Decision Trees and logistic regression in particular was assessed. Examining the dataset visually was the first step. To understand the distribution and properties of the features, a variety of data visualization approaches were used. This involved utilizing correlation matrices, scatter plots, and histograms to comprehend the connections between various variables and how they affect the categorization of network traffic.

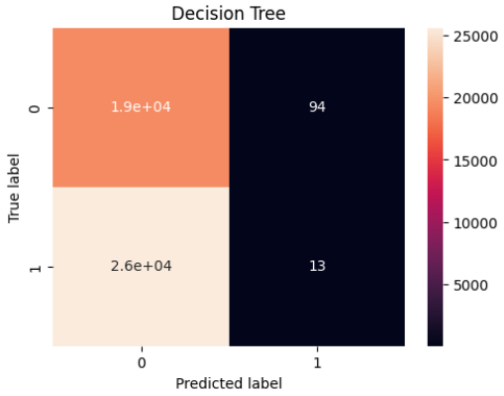

Fig. 2 - Confusion Matrix of Decision Tree

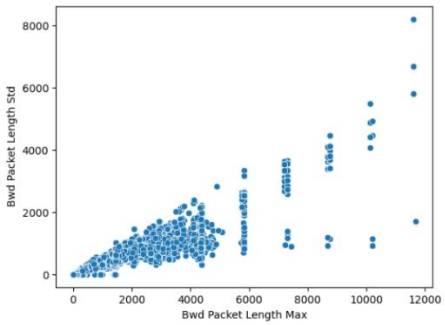

Fig. 3 - Max & Std Packet Lengths

The machine learning models' performance was assessed once the data was visualized. The Decision Tree model's astounding 99.96% accuracy was attained. The Decision Tree model was very successful in categorizing network traffic as "Normal" or "Anomaly," as evidenced by its high accuracy. But accuracy by itself could not give a whole picture of how well the model performs. By comparison, the accuracy of the Logistic Regression model was 43%. Even though it performed far worse than the Decision Tree model, it was nevertheless rather helpful in classifying network traffic. The intricacy of the dataset and the logistic regression model's linearity assumption may be the causes of the model's reduced accuracy. Confusion matrices were created in order to examine both models' performance in more detail. Confusion matrices offer a thorough analysis of how the model's predictions and the real labels differ. The confusion matrix for the Decision Tree model demonstrated a high proportion of true positives and true negatives, suggesting that the model performed well in distinguishing between "normal" and "anomaly" network data. The confusion matrix for the Logistic Regression model, however, showed a greater quantity of false positives and false negatives.

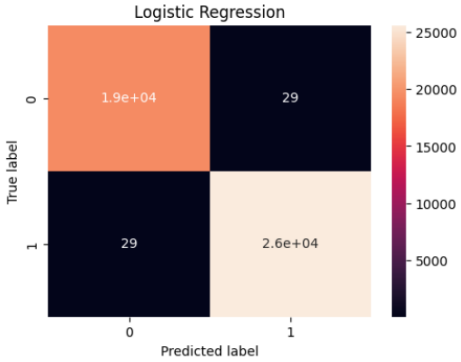

Fig. 4 - Confusion Matrix of Logistic Regression

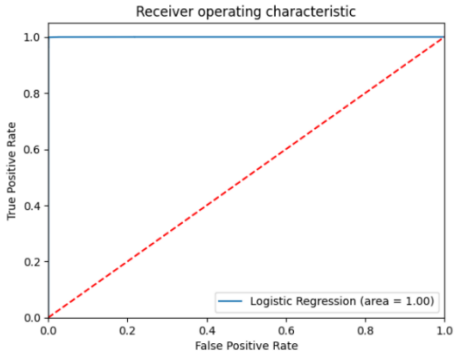

Fig. 5 - ROC AUC Curve

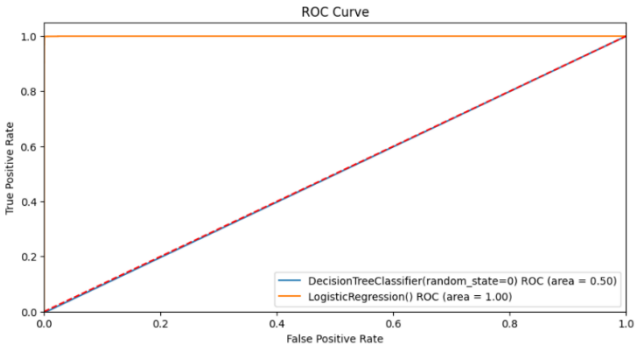

Fig. 6 - ROC AUC Curve for Both Models

For both models, ROC (Receiver Operating Characteristic) curves were additionally displayed. When the discriminating threshold of a binary classification system is changed, ROC curves show how diagnostic the system is. The Decision Tree model's ROC curve had a sharp climb, signifying good sensitivity and specificity. However, in contrast to the Decision Tree model, the Logistic Regression model's ROC curve had reduced sensitivity and specificity, indicating a lower level of overall performance. When categorizing network traffic as "Normal" or "Anomaly," the Decision Tree model fared better than the Logistic Regression approach. The Logistic Regression model failed to identify network traffic, obtaining an accuracy of just 43%, whereas the Decision Tree model displayed strong sensitivity and specificity, attaining a high accuracy of 99.96%. The confusion matrices and ROC curves, among other visualizations, provide insightful information about the effectiveness of both models and emphasized the advantages and disadvantages of each strategy.

6. Conclusion

The research showed that when it came to categorizing network traffic as "Normal" or "Anomaly," the Decision Tree model performed noticeably better than the Logistic Regression model. The Decision Tree model demonstrated strong sensitivity and specificity, successfully differentiating between the two groups, with an astounding accuracy of 99.96%. On the other hand, the Logistic Regression model had trouble correctly classifying network traffic, with an accuracy of only 43%. Confusion matrices and ROC curves, among other visualizations, highlighted the Decision Tree model's advantage in classifying network traffic and offered more insight into the performance disparity between the two models. These results highlight the significance of choosing suitable models of machine learning for given tasks, with Decision Trees proving to be an excellent option for this particular categorization issue.

References

[1] Tao, Fei, Jiangfeng Cheng, and Qinglin Qi. "IIHub: An industrial Internet-of-Things hub toward smart manufacturing based on cyberphysical system." IEEE Transactions on Industrial Informatics 14.5 (2017): 2271-2280.

[2] Bansal, Sharu, and Dilip Kumar. "IoT ecosystem: A survey on devices, gateways, operating systems, middleware and communication." International Journal of Wireless Information Networks 27.3 (2020): 340-364.

[3] Moreno, Valeria Casson, et al. "Analysis of physical and cyber security-related events in the chemical and process industry." Process Safety and Environmental Protection 116 (2018): 621-631.

[4] Verma, Shikhar, et al. "A survey on network methodologies for real-time analytics of massive IoT data and open research issues." IEEE Communications Surveys & Tutorials 19.3 (2017): 1457-1477.

[5] Shevchenko, Pavel V., et al. "The nature of losses from cyber-related events: risk categories and business sectors." Journal of Cybersecurity 9.1 (2023): tyac016.

[6] Preble, Keith A., and Bryan R. Early. "Enforcing economic sanctions by tarnishing corporate reputations." Business and Politics 26.1 (2024): 102-123.

[7] Shahin, Mohammad, et al. "A novel fully convolutional neural network approach for detection and classification of attacks on industrial IoT devices in smart manufacturing systems." The International Journal of Advanced Manufacturing Technology 123.5 (2022): 2017-2029.

[8] Shahin, Mohammad, et al. "Implementation of a novel fully convolutional network approach to detect and classify cyber-attacks on IoT devices in smart manufacturing systems." International Conference on Flexible Automation and Intelligent Manufacturing. Cham: Springer International Publishing, 2022.

[9] Tuptuk, Nilufer, and Stephen Hailes. "Security of smart manufacturing systems." Journal of manufacturing systems 47 (2018): 93-106.

[10] Mrabet, Hichem, et al. "A secured industrial Internet-of-things architecture based on blockchain technology and machine learning for sensor access control systems in smart manufacturing." Applied sciences 12.9 (2022): 4641.