Smartsentry Cyber Threat Intelligence in IOT

Dr. G. Sanjay Gandhi[1], Professor, Department of CSE,

Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.

Vankayala Bala Harshitha [2], Shaik Nayum Akthar [3], Yaragopu Mukesh [4], Shaik Abdul Rehaman [5]

UG Students, Department of CSE,

Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.

[1]sanjaygandhi.g@gmail.com, [2]balaharshitha03@gmail.com, [3]nayunayum18@gmail.com,

[4]yaragopumukesh@gmail.com, [5]skabdul123412@gmail.com

Abstract: The SmartSentry system has established itself as the Cyber Threat Intelligence (CTI) system in the Industrial Internet of Things (IIoT) area owing to its possibility to identify threats right down to the source of the data. To do this, additional detectors are applied along with more sophisticated machine learning and deep learning algorithms that carry out comprehensive and detailed analyses of the cyber threats. A team of researchers has put together an entire collection of algorithms to single out the threat and to determine its severity based on the type of information, which is composed of the following: Decision Tree (DT), Extra Trees Classifier (ETC), Support Vector machine (SVM), k-Nearest Neighbor (KNN), and Deep Neural Network (DNN). As a part of our IIoT anomaly detection study, we handled the issue of data imbalance with the integration of the Synthetic Minority Over-sampling Technique (SMOTE) into our method. The Decision Tree and Random Forest models got accuracy scores of 0.9979, while the SVM model was at 0.7644, the Extra Trees model at 0.9961, and KNN at 0.9196. The SmartSentry's capability to find and stop anomalies, which is the latter of the two cyber threat tactics - the proactive one thereby ensuring IIoT systems stability and safety, will be a great leap forward in the war to protect IIoT.

Keywords:
IIoT, Cyber Threat Intelligence, Machine Learning, Deep Learning, Anomaly Detection, Random Forest, Decision Tree, Support Vector Machine, k-Nearest Neighbor

I. INTRODUCTION

The industrial Internet of Things (IIoT) has undoubtedly become the main player in the industries as it was the main driver behind the rapid interconnection and expansion of various devices and systems. But still, introducing more devices will definitely mean increasing the chance of critical infrastructure being compromised. The IIoT system threshold to handle and lighten the really big data together with the constant advent of cyber threats are what making the traditional security measures useless. To tackle the issues raised above, we have devised SmartSentry, which is an awesome Cyber Threat Intelligence (CTI) solution that is able to fulfill the particular security needs of the IIoT ecosystem. The introduction of such a solution can be demonstrated through the SmartSentry case which is not merely a model depending on machine learning (ML) and deep

learning (DL) being used to improve the standard of the threat detection capabilities. The system, additionally, is applying Synthetic Minority over-Sampling Technique (SMOTE) to deal with the issue of data imbalance and to make the discovery of low-frequency and high-prevalence intrusions simpler.

SmartSentry will be the one that not only indicates the instant threat but also proposes the solution to it that could be harmful to the IIoT systems and the processes. The firm's preemptive security actions along with SmartSentry's help make the firm safe and give it a favorable position amidst the constant global ties. The infrastructure of the system will be set up on a pedestal which will not only permit the discovery of a new cyber threat but also facilitate the rapid reaction thus guaranteeing that the infrastructure will not be shut down in the IIoT areas.

A. Objective of The Study:

The proposal focuses on the development and integration of a smart yet complex CTI that is capable of securing the IIoT ground and it gets the name SmartSentry. Project implementation is expected to apply various techniques of machine learning and deep learning for the trust and security of critical infrastructures which means that the risks will be prevented and reduced during the whole process as well. The performance of the detection and the identification of security violations will be evaluated by SmartSentry with a particular focus on the issue of imbalanced data which will be solved using the SMOTE technique. The second goal is to introduce security services that will not only be a cyber smart way to attack the IIoTs thus securing them from the soaring number of cyber attacks but also be such that it is a smart way to attack the IIoTs.

B. Problem Statement:

The trading list of the connecting devices used IIoT is also on the rise, thus making all the big buildings targets also of cyber attacks. This form of IIoT architecture is dynamic, multi-layered that cannot be applied to the more traditional paradigms of security and cannot be responsive to the real-time demands of IIoT design. The current responses are to a degree frozen and non-follow-up and response to evolving cyber threats. It indicates a requirement in a new developed CTI topology IIoT definite. This is the solution to which SmartSentry, which is the machine and machine learning interface, is claimed to provide and which would be capable of sensing and accordingly responding to the cyber attacks in the real time even without disrupting the process of actual functioning of the IIoT systems.

II. RELATED WORKS

It is this added complexity and the added number of gadgets being added to the internet of things (IoT) that has been seeing the implementation of machine learning (ML) in the solution and the deep learning itself in the identification of anomalies in the systems. Other sources stated that machine learning algorithms (decision tree, k-Nearest Neighbors (KNN), Support Vector machine (SVM)) had been utilized in domain of detection of cyber threats in an IoT contention. They are normally employed together with Synthetic Minority Over-sampling Technique (SMOTE) in correcting the bias information as well as simplifying the process of recognizing part of the rare dangers other than being identified [1].

A clearer review of the existing intrusion detection systems (IDSA) in respect of the non-ecological networks on an analytic analysis, reveals that one of the salient issues is the capability to process a high volume of information and sheer throughput of the non-ecological networks. The article also reported that the necessity of more scalable and more efficient detecting processes may mean that blockchain and federated learning may be employed to enhance the system security and scalability which will still persist as one of the largest discussion on the topic or at least the subject area on the security of IoT in the future [2].

cyber threats is a dynamic setting and machine learning and deep learning are the prospects that are not explicitly suggested in the paper but taken into account by other sources when the conditions of the development of the IoT are relied upon in the context. The sides interested in the statement of an example, sign-onedly, incorporate the authors who identify deficiencies of the traditional systems that do not allow the flow of information and new transfer of attacks flowing. Within these lines, the exploration with the random Forest model and SVM experiments can also possibly manage and control the risks lingering in the IoT systems in a more deft manner[3]. Also, models, as a process involving deep learning (Deep Neural Networks or DNNs), have also been demonstrated to be capable of learning more complex structures of the IoT traffic and other types of networks, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) and Long Short-Term Memory networks (LSTMs) are also more

effective in the detection of anomalies than the traditional ones[4].

Alternatively, one may find other publications on optimization of IDS, which would also be applied to the IoT systems once the deep learning-based approaches are applied. Another aspect mentioned by the authors is that the feature deletion method may serve to track the changes in the network traffic pattern that would guarantee reoccurrence of already familiar cyber attacks such as the Distributed Denial of Service (DDoS) and the Man-in-the-Middle (MITM). In particular, CNNs can be described as relatively useful in ascertaining an attacking mode as a multinomial within the IoT networks [5]. Otherwise, they have simulated the spread of attack in IIoT using the support of Graph Neural Networks (GNNs). The models are capable of tracking the existence of the threat spread in the system by virtue of the fact that they have the geometrical nature of the IoT networks that can enable the actualisation of the process of anomalies detection [6].

The other categories of the machine learning that will be incorporated in the project of securing the maintenance of the IoT to be used in the process of performing comparison and contrast of the decision tree involve the decision tree, the SVMs and the neural networks. It has also recognized the application of ensemble approaches that can be borrowed by Facebook in a bid to optimize the intrusion detect system by introducing a number of algorithms. It has been established that ensemble techniques represent a more straightforward approach to decision making on behalf of the different models in superficial resolution and a low false-positive rate when working in depth resolution [7]. Specifically, they have learned that among the issues, such as an enormous and unforeseeable IoT traffic, they can be solved by using such an algorithm as Random Forest, SVM and KNN [8].

The other frameworks to which they refer to the IoT IDS were the Generative Adversarial Networks (GANs), and Generative Adversarial Networks (XAI). The XAI will help transform the process of the decision-making to the one much more transparent, and the generation of the very scenario of the attack will be made possible with the help of GANs. The connotation with regards to the security teams is staggering and a person can even become familiar with the motives, which are liable to criticism, framing of threats. Responses to the question of the acceptability of the identified threat to the eyes of the security experts is not so secretive in regards to the decisions made by the system capable of linking and utilizing such tools as LIME and SHAP[9].

Despite the above, the other literature materials state that the feature selection instruments, i.e., the Random Forest and Decision Trees, should be utilized to enhance the functionality of the anomaly detectors of the giant IoT networks. The tools are also effective in ensuring that any data that are irrelevant are filtered out and only the relevant data are included in the process of laying down the threats that may be utilized in the generation of the best of the entire system[10]. The discussion of the Id of Artificial Intelligence AI in the Internet of Things systems which comprise the deep-learning, reinforcement-learning and natural language processing (NLP) that is being cultivated to address the dynamic and complex threats have also been addressed in the study. It can be stated that one of the successful strategies to adjust the existing models to the pursued forms of cyber threats was the transfer learning that could be scaled more and offered to the IDS systems[11].

The BiGRU network and Long Short-Term Memory model are attacked with the aid of the hybrid networks and proved instrumental in identifying the threats of the dynamical IoT traffic. The examples of the models presented in these scholarly sources give sufficient details on how one can identify any type of cyber-attack by the specifics of certain patterns in the traffic and consequently simplify the identification infrastructure thus rendering it more logical [12]. Also, the models CST-AFNet that is based on CNN and CoBiGRU are introduced to explore the time-space of the IoT traffic and create the more fitting definition of minor anomaly in the network [13].

It will use the system integration in the IOT systems, which will identify the cyber attacks, ensemble method, by inscribing the effectiveness of Rand Forest SVM differentiated and XGBoost techniques. In the minimisation of the false-negatives too can be found to be sufficiently true when a large number of classifiers are pooled and reintroduction of high accuracy attack-detection, high-detection-accuracy and low probability of false negative, which can be extremely malevolent to the IoT well being [14]. ast but not the least, machine learning and blockchain have been suggested to adopt the blockchain as a means of improving the security and integrity of the IoT systems. This will provide authentication, access control and decentralization of the blockchain trying to advance the level of trust and integrity in relaying the information in the network of looting IoT[15].

In much as the machine learning and deep learning solutions offered to the security of the iot have registered improvements as of now, it has its fair share of issues. The highest point of disjunction that is presented is that the models provided are not applicable to the range of different types of attack under the conditions of the IIoT, especially the ones that operate with highly dynamical and unstable streams of information. Other existing structures are also capable of being unhelpful to the area of an imbalanced data, and thus lead to the decreasing imposition of rate on the detection of infrequent but extreme dangers. Among the SmartSentry systems which can be offered in the current paper will seal the proposed gaps and integrate the type of machine learning and deep learning that include: Random Forest, SVM and DNN, SMOTE, and will enhance the IIoT network efficiency with regards to detections. The system will also provide a viable alternative in the detection and response to cyber threat in the stability and security of majority of the central IIoT structures..

III. METHODOLOGY

The proposed system of the same is known as SmartSentry, which is approximated to be a viable solution of IIoT systems within the landscape of the Cyber Threat Intelligence (CTI). The system will be prone to system working process that will start with the convergence of data on the entire IIoT equipments and aggregation data on the entire sensors and other sections of the infrastructure. This is succeeded by information processing that involves purifying the data anomalies including missing information, outliers and normalization as the most appropriate to be processed and inputting the machine learning procedures with processed information and simplified information to compute.

A. Dataset

SmartSentry is configured as a network traffic metadata and traffic of the different protocols including the ARP, ICMP, HTTP, TCP/UDP, the DNS and the MQTT. Based on these descriptive data, it can be concluded that frame-level data, IP address (source and destination) and checksum using ICMP, message content sent using the HTTP, TCP connection flags, query type in DNS query, and MQTT data are the most important attributes. These features are used in the making of a distinction and categorization of different types of attacks within the cyber-plane, which entail DDoS, port scanning, ransomware, and vulnerability scanning. Supposedly determined variable, which presupposes the type of attacks, is Attack Type An, which involves: DDoS, UDP, Port Scanning and Ransomware. The system also undergoes training with the information to cause the system to be mindful of the cyber threats presented by the IoT because it is based on the identification of the aberrant patterns in the network traffic.

B. Preprocessing Steps

- **Encoding of Categorical Variables**

Scaling of the categorical data is accomplished with scikit-learn library LabelEncoder which is of type numeric. The information may be converted to the machine learning models. Such an explanation is known as numerical label placed on each of the set of data that are categorical.

- **Outlier Detection and Removal**

The Interquartile range (IQR) is the method that is used to find the outliers. The 25 th percentile as well as the 75 th percentile of the data are computed in an effort to establish the IQR and the rest of the data that falls outside the limit that has been calculated and is outliers are got rid of. This is so that the impacts of extreme values that can be polluted are reduced to contaminating the training of the model.

Formula:

$$\text{IQR} = Q_3 - Q_1 \qquad (1)$$

where $\boldsymbol{Q_1}$and $\boldsymbol{Q_3}$are the 25th and 75th percentiles, respectively, and **IQR** is the difference between $\boldsymbol{Q_3}$and $\boldsymbol{Q_1}$.

- **SMOTE (Synthetic Minority Over-sampling Technique)**

The skewness of the dataset is being eliminated with the assistance of SMOTE that generates fake samples of the minorities. The approach will be central towards ensuring that the machine learning models will not overlook a single type of attack, but which are quite numerous and pertinent, and that the detection system will be easier to balance and more precise.

- **Feature Selection**

It utilizes the mutual information to assist in the selection of the features to be experienced so that

it is guaranteed that the SelectKBest algorithm that selects the most significant features to predict the target variable, Attack Type are implemented. Attributes that have the highest potential of co-occurring with the highest level of gold are retained and the model simplifies to be trained besides approximating the truth.

C. Model Training

DT:

In the assigned case, we can talk about Decision Tree as a supervised learning algorithm, however, this time the algorithm will be implemented taking into account the classification of the data by the number of the limited numbers, with the assistance of the features to reach the decision-tree. The optimal anthropomorphic choice of the separation is done at every node of the separation of the impurity or the loss of information that is usually the Gini Index or the Entropy. The new data can be distributed to the leaf node in the tree that displays the splits of the tree and the category that is forecasted is represented there.

Formula:
Gini Index:

$$G = 1 - \sum_{i=1}^{C} p_i^2 \qquad (2)$$

RF:

Random Forest It refers to an ensemble algorithm, which was more or less first introduced as bootstrapping, randomization choice tree (also called bootstrapping, and randomizing samples) and randomization features and averages (logistics), and randomization features and averages (classification). It is founded on the opportunity sampling of the information utilized in the creation of the trees on the forest and ultimately overlaid onto each and every tree.

Formula:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} h_t\,(x) \qquad (3)$$

where $h_t(x)$ is the prediction of the t-th tree.

Extra Tree Classifier (ETC):

Extra Tree Classifier- It can be stated that it is largely comparable to decision trees as it is trying to model the idea of random Forest since it will not assign it based on optimization but only haphazardly chooses a split thresholds with low variance and high generalization that it utilizes. The criteria of the split are likewise too arbitrary, also, and this is why they are likewise at random prior to training the features hence, does not drain much time on the training aspect.

Formula:

$$\hat{y} = \text{majority vote}\{h_1(x), h_2(x), \ldots, h_T(x)\} \qquad (4)$$

SVM:

A support vector machine could also be considered a training model because that categorizes the data according to a hypotheticalized linear separation line upon two groups of objects by creating maximum dichotomy between the support vectors and vectors that could be assisted by the nonlinear division of the categorization through the assistance of the kernel functions. SVM is founded on the bigger mapping of the elements of the data points map onto which the explanatory line may have been guided. This is what helps SVM to be particularly strong in the situation where the data on the original feature space is hard to be differentiated in the original feature space.

Formula:

$$\min_{w,b} \frac{1}{2} \| w \|^2 \quad \text{subject to } y_i(w \cdot x_i + b) \geq 1 \qquad (5)$$

KNN:

The k-Nearest Neighbor is a hypothesis that is not parameterized, which assumes that there is a central point of the data set, which theorizes most of the nearest k points where a summation distance represented by a metric. This algorithm calculates the disparity between the input point and other information, which are included in the training set, and are mature and categorize the most recurrent one among the first k nearest neighbours.

Formula (Euclidean Distance):

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \qquad (6)$$

DNN:

Deep Neural Network has far more layers undercover and, in reality, it is provoked by rather complex statements of the information by weighted terms of connection and nonlinear activation activators and the fact that it is trained, through the help of backpropagation. The layers isolate the rest of the dormant functions of the information and the network

can carry out several operations such as natural language processing.

Formula:

$$a^{(l)} = f\left(W^{(l)}a^{(l-1)} + b^{(l)}\right) \quad (7)$$

IV. RESULTS AND DISCUSSION

Decision Tree Classifier:

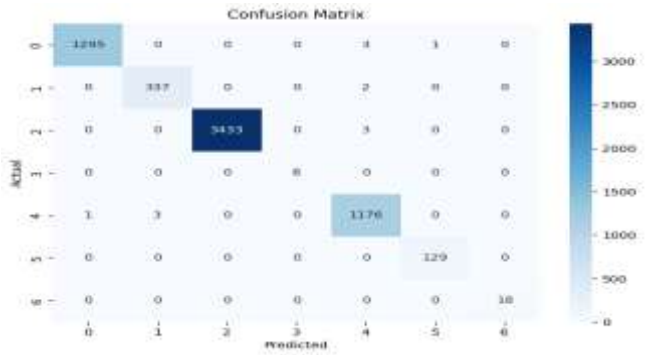

Figure 1: DT confusion matrix

Compared to it, Decision Tree (DT) model had a phenomenal accuracy of 0.9979 and it suggests that the model can be applied in the development of the appropriate categorization of network traffic information. The F1 score of 0.9973 which is associated with a high precision and high recall means that the model will be at the threshold of focusing on cyber threats at a high rate and with any reduction in the false positive results. It means that the 0.9966 accuracy and 0.9981 recall proves that most of the positive predictions made by the model are correct and it is rather likely to be wrong in the prediction it makes as is demonstrated by recall. The confusion matrix also indicates to show that the model is doing well in terms of misclassifying some of the category that has low levels of confusion in specific the DDoS and the ransomware. Its success, according to the results, is rather high, which allows concluding that Decision Tree can be used in order to classify and detect cyber threats in IIoT systems.

Random Forest Classifier:

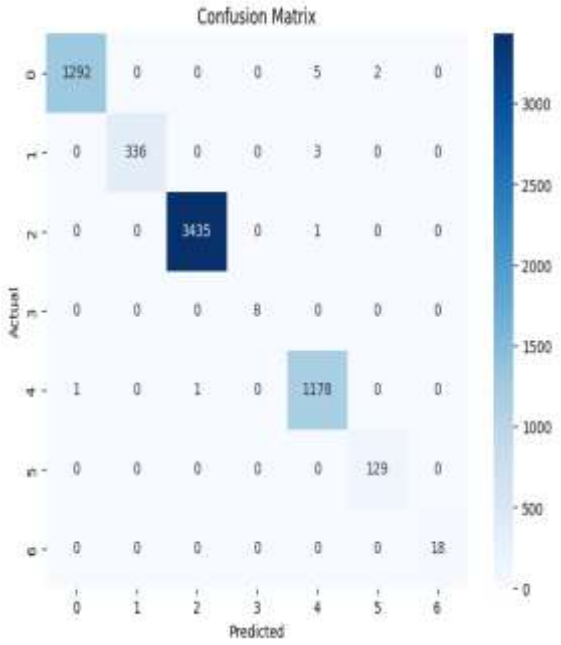

Figure 2: RF Confusion Matrix

RF model was not bad either than the Decision Tree since it had a value of 0.9979. The model had F1 score of 0.9971 that was a strong indication that there was no low level of accuracy and recall. The model accuracy of the random forest model is 0.9966 and model recall is 0.9977 that implies that the random forest model possesses high boot of identifying cyber-attacks of diverse types. As per the confusion matrix, the model has been able to handle most of the types of attacks with some inalienable errors being experienced in the case of the DDoS and the process of port scanning. A viable and efficient choice of the Random Forest model would be the network-based threats detection of the IIoT systems.

SVC:

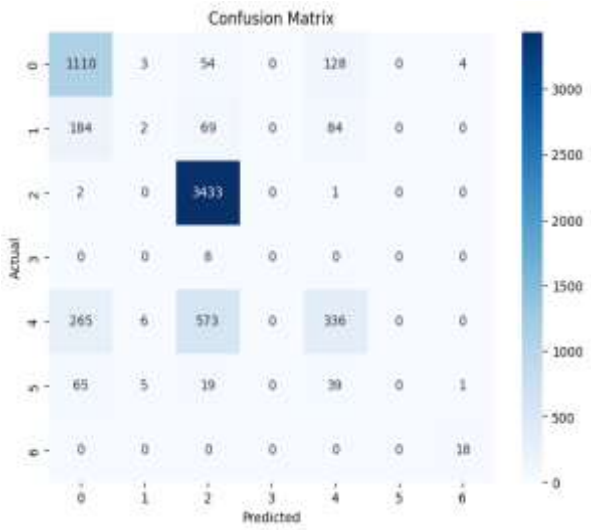

Figure 3: SVC Confusion Matrix

The estimating value of Support Vector Classifier (SVC) model is also quite low, 0.7644, and it is the sign that the data of the network traffic cannot be easily denoted. The inability of the model to discriminate the type of the attacks is supported by F1 score of 0.419, and by the precise and the recall scores of 0.427 and 0.449 respectively. The misclassifications are intensive as indicated in the confusion matrix within that category like the DDoS, vulnerability scanning and ransomware. Based on the results, it is not possible to apply the SVC model to such data, without subsequent modification or parameter performance, and would probably need additional changes.

ExtraTreesClassifier:

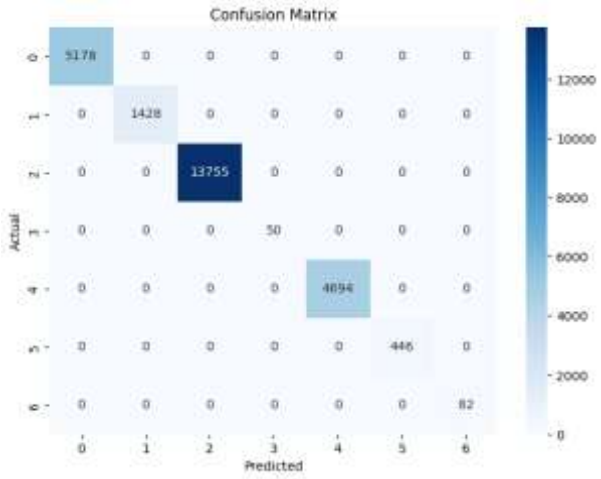

Figure 4: ExtraTreesClassifier Classification Matrix

Extra Trees Classifier (ETC) model possesses a good rate of accuracy along with trial (.9961) and the recall (.9950) with an accuracy and F1 rate of 0.9954 and 0.9961 respectively. The confusion matrix revealed such a succession that indicates that this model was satisfactory in classifying cyber-attacks and its misclassification rate is very minimal. The reason to this is that the Extra Trees model can be an alternative significant to network security in IIoT systems in that it is able to manage a great number of attacks with the minimal amount of error. The other strength is that it is recallable and high precision because it is also appropriate in terms of probability of detecting aberrations in cyber threat classification.

KNN :

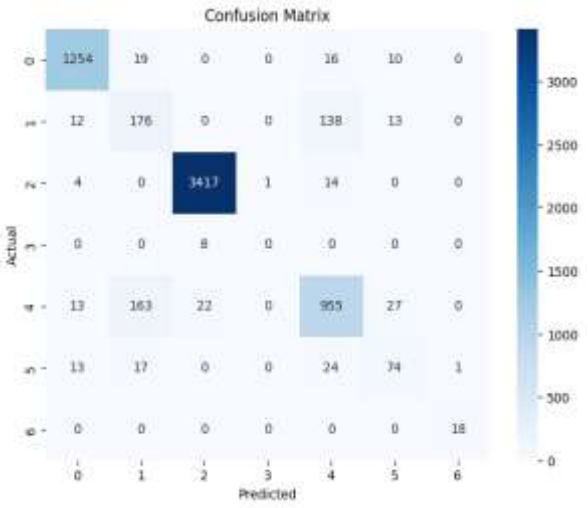

Figure 5: KNN Confusion matrix

The accuracy of k nearest neighbor model (KNN) is 0.9196 which is less than the accuracy of the tree based models but higher than the accuracy of the SVC model. The value of F1 (0.6902) is average performance where precision (0.6864) and the values of recall (0.6946) are considered as this means that this model is doing decently on the measure of precision and average on the measure of recall, but not of Decision Tree nor of the Random Forest models. In the confusion matrix, few of the attacks were missed by KNN particularly DDoS and ransom. This, however, is not the best in this regard, yet KNN, in this case, has some value to the threat detection, though in that case, the trees-based models beat it.

DNN:

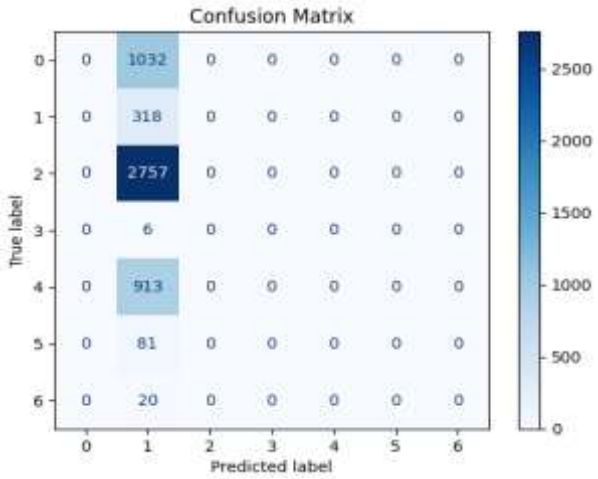

Figure 6: DNN Confusion Matrix

The Deep Neural Network (DNN) was not an effective model because the accuracy resulted to be 0.0620. This shows that DNN model was not doing a good job in the network traffic data category and it might be due to its

overfitting, insufficient data on training or even inappropriate structure. The fact that such rates of errors are rather high and unreliable signifies that tuning, training or architecture of this activity should be made one of the key areas of the DNN model improvements.

Table 1: classification model results

Model	Accuracy	Precision	Recall	F1 Score
Decision Tree	0.9979	0.9966	0.9981	0.9973
Random Forest	0.9979	0.9966	0.9977	0.9971
SVC	0.7644	0.4268	0.4492	0.419
Extra Trees	0.9961	0.9959	0.9950	0.9954
KNeighbors	0.9196	0.6864	0.6946	0.6902

Discussion:

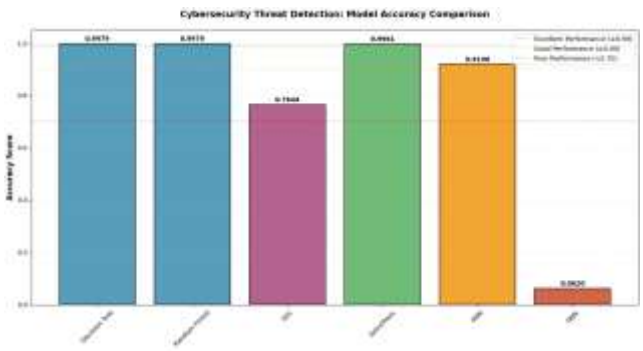

Figure 7: Results Comparison Graph

The decision tree and the random forest model were superior among all the other models and the accuracy of 0.9979 with high figure of precision and recall. The misclassifications are of a small scale as well, and both models could reasonably fit in the classification of such a cyber-attack to DDoS and ransomware. They are very dependable models that may offer security of IIoT network in addition to being effective in monitoring and classifying abnormal network traffics.

Extra Trees is the other good model whose accuracy of 0.9961 and F1 of high were high at 0.9954. The precision of model is not as high as of decision tree and the random Forest model, but even in the threat detection case the model competes with the accuracy and recall.

On the other hand, SVC was very imprecise with 0.7644. It is not that high and thus the accuracy and the recall is large and hence, we cannot apply SVC on such data without some form of a trick or parameter adjustment. The model that predicted by the value of 0.9196 of accuracy of the KNN model had no significant difference in the value but was unlucky compared to another model that happened to be the tree based models at least in terms of accuracy and the recall.

Last but not the least was the Deep Neural Network (DNN) model which performed the poorest by the accuracy of 0.0620 because this model could not be used in the given job because of the way it was implemented in the existing system. This could have been the overfitting or underfitting of the data and further optimizations of the model would be required to achieve a more enhanced performance of the model.

V. CONCLUSION

In conclusion, the SmartSentry can be described as very effective as a Cyber Threat Intelligence (CTI) framework of the security of the Industrial Internet of Things (IIoT). The system also has a number of machine learning /deep learning models (Extra Trees Classifier, Support Vector machine, Lesson 7, 2020): Decision Tree, Extra Trees Classifier, Lesson 7, 2020, August 22, 2020, Review, Lesson 7, 2020, Lesson 7, 2020). Such discontinuity of a problem as the imbalance of the classes handling IIoT information may be guarded through the help of a method like cross-Synthetic Minority Over-Sampling Technique (SMOTE) which would enable the system to track the cyber threat but improperly in velocity and with an extremely high probability. The opportunity to retrieve the anomaly detection of the system in the active form which will support the discovery of the potential security breach in the early stages of the given process which is the most significant input into the security and integrity of the IIoT infrastructures.

The ancient phenomena called cyber-attacks are dynamic, hence, making SmartSentry a significant part of the verification of the presence of the plan, and the industrial environment is feasible. In order to achieve the resistance of the Internet technologies to the recent generation of Internet attacks and implement the demanded infrastructure without the damage or disadvantages, it is possible to utilize SmartSentry it is a grand and versatile program of CTI.

VI. FUTURE ENHANCEMENT

Being a potential prospect, there are a number of improvements which can be applied to enhance the functionality of SmartSentry. First, the more adaptive and dynamic responses to the new threats can be implemented as dynamic reinforcement learning since the system would be at a position to reacts to the new trends of new attacks in the real- time. The specified data analysis may also be simplified with references to the Natural Language Processing (NLP) that interprets the unstructured information posted by the social media or the Dark Web forums that might host the gap in the Threat Intelligence that fulfills the role of determining the threat development.

It may also be implemented with the use of the Federated Learning that will provide an additional layer of safeguarding the information privacy, but, still, will enable the system to enjoy the fruits of the intelligence sharing in various IIoT settings. This would enhance scalability and privacy of SmartSentry that would be more acceptable to use it in other industries. The remaining one which can be optimized is user interface (UI) and notification system which can be optimized to show more relevant visualization and real-time notification.

Finally, the extension of the framework to the cross-industry applications will allow the transport of the knowledge to the other industries, which would make the threat intelligence resistance to the global levels, and the IIoT ecosystem security-proofs.

VII. REFERENCES

[1] L. Hazzam and S. Fenanir, “Anomaly Detection for the Internet of Things Using Machine Learning Techniques,” *International Journal of Computational and Experimental Science and Engineering*, vol. 11, no. 4, Dec. 2025, doi: 10.22399/IJCESEN.4166.

[2] M. Berhili, O. Chaieb, and M. Benabdellah, “Intrusion Detection Systems in IoT Based on Machine Learning: A state of the art,” *Procedia Comput Sci*, vol. 251, pp. 99–107, Jan. 2024, doi: 10.1016/J.PROCS.2024.11.089.

[3] M. M. Rahman, S. Al Shakil, and M. R. Mustakim, “A survey on intrusion detection system in IoT networks,” *Cyber Security and Applications*, vol. 3, p. 100082, Dec. 2025, doi: 10.1016/J.CSA.2024.100082.

[4] A. Aldhaheri, F. Alwahedi, M. A. Ferrag, and A. Battah, “Deep learning for cyber threat detection in IoT networks: A review,” *Internet of Things and Cyber-Physical Systems*, vol. 4, pp. 110–128, Jan. 2024, doi: 10.1016/J.IOTCPS.2023.09.003.

[5] M. A. Hossain, “Deep learning-based intrusion detection for IoT networks: a scalable and efficient approach,” *EURASIP J Inf Secur*, vol. 2025, no. 1, pp. 28-, Dec. 2025, doi: 10.1186/S13635-025-00202-W/FIGURES/11.

[6] S. Yang *et al.*, “Industrial Internet of Things Intrusion Detection System Based on Graph Neural Network,” *Symmetry 2025, Vol. 17, Page 997*, vol. 17, no. 7, p. 997, Jun. 2025, doi: 10.3390/SYM17070997.

[7] A. J. Aparcana-Tasayco, X. Deng, and J. H. Park, “A systematic review of anomaly detection in IoT security: towards quantum machine learning approach,” *EPJ Quantum Technology 2025 12:1*, vol. 12, no. 1, pp. 112-, Sep. 2025, doi: 10.1140/EPJQT/S40507-025-00414-6.

[8] M. Z. Mahmud, S. Islam, S. R. Alve, and A. J. Pial, “Optimized IoT Intrusion Detection using Machine Learning Technique,” *2024 IEEE 3rd International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things, RAAICON 2024 - Proceedings*, pp. 167–172, Dec. 2024, doi: 10.1109/RAAICON64172.2024.10928532.

[9] T. Hasan and S. Tasnim, “Real-time explainable IoT security with machine learning and CTGAN-enhanced detection for resource-constrained devices,” *Ad Hoc Networks*, vol. 178, p. 103937, Nov. 2025, doi: 10.1016/J.ADHOC.2025.103937.

[10] “(PDF) Anomaly Detection in IoT Networks Using Machine Learning Techniques.” Accessed: Dec. 20, 2025. [Online]. Available: https://www.researchgate.net/publication/394808458_Anomaly_Detection_in_IoT_Networks_Using_Machine_Learning_Techniques

[11] A. Villafranca, K. M. Thant, I. Tasic, and M.-D. Cano, “AI-Enabled IoT Intrusion

Detection: Unified Conceptual Framework and Research Roadmap," *Machine Learning and Knowledge Extraction 2025, Vol. 7, Page 115*, vol. 7, no. 4, p. 115, Oct. 2025, doi: 10.3390/MAKE7040115.

[12] A. Gueriani, H. Kheddar, A. C. Mazari, and M. C. Ghanem, "A Robust Cross-Domain IDS using BiGRU-LSTM-Attention for Medical and Industrial IoT Security," Aug. 2025, Accessed: Dec. 20, 2025. [Online]. Available: https://arxiv.org/pdf/2508.12470

[13] W. Ishtiaq, A. Zannat, A. H. M. Shahariar Parvez, M. Alamgir Hossain, M. Hasan Kanchan, and M. Masud Tarek, "CST-AFNet: A dual attention-based deep learning framework for intrusion detection in IoT networks," *Array*, vol. 27, Sep. 2025, doi: 10.1016/j.array.2025.100501.

[14] S. L. Qaddoori and Q. I. Ali, "An Efficient Security Model for Industrial Internet of Things (IIoT) System Based on Machine Learning Principles," *Al-Rafidain Engineering Journal (AREJ)*, vol. 28, no. 1, pp. 329–340, Feb. 2025, Accessed: Dec. 20, 2025. [Online]. Available: https://arxiv.org/pdf/2502.06502

[15] S. H. Sababe and E. L. Ghasab, "Iterative Splitting Methods for Stochastic Dynamic SVIs," May 2025, Accessed: Dec. 20, 2025. [Online]. Available: https://arxiv.org/pdf/2505.06570