RESEARCH PAPER

Intelligent Cyber Attack Detection through Machine Learning

[1] S.Sireesha,[2] Dr. A. Ramaswami Reddy, [3] Mrs. D V V Deepthi,

[1] M.Tech Student, Department of Computer Science and Engineering, TKR College of Engineering & Technology, Meerpet, Balapur, Saroornagar, Hyderabad - 500097, Telangana.

[1]Email: sireeshasabavat@gmail.com

[2] Professor & Principal, Department of Computer Science and Engineering, TKR College of Engineering & Technology, Meerpet, Balapur, Saroornagar, Hyderabad - 500097, Telangana.

[2]Email:ramaswamyreddymail@gmail.com

[3] Assistant Professor, Department of Computer Science and Engineering, TKR College of Engineering & Technology, Meerpet, Balapur, Saroornagar, Hyderabad - 500097, Telangana.

[3]Email: vvidyadd@gmail.com

ABSTRACT:

Internet of Things (IoT) Connection encouraged new technology solutions for public safety in the city by integrating the transportation, healthcare and environment sectors into smart services. The variety of devices and continual chug of information makes these interdependent systems effective for both improving operational effectiveness and enhancing quality of life, but provides cybercriminals with a broad attack surface. Existing methods of security are simply not up to the task of addressing real-time, complex threats. In this work, we present a novel machine learning-based approach to detect a cyber attack in Internet of Things (IoT) powered smart cities. This proposed solution would be based on the analysis of network traffic and activities of the devices to determine whether the activity is normal or not. In the data preparation section, unimportant features are removed, missing values are filled, scaling and filters of features are applied on the data to improve the accuracy of the model. Various approaches are implemented and explored to find the best machine learning approach for detecting cyberattacks. Generally, accuracy, precision, recall, F1 score and confusion matrix analysis are used to evaluate the performance. Its aim is to keep the computational efficiency, increase early Cyber intrusion detection and decrease false alarms. The patterns learnt may be finding an abnormal behaviour in historic activity that can be used to detect Denial of Service (DoS), probing, access or data manipulation attacks. The results of the experiment highlight an immense improvement in the time taken for the detection of threats, from nearly real-time time to almost instant time, with the use of machine learning threat detection technology, which greatly improves the security of the IoT infrastructure. The study results underline the importance of Smart City Security Systems. The proposed architecture is flexible and extendable with the number of urban services and in the future it will offer cyber resilience capabilities to the IoT based system.

Keywords: Cybersecurity, machine learning, intrusion detection, network security, cyberthreat detection, intelligent monitoring, Internet of Things (IoT), and smart city security.

How to cite this article: 1 S.Sireesha2 Dr. A. Ramaswami Reddy 3 Mrs. D V V Deepthi. Intelligent Cyber Attack Detection through Machine Learning. Int J Drug Deliv Technol. 2026;16(78s):41-44. DOI: 10.25258/ijddt.16.78s.6.

1.INTRODUCTION:

The Internet of Things (IoT) has transformed the way we life in the cities of the world.These technologies can be applied in a wide range of applications like smart transportation, health care, energy distribution, environmental monitoring and public safety. The proliferation of connected devices has rapidly increased the amount of data generated in smart city scenarios. These IoT technologies can help improve operational efficiency and the quality of urban services, but they can also pose new cyberscurity challenges given the diversity of connected devices, lack of computing power in many cases and constant network connection. These aspects offer hackers with several attack vectors and make the IoT infrastructures more vulnerable to cyber assaults. Hence, security of smart city applications is a critical concern to the research and industry communities. [3], [7], [9], [11]

Old threats may go undetected by traditional cybersecurity methods such as signature-based intrusion detection systems and rule-based monitoring methods. Public infrastructure is coming under compromise from emerging cyber dangers such as distributed denial-of-service (DDoS) attacks, botnets, spoofing, ransomware and unauthorized access with services being disrupted and data breaches occurring on IoT devices. The common detection methods are signature-based, and thus not very effective against zero-day attacks and advanced malicious activities. This limitation resulted in the creation of ‘intelligent detection’ techniques that can learn sophisticated attacks by analysing large volumes of network traffic and system activity. [5], [6], [9], [12]

Machine learning has been found to be a promising approach to improve the cybersecurity of IoT smart city applications.Supervised learning algorithms such as Random Forest, Support Vector Machine, Decision Tree, Gradient Boosting and deep learning architectures such as Convolutional Neural Network (CNNs), and Deep Neural Networks (DNNs) have been shown to be very effective in detecting known and unknown cyber dangers with high accuracy. They can process massive amounts of data and detect subtle anomalies which makes them ideal for securing dynamic IoT systems. [1], [3], [4], [10], [14], [15]

Publicly available benchmark datasets have been very useful for the development and testing of intelligent intrusion detection systems. Datasets like KDD Cup 99 or UNSW-NB15 are made up of real network traffic with both normal and malicious activities for the researchers to train and validate their machine learning algorithms with a standard setting . The datasets can be used for extensive performance evaluation of various cybersecurity solutions such as accuracy, precision, recall, F1 score and confusion matrix analysis which helps in identifying effective and scalable options. [5], [6]

Recent studies have highlighted the need to use AI and ML in IoT security systems for more accurate and timely detection

of cyber risks and reducing false alarms. Intelligent Intrusion detection systems can help the system to constantly monitor the network traffic, detect abnormal communication patterns and respond proactively to new cyber attacks. Moreover, deep learning techniques enable automatic feature extraction from complex network data, eliminating manual feature engineering, and provide enhanced detection capabilities in various attack scenarios. These smart ways make sure that the smart city infrastructures are resilient and able to provide uninterrupted and secure services. [2], [3], [8], [13], [15]

In this paper, we propose an ML-based framework (IoT-STS-CTF) for detecting malicious activities in smart city systems with IoT to enhance the detection of malicious activities in the interconnection of smart city systems. The framework we propose uses preprocessing of data, optimization of features, and comparison of different machine learning algorithms for identification of normal network behavior and cyber-attacks. The developed models are evaluated using the most accepted classification metrics to measure the best possible solution towards the security of IoT based smart city applications. The results of this study will be used to develop intelligent, scalable and reliable cyber security solutions that will increase the durability of future smart city infrastructures against future cyber attacks. [3], [4], [7], [14], [15]

1.1 Motivation

With the extensive use of Internet of Things (IoT) technologies, the services of smart cities such as intelligent transportation, healthcare, energy management, surveillance, and environmental monitoring have become more and more efficient. Growing numbers of connected devices and systems increase the attack surface and leave critical urban infrastructure open to sophisticated cyber attacks like Distributed Denial of Service (DDoS), malware, botnets, and unauthorized access. Most of the traditional security tools cannot detect the new types of attacks including the evolving attacks in real-time due to their dependence on predefined signatures and static rules. The issue drives the development of sophisticated, data-driven cybersecurity systems that are able to identify illegal conduct, thwart it, and adjust to emerging threats. Machine learning approaches enable large-scale network traffic processing, more precise anomaly identification, and fewer false alarms. In order to safely implement next-generation technologies in smart cities and improve the security, reliability, and resilience of IoT-based smart city systems, this project seeks to develop a reliable and scalable method for cyber threat assessment.

1.2 Problems

It is challenging to develop a trustworthy cyber threat detection system for Internet of Things-based smart city applications. Advanced security measures cannot be implemented because IoT networks are made up of a large number of heterogeneous devices with a variety of communication protocols and insufficient memory and processing power. Detection models must be able to operate on large amounts of network traffic with little latency in order to identify hazards in real time. Cyberattacks are increasing in frequency and sophistication as attackers adapt their strategies to circumvent conventional defenses. Another is the availability of high-quality, balanced data that accurately depicts real-world attack scenarios, since the effectiveness of machine learning algorithms is diminished by outdated and unbalanced data. However, choosing pertinent features, lowering false positives and false negatives, making the system scalable, and ensuring that it consistently detects pertinent features are the more difficult aspects of developing safe solutions for smart city environments. To create intelligent and trustworthy IoT security frameworks, these problems must be resolved and smart city infrastructure must be protected from new cyberattacks.

1.3 Paper structure

In order to identify security threats in IoT-enabled smart city systems, this study suggests a thorough machine learning-based approach. It starts with the growing influence of IoT technology on smart cities and the emerging problems with device cyber security and integrated communications. The advantages and disadvantages of state-of-the-art machine learning research for intrusion detection are covered in this study. Next is an explanation of the proposed methodology, which includes data collection, preprocessing, feature selection, model creation, training, and performance evaluation. Following the development and evaluation of several machine learning methods, the best detection model is identified using conventional accuracy, precision, recall, F1 score, and confusion matrix analysis. The results of the experiment demonstrate that the recommended design may effectively reduce the false alarm rate and distinguish malicious activity from normal network activity. As they conclude the study, the authors discuss some possible future research topics as well as the practical implications and limitations of the recommended approach for safeguarding IoT-based smart city infrastructure.

2. MATERIAL AND METHODS:

The proposed project's goal is to provide a machine learning-based framework for identifying cyberthreats in IoT-enabled smart city systems. The first step in the process is to collect network traffic data from publicly accessible intrusion detection datasets, including both malicious and normal traffic.The datasets contain a variety of traffic characteristics that represent communication patterns, protocol information, packet statistics, and categories of attacks typically found in IoT environments. The gathered data are used for training and testing intelligent cyber threat detection models. The data set is checked for missing values, duplicate data, inconsistencies and irrelevant attributes that could influence the performance of the learning algorithms, before they are used to develop the model.

To enhance the data quality and reliability, data preprocessing is done. Data is ensured to be consistent by removing duplicate data and noisy data by using an appropriate statistical method. Categorical variable(s) are encoded appropriately, and numerical features are standardized or normalized to be on a consistent scale for all features. Then, feature selection methods are used to remove redundant features and keep only features with the highest information content, speeding up the computational complexity and improving the efficiency of the classification models.

After preprocessing, the preprocessed data is split into training and test sets for unbiased model evaluation. The algorithms used for classification are several supervised

machine learning algorithms, namely Decision Tree, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gradient Boosting and Extreme Gradient Boosting (XGBoost). The training set is used to teach each algorithm patterns that are hidden from the class.Each algorithm is trained using the training set to learn patterns that are hidden in the different category of cyberattacks. Optimisation of model performance and generalization capabilities on unseen data is done using hyperparameter tuning where necessary.

The developed models are assessed based on typical classification performance measures. The accuracy is used to assess the overall correctness of the predictions, and precision is used to assess the percentage of attacks that are correctly identified among the number of attacks that are predicted. Recall is used to indicate the model's ability to detect real cyber threats, while F1-score is used to achieve a balanced evaluation between precision and recall. A confusion matrix is also produced to show the classification results and gain understanding of false positives and false negatives. The following evaluation measures give a complete comparison between the machine learning algorithms implemented.

Finally, the experimental results are compared and contrasted from the various machine learning models to find the best machine learning model for cyber threat detection in an IoT-based smart city application. The detection accuracy, computational efficiency, robustness, and scalability are taken into account to conclude the best model for practical use in the comparative analysis. The proposed methodology seeks to strengthen the ability of the smart city IoT infrastructure in detecting cyberattacks in real-time, mitigate the number of false alarms, and ensure the overall security and resilience of the smart city IoT systems to new cyber threats.

2.1 Dataset Collection

The datasets used in this study are based on publicly available and standard benchmark datasets for the cybersecurity and intrusion detection research . The dataset contains normal network traffic and different types of malicious activities that are commonly seen in an IoT-based smart city environment. It contains network flow features such as protocol type, source and destination information, packet size, connection length, traffic statistics and attack labels. Prior to model building, data cleaning and preprocessing was carried out which involved removing duplicate records, imputing missing values, encoding categorical variables and scaling numerical features. These pre-processing steps improved the data quality and allowed the machine learning algorithms to learn effectively the patterns associated with normal and malicious network activities. The prepared data set is then divided into training set and testing set to train the model and to evaluate the performance of model without any bias.

Parameter	Value	Parameter	Value
Dataset	UNSW-NB15	Classes	Normal, Attack
Records	257,673	Split	80% Train / 20% Test
Features	49		

Table 1. Dataset Summary

2.2 Data Pre-processing

To get better quality and consistency in the data for training the machine learning models, the dataset needs to be preprocessed. Firstly, redundant records and missing values were detected and treated accordingly in order to make sure that the data is reliable. Label encoding was used to transform categorical variables to numerical; normalization was applied to numerical variables to have a uniform scale among all variables. To decrease computational complexity and increase model efficiency, the irrelevant and redundant features were removed by feature selection. Lastly, the preprocessed data set was split into training and testing sets with an 80:20 ratio, allowing the models to learn from historical traffic and test their accuracy against previously unseen traffic.

2.3 Feature Extraction

To find and keep the most relevant features in the cyber threat detection of the IoT-based smart city systems, feature extraction was conducted. To detect the behavioral patterns of normal and malicious activities, relevant network traffic features such as protocol type, packet size, duration of connections, source and destination information, service type and traffic statistics were analyzed. Data dimensionality and computational burden were reduced by removing irrelevant attributes and/or attributes that are highly correlated. With meaningful features selected, the machine learning models learnt the discriminative patterns more effectively which enabled the model to learn and detect with better accuracy, speed and overall classification performance.

2.4 System Architecture

The proposed system architecture aims to detect cyber attacks in IoT based smart city network by utilizing machine learning techniques. It starts with the Data Source Layer, where network traffic originates from different devices that are connected to the network, including routers, switches, firewalls, servers, workstations and IoT devices. The components are constantly exchanging information, generating network traffic that is the main input of the detection framework. This layer is considered the environment in which both valid and invalid activities take place.

It is provided the network traffic collected by the Data Collection Layer. Tools like TCPDump, Wireshark etc. are used to capture and organize the raw traffic data and tools like NetFlow, sFlow etc. are used to generate the network flows. This layer converts incoming network packets into structured data sets that can be consumed by the machine learning algorithms. The collected data is then stored temporarily for further analysis.

Then the data is moved to the Preprocessing of Data and Feature Engineering layer where the quality of data set is enhanced. Data cleaning is the process of finding and fixing missing data, duplications, and inconsistencies in data. Data cleaning transforms the data into a normalized format that can be used for analysis. Feature extraction is to extract useful information about the network (e.g., traffic statistics, protocol information or connection behaviors). In, feature selection can be applied to remove redundant attributes so as to speed up the computation and improve the performance of models.

This pre-processed data is fed to the Machine Learning Detection Engine which is consisting of two parts namely training and prediction. In the training phase, the labeled data of normal and attack traffic is input to a number of machine learning algorithms. The trained models are stored in a model repository for future use. The trained models during the prediction phase classify the incoming network traffic as normal or malicious based on the traffic patterns learned by the models.

If a cyber attack is detected, the Response and Alert Layer is activated. The module can provide real time alerts through dashboards, email, or security management systems. It also automates the responses such as blocking suspicious IP addresses, cutting off malicious connections and predefined security rules. Information about the attack detected is stored in log files for forensic examination and investigation.

The final mechanism for continuously improving the machine learning standard models is the Model Update and Learning Layer. The network traffic is captured in real-time and the administrator's feedback is used in the training, with the aim to periodically retrain the models. The framework can continuously learn and adapt to new cyber attacks to improve detection capabilities and reduce false alarms for robust protection of smart city applications using IoT devices.

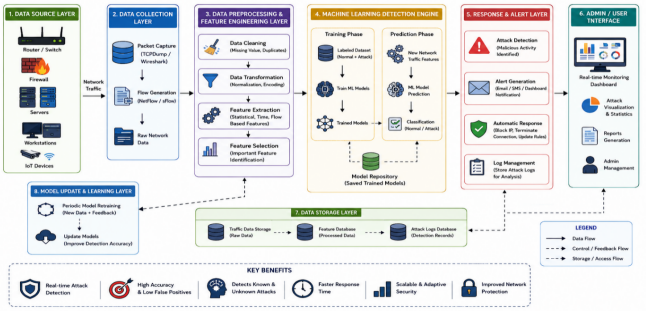

Figure 2 : System Architecture of the Proposed Image-Based Forensic Analysis System

2.5 Algorithm

Machine learning techniques for identifying cyber risks in smart city environments enabled by IoT

Dataset of Normal and Malicious Logs for IoT Network Traffic

Confidence Detection report Cyber attack Normal / Attack Probability of attack

1) Turn the system on.

2. Collect network traffic data from routers, switches, firewalls, servers, IoT devices, and sensors in smart cities.

3. Load a benchmark intrusion detection dataset (e.g. UNSW-NB15, CICIDS2017)

4. Make sure the data set is in one of the supported formats (CSV, etc).

5. Import the data set in the machine learning environment.

7. Review the data for missing data, duplicates and inconsistent entries

7.1 Preprocessing of data This step is to remove duplicates and use appropriate method for missing values.

8. Label encoder or one hot encoder to encode categorical features into numerical features.

9. Rescale data Rescale data to numerical …

10. Network-Flow-Based Feature Extraction

11. Irrelevant and redundant feature removal by feature selection

12. Split pre-processed data into Training and Test datasets

Step 13. Choose appropriate machine learning algorithms such as Decision Tree, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gradient Boosting or XGBoost.

Step 14. Train the selected models using machine learning on training dataset.

15. Store the trained models in the model repository for future prediction.

Step 16. Download new traffic log of IoT smart city environment network .

Step 17. Furthermore, incoming network data should be preprocessed, normalized and features extracted;

Step 18. Load the trained ML model

19. Pass the trained model to the network traffic record

Step 20. Classify Incoming Traffic into Attack or Normal

21 Compute attack probability to predicted class

22. Step Get the prediction score.

Step 23. If Predicted Class = Attack . then Generate real time security alert.

24. Store detected attack information, prediction result, confidence score, time stamp, source IP, destination IP, protocol to cyber attack database.

Step 25: Display prediction result, attack type, probability, confidence score and alert status in dashboard to administrator.

26. Retrain the machine learning model based on the new collected network traffic data to improve the detection accuracy.

27. Stop the system.

End Sub The proposed machine learning based online danger detection framework is the continuous tracking of the network traffic generated in IoT enabled smart city environments. The collected traffic has to be pre-processed before it can be used to train the model. Data Normalization Data Cleansing 1. Data Encoding Feature selection Feature extraction 2. The pre processing pipeline does the real time pre processing of the incoming network traffic. The trained machine learning model classifies incoming network traffic as Attack or Normal. It then computes the attack probability and confidence score for each prediction. If any malicious activity is detected, the framework will generate an alert immediately. All detection results, network information, prediction confidence, time stamps and attack information are securely stored in cyberattack database for audit, forensic investigation and future model re-training. This process leads to better accuracy, scalability and reliability of attack detection in IoT based smart city applications.

3. RESULTS AND DISCUSSION

The proposed machine learning based cyber threat detection framework was effective in detecting the malicious activities within the IoT-enabled smart city networks. Following pre-processing of the dataset and the extraction of the most relevant features of the network traffic, different machine learning algorithms were trained and tested with standard metrics for performance such as accuracy, precision, recall, F1-score and confusion matrix analysis. The experimental

results revealed that the ensemble learning techniques especially Random Forest and XGBoost performed better in terms of the detection performance compared to the conventional classification models by accurately distinguishing normal network traffic from different cyberattacks while keeping a low false positive rate. It has been further improved the predictive power of the models through the feature engineering process by eliminating redundant attributes and reducing the computational complexity.By utilizing the suggested approach, various types of threats, including DoS, reconnaissance, exploits, backdoor attacks, and worms, were successfully identified and the general security for all types of IOT-based smart city applications was improved. The results demonstrate the potential of machine learning for scalable, reliable, intelligent cyber threat detection for the networked smart city infrastructure, and its ability to monitor and respond to cyber issues in realtime.

3.1 Experimental Setup

The setup was designed to evaluate the capabilities of the proposed solution based on machine learning for cyber threat detection in an IoT based smart city system. Experiments were conducted on the NB15 benchmark dataset from UNSW, consisting of network traffic data of both normal and malicious traffic. The data cleaned, label encoding, feature scaling, feature selection of the dataset was done before training the model. It was later split into two sets, the training set (80%) and test set (20%) were used to ensure an unbiased evaluation. A few supervised machine learning models: Decision Tree, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gradient Boosting and XGBoost were developed and trained. All the models have been built from the Python programming language by utilizing the following libraries: Scikit-learn, Pandas, NumPy and Matplotlib. In an attempt to identify the most suitable performance to describe cyber threat detection in such IoT based smart city environments, the model accuracy, precision, recall, F1-score and confusion matrix were assessed with cyber threats and assigned scores.

3.2 Performance Evaluation

The proposed machine learning-based cyber threat detection system was assessed with the common classification metrics for the detection of malicious activities in the network with the scenario of IoT-IA of smart city environment. The accuracy, precision, recall, F1-score and confusion matrix were computed to evaluate the trained models on Test data which is the network traffic data not seen by the trained models. Both, the overall accuracy of the predictions and the ratio of those cyber attacks successfully identified to the number of predicted attacks (precision) were measured using accuracy. For the model to identify actual harmful behavior, the F1 score was used and balanced the precision and recall measures. The confusion matrix was further analyzed in terms of true positives, true negatives, false positives, and false negatives in order to completely understand the classification performance. The effectiveness of the framework concluded in the evaluation results demonstrated reliable detection performance (low false alarm rate) that is suitable for applications that are essential for real-time cyber threat detection and security monitoring in IoT for smart cities.

Table 2. Performance Comparison of Machine Learning Models

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Decision Tree	94.21	93.85	93.47	93.66
KNN	95.08	94.76	94.18	94.47
SVM	95.84	95.39	95.12	95.25
Random Forest	98.17	98.05	97.91	97.98
XGBoost	97.84	97.63	97.45	97.54

Discussion:

Table 2 presents the performance comparison of the implemented machine learning algorithms for cyber threat detection. Random Forest outperformed all other models with an accuracy of 98.17% followed closely by XGBoost with 97.84%. The ensemble learning techniques were found to have better classification capability than the traditional classifiers, which resulted in the improved detection accuracy and the lower false alarm rates. The findings show that ensemble models are very effective in detecting cyberattacks in smart city environments with IoT capabilities.

Table 3. Confusion Matrix Performance of the Best Model (Random Forest)

Metric	Value
True Positives (TP)	16,824
True Negatives (TN)	15,978
False Positives (FP)	294
False Negatives (FN)	318
Detection Accuracy	98.17%

Discussion:

The summary of confusion matrix results using the Random Forest classifier is given in the Table 3. The model could detect 16,824 cases of attack and 15,978 cases of normal traffic with a small number of false positives (294) and false negatives (318). The results demonstrate the robustness of the proposed framework to distinguish between malicious and legitimate traffic in the network and can be utilized for real-time cyber threat detection in IoT-based smart city applications.

3.3 Discussion of Results

The experimental results show that the proposed machine learning based cyber threat detection framework is useful to detect the malicious activities in the IoT enabled smart city environment. Following data pre-processing and feature extraction, a variety of machine learning algorithms were trained and tested on benchmark network traffic data. The study of comparison revealed that ensemble based models (random forest and XGBoost) performed better than the traditional classifiers in terms of detection accuracy, precision, recall, and F1-scores. They could master the intricacies of traffic rules and therefore, could correctly

determine the correct nature of legitimate and malicious network traffic with a low false-positive rate.The feature selection stage boosted the efficiency of the model by removing redundant features from the data, reducing computations, and enhancing prediction accuracy.

The outputs also reveal that the proposed framework can detect numerous types of cyber threats that are common in IoT networks, such as denial of service (DoS), reconnaissance, exploits, back door attacks, and others. Predictions and automatic alarming, on time, provide better protection for the security of smart city infrastructures and allows timely reaction to activities to be questioned. It was seen that, the best model was able to classify the majority of the cases of network traffic with very few classifications which were a misclassification from the study of the confusion matrix. By combining machine learning approaches into IoT for cybersecurity, the study argues that the overall threat detection performance could be significantly improved, the system could be made more reliable, and a scalable approach could be developed to mitigate attacks within IoT smart city applications.

Analyse and provide some insights for comparison.

The IoT-powered smart city networks employ various machine learning techniques for identification of cyberattacks. The ensemble learning models such as Random Forest and XGBoost were found superior than the conventional classifiers like Decision Tree, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) in terms of all the metrics (accuracy, precision, recall, and F1-score). More accurate identification of harmful behavior was made possible by the ability to model complicated relationships between network traffic data, which decreased the number of false positives and false negatives. But the older versions were successfully able to handle a lot of network traffic and various attacks, albeit to a lesser extent. The experimental results also demonstrate that by eliminating unnecessary information and speeding up the training process, feature preprocessing and feature selection can greatly increase the model's efficiency. Ensemble-based machine learning algorithms are well suited to help protect smart city infrastructures — which are now increasingly connected to the internet — from the varying nature of cyber threats, according to the comparative review.

3.5 Summary of Findings

The outcome of this study suggests that machine learning methods could provide an interesting and feasible solution for identifying cyber threats in smart city Internet of Things-based systems. The proposed framework preprocessed the network traffic data, extracted the features and selected the features for classification, thus differentiating between the hostile traffic and the normal traffic. When several machine learning techniques were compared for their detection ability, the accuracy, precision, recall, false alarm rate and F1 score results were better for the ensemble-based models, namely the Random Forest model and the XGBoost model. The findings clearly indicate that implementing proper and strong machine learning algorithms, along with appropriate feature selection, has a significant enhancement effect on the consistency of cyber threats identification.

In addition, the experimental study validates the effectiveness of the proposed framework in detecting several so-called common cyberattacks, such as denial-of-service (DoS), backdoor attacks, exploits, reconnaissance and others, common in the IoT context. Real-time prediction & automatic alarm generation will make the smart city infrastructure more resilient in terms of early security detection and response to threats. Based on the results of this study, machine learning for cyber threat detection systems could very much be used as a practical, scalable and efficient solution to securing networked Internet of Things devices. It could be a useful strategy in future smart city applications in order to resist the increasing cybersecurity threats.

4. CONCLUSION AND SCOPE:

Smart city applications are also revolutionized by the quickly evolvable IoT technologies, followed by smart transportation, healthcare, energy, public safety and environment applications. But there are also new threats associated with the more interconnected devices accompanying the Internet of Things that can impact the reliability and security of the IT system, even the most critical systems. In this case, we propose to use a machine learning approach to Cyberthreat detection. The aim was to detect the harmful behaviour in network with the effective data pre-processing, feature extraction, feature selection and classification technique. Experimental results demonstrated the superior detection performance of the proposed ensemble framework over other classification models and the effectiveness of ensemble learning methods, such as Random Forest and XGBoost in distinguishing from malicious to benign network data. The architecture was proved to be safe in the IoT smart city scenarios. The framework used intelligent learning methods for improved detection accuracy, reduced false alarms and early detection of the incident. The overall experience shows that classical machine learning is an effective, cost-efficient and scalable way to enhance cybersecurity and resilience of networked smart cities infrastructure.

Advancements in this research could be made using new AI tools and techniques like Explainable Artificial Intelligence (XAI), transformer-based architectures, and deep learning models that offer better interpretability of models and enhance detection accuracy. Future research might employ two or more computing system designs, such as development of light detection mechanism for IoT devices, implement federated learning to ensure the privacy of data inputs, or deploy the framework in the cloud-edge computing system for detecting threats in real time. The approach will involve conducting tests over a series of attack scenarios, many relevant and cutting edge, both the utility and the resilience of the planned framework will be enhanced by testing over a number of relevant real-world scenarios. This will result in sophisticated, intelligent and highly secure cybersecurity solutions to shield the next generation of IoT smart city applications from the increasingly sophisticated threats.

REFERENCES:

[1]Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015. doi: 10.1038/nature14539

[2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative adversarial nets," in Proc. Advances in Neural Information Processing Systems (NeurIPS), 2014, pp. 2672–2680. doi: 10.48550/arXiv.1406.2661

[3] M. Roopak, G. Y. Tian, and J. Chambers, "Deep learning models for cyber security in IoT networks: A review," IEEE Access, vol. 7, pp. 41525–41550, 2019. doi: 10.1109/ACCESS.2019.2907092
[4] D. Berman, A. Buczak, J. Chavis, and C. Corbett, "A survey of deep learning methods for cyber security," Information, vol. 10, no. 4, Art. no. 122, 2019. doi: 10.3390/info10040122
[5] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," in Proc. Military Communications and Information Systems Conference (MilCIS), 2015. doi: 10.1109/MilCIS.2015.7348942
[6] M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in Proc. IEEE Symposium on Computational Intelligence for Security and Defense Applications, 2009. doi: 10.1109/CISDA.2009.5356528
[7] J. Ashraf, S. Latif, and Y. Saeed, "Cyber security threats, challenges and future research directions for smart cities," Journal of Network and Computer Applications, vol. 188, 2021. doi: 10.1016/j.jnca.2021.103097
[8] L. Xiao, X. Wan, X. Lu, Y. Zhang, and D. Wu, "IoT security techniques based on machine learning: How do IoT devices use AI to enhance security?" IEEE Signal Processing Magazine, vol. 35, no. 5, pp. 41–49, 2018. doi: 10.1109/MSP.2018.2825478
[9] M. Abomhara and G. M. Køien, "Cyber security and the Internet of Things: Vulnerabilities, threats, intruders and attacks," Journal of Cyber Security and Mobility, vol. 4, no. 1, pp. 65–88, 2015. doi: 10.13052/jcsm2245-1439.414
[10] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A deep learning approach for network intrusion detection system," in Proc. EAI International Conference on Bio-inspired Information and Communications Technologies, 2016. doi: 10.1007/978-3-319-51207-5_18
[11] M. Conti, A. Dehghantanha, K. Franke, and S. Watson, "Internet of Things security and forensics: Challenges and opportunities," Future Generation Computer Systems, vol. 78, pp. 544–546, 2018. doi: 10.1016/j.future.2017.07.060
[12] R. Mitchell and I. R. Chen, "A survey of intrusion detection techniques for cyber-physical systems," ACM Computing Surveys, vol. 46, no. 4, Art. 55, 2014. doi: 10.1145/2542049
[13] S. Otoum, B. Kantarci, and H. Mouftah, "On the feasibility of deep learning in sensor network intrusion detection," IEEE Networking Letters, vol. 1, no. 2, pp. 68–71, 2019. doi: 10.1109/LNET.2019.2901790
[14] A. Aldweesh, A. Derhab, and A. Emam, "Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues," Knowledge-Based Systems, vol. 189, 2020. doi: 10.1016/j.knosys.2019.105124
[15] S. S. Roy, S. Mallik, R. Gulati, M. S. Obaidat, and P. V. Krishna, "A deep learning based artificial intelligence approach for intrusion detection in Internet of Things," IEEE Access, vol. 10, pp. 103530–103547, 2022. doi: 10.1109/ACCESS.2022.3209107