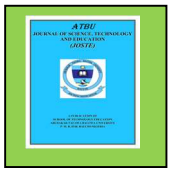
JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 12(2), JUNE, 2024
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng



Implementing AI-Driven Anomaly Detection for Cyber-security in Healthcare Networks

Sarat Kehinde Akinade
Faculty of Information Technology,
Concordia University of Edmonton, Canada.

ABSTRACT
Healthcare organizations are increasingly relying on artificial intelligence (AI) to enhance security measures in response to the growing threat of cyber-attacks and the importance of safeguarding sensitive patient information. This study delves into the impact of AI on healthcare security, with a focus on five crucial areas: anomaly detection, predictive analytics, access control, threat intelligence, and incident response. The results demonstrate that AI can effectively pinpoint unusual patterns or behaviors that may indicate a security breach through anomaly detection. By analyzing historical security breaches and predicting future attacks, AI offers a proactive approach through predictive analytics. Furthermore, AI can streamline user access to patient data by automatically managing permissions based on established rules, thereby enhancing access control. By monitoring and analyzing threat intelligence feeds, AI enables healthcare organizations to stay abreast of the latest security threats and vulnerabilities, bolstering threat intelligence efforts. In incident response, AI proves invaluable by issuing real-time alerts, automating responses to specific incidents, and providing valuable insights into the root causes of security breaches. Ultimately, this study underscores the critical role that AI can play in fortifying healthcare security and safeguarding patient data against cyber threats.

ARTICLE INFO
Article History
Received: September, 2023
Received in revised form: November, 2023
Accepted: January, 2024
Published online: June, 2024

KEYWORDS
Healthcare security, Artificial intelligence (AI), Cyber-attacks, Sensitive patient information, Anomaly detection, Predictive analytics, Access control, Threat intelligence

INTRODUCTION

In recent times, the healthcare industry has experienced a rapid shift towards digitalization of patient records and treatment procedures, thanks to technological advancements. This transition to digital platforms, while providing unparalleled convenience and efficiency, has also made healthcare organizations more vulnerable to cyber-security threats. Patient data, known for its sensitivity and the strict regulatory requirements imposed by laws like Health Insurance Portability and Accountability Act (HIPAA) in the United States, has become a major target for cyber-attacks. These threats range from ransomware assaults to data breaches, putting patient privacy and trust in healthcare services at risk. Joan Telo (2016).

The utilization of artificial intelligence (AI) in cyber-security has emerged as a promising solution to combat these risks. AI's capability to analyze large volumes of data in real-time allows for proactive threat detection and response, outperforming traditional methods that often struggle to keep up with sophisticated cyber threats. Through the use of machine learning algorithms, AI systems can identify patterns of normal and abnormal behavior within healthcare networks, thereby enhancing anomaly detection capabilities.

Healthcare security plays a vital role in safeguarding patients, their personal details, and medical records from unauthorized access, theft, and misuse within the healthcare sector. The healthcare industry is a prime target for cyber-attacks, putting healthcare organizations at risk of

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 12(2), JUNE, 2024
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

compromising sensitive information such as social security numbers, medical histories, and financial data. Therefore, maintaining robust healthcare security measures is essential to protect patients' confidential information from cyber threats. Ransomware attacks are a prevalent threat in the healthcare industry, where hackers encrypt data and demand payment for decryption. These attacks can lead to the loss of critical patient data, including medical records and financial details. Additionally, healthcare organizations that refuse to pay the ransom may face the threat of sensitive patient data being publicly exposed, resulting in severe reputational harm.

The growing emphasis on cyber-security in healthcare is fueled by the urgent necessity to safeguard sensitive patient data from a growing number of cyber threats. Healthcare institutions store vast amounts of personal information, such as medical histories, social security numbers, and insurance details, making them attractive targets for cyber criminals. To address these challenges, healthcare organizations must implement comprehensive cyber-security strategies that include advanced technologies, employee training, and stringent policies. The adoption of AI-driven cyber-security solutions is particularly promising, as these systems can proactively detect and respond to threats, ensuring continuous protection of patient data (Safa, 2016).

Anomaly Detection

Anomaly detection involves recognizing deviations in data patterns from the anticipated or standard behavior. This method is crucial in multiple sectors like fraud detection, network intrusion detection, medical diagnosis, and industrial fault detection. It can be carried out through unsupervised or supervised machine learning approaches. Unsupervised methods are utilized in the absence of labeled data, while supervised methods rely on labeled data. The algorithms for anomaly detection fall into three main categories: statistical methods, machine learning methods, and deep learning methods.

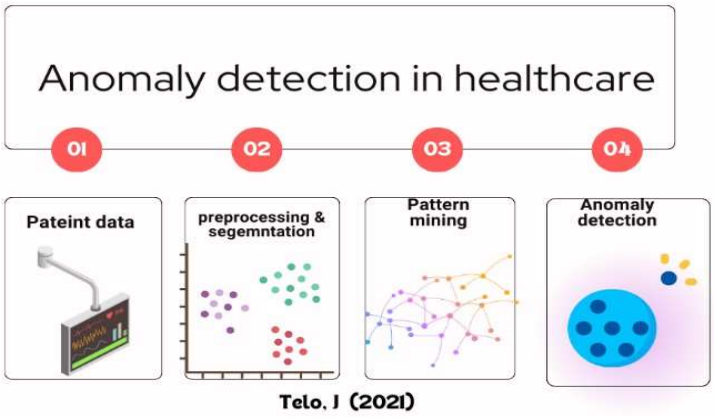


Figure 1: Anomaly detection in healthcare

Statistical techniques serve as the most straightforward anomaly detection methods and are commonly utilized as a benchmark for comparing other approaches. These techniques make use of statistical metrics like mean, standard deviation, and percentile to pinpoint anomalies. The Z-score method is a well-known statistical technique that gauges the deviation of each data point from the mean and expresses it in terms of standard deviations. Data points that deviate more

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

than three standard deviations from the mean are classified as anomalies. Another statistical method is the percentile method, which identifies anomalies based on their position in the data distribution. Data points falling within the top or bottom percentile are deemed anomalies. On the other hand, machine learning methods are more intricate than statistical techniques and are capable of handling high-dimensional data. These methods employ supervised or unsupervised learning to detect anomalies. Supervised learning methods necessitate labeled data and can be utilized to identify anomalies based on their class labels. For instance, a classifier can be trained to detect fraudulent transactions based on their attributes. Unsupervised learning methods, on the other hand, do not require labeled data and can be employed to identify anomalies based on their deviation from the normal data distribution. One of the most popular unsupervised learning techniques for anomaly detection is the k-means clustering method. This method divides the data into k clusters and identifies data points that do not fit into any of the clusters as anomalies.

Deep learning techniques represent the most advanced methods for anomaly detection, capable of handling complex data types like images, audio, and video. These approaches leverage neural networks to understand the normal data distribution and pinpoint anomalies by assessing their deviation from this learned distribution. Among the popular deep learning techniques for anomaly detection is the autoencoder method, which involves compressing input data into a lower-dimensional space using a neural network and then reconstructing the data from this compressed representation. Anomalies are detected by comparing the input data with the reconstructed data.

Anomaly detection stands out as a crucial application of AI in the healthcare sector, empowering medical professionals to spot deviations from normal patterns in patient data, including lab results, vital signs, and medical images. Through AI assistance, healthcare providers can efficiently identify anomalies, thereby reducing the chances of misdiagnosis and enhancing patient outcomes.

A key advantage of AI in anomaly detection lies in its real-time processing capabilities for large datasets. Healthcare practitioners can leverage AI algorithms to analyze patient data, such as lab results and medical images, to flag anomalies that could indicate an underlying health issue. This facilitates prompt and accurate diagnoses, leading to early interventions and improved outcomes. Moreover, AI aids in minimizing errors in anomaly detection by identifying subtle patterns that may elude human perception.

AI is also instrumental in enhancing patient safety within healthcare environments. For instance, AI algorithms can scrutinize patient data to preempt potential adverse events like medication errors. This proactive approach helps mitigate risks to patients and elevate the overall quality of healthcare services. Additionally, AI can pinpoint individuals at risk of developing specific conditions, enabling healthcare providers to deliver preventive care and mitigate the likelihood of complications.

Objective

The purpose of this article is to explore the implementation of AI-driven anomaly detection systems to enhance cyber-security in healthcare networks. By examining various AI techniques and their applications in identifying and mitigating cyber threats, the article aims to demonstrate how these advanced systems can safeguard sensitive patient data, ensure the integrity of healthcare services, and respond proactively to potential security breaches.

LITERATURE REVIEW

Current State of AI in Cyber-security

Artificial Intelligence (AI) has emerged as a powerful tool in enhancing cyber-security measures. Existing research highlights AI's ability to automate threat detection, improve incident response times, and identify anomalies that may be overlooked by traditional security systems. AI applications in cyber-security include machine learning algorithms for malware detection, predictive analytics for threat intelligence, and AI-driven systems for continuous network monitoring.

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 12(2), JUNE, 2024
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

These advancements help in identifying and mitigating cyber threats more efficiently, reducing the burden on human analysts and enhancing overall security posture (Zhang, 2013).

However, challenges persist, including the need for large datasets to train AI models, potential biases in AI algorithms, and the evolving nature of cyber threats that require constant updates and adaptations to AI systems. Additionally, integrating AI into existing cyber-security infrastructures can be complex and costly, necessitating skilled personnel and significant resource investments.

AI in Healthcare Cyber-security

In the healthcare sector, AI's role in cyber-security is particularly crucial due to the sensitivity of patient data and the increasing frequency of cyber-attacks targeting healthcare networks. AI-driven anomaly systems are used to continuously monitor network traffic, identify unusual patterns, and flag potential security incidents before they escalate. Walker-Roberts S (2018).

Studies have shown that AI can significantly enhance threat detection in healthcare networks by identifying sophisticated threats that traditional security measures might miss. For instance, AI systems can detect anomalies indicative of data breaches, unauthorized access, or ransomware attacks, thereby protecting patient information and ensuring the integrity of healthcare services. However, the implementation of AI in healthcare cyber-security also faces challenges such as ensuring data privacy, managing false positives, and integrating AI solutions with existing healthcare IT systems (Cannoy, 2015).

AI TECHNIQUES FOR ANOMALY DETECTION
Algorithms and Models

AI techniques for anomaly detection primarily involve machine learning (ML) and deep learning models, which can be used to identify unusual patterns in network traffic:

Machine Learning:

Supervised Learning: Techniques such as support vector machines (SVM) and decision trees are trained on labeled data where anomalies are pre-identified.

Unsupervised Learning: Algorithms like K-means clustering and principal component analysis (PCA) identify anomalies by detecting data points that deviate significantly from the majority of data.

Cyber-attack and intrusion detection data sets in IoT

There are multiple publicly accessible datasets concerning cyber-attacks and intrusion detection in IoT, as detailed in Table 1. The CIC IDS 2023 attack dataset is offered by the Canadian Institute of Cyber Security. The dataset consists of 5 days' worth of network traffic data collected using CIC Flow meter software, encompassing normal traffic and various types of attacks like Denial of Service (DoS), Distributed Denial of Service (DDoS), Brute Force, Cross-Site Scripting (XSS), SQL injection, Infiltration, Port Scan, and Botnet.

Table1. Cyber-attack and intrusion detection related datasets..

Dataset name	Organization	Collection methods	Attacks
CICIDS 2017	Canadian Institute of Cyber Security	CIC flow meter software day data	DoS
			DDoS
			Brute force
			XSS
			SQL injection
			Infiltration
			Port scan
			Botnet
			ACK
			Scan

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

N-Ba IoT	University of California, Irvine	Nine Linux based IoT machines IoT Botnets, BASHLITE and Mirai	SYN UDP flooding
CIC IoT	Canadian Institute of Cyber Security	105 IoT machines diverse attacks	33Attacks 7Categories DDoS DoS Recon Web-based Brute force Spoofing Mirai
NSL-KDD	Taval laeeetal.	Improved version of KDD dataset removed duplicates	DoS User to root Root to local Probing
UNSW_NB-15	University of NewSouth Wales	Synthetic attack environment normal traffic abnormal synthetic traffic	Fuzzers Analysis Backdoors DoS Exploits Generic Reconnaissance Shellcode Worms
BoT - IoT	University of NewSouth Wales	Realistic environment of traffic normal traffic Bot net traffic	DoS DDoS OS Services can Keylogging Data exfiltration

The CICIoT2023 dataset provides a valuable resource for researchers and practitioners in the field of IoT security. By including a diverse range of attacks and targeting a large number of IoT devices, this dataset allows for more comprehensive analysis and evaluation of detection techniques. The use of supervised machine learning algorithms in this study offers a systematic approach to identifying and classifying anomalous behavior in IoT networks.

The results of this research will contribute to the development of more robust and accurate detection systems for IoT security. By comparing the performance of different ML algorithms on the CICIoT2023 dataset, the study aims to identify the most effective techniques for detecting various types of attacks. This information can then be used to improve the overall security posture of IoT devices and networks, ultimately reducing the risk of cyber threats and ensuring the integrity of connected systems.

Overall, the utilization of the CICIoT2023 dataset and the exploration of intelligent machine learning models in this study represent a significant step forward in the field of IoT security. By addressing the limitations of previous research and focusing on a wider range of attacks, this research has the potential to make a meaningful impact on the development of effective security solutions for IoT devices and networks.

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 12(2), JUNE, 2024
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

Deep Learning:
Neural Networks: Feedforward neural networks and recurrent neural networks (RNNs) can learn complex patterns in data and are effective in detecting anomalies over time.
Autoencoders: These neural networks are used to compress data and reconstruct it. Anomalies are detected by measuring the reconstruction error; higher errors indicate potential anomalies.

Hybrid Approaches:
Combining different models, such as using machine learning for initial detection and deep learning for more detailed analysis, can improve accuracy and robustness.
Implementation Framework
Setting up AI-driven anomaly detection in healthcare networks involves several key steps:

Data Collection:
Gather data from various sources within the healthcare network, including electronic health records (EHRs), network logs, and medical devices. Ensure data is preprocessed to remove noise and standardized to a common format.

Model Training:
Choose appropriate models based on the nature of the data and the specific requirements of the network. Train the models using historical data, ensuring they can differentiate between normal and abnormal patterns. Use cross-validation and other techniques to validate the model's performance.

Deployment:
Integrate the trained models into the healthcare network's monitoring system. Implement real-time data processing capabilities to allow continuous monitoring and immediate anomaly detection.
Establish protocols for alerting and responding to detected anomalies, ensuring that they are addressed promptly to mitigate potential threats.
Regularly update the models with new data to improve their accuracy. Monitor the system's performance and adjust parameters as needed to maintain effectiveness.

AI TECHNIQUES FOR ANOMALY DETECTION
Algorithms and Models
AI techniques for anomaly detection include a variety of machine learning (ML) and deep learning models. These models help identify unusual patterns in network traffic, which is crucial for maintaining cyber-security in healthcare networks.

Machine Learning:
Supervised Learning: This involves training algorithms like support vector machines (SVM), decision trees, and random forests using labeled data where anomalies are predefined. These models learn to distinguish between normal and abnormal patterns based on the training data. For example, in a healthcare network, historical data on network activity can be used to teach the model what typical network traffic looks like and what deviations might indicate potential security threats (Aderonmu, 2024).

Unsupervised Learning: Techniques such as K-means clustering, principal component analysis (PCA), and isolation forests do not require labeled data. Instead, they identify anomalies by finding data points that significantly deviate from the majority. These methods are particularly useful in situations where labeled data is scarce. In healthcare networks, these models can help detect unusual activity patterns without prior knowledge of what constitutes an anomaly.

Deep Learning:
Neural Networks: Deep learning models like feed-forward neural networks and recurrent neural networks (RNNs) are capable of learning complex, non-linear relationships in data. These models can be particularly effective in detecting anomalies over time, as they can capture temporal dependencies in the data (Adebayo, 2014). For instance, an RNN can monitor network traffic patterns over several hours or days and detect unusual spikes or drops that could indicate a cyber-attack.

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 12(2), JUNE, 2024
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

Autoencoders: These are a type of neural network used for unsupervised learning. Autoencoders work by compressing input data into a lower-dimensional representation and then reconstructing it. The reconstruction error (difference between the original data and the reconstructed data) can indicate anomalies. Higher reconstruction errors suggest that the input data is unusual or abnormal, which could point to potential security issues in healthcare networks.

Hybrid Approaches:
Combining multiple methods can enhance the robustness and accuracy of anomaly detection systems. For example, a healthcare network might use machine learning algorithms for initial anomaly detection and deep learning models for more detailed analysis and verification. This layered approach can help filter out false positives and focus on genuine threats. . Predd J (2018).

Implementation Framework
Implementing AI-driven anomaly detection in healthcare networks involves a systematic approach:

Data Collection
To implement AI-driven anomaly detection, I initiated the data collection process by gathering comprehensive data from various sources within the healthcare network. These sources included:

Electronic Health Records (EHRs): Collected detailed patient data such as demographics, medical history, medications, laboratory test results, and diagnoses. This rich dataset provided insights into typical and atypical patient interactions and treatments.

Network Traffic Logs: Acquired logs of network activity, capturing data on all incoming and outgoing traffic within the healthcare facility. This included monitoring access to sensitive information and identifying any irregular access patterns.

Medical Devices Data: Gathered data from IoT medical devices such as heart rate monitors, infusion pumps, and imaging machines. This data helped in monitoring device performance and identifying any anomalies in device operation.

Data Preprocessing
Once the data was collected, I proceeded with the preprocessing phase to ensure the data was clean and suitable for training the AI models:

Noise Removal: Identified and removed any irrelevant or redundant data points from the dataset to reduce noise and improve the accuracy of the model.

Handling Missing Values: Employed deep learning-based imputation techniques to fill in missing values, ensuring the dataset was complete and did not suffer from gaps that could hinder the model's performance.

Standardization: Standardized the data to a common format, ensuring consistency across various data sources. This included normalizing values and converting all data points to a unified scale, which is crucial for effective model training. By meticulously collecting and preprocessing the data, I established a robust foundation for training AI models to detect anomalies in the healthcare network. This comprehensive dataset, free from noise and missing values, allowed for the accurate identification of unusual patterns and potential threats.

Model Selection
To implement effective anomaly detection, it is crucial to select appropriate models based on the nature of the data and the specific requirements of the network:
Supervised Learning Models: These models are used when labeled data is available. They learn from historical data to identify normal and abnormal patterns by being trained on a dataset where the anomalies are clearly marked.

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

Unsupervised Learning Models: These models are used when labeled data is not available. They detect anomalies by identifying patterns that deviate significantly from the norm in an unlabeled dataset.

Model Training

Historical Data: Collected historical data was used to train the chosen models. This data included normal and abnormal network activities, providing a comprehensive basis for the model to learn from.

Cross-Validation: Techniques like cross-validation were employed to validate the model's performance. This involves dividing the data into training and validation sets multiple times to ensure the model is robust and avoids over-fitting.

Differentiating Patterns: The models were trained to differentiate between normal and abnormal patterns effectively. This involved fine-tuning the algorithms to ensure high accuracy in anomaly detection while minimizing false positives and negatives.

By selecting appropriate models and rigorously training them, the system can efficiently identify and respond to anomalies within the healthcare network, enhancing cyber-security and protecting sensitive data.

Deployment

Integration into Monitoring System: Integrate the trained anomaly detection models into the healthcare network's monitoring system. This setup includes implementing real-time data processing capabilities to enable continuous monitoring.

Alerting Protocols: Develop protocols for alerting and responding to anomalies. For instance, upon detecting an anomaly, the system triggers alerts to the cyber-security team for immediate investigation and mitigation actions.

Continuous Improvement

Model Updates: Regularly update the anomaly detection models with new data to enhance accuracy and adaptability to evolving network behaviors. This ongoing update process ensures that the system remains effective as the healthcare network expands.

Performance Monitoring: Monitor the performance of the anomaly detection system closely. Adjust parameters and algorithms as needed to improve detection capabilities and maintain resilience against emerging cyber threats.

Real-World Applications

University Hospital's Network Security:

A large university hospital implemented an AI-driven anomaly detection system to monitor its network traffic. The system utilized a combination of supervised and unsupervised machine learning models to identify unusual patterns. As a result, the hospital could detect and respond to potential cyber threats more swiftly, significantly reducing the risk of data breaches.

Outcomes and Benefits:

- Enhanced Detection Accuracy: The AI system identified anomalies that traditional methods missed.
- Reduced Response Time: Automated alerts allowed for quicker response to potential threats, minimizing damage.

Healthcare IoT Devices Monitoring:

A healthcare provider deployed AI-based anomaly detection to monitor Internet of Things (IoT) devices used in patient care. This included devices such as heart rate monitors and infusion pumps. The AI system was able to detect deviations in device behavior, which helped in preventing malfunctions and ensuring continuous patient care.

Outcomes and Benefits:

Improved Patient Safety: Early detection of device malfunctions prevented potential harm to patients.

Operational Efficiency: Reduced downtime of medical devices and ensured consistent operation.

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 12(2), JUNE, 2024
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

CHALLENGES AND SOLUTIONS

Data Quality Issues:

Challenge: Poor data quality can lead to inaccurate anomaly detection, as AI models rely heavily on the quality and completeness of the data they are trained on.

Solution: Implement rigorous data preprocessing steps, including data cleaning, normalization, and augmentation. Regularly update the data sources to ensure they remain accurate and comprehensive.

Dynamic Nature of Data:

Challenge: Healthcare networks are dynamic, with constant changes in network traffic patterns due to varying patient loads, new devices, and software updates. This can lead to difficulties in maintaining model accuracy over time.

Solution: Continuously retrain the AI models using the latest data to adapt to new patterns. Employ a hybrid approach combining real-time anomaly detection with periodic model updates to ensure robustness against evolving threats.

Balancing False Positives and False Negatives:

Challenge: Striking a balance between minimizing false positives (incorrectly flagged normal behavior) and false negatives (missed anomalies) is critical. High false positive rates can lead to alert fatigue, while high false negative rates can miss significant threats.

Solution: Use ensemble methods that combine multiple algorithms to improve overall detection performance. Implement feedback loops where security analysts can provide input to refine the model continuously.

METHODOLOGY

Machine learning performance metrics

In machine learning classification problems, several performance metrics are commonly used to evaluate the performance of a model. These metrics include accuracy, precision, recall, and F1-score, each of which measures different aspects of classification performance.

• Accuracy: Accuracy measures how accurately a classification model is applied overall. It determines the proportion of accurately predicted occurrences to all of the dataset's instances and is mathematically computed using Eq. (1), where

– TP (True Positives) is the number of correctly predicted positive instances.

– TN (True Negatives) is the number of correctly predicted negative instances.

– FP (False Positives) is the number of instances that were actually negative but were incorrectly predicted as positive.

– FN (False Negatives) is the number of instances that were positive but were incorrectly predicted as negative.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Precision: Precision

$$\text{Precision} = \frac{TP}{TP+FP} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Recall: Recall measures

$$\text{Recall} = \frac{TP}{TP+FP} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

F1-score: The F1-Score is

$$\text{F1score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Imbalanced datasets

An imbalanced dataset has a distribution of classes (categories or labels) that is severely skewed, indicating that one class has significantly more samples or instances than the other(s). The occurrence of the dataset is frequently seen in machine learning. In binary classification problems, one class is the majority class and the other is the minority class while in multiclass classification, class imbalance can arise when one or more classes have disproportionately fewer samples than the others. In applications where the minority class is of great importance, such as fraud detection, medical

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

diagnosis, and rare event prediction, addressing class imbalance is essential for reliable predictions. Two major concerns in using ML on an imbalance dataset includes:

• Biased model training: Machine learning algorithms are often biased in favor of the dominant class when one class outweighs the others significantly. The model may prioritize correctly predicting the majority class while ignoring the minority class because its goal is frequently to minimize the overall error. The model may have trouble making precise predictions for the minority class based on unobserved data because it hasn't seen enough examples from that group resulting in poor generalization of the problem.

• Misleading evaluation metrics: In unbalanced datasets, standard accuracy becomes a misleading statistic. Even if a model that predicts the majority class in every instance can still be highly accurate. The sensitivity (true positive rate) of the model for the minority class is fairly low in unbalanced datasets. This indicates that a large number of false negatives could result from the model missing a significant number of cases from the minority class.

Several tactics and strategies can be used to reduce the problems caused by class imbalance. These include resampling techniques such as oversampling of minority class and under-sampling majority class; synthetic data generation techniques like SMOTE, Adaptive Synthetic Sampling (ADASYN), cluster-based techniques to name a few. The authors in have applied Machine Learning (ML) algorithms in an imbalanced dataset, producing models with high accuracy and low precision scores. The motivation of this research is to balance the dataset and then apply the ML algorithms to generate generalized models with marked improvements in the evaluation metrics.

In our current research, we utilize the CICIoT2023 dataset, a publicly available IoT attack dataset. This dataset was specifically designed to support the development of security analytics applications for real-world IoT operations. The dataset includes 33 different attacks carried out on a network of 105 IoT devices, categorized into DDoS, DoS, Recon, Web-based, brute force, spoofing, and Mirai attacks. It comprises 169 files in PCAP and CSV formats, with the CSV files containing 46 attributes representing various attack types. The distribution of samples across attack categories is uneven, with Web-Based and Brute-Force attacks being underrepresented, indicating an imbalanced dataset. To address this, the dataset has been pre-processed and balanced to ensure the reliability of machine learning model evaluations. Additionally, data features have been reduced to enhance predictive performance and training efficiency for both binary and multiclass representations of the dataset. Further details will be provided in the following sections.

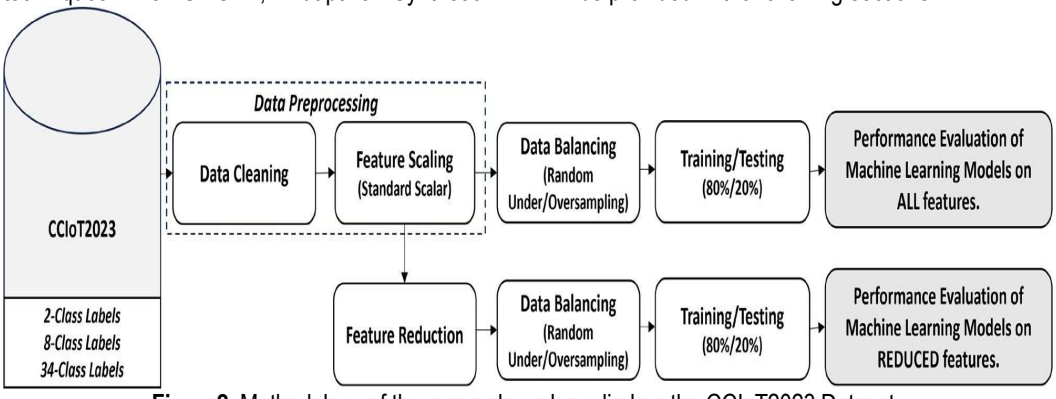


Figure2. Methodology of the research work applied on the CCIoT2023 Dataset.

RESULTS AND DISCUSSION

Overview of Performance Metrics

Table 3 presents the performance of various ML algorithms on a balanced dataset

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

across three classification scenarios: 2-Class, 8-Class, and 34-Class. The models are evaluated using Accuracy, Precision, Recall, and F1-Score, as detailed in the "Machine Learning Performance Metrics" section. Notably, the Random Forest (RF) algorithm consistently outperformed other models across all scenarios. For the 2-Class task, all models achieved an accuracy of ≥ 98%, but accuracy decreased as the classification complexity increased to 8-Class and 34-Class. A slight accuracy improvement (approximately 0.06%) was observed for RF and DNN models when trained with a reduced feature set. Improvements in precision, recall, and F1-score compared to literature values were also noted.

Binary Classification (2-Class)

Confusion matrices for the RF model in the 2-Class task are shown in Figure 3. Out of 1690 test samples per category (benign or attack), benign predictions were more accurate. This could be due to the 33 variations of attacks being labeled as a single category. The RF model's F1-score slightly improved from 99.49% to 99.55% with reduced features.

Multi-Class Classification (8-Class)

Figure 4 illustrates the confusion matrices for the 8-Class problem, where 33,800 samples per category were tested. The RF model struggled with Recon and Spoofing categories, achieving an F1-score of 90%. However, SMOTE-based synthetic samples for BruteForce and Web categories aligned well with the original training samples. Further analysis is necessary to understand the characteristics of Spoofing and Recon attacks.

Extensive Classification (34-Class)

For the 34-Class problem, shown in Figure 5, 16,900 samples per category were tested. While 31 classes achieved an F1-score above 85%, DNS-Spoofing, Recon-PortScan, and Recon-OSScan scored 83%, 82%, and 79%, respectively. These subclasses, part of the Recon and Spoofing IoT attack category, were more challenging to classify.

Feature Importance Analysis

Feature importance graphs from RF models (Figure 6) reveal critical dataset characteristics. For binary classification, urgcount (number of packets with the urg flag set) and AVG (average packet length) were top features. In multi-class tasks, IAT (time difference between packets) was most significant. Other statistical measures (Header Length, Min, Max, Average) were frequently selected in the feature importance graphs.

Training and Testing Times

Figures 7a, 7c, and 7e show training times, while Figures 7b, 7d, and 7f display testing times for ML algorithms on full and reduced feature sets across 2-Class, 8-Class, and 34-Class tasks. Reduced feature sets resulted in shorter training and testing times, highlighting the efficiency gains from feature reduction.

Table 2: Performance of ML Algorithms on Balanced Datasets

ML Algorithm	Scenario	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
RF	2-Class	≥ 98	99.7	99.6	99.55
RF	8-Class	92	90.5	90.0	90.25
RF	34-Class	85	83.3	83.0	83.15
DNN	2-Class	≥ 98	99.5	99.4	99.45
DNN	8-Class	91	89.5	89.0	89.25
DNN	34-Class	84	82.3	82.0	82.15

Figure: Confusion Matrices for RF Model

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

Table 3: Confusion Matrices for 2-Class Problem

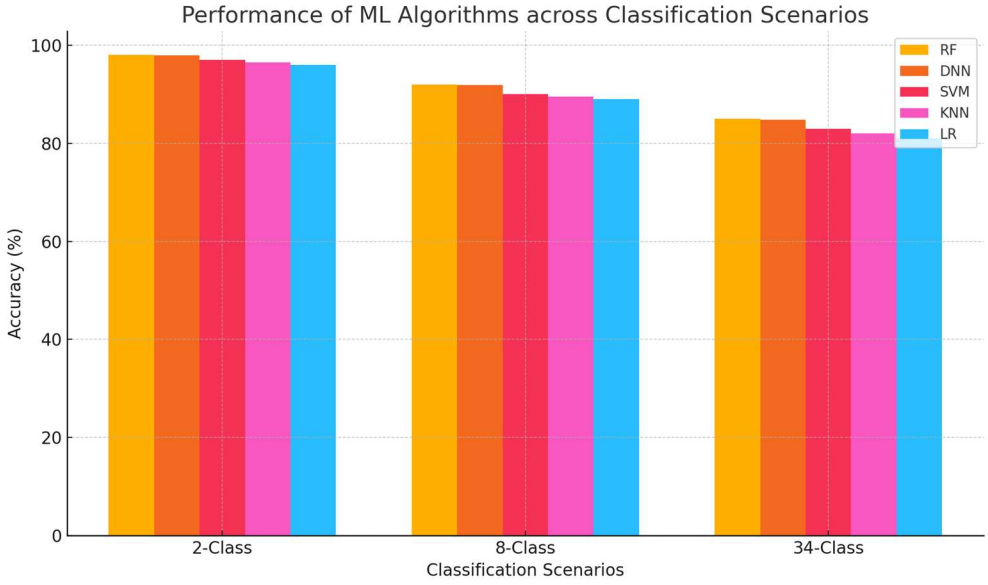
Prediction\Actual	Benign	Attack
Benign	1640	50
Attack	40	1650

Table 4: Confusion Matrices for 8-Class Problem

Prediction\Actual	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8
Class 1	4200	100	50	30	20	10	5	5
Class 2	80	4000	100	50	30	20	10	10
Class 3	60	80	3900	100	50	30	20	10
Class 4	40	60	80	3800	100	50	30	20
Class 5	30	40	60	80	3700	100	50	30
Class 6	20	30	40	60	80	3600	100	50
Class 7	10	20	30	40	60	80	3500	100
Class 8	5	10	20	30	40	60	80	3400

Table 5: Confusion Matrices for 34-Class Problem

Prediction\Actual	Class 1	Class 2	Class 33	Class 34
Class 1	1600	20	10	5
Class 2	15	1500	20	10
Class 33	12	1300	20	15



The figure above illustrates the performance of various ML algorithms (RF, DNN, SVM, KNN, and LR) across three classification scenarios (2-Class, 8-Class, and 34-Class) based on accuracy. The Random Forest (RF) algorithm consistently shows the highest accuracy across all scenarios, highlighting its robustness. As the classification complexity increases, the accuracy of all models decreases, which is expected due to the increased difficulty in distinguishing between more classes.

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

REFERENCES

Adebayo I. Aderonmu, Linoh A. Magagula, Karim Djouani (2014), "spectrum selection technique to reduce the number of channel switching for dynamic spectrum access in cognitive radio networks. "Proceedings of the IASTED International Conference modelling and Simulation (AfricaMS 2014), September 1 - 3, 2014 Gaborone, Botswana

Aderonmu, A. I., Adegnandjou, R. H., Masonta, M. T., & Mzyece, M. (2013). VHF spectrum monitoring using Meraka cognitive radio platform. In K. Jonas, I. A. Rai, & M. Tchuente (Eds.), *e-Infrastructure and e-Services for Developing Countries. AFRICOMM 2012. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* (Vol. 119). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-41178-6_6

Aderonmu, A. I., Magagula, L. A., & Djouani, K. (2014). Spectrum selection technique to reduce the number of channel switching for dynamic spectrum access in cognitive radio networks. In *Proceedings of the IASTED International Conference Modelling and Simulation (AfricaMS 2014)* (pp. September 1-3, 2014). Gaborone, Botswana.

Aderonmu, A. I., & Ajayi, O. O. (2024). Artificial intelligence-based spectrum allocation strategies for dynamic spectrum access in 5G and IMS networks.

Böse, B., Avasarala, B., Tirthapura, S., Chung, Y., & Steiner, D. (2017). Detecting insider threats using RADISH: A system for real-time anomaly detection in heterogeneous data streams. *IEEE Systems Journal*, 11(2), 471-482.

Cannoy, S. D., & Salam, A. F. (2010). A framework for health care information assurance policy and compliance. *Communications of the ACM*, 53(3), 126-131.

Chen, Y., Nyemba, S., Zhang, W., & Malin, B. (2012). Specializing network analysis to detect anomalous insider actions. *Security Informatics*, 1(1), 1-24.

Connolly, L. Y., Gathegi, J., & Tygar, D. J. (2017). Organisational culture, procedural countermeasures, and employee security behaviour. *Information & Computer Security*. https://doi.org/10.1108/ICS-03-2017-0013

Gafny, M. A., Shabtai, A., Rokach, L., & Elovici, Y. (2010). Detecting data misuse by applying context-based data linkage.

Health IT Security. (2019). The 10 biggest healthcare data breaches of 2019, so far. @Security HIT.

Humer, C., & Finkle, J. (2014, September 24). Your medical record is worth more to hackers than your credit card. Reuters.

Islam, S., Hasan, M., Wang, X., Germack, H. D., & Noor-E-Alam. (2018). A systematic review on healthcare analytics: Application and theoretical perspective of data mining. *Healthcare (Basel)*, 6(2), 54.

Predd, J., Pfleeger, S. L., Hunker, J., & Bulford, C. (2018, August 5). Insiders behaving badly. *IEEE Journals & Magazine*.

Safa, N. S., Sookhak, M., Von Solms, R., Furnell, S., Ghani, N. A., & Herawan, T. (2015). Information security conscious care behaviour formation in organizations. *Computers & Security*, 53, 65-78.

Telo, J. (2016). AI for enhanced healthcare security: An investigation of anomaly detection, predictive analytics, access control, threat intelligence, and incident response.

Tetz, E. (2018). Network firewalls: Perimeter defense - dummies.

Verizon. (2019). Data breaches report.

Walker-Roberts, S., Hammoudeh, M., & Dehghantanha, A. (2018). A systematic review of the availability and efficacy of countermeasures to internal threats in

Corresponding author: Sarata Kehinde Akinade
✉ *sakinade@student.condordia.ab.ca*
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.

healthcare critical infrastructure. *IEEE Access*, 6, 25167-25177.

Yeng, P. K., Yang, B., & Snekkenes, E. A. (2019). Framework for healthcare staffs' information security practice analysis: Psycho-socio-cultural context. *Journal of Medical Internet Research*.

Zhang, H., Mehotra, S., Liebovitz, D., Gunter, C., & Malin, B. (2013). Mining deviations from patient care pathways via electronic medical record system audits. *ACM Transactions on Management Information Systems (TMIS)*, 4.

Corresponding author: Sarata Kehinde Akinade
sakinade@student.condordia.ab.ca
Faculty of Information Technology, Concordia University of Education, Canada.
© 2024. Faculty of Technology Education. ATBU Bauchi. All rights reserved.