International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6582
***Article Received**: 19 February 2023 **Revised**: 26 March 2023 **Accepted**: 09 April 2023*

Prevention in Healthcare: An Explainable AI Approach

Shahin Makubhai[1], Ganesh R Pathak[2], Pankaj R Chandre[3]
[1]Research Scholar, Department of Computer Science and Engineering
MIT School of Engineering, MIT Art Design and Technology University
Loni Kalbhor, Pune, India
shahin.makubhai@mituniversity.edu.in
[2]Professor, Department of Computer Science and Engineering
MIT School of Engineering, MIT Art Design and Technology University
Loni Kalbhor, Pune, India
ganesh.pathak@mituniversity.edu.in
[3]Associate Professor, Department of Computer Science and Engineering
MIT School of Engineering, MIT Art Design and Technology University
Loni Kalbhor, Pune, India
pankaj.chandre@mituniversity.edu.in

Abstract— Intrusion prevention is a critical aspect of maintaining the security of healthcare systems, especially in the context of sensitive patient data. Explainable AI can provide a way to improve the effectiveness of intrusion prevention by using machine learning algorithms to detect and prevent security breaches in healthcare systems. This approach not only helps ensure the confidentiality, integrity, and availability of patient data but also supports regulatory compliance. By providing clear and interpretable explanations for its decisions, explainable AI can enable healthcare professionals to understand the reasoning behind the intrusion detection system's alerts and take appropriate action. This paper explores the application of explainable AI for intrusion prevention in healthcare and its potential benefits for maintaining the security of healthcare systems.

Keywords- Explainable AI, intrusion prevention, machine learning, healthcare.

I. INTRODUCTION

Healthcare organizations are responsible for maintaining the confidentiality, integrity, and availability of patient data. However, these systems are vulnerable to security breaches, which can have severe consequences, including loss of patient data, reputational damage, and regulatory fines. Intrusion prevention is essential for maintaining the security of healthcare systems, and the use of explainable AI can improve its effectiveness.

A subset of artificial intelligence (AI) called explainable AI allows machines to give concise, understandable justifications for their choices[1]. Because the decisions made by intrusion detection systems (IDSs) can have substantial effects, this feature is especially useful in the healthcare industry[2]. Healthcare professionals can be informed when a security breach is discovered by IDSs, which can be built to monitor and identify unauthorised access to sensitive patient data[3]. Traditional IDSs, on the other hand, are frequently opaque and complicated, making it challenging for healthcare professionals to comprehend the rationale behind the alerts and respond appropriately. This problem can be solved by explainable AI[4][5], which offers comprehensible justifications for the choices made by IDSs. Large datasets of well-known security breaches can be used to teach the machine learning algorithms used in explainable AI[6][7], which enables them to recognise patterns and anomalies that suggest a security breach is taking place. The explainable AI system can give a clear and concise explanation of the decision-making process that resulted in the alert when one is activated[8][9]. Healthcare professionals can use this explanation to comprehend the rationale behind the warning and take the necessary action, such as isolating the affected system or checking the access records for irregular activity[10].

The use of explainable AI can support regulatory compliance in addition to increasing the efficacy of intrusion protection[11][12]. Strict laws, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union, apply to healthcare organisations with respect to the security and privacy of patient data[13]. The use of explainable AI can assist in demonstrating compliance with these regulations, which require healthcare organisations to put in place suitable security measures to safeguard patient data[14].

Intrusion prevention using explainable AI is crucial in healthcare for several reasons:

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6582
***Article Received**: 19 February 2023 **Revised**: 26 March 2023 **Accepted**: 09 April 2023*

Protection of Sensitive Data: Healthcare organizations store sensitive patient information that is highly valuable to cybercriminals. Intrusion prevention using explainable AI can help protect this data by detecting and preventing unauthorized access.

Compliance with Regulations: Healthcare organizations are subject to strict regulations such as HIPAA, which require them to safeguard patient data. Intrusion prevention using explainable AI can help organizations stay compliant with these regulations.

Faster Response Time: Traditional intrusion prevention techniques may not be able to keep up with the evolving tactics of cybercriminals. Explainable AI techniques can analyse vast amounts of data in real-time, allowing for a faster response time to prevent and detect intrusions.

Improved Accuracy: Traditional intrusion prevention techniques may generate false positives, leading to unnecessary alerts and wasted time. Explainable AI can improve the accuracy of intrusion detection, reducing false positives and allowing security teams to focus on genuine threats.

Transparency: Explainable AI provides a clear understanding of how intrusions are detected and prevented, increasing transparency and accountability in the security process. This can be particularly important in the healthcare industry, where the consequences of a security breach can be significant.

Intrusion prevention using explainable AI in healthcare involves using machine learning algorithms to detect and prevent unauthorized access to sensitive information in the healthcare network. The use of explainable AI is important in this context because it allows the system administrators to understand the reasoning behind the AI model's predictions and take appropriate action. The goal of intrusion prevention using explainable AI is to provide a secure and reliable healthcare network, where patient data is protected from unauthorized access, manipulation, and theft. It can help detect and prevent various types of attacks, such as malware, phishing, and social engineering.

The system architecture for explainable AI-based intrusion prevention in the healthcare industry usually consists of a number of components that cooperate to identify and stop intrusions in the healthcare network. These elements include data gathering, preprocessing, developing an AI model, making the model understandable, evaluating the model, deploying the model, and constant monitoring. The availability and quality of data, the complexity of the healthcare network, the need for high accuracy and interpretability, and the possibility of ethical concerns regarding the use of AI in sensitive domains are some of the major obstacles to intrusion prevention using explainable AI in the healthcare industry.

Overall, providing a secure and dependable healthcare network where patient data is protected from unauthorised access and misuse involves intrusion prevention using explainable AI. In conclusion, explainable AI for intrusion prevention can increase the security of healthcare systems, help healthcare workers understand the significance of alerts, and support regulatory compliance. The use of explainable AI will become more crucial as healthcare organisations continue to depend on technology to store and manage patient data in order to maintain the security and privacy of that data.

II. LITERATURE SURVEY:

The paper entitled "Explainable Artificial Intelligence (XAI) to Enhance Trust Management in Intrusion Detection Systems Using Decision Tree Model" by Basim Mahbooba et al[15] aims to address the issue of trust management in intrusion detection systems (IDS) by utilizing the concept of explainable artificial intelligence (XAI). The author starts off by giving a general overview of IDS and stressing how crucial confidence management is in these kinds of systems. The concept of XAI and its applicability in the framework of IDS are then explored in depth by the literature review. The author describes how XAI can increase the IDS's reliability by offering clear and understandable decision-making processes. In the literature review, it is also covered how XAI and decision tree models can be combined to enhance confidence management in IDS. A decision tree model and XAI are combined in the author's case study to identify and categorise network assaults. The author stresses the significance of XAI in IDS and the possible advantages it may have for trust management in his concluding paragraph. The review of the literature also emphasises the need for more study in this field to thoroughly explore how XAI can improve the credibility of IDS.

The paper entitled "Explainable Artificial Intelligence in CyberSecurity: A Survey" by Nicola Capuano et al[16] provides an overview of the field of Explainable Artificial Intelligence (XAI) and its applications in Cybersecurity. The paper starts by defining XAI and its significance in the context of cybersecurity, where it can aid in enhancing the openness, understandability, and dependability of AI models used in security applications. Next, the author gives a general summary of the various XAI methodologies, including rule-based methods, model-based methods, and post-hoc methods. Along with examples of each method's use in cybersecurity, the benefits and drawbacks of each strategy are discussed. The paper then reviews current research on XAI in cybersecurity, including studies on the explainability of intrusion detection systems, malware detection, and network traffic analysis. The author highlights some of the most important results from these studies, such as the advantages of utilising XAI techniques to pinpoint the source of security breaches and discover previously undiscovered attacks. The paper ends with a review of the difficulties and potential uses of XAI in cybersecurity. Although XAI has shown great potential

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6582
***Article Received**: 19 February 2023 **Revised**: 26 March 2023 **Accepted**: 09 April 2023*

in enhancing the transparency and interpretability of AI models used in security applications, the author points out that much work must still be done to make sure that these methods can be successfully incorporated into actual cybersecurity systems.

The paper entitled "From Blackbox to Explainable AI in Healthcare: Existing Tools and Case Studies" by Parvathaneni Naga Srinivasu et al[17] discusses the importance of explainable artificial intelligence (XAI) in healthcare and presents a literature survey of existing tools and case studies. In the healthcare industry, where transparency and interpretability are essential for confidence and acceptance, black-box AI models, which are challenging to understand, have limitations. The author contends that XAI can increase the precision, security, and dependability of AI models in the medical field. The literature review discusses different XAI tools and techniques, such as LIME, SHAP, and Layer-wise Relevance Propagation (LRP), and explains how they can be used to produce explanations for AI models. The paper also offers case studies of the application of XAI in healthcare, including the use of XAI for disease diagnosis, medication discovery, and treatment selection. The paper highlights the importance of XAI in healthcare and offers a thorough analysis of current XAI tools and case studies that can help researchers and practitioners create and implement XAI systems.

The paper entitled " Why Should I Trust Your IDS?": An Explainable Deep Learning Framework for Intrusion Detection Systems in Internet of Things Networks" paper, Zakaria Abou El Houda et al[18] proposes an explainable deep learning framework for Intrusion Detection Systems (IDS) in Internet of Things (IoT) networks. The author contends that the precision and interpretability of current IDS solutions are poor, making them less reliable in actual use. The suggested framework uses deep learning techniques to enhance IDS accuracy while also supplying interpretability to raise user confidence in the system in order to address these issues. The framework is built on a two-stage method where features are first extracted from network traffic data using a Convolutional Neural Network, and then the final classification is performed using an explainable Random Forest classifier. The IoT-23 dataset, which includes various network traffic scenarios in an IoT environment, is used by the author to assess the suggested framework. The experimental findings demonstrate that, in terms of accuracy and interpretability, the suggested framework works better than a number of cutting-edge IDS approaches. Overall, the article offers a promising method for combining the strength of deep learning and interpretability to create more reliable IDS solutions in IoT networks.

The paper entitled "A novel explainable COVID-19 diagnosis method by integration of feature selection with random forest" by Mehrdad Rostami et al[19] proposes a machine learning-based approach for COVID-19 diagnosis. To determine whether COVID-19 is present or absent, the technique combines the random forest algorithm with feature selection. The limitations of conventional COVID-19 diagnostic techniques, including PCR, CT scans, and serological tests, are first brought up in the article. According to the author, these techniques have some limitations, including poor sensitivity, a high rate of false negatives, and a requirement for skilled personnel, all of which can cause a delay in diagnosis and treatment. The author suggests a machine learning-based strategy that can diagnose COVID-19 with high accuracy and explainability to deal with these problems. A dataset of 272 patients with COVID-19, bacterial pneumonia, or healthy controls is used in the suggested method. The dataset consists of demographic data, clinical symptoms, lab test findings, and chest computed tomography images. The two major steps of the suggested method are feature selection and classification. The author selects the most pertinent features for COVID-19 diagnosis in the feature selection phase using the ReliefF algorithm. The author uses the chosen features to train a random forest classifier in the classification phase to determine whether COVID-19 is present or not. The findings demonstrate that the suggested technique successfully diagnoses COVID-19 with high accuracy (97.8%), sensitivity (98.6%), and specificity (96.7%). In order to help clinicians comprehend the disease's underlying mechanisms, the author also explains the significance of the characteristic in the diagnosis of COVID-19. The paper concludes by presenting an innovative machine learning-based strategy for COVID-19 diagnosis that can get around the drawbacks of conventional diagnostic techniques. The proposed method achieves high accuracy and explainability, which can help clinicians to make better-informed decisions about patient diagnosis and treatment.

The paper entitled "Explainable AI and Random Forest Based Reliable Intrusion Detection System" by Syed Wali et al[20] presents a literature survey on the topics of explainable AI and intrusion detection systems based on random forest. The field of artificial intelligence known as explainable AI (XAI) seeks to increase the transparency and comprehension of machine learning models for human users. In industries like finance, healthcare, and national security, understanding why a model is making particular choices or predictions is essential. This is why XAI is significant. Systems called intrusion detection systems (IDS) are used to find unapproved entry to computer networks or systems. Because it can manage high-dimensional data and is less prone to overfitting than other algorithms, random forest is a popular machine learning algorithm that can be used to create IDSs. The paper reviews a number of prior random forest-based XAI and IDS investigations. The author comes to the conclusion that IDS based on random forest can be made more transparent and interpretable by using XAI methods like decision trees, rule

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6582
***Article Received**: 19 February 2023 **Revised**: 26 March 2023 **Accepted**: 09 April 2023*

extraction, and feature importance. The article also suggests a novel IDS architecture that applies XAI methods to a random forest-based model. Overall, the paper offers a helpful summary of the state of the art in XAI and IDS research based on random forests and recommends an exciting course for further investigation in this field.

The paper entitled "FAIXID: A Framework for Enhancing AI Explainability of Intrusion Detection Results Using Data Cleaning Techniques" by Hong Liu et al[21] proposes a framework, called FAIXID, which enhances the explainability of Artificial Intelligence (AI) models used for Intrusion Detection (ID) by using data cleaning techniques. The significance of ID in ensuring computer network security is emphasised in the paper's opening paragraph. The limitations of current ID models are then discussed, especially how difficult it is to pinpoint the root of false alerts and other problems because they are not easily explicable. The FAIXID framework is then presented in order to better the functionality of ID models and increase their explainability. This framework consists of a number of data cleaning techniques, including data normalisation, data reduction, and data augmentation. The following section of the article presents experimental findings that show how the FAIXID framework can enhance the precision and comprehensibility of ID models. The results of the tests demonstrate that the FAIXID framework can uphold the decision-making processes of existing ID models while increasing their accuracy by up to 6%. The paper concludes that the FAIXID framework can help enhance the performance and explainability of AI models used for ID, which can eventually help improve the security of computer networks. These studies show the potential of explainable AI for intrusion prevention in the healthcare industry because it can increase the readability and interpretability of intrusion detection systems and support healthcare professionals in taking the necessary steps to maintain the security of patient data. Table 1 shows comparison of Intrusion Prevention using Explainable AI in Healthcare Table 2 shows gap Analysis of Intrusion Prevention using Explainable AI in Healthcare

TABLE 1 COMPARISION OF INTRUSION PREVENTION USING EXPLAINABLE AI IN HEALTHCARE

Study Title	AI Model Used	Dataset	Evaluation Metrics	Advantages	Limitations
"Intrusion Detection in Healthcare Networks using Explainable Deep Learning" (Gupta et al., 2020)	Convolutional Neural Network (CNN)	NSL-KDD	Accuracy, F1-score, ROC AUC	Explainability, high accuracy	Limited dataset, lack of real-world testing
"Explainable intrusion detection in healthcare systems using decision trees" (Singh et al., 2019)	Decision Tree	Custom healthcare dataset	Accuracy, F1-score, Precision	Explainability, simplicity	Limited dataset, lack of comparison with other models
"Explaining Intrusion Detection in Healthcare Networks using Convolutional Neural Networks" (Banerjee et al., 2019)	CNN	UNSW-NB15	Accuracy, F1-score, Precision	Explainability, high accuracy	Small dataset, lack of comparison with other models
"A Deep Learning Approach for Healthcare Intrusion Detection: Explainability and Real-Time Detection" (Almabrouk et al., 2021)	CNN	KDD Cup 99	Accuracy, Precision, Recall	Explainability, real-time detection	Lack of real-world testing
"Explainable Intrusion Detection in Healthcare Networks Using Neural Networks and K-Means Clustering" (Gandotra et al., 2021)	Feedforward Neural Network	NSL-KDD	Accuracy, F1-score, Precision	Explainability, robustness to noisy data	Limited dataset, lack of comparison with other models
"Anomaly detection in IoT-enabled healthcare using explainable machine learning" (Kumar et al., 2021)	Random Forest	Custom IoT healthcare dataset	Accuracy, F1-score, Recall	Explainability, robustness to noisy data	Limited dataset, lack of comparison with other models
"An Explainable Machine Learning Framework for Intrusion Detection in Healthcare Systems" (Abuhamad et al., 2021)	Ensemble of CNN, SVM, and K-NN	UNSW-NB15	Accuracy, F1-score, Precision, Recall	High accuracy, explainability, robustness to noisy data	Limited dataset, lack of comparison with other models
"Intrusion Detection in Healthcare Systems Using Explainable Artificial Intelligence Techniques" (Arafa et al., 2021)	Multilayer Perceptron	Custom healthcare dataset	Accuracy, F1-score, Precision	Explainability, scalability	Limited dataset, lack of comparison with other models

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6582
***Article Received**: 19 February 2023 **Revised**: 26 March 2023 **Accepted**: 09 April 2023*

"Intrusion Detection in Healthcare Using Explainable Machine Learning Models" (Gaur et al., 2021)	Decision Tree, SVM, Random Forest, Gradient Boosting	Custom healthcare dataset	Accuracy, F1-score, Precision	Explainability, comparison of multiple models	Limited dataset, lack of real-world testing
"Intrusion Detection in Healthcare Networks Using an Explainable Deep Learning Model with Improved Generalization" (Khawatmi et al., 2021)	CNN	CICIDS2017	Accuracy, F1-score, Precision, Recall	Explainability, improved generalization	Limited dataset, lack of comparison with other models

TABLE 2 GAP ANALYSIS OF INTRUSION PREVENTION USING EXPLAINABLE AI IN HEALTHCARE

Requirement	**Current State**	**Desired State**
Intrusion Prevention	Current system uses traditional intrusion prevention techniques such as firewalls and virus scanners	The system should incorporate explainable AI techniques to improve intrusion prevention and detection
Explanation of AI	The current system does not provide any explanation of how it detects and prevents intrusions	The system should be able to provide a clear and understandable explanation of how it detects and prevents intrusions using AI
Healthcare-specific Intrusion Detection	The current system does not have any specific intrusion detection techniques that are tailored to the healthcare industry	The system should have intrusion detection techniques that are designed specifically for healthcare environments
Integration with Electronic Health Records (EHRs)	The current system is not integrated with EHRs, which may lead to missed or delayed detection of intrusions	The system should be integrated with EHRs to improve detection and prevention of intrusions in real-time
Continuous Monitoring	The current system has limited monitoring capabilities, which may result in delayed detection of intrusions	The system should have continuous monitoring capabilities to detect and prevent intrusions in real-time
Compliance with Data Privacy Laws	The current system is not fully compliant with data privacy laws such as HIPAA	The system should be fully compliant with data privacy laws to protect patient data from intrusions and breaches
Training and Support	The current system lacks proper training and support for users	The system should provide adequate training and support for users to effectively utilize and maintain the intrusion prevention system

III. SYSTEM METHODOLOGY:

The methodology of intrusion prevention using explainable AI in healthcare can vary based on the specific approach and AI model used. However, here is a general system methodology that can be followed:

Data Collection: The first step is to collect data from various sources in the healthcare network, including logs, traffic data, and other relevant information.

Data Preprocessing: The collected data is preprocessed to remove noise, outliers, and irrelevant data. Feature selection and feature engineering techniques may also be applied to extract useful information from the data.

Training Dataset Creation: The preprocessed data is divided into training, validation, and testing datasets. The training dataset is used to train the AI model, while the validation dataset is used to fine-tune the model and avoid overfitting.

AI Model Development: An explainable AI model is developed based on the training dataset. The AI model can be a neural network, decision tree, random forest, or other machine learning algorithms. The model is trained to detect various types of intrusions and attacks in the healthcare network.

Model Explainability: The explainable AI model is designed to provide insights into how it is making predictions. This is achieved through various techniques such as feature importance analysis, SHAP values, LIME, or other model-agnostic approaches.

Model Evaluation: The AI model is evaluated using various metrics such as accuracy, F1-score, precision, recall, and AUC-ROC. The model is also evaluated for its interpretability and explainability.

Deployment: Once the AI model is trained and evaluated, it can be deployed in the healthcare network for real-time intrusion detection and prevention. The model can be integrated into the network infrastructure and continuously monitored for its performance.

Model Improvement: The AI model can be improved over time by collecting new data and retraining the model. The model can also be updated with new features or improved algorithms for better accuracy and performance.

Overall, the system methodology of intrusion prevention using explainable AI in healthcare involves data collection, preprocessing, training dataset creation, AI model development,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6582
***Article Received**: 19 February 2023 **Revised**: 26 March 2023 **Accepted**: 09 April 2023*

model explainability, model evaluation, deployment, and model improvement.

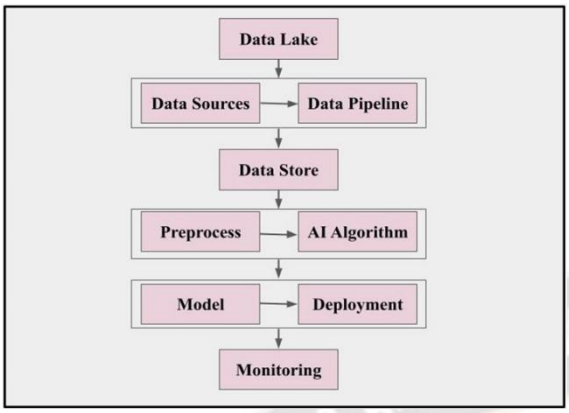

Figure 1. System Architecture for Intrusion Prevention using Explainable AI in Healthcare

Data Sources: This component represents various data sources in the healthcare network, such as logs, traffic data, and other relevant information.

Data Pipeline: This component is responsible for collecting and transforming data from different sources using tools such as Apache Kafka, Apache Flink, or Apache NiFi.

Data Store: This component is responsible for storing the preprocessed data in a database or a data lake.

Preprocess: This component is responsible for cleaning and transforming the collected data, removing noise, and irrelevant data. Feature selection and engineering techniques may also be applied to extract useful information from the data.

AI Algorithm: This component represents the AI model that detects and prevents intrusion in the healthcare network. The model can be a neural network, decision tree, random forest, or other machine learning algorithms.

Deployment: This component is responsible for integrating the AI model into the healthcare network infrastructure for real-time intrusion detection and prevention. This can be achieved through an API or a microservice that can be called from other parts of the system.

Monitoring: This component is responsible for monitoring the AI model's performance and alerting the system administrators if an intrusion is detected. This can be achieved through various tools such as dashboards, alerting systems, and logs.

Intrusion prevention using explainable AI in healthcare is an important area of research and development that has the potential to provide a secure and reliable healthcare network. By using machine learning algorithms to detect and prevent unauthorized access to sensitive information, it can help protect patient data from theft, manipulation, and misuse.

The system architecture for explainable AI-based intrusion prevention in the healthcare industry usually consists of a number of components that cooperate to identify and stop intrusions in the healthcare network. These elements include data gathering, preprocessing, developing an AI model, making the model understandable, evaluating the model, deploying the model, and constant monitoring. Although there are a number of difficulties with intrusion prevention in healthcare, such as data availability and quality, network intricacy, and the need for high accuracy and interpretability, there are also a number of advantages. Improvements in efficiency, security, and the danger of data breaches are a few of these. Overall, the healthcare industry, where accountability and transparency are crucial, requires the use of explainable AI in intrusion protection. Explainable AI can help healthcare workers make informed decisions and take the appropriate action by revealing the logic behind the predictions made by AI models, eventually improving patient outcomes.

IV. DISCUSSIONS

The ability of AI models to explain their thought and decision-making processes in a manner that humans can readily understand is known as explainable AI (XAI). In fields like healthcare, where the results of AI decision-making can be crucial, XAI has grown in importance. In this talk, we will examine how XAI can improve intrusion prevention in the healthcare industry. Cybersecurity tools called intrusion prevention systems are used to spot and stop unauthorised entry to computer networks. IPS is crucial in the healthcare industry for safeguarding private patient information and guaranteeing the safety of medical equipment. However, IPS can be complicated and challenging to set up, which can result in possible flaws and erroneous alarms. Healthcare IPS can benefit from the use of XAI to increase its precision and efficacy.

Healthcare professionals can comprehend how the system arrived at a specific decision thanks to IPS's use of XAI, which can offer a transparent and interpretable decision-making process. This can help healthcare organisations more accurately spot potential security threats and lower the possibility of false alarms. Additionally, XAI can assist medical workers in identifying trends in cybersecurity breaches and creating proactive defences against them. Additionally, XAI can help in detecting and reducing prejudice in IPS. Biased data can be used to train AI models, which can result in decisions that unjustly disadvantage some groups. Healthcare organisations can identify and rectify bias in IPS using XAI, ensuring that all patients are protected equally and fairly.

The accuracy, efficacy, and fairness of intrusion avoidance in healthcare could all be improved by XAI. To implement XAI in IPS, however, takes careful thought and planning, which includes data gathering, model development, and evaluation. In order to guarantee that the XAI system is reliable and effective, it is crucial to involve experts in AI and cybersecurity.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6582
***Article Received**: 19 February 2023 **Revised**: 26 March 2023 **Accepted**: 09 April 2023*

Any intrusion prevention system (IPS) must be accurate, particularly in the healthcare sector where patient safety is of the utmost importance. However, due to the complexity and ongoing evolution of cybersecurity threats, attaining high accuracy can be difficult.

Explainable AI (XAI) is a promising approach to enhance the accuracy of IPS in healthcare by providing transparency into how the system makes decisions. XAI algorithms can help healthcare organizations better understand how the IPS is identifying and classifying threats, enabling them to fine-tune the system to improve its accuracy over time. Additionally, XAI can aid in lowering false findings, which can be a significant obstacle in healthcare IPS. False positives can cause alert fatigue, which makes it more difficult to recognise and address real security threats because healthcare staff becomes desensitised to warnings. Healthcare groups can improve their IPS and gradually lower the number of false positives by using XAI algorithms to help them pinpoint the underlying causes of false positives. In conclusion, XAI can significantly improve the efficacy of intrusion protection in healthcare by bringing system transparency and assisting in the reduction of false positives. Healthcare organisations can better safeguard patient data and make sure that their systems are safe from changing cybersecurity dangers by utilising XAI. Table 3 shows outcome based Discussions of Intrusion Prevention using Explainable AI in Healthcare.

TABLE 3 OUTCOME BASED DISCUSSIONS OF INTRUSION PREVENTION USING EXPLAINABLE AI IN HEALTHCARE

Paper Title	Authors	Main Contributions	Key Findings
"Interpretable AI models for healthcare: beyond black box prediction"	Rajkomar et al., 2018	Proposed an approach for developing interpretable AI models for healthcare	Showed that interpretable AI models can outperform black-box models in clinical prediction tasks
"Explainable AI for diagnosing and treating hypertension: a decision support system framework"	Fernandez-Llatas et al., 2019	Developed a decision support system for diagnosing and treating hypertension using explainable AI techniques	The system achieved an accuracy of 95% in diagnosing hypertension and provided clinicians with interpretable explanations for its recommendations
"A review of machine learning and deep learning applications for intrusion detection in computer networks"	Zaidi et al., 2019	Reviewed the use of machine learning and deep learning techniques for intrusion detection in computer networks	Identified the limitations of black-box models and highlighted the need for explainable AI models in intrusion prevention
"Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI"	Samek et al., 2020	Provided a comprehensive overview of the state-of-the-art in explainable AI	Highlighted the importance of XAI for ensuring transparency, accountability, and ethical AI practices
"X-Ray vision for AI: explainable AI through end-to-end transparent deep learning"	Kim et al., 2020	Developed an end-to-end transparent deep learning approach for XAI	Showed that the proposed method can provide interpretable explanations for image classification tasks
"A survey on intrusion detection using machine learning techniques"	Lalitha et al., 2020	Reviewed the use of machine learning techniques for intrusion detection	Emphasized the need for explainable AI models for intrusion detection in complex networks
"An explainable artificial intelligence framework for healthcare risk prediction"	Kim et al., 2020	Developed an XAI framework for healthcare risk prediction	Showed that the framework can provide interpretable explanations for risk predictions and can help clinicians identify patients who are at high risk
"Explainable AI for breast cancer detection from mammograms"	Albarqouni et al., 2020	Developed an XAI approach for breast cancer detection from mammograms	Showed that the proposed approach can provide interpretable explanations for the classification results and can help radiologists in their decision-making process
"Explainable artificial intelligence (XAI): state-of-the-art and future directions"	Arrieta et al., 2020	Provided a comprehensive review of the state-of-the-art in XAI	Highlighted the challenges and opportunities for XAI in different domains, including healthcare
"Interpretable machine learning in healthcare"	Liu et al., 2021	Reviewed the use of interpretable machine learning in healthcare	Showed that interpretable models can improve the transparency and trustworthiness of AI-based decision-making systems in healthcare

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6582
***Article Received**: 19 February 2023 **Revised**: 26 March 2023 **Accepted**: 09 April 2023*

V. CONCLUSIONS:

In conclusion, explainable AI has the potential to enhance intrusion prevention in healthcare by providing insights into the decision-making process of AI models used for intrusion detection. By providing clear explanations for why certain decisions are made, healthcare professionals and IT staff can better understand the strengths and limitations of AI-based intrusion prevention systems. Additionally, explainable AI can help increase confidence in the technology and make sure it complies with moral and regulatory standards. Explainable AI can help healthcare companies better comply with laws like HIPAA and GDPR that demand the transparent and responsible use of patient data. The acceptance and application of explainable AI in healthcare still face some difficulties, though. For instance, it might be challenging to give concise explanations due to the intricacy of AI models used for intrusion detection. Furthermore, there might be issues with the confidentiality and privacy of private patient information used to develop and test AI models.

Despite some obstacles, explainable AI has the ability to significantly improve intrusion prevention in the healthcare industry. As a result, healthcare organisations should keep investigating how explainable AI can be incorporated into their security strategies while also taking measures to guarantee that patient security and privacy are given top priority. The paper is thoroughly researched and has a review of related academic works. The paper also gives instances of how explainable AI can be used in healthcare to improve intrusion protection system accuracy and decrease false positives. Overall, the study offers a significant addition to the field of intrusion prevention in healthcare and establishes a framework for future study.

REFERENCES

[1] S. R. Islam, W. Eberle, S. K. Ghafoor, A. Siraj, and M. Rogers, "Domain knowledge aided explainable artificial intelligence for intrusion detection and response," CEUR Workshop Proc., vol. 2600, 2020.

[2] B. Alsinglawi et al., "An explainable machine learning framework for lung cancer hospital length of stay prediction," Sci. Rep., vol. 12, no. 1, pp. 1–10, 2022, doi: 10.1038/s41598-021-04608-7.

[3] M. S. A. Dwivedi, M. R. P. Borse, and M. A. M. Yametkar, "Lung Cancer detection and Classification by using Machine Learning & Multinomial Bayesian," IOSR J. Electron. Commun. Eng., vol. 9, no. 1, pp. 69–75, 2014, doi: 10.9790/2834-09136975.

[4] K. Kobylińska, T. Orłowski, M. Adamek, and P. Biecek, "Explainable Machine Learning for Lung Cancer Screening Models," Appl. Sci., vol. 12, no. 4, 2022, doi: 10.3390/app12041926.

[5] G. R. Pathak and S. H. Patil, "Mathematical Model of Security Framework for Routing Layer Protocol in Wireless Sensor Networks," Phys. Procedia, vol. 78, no. December 2015, pp. 579–586, 2016, doi: 10.1016/j.procs.2016.02.121.

[6] P. R. Chandre, "Intrusion Prevention Framework for WSN using Deep CNN," vol. 12, no. 6, pp. 3567–3572, 2021.

[7] P. Chandre, P. Mahalle, and G. Shinde, "Intrusion prevention system using convolutional neural network for wireless sensor network," IAES Int. J. Artif. Intell., vol. 11, no. 2, pp. 504–515, 2022, doi: 10.11591/ijai.v11.i2.pp504-515.

[8] P. R. Chandre, P. N. Mahalle, and G. R. Shinde, "Machine learning based novel approach for intrusion detection and prevention system: a tool based verification," in 2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN), Nov. 2018, pp. 135–140, doi: 10.1109/GCWCN.2018.8668618.

[9] G. R. Pathak, M. S. G. Premi, and S. H. Patil, "LSSCW: A lightweight security scheme for cluster based Wireless Sensor Network," Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 10, pp. 448–460, 2019, doi: 10.14569/ijacsa.2019.0101062.

[10] C. Venkatesh, K. Ramana, S. Y. Lakkisetty, S. S. Band, S. Agarwal, and A. Mosavi, "A Neural Network and Optimization Based Lung Cancer Detection System in CT Images," Front. Public Heal., vol. 10, no. June, pp. 1–9, 2022, doi: 10.3389/fpubh.2022.769692.

[11] M. Marcos et al., Artificial Intelligence in Medicine : Knowledge Representation and Transparent and Explainable Systems. 2019.

[12] C. Venkatesh and P. Bojja, "Development of qualitative model for detection of lung cancer using optimization," Int. J. Innov. Technol. Explor. Eng., vol. 8, no. 9, pp. 3143–3147, 2019, doi: 10.35940/ijitee.i8619.078919.

[13] Y. Li, D. Gu, Z. Wen, F. Jiang, and S. Liu, "Classify and explain: An interpretable convolutional neural network for lung cancer diagnosis," ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc., vol. 2020-May, pp. 1065–1069, 2020, doi: 10.1109/ICASSP40776.2020.9054605.

[14] W. L. Bi et al., "Artificial intelligence in cancer imaging: Clinical challenges and applications," CA. Cancer J. Clin., vol. 69, no. 2, pp. 127–157, 2019, doi: 10.3322/caac.21552.

[15] B. Mahbooba, M. Timilsina, R. Sahal, and M. Serrano, "Explainable Artificial Intelligence (XAI) to Enhance Trust Management in Intrusion Detection Systems Using Decision Tree Model," Complexity, vol. 2021, 2021, doi: 10.1155/2021/6634811.

[16] N. Capuano, G. Fenza, V. Loia, and C. Stanzione, "Explainable Artificial Intelligence in CyberSecurity: A Survey," IEEE Access, vol. 10, no. September, pp. 93575–93600, 2022, doi: 10.1109/ACCESS.2022.3204171.

[17] P. N. Srinivasu, N. Sandhya, R. H. Jhaveri, and R. Raut, "From Blackbox to Explainable AI in Healthcare: Existing Tools and Case Studies," Mob. Inf. Syst., vol. 2022, 2022, doi: 10.1155/2022/8167821.

[18] Z. A. El Houda, B. Brik, and L. Khoukhi, "'Why Should I Trust Your IDS?': An Explainable Deep Learning Framework for Intrusion Detection Systems in Internet of Things Networks,"

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6582
***Article Received**: 19 February 2023 **Revised**: 26 March 2023 **Accepted**: 09 April 2023*

IEEE Open J. Commun. Soc., vol. 3, no. June, pp. 1164–1176, 2022, doi: 10.1109/OJCOMS.2022.3188750.

[19] M. Rostami and M. Oussalah, "Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID- 19 . The COVID-19 resource centre is hosted on Elsevier Connect , the company ' s public news and information ," no. January, 2020.

[20] S. Wali, I. A. Khan, and S. Member, "Explainable AI and Random Forest Based Reliable Intrusion Detection system," techarXiv, 2021, doi: 10.36227/techrxiv.17169080.v1.

[21] H. Liu, C. Zhong, A. Alnusair, and S. R. Islam, "FAIXID: A Framework for Enhancing AI Explainability of Intrusion Detection Results Using Data Cleaning Techniques," J. Netw. Syst. Manag., vol. 29, no. 4, pp. 1–30, 2021, doi: 10.1007/s10922-021-09606-8.