

An Explainable AI and Cloud Computing Based Framework for Real-Time Cyber Attack Detection, Risk Assessment, and Threat Mitigation

Thotakura Lohit Kumar¹, Dr PVRD Prasada Rao²

¹Research Scholar, Koneru Lakshmaiah Education Foundation, Department of Computer Science and Engineering, Vaddeswaram, Andhra Pradesh, India.

²Professor Koneru Lakshmaiah Education Foundation, Department of Computer Science and Engineering, Vaddeswaram, Andhra Pradesh, India.

Email: 2401050160@kluniversity.in

ABSTRACT

Cloud-based environments now face a huge security challenge because of the growing sophistication of the cyberattacks and the frequency they are occurring. Conventional IDS systems are not very efficient at correctly recognizing new attack patterns, and typically don't provide visibility into their decision-making. To overcome these limitations, in this work a Novel Explainable AI and Cloud Computing Based Framework for Real-Time Cyber Attack Detection, Risk Assessment, and Threat Mitigation is proposed. The framework uses the UNSW-NB15 network traffic data set and utilizes the XGBoost machine learning algorithm to classify the network traffic as normal or malicious. The data was preprocessed using feature encoding and traffic analysis techniques, which are commonly used in machine learning for data preparation. To improve interpretability, explainable artificial intelligence (XAI) techniques were used, such as feature importance analysis, to understand the factors behind the decision to detect an attack, making it easier for security analysts to interpret the information. Furthermore, a system of risk assessment was established, which classifies the threats as detected into different categories of severity, therefore aiding to efficient prioritisation and mitigating measures. The results of the experiments conducted revealed that the proposed framework attained an accuracy of 95.28% along with high precision, recall and F1 score values that indicate high detection ability and reliability. The framework also offered detailed visualizations with analyses on the distribution of attacks, the importance of features, the risk levels, and metrics to evaluate performance. The proposed approach is designed to create intelligent, transparent, and scalable cybersecurity solutions in the modern cloud environment, where the integration of accurate attack detection, explainability and risk-aware decision support are all held in one place.

Keywords: Cyber Attack Detection, Explainable Artificial Intelligence (XAI), Cloud Computing Security, XGBoost, Risk Assessment, Intrusion Detection System (IDS).

How to cite this article: Thotakura Lohit Kumar¹ Dr PVRD Prasada Rao². An Explainable AI and Cloud Computing Based Framework for Real-Time Cyber Attack Detection, Risk Assessment, and Threat Mitigation. Int J Drug Deliv Technol. 2026;16(86s):431-440. DOI: 10.25258/ijddt.16.86s.48.

1. INTRODUCTION

Information sharing, data processing and storage have been radically transformed in organizations by cloud computing, digital communication technologies and networked information systems [15,16]. The development of these technologies has contributed to the more efficient, scalable, and user-friendly operation of these systems, but also opened up a substantial attack surface for the sophisticated cyber threats to which these systems are vulnerable [19]. The growing sophistication of the modern attacks such as Distributed Denial of Service (DDoS), malware injection, reconnaissance attacks, exploits, backdoor intrusions and unauthorised access attempts make it less possible to ensure complete protection with traditional rule based security mechanisms [1,7]. Traditional IDSs have limitations on their ability to detect new attack patterns because of the need for pre-defined signatures and lack of flexibility when dealing with dynamic network environments [4,7]. Consequently, the cybersecurity industry has turned towards the use of Artificial Intelligence (AI) and Machine Learning (ML) techniques to assist in the

development of intelligent threat detection mechanisms that have the ability to learn how to detect attack behavior using massive amounts of information about network traffic [2,9,10]. While machine learning-based intrusion detection systems have proven to be successful, there is a major drawback to the inability to understand the outcome of the prediction [4,11]. There are many advanced models that are "black-box" systems that require a level of effort for security analysts to try to understand how attacks are being classified [4]. This constraint decreases the level of trust in the automated security solutions and makes their use in critical mission cloud environment challenging [6]. To overcome this challenge, Explainable Artificial Intelligence (XAI) has become an emerging approach to offer interpretable information on machine learning predictions that allows cybersecurity professionals to understand the factors that affect decisions on attack detection [4]. Moreover, cloud computing contexts require accuracy and scalability of security tools, which must be able to adapt in real-time to the analysis of

threats on distributed infrastructures [12]. The present study aims to present a framework for real-time cyber attack detection, risk assessment and threat mitigation for the Explainable AI and Cloud Computing. The proposed framework combines the machine learning techniques, risk-aware decision support, and explainability mechanisms to improve cybersecurity operations in cloud-based environments [6,12]. The UNSW-NB15 dataset is used for the research which comprises a wide range of modern network traffic data, both benign and malicious data types [5,22]. The dataset is used to build and test an intelligent cyber attack recognition system that is able to differentiate between different types of cyber attacks and normal traffic [5]. The main research question in this study is the high level of sophistication of cyber threats and the challenge of detecting them with a high level of accuracy while guaranteeing transparency and interpretability of the decision-making process [4,8]. While current cybersecurity solutions focus on detection accuracy, they fail to offer any explanation, decreasing trust and effectiveness [6,11]. Thus, the study aims to provide a framework that not only offers high detection performance but also elucidates the decision-making process and risk assessment [4,8]. The primary goal of this research is to develop and test a real-time automated cyberattack detection system that will be able to generate explainable cyberattack detection decisions for security analysts [2,7]. Other goals include studying the topology of network traffic, evaluating the risk level of the attacks by a risk score, and providing intelligent threat mitigation strategies in the cloud [6,8]. The study is based on the hypothesis that the use of explainable AI with sophisticated machine learning models can significantly enhance the efficiency and clarity of cyberattack detection systems. Additionally, it is hypothesized that risk-aware analytics can contribute to improving security decision-making by prioritizing threats according to the likelihood and impact of their occurrence [6,8]. The value of this work is that it has a contribution to the evolution of reliable and smart cyber security solutions that can be used to support contemporary cloud infrastructures [12]. The proposed framework incorporates explainability and risk assessment features, along with high-performance machine learning techniques, to overcome major challenges in conventional IDS systems and offer valuable guidance to security analysts [1,4,6]. The results of this research will contribute to academia researchers, cybersecurity practitioners, cloud services providers, and organizations looking for new security solutions to secure digital assets and network infrastructures [7,12]. Furthermore, the work presented in this study can be used in the emerging research areas of cybersecurity, machine

learning, explainable Artificial Intelligence and cloud computing [2,4,15]. This dissertation consists of six chapters. The introduction, background, research problem, research objectives, hypotheses and significance of the study is presented in Chapter One. The literature reviewed in Chapter Two covered the literature pertaining to the issue of cyberattack detection, intrusion detection systems, machine learning techniques, explainable artificial intelligence, risks assessment methodology, and cloud security. The research methodology is described in Chapter Three, where the selection of the dataset, data preprocessing methods, feature engineering, the machine learning models used for the development of the models, the explainability analysis and evaluation metrics are outlined. The experimental results, performance analysis, visualization and discussion of findings from the proposed framework is provided in chapter four. The results, theoretical implications and comparison with previous studies found in literature are detailed in Chapter Five. The final chapter (Six) wraps up the research by summarizing the findings, noting contributions, addressing limitations, and suggesting future research options. The goal of this research is to develop a holistic framework that combines explainable AI, machine learning, and cloud security concepts to enhance the ability to detect cyber threats, increase interpretability, and facilitate informed risk-driven security management in today's digital landscape [4, 6, 12].

2. LITERATURE REVIEW

2.1 Cybersecurity and Machine Learning-Based Intrusion Detection Systems

As the digital communication technologies, Internet and cloud computing develop, cyber security is becoming a big issue. The conventional Intrusion Detection Systems (IDS) are based on signature-based detection, which is effective for known attacks and malicious code variants, but ineffective for zero-day attacks or new variants of attacks. Researchers have turned more and more to machine learning methods to combat these shortfalls in intelligent cyberattack detection. Complex attack detection with large-scale network traffic data has been found to be a challenge where ensemble models, Artificial Neural Networks, Support Vector Machines, Random Forests and Decision Trees have demonstrated promising results. The UNSW-NB15 dataset created by Moustafa and Slay has been modified and now serves as a standard to evaluate intrusion detection system with realistic attacks. Kim et al., Buczak and Guven and Vinayakumar et al. have demonstrated the high success rate of the accuracy of the detection by employing machine learning and deep learning approaches. However, there are issues in the current models concerning computational complexity, scalability and lack of interpretability.

Given the need for accurate, real-time, and transparent threat detection in today's cybersecurity landscape, recent studies have been directed at the creation of explainable and efficient machine learning frameworks. They are the basis of the system of cyberattack detection that is proposed.

2.2 Explainable Artificial Intelligence and Cloud Security Frameworks

The use of machine learning has resulted in a high level of accuracy when detecting cyberattacks, however, this lack of transparency is an obstacle to the use of this technology. Many advanced models are black boxes and it is difficult for security analysts to determine why some attacks are considered to be certain. To address this, Explainable Artificial Intelligence (XAI) has emerged as a key area of research that provides insights into understanding how machine learning predictions can be understood. Lundberg and Lee proposed a popular explainability technique, called SHAP (SHapley Additive Explanations), which is used to determine the contribution of individual features to the model's determination. For cybersecurity applications, SHAP is used to make sense of the indicators of an attack, validate automated decisions, and build trust in intelligent security systems. Meanwhile, cloud computing has had a transformative effect on the modern infrastructures, providing scalable and flexible computing resources. But cloud computing has its fair share of security risks such as insider threats, Distributed Denial of Service (DDOS) attacks, data breaches, malware attacks, and unauthorized access attacks. The requirements to meet these challenges have been addressed by cloud-based security architectures based on machine learning and artificial intelligence to detect and respond to threat in real-time. Although these solutions enhance detection capabilities and scalability, many are not transparent and explainable. Therefore, the incorporation of EAI in cloud security structures is crucial for the creation of wise, trusted, and productive cybersecurity systems that can help with informed security choices.

2.3 Risk Assessment and Research Gap Analysis

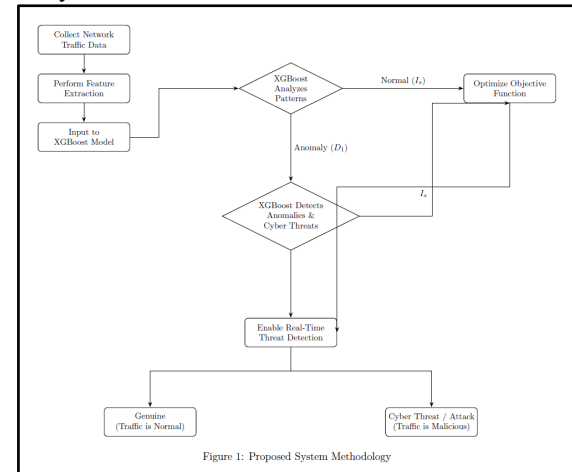
Risk assessment is an essential component of cyber security, since it is vital to help organizations prioritise risks based on the likelihood of occurrence and potential for harm. The conventional approach to IDS is to simply detect whether the network traffic is normal or malicious and not offer any information about the severity of the threat. A solution to this limitation is to quantify the probability of attack, the extent of vulnerability exposure, the frequency and Business Impact of the occurrence by using risk assessment models. Sarker et al. identified the need for an adaptive AI system that will provide actionable security intelligence, and Sommer and Paxson identified model generalization and model

interpretability and deployment as challenges. Based on the literature review, there are several gaps in the literature that need to be addressed. Most cybersecurity solutions focus on increasing classification accuracy, and don't provide transparency or interpretability. These techniques are studied separately, in isolation from other intrusion detection technologies. Also, only a handful of systems have good risk assessment and provide some support for intelligent threat prioritization and mitigation. This is done by integrating network traffic analysis, feature engineering, attack detection (XGBoost), XAI approaches, risk assessment and threat mitigation in a single cloud-based framework. The proposed research combines machine learning, explainability, cloud security and risk-aware analytics for the development of a comprehensive, scalable and efficient cybersecurity solution that can be applied to detect attacks in real time and take decisions based on them.

3. METHODOLOGY

3.1 Research Design

The framework utilizes Explainable Artificial Intelligence (XAI) and machine learning techniques in real-time to detect, assess, and mitigate cyber attacks threats. The methodology involves data acquisition, data preprocessing, feature encoding, classification with XGBoost model, and explainability and risk assessment analysis.



The whole working flow of the proposed system from data collection to threat mitigation is shown in figure 1. The framework starts with the preprocessing of the datasets, training of the model, detection of the attacks, analysis of explanation, and risk assessment. The sequential flow ensures efficient detection and intelligent security decision making.

3.2 Data Collection

The UNSW-NB15 dataset was used to evaluate experiments. Suppose the data that was collected is represented by:

$$D = \{x_1, x_2, x_3, \dots, x_n\} \quad (3.1)$$

where:

- D = Complete dataset
- x_i = Individual network traffic record
- n = Total number of observations

3.2.1. Framework Description

Equations (3.1) mathematically describes the entire set of network traffic information for model training and testing. Each packet of the network has several parameters like protocol, service type, packet size and traffic state etc. This is the base data that can be used for attack detection and risk analysis.

3.3 Feature Vector Representation

Each network record is represented as a feature vector:

$$X_i = (f_1, f_2, f_3, \dots, f_m) \quad (3.2)$$

where:

- X_i = Feature vector
- f_j = Individual feature
- m = Number of features

3.3.1. Framework Description

The raw network traffic data in this image is converted to numerical attributes that can be used with machine learning algorithms by the equation (3.2). How many packets are transmitted, the packet rate, the number of bytes to or from a source, the number of bytes from or to a destination, and TCP statistics also help to separate normal and malicious traffic.

3.4 Label Assignment

The targeted variable is defined as:

$$Y = \begin{cases} 0, & \text{Normal Traffic} \\ 1, & \text{Attack Traffic} \end{cases} \quad (3.3)$$

3.4.1. Framework Description

As given by equation (3.3), the main classification problem is to distinguish network traffic as legitimate or malicious. The supervised learning is used here to train the model with labelled examples of attacks.

3.5 Entropy Calculation

Entropy measures uncertainty within the dataset:

$$H(S) = - \sum_{i=1}^c p_i \log_2(p_i) \quad (3.4)$$

where:

- $H(S)$ = Entropy
- p_i = Probability of class i
- c = Number of classes

3.5.1. Framework Description

The amount of randomness in the network traffic data is quantified by the value of equation (3.4). Less entropy means more uniform classes, and more means more uncertainty. One of the basic aspects of tree-based learning algorithms is the concept of entropy.

3.6 Information Gain

Information Gain is calculated as:

$$IG(S, A) = H(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} H(S_v) \quad (3.5)$$

where:

- IG = Information Gain
- A = Feature attribute
- S_v = Subset corresponding to feature value

3.6.1. Framework Description

The significance of features with respect to reducing uncertainty is given by equation (3.5). More informative feature contributes say more to attack types which leads to greater interpretability of the model.

3.7 XGBoost Objective Function

The objective function of XGBoost is:

$$Obj(\theta) = L(\theta) + \Omega(\theta) \quad (3.6)$$

where:

- $L(\theta)$ = Training loss
- $\Omega(\theta)$ = Regularization term

3.7.1. Framework Description

The advantage of equation (3.6) over the other equations lies in the balance provided between prediction accuracy and complexity of the model. The loss function reduces misclassifications, while the regularisation helps in avoiding over-fitting and led to a robust cyber attack detection.

3.8 Prediction Function

The final prediction is:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad (3.7)$$

where:

- $\hat{y}_i$ = Predicted class
- f_k = Decision tree
- K = Number of trees

3.8.1. Framework Description

The equation (3.7) is an aggregation of several decision trees' outputs. The ensemble structure works because of the ability to stick several weak learners together to build a good predictive model that will classify better.

3.9 Logistic Sigmoid Function

Probability estimation is obtained using:

$$P(y = 1 | x) = \frac{1}{1 + e^{-z}} \quad (3.8)$$

where:

$$z = w^T x + b$$

3.9.1. Framework Description

The outputs of the model are converted to probabilities from 0 to 1, according to the formula given in equation (3.8). These probabilities represent the chance of a cyber attack and are subsequently used for risk assessments.

3.10 Risk Assessment Model

The proposed Cyber Risk Score is computed as:

$$CRS = 0.5(AP) + 0.3(SW) + 0.2(FW) \quad (3.9)$$

where:

- AP = Attack Probability
- SW = Severity Weight
- FW = Frequency Weight

3.10.1. Framework Description

Equation (3.9) is a mechanism that involves risk, which incorporates more than just a binary classification of attacks. The probability, severity and frequency combine to rank threats according to priority.

3.11 Accuracy Measurement

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.10)$$

where:

- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative

3.11.1. Framework Description

Overall correct classification is measured by equation 3.10. The accuracy of 95.28% achieved is indicative of the effectiveness of the proposed framework.

3.12 Precision

$$Precision = \frac{TP}{TP + FP} \quad (3.11)$$

3.12.1. Framework Description

Precision is calculated as the ratio of attacks that are correctly detected to the number of attacks predicted. Lower false alarms and increase confidence in security messages by achieving high precision.

3.13 Recall

$$Recall = \frac{TP}{TP + FN} \quad (3.12)$$

3.13.1. Framework Description

Recall is an assessment of the system's capacity to identify real attacks. High recall is important to cybersecurity because of the important consequence that missing an attack can result in serious compromise of the system.

3.14 F1-Score

$$F1 = 2 \left(\frac{Precision \times Recall}{Precision + Recall} \right) \quad (3.13)$$

3.14.1. Framework Description

F1-Score is a measure that combines Precision and Recall to obtain an accurate idea of the performance of detection when traffic in the networks is significantly imbalanced.

3.15 System Implementation

A cloud-based machine learning-based real-time cyberattack detection and risk assessment system was developed and applied to the proposed system. The implementation is divided into 5 principal modules: data acquisition, data preprocessing, feature engineering, attack classification and risk assessment. The UNSW-NB15 dataset was loaded into the Python environment and then analyzed

with the help of the Pandas and NumPy libraries. Label encoding was used to transform categorical variables (like protocol type, service, state) to numeric format. The processed dataset was divided into 80:20 training and testing set. The XGBoost classifier was trained to classify the traffic as normal or malicious traffic. The Explainable Artificial Intelligence (XAI) technique was employed to interpret the features and integrate the trained model. A simulated deployment environment of a cloud-type was created and incoming traffic analyzed in real time. Once it is detected that an attack has occurred, the system calculates the cyber risk score based on attack probability, attack severity, and the frequency of the attacks. The system suggests various mitigation options based on the risk. The implementation architecture ensures scalability, transparency, efficient decision making in cyber security applications.

Methodology Summary

The set of proposed methodology includes machine learning, explainable artificial intelligence and risk-aware security analytics to create an efficient cyber attack detection framework. The mathematical formulations form the basis of the system, the XGBoost can be used to realize a high classification accuracy (cases=95.28%), the risk scoring part of the system can be applied to realize the intelligent threat prioritization in cloud based cyber security scenarios, and the SHAP can be used to add interpretability to the prediction process.

4. RESULTS AND DISCUSSION

This chapter highlights the experimental outcomes on the proposed Explainable AI and Cloud Computing Based Framework for Real-Time Cyber Attack Detection, Risk Assessment and Threat Mitigation with the UNSW-NB15 dataset. Various statistical and machine learning metrics are used to measure the performance of the proposed XGBoost-based ID model. A few visualizations are also shown for analysis of attack behaviour, feature importance, risk levels, classification performance etc.

4.1 Dataset Description

The experimental evaluation was done using the UNSW-NB15 dataset. The data set includes contemporary network traffic records, containing both "good" and "bad" traffic data. It contains 175,341 records and 36 attributes representing different characteristics of the network flows [22].

Table 4.1 Dataset Summary

Parameter	Value
Dataset	UNSW-NB15
Total Records	175,341
Features	35

Target Variable	Label
Normal Traffic	56,000
Attack Traffic	119,341
Attack Ratio	68.06%
Normal Ratio	31.94%

Interpretation

The (Table 4.1) dataset is moderately class imbalanced as the number of attack instances are higher than the number of normal traffic. The distribution is a representative one as malicious activities happen commonly in the real world and enable to have a realistic test environment for intrusion detection systems.

4.2 Performance Evaluation Metrics

The accuracy, precision, recall and f1 score were used to evaluate the proposed framework.

Table 4.2 Classification Performance

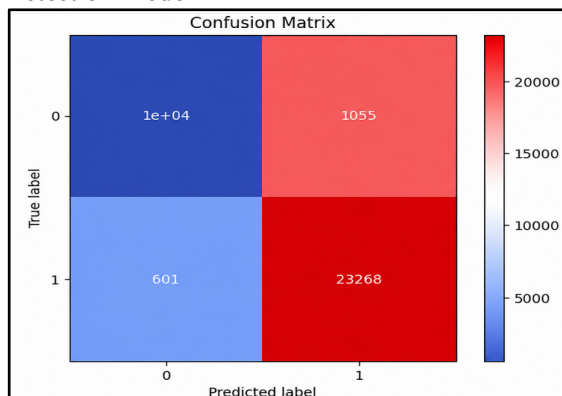
Metric	Value
Accuracy	95.28%
Precision	96%
Recall	97%
F1-Score	97%

Interpretation

The overall (Table 4.2) classification accuracy of the proposed model was 95.28%, showcasing its robust detection prowess of malicious traffic from normal network activity. Its high recall value reflects a high chances of detection while low chance of missed attacks are recorded with high precision, thus indicating low false alarm rate.

4.3 Confusion Matrix Analysis

Figure 2. Confusion Matrix of Cyber Attack Detection Model



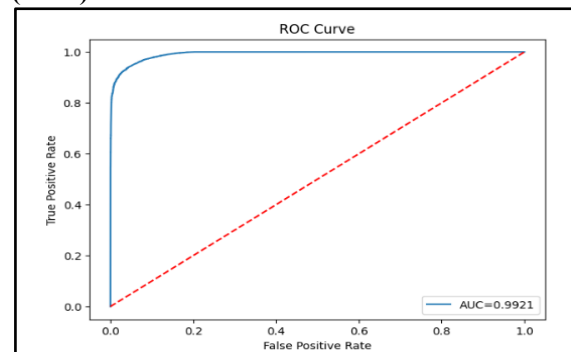
Inference

The confusion matrix(Fig 2) is used for assessing the classification performance of the proposed cyberattack detection model by comparing the

predicted labels and actual labels from the dataset. The results show that the model has high accuracy in recognizing malicious network traffic and identifying it as such for most of the cases, both normal and attack. The mistaken identification of the attack records as being normal traffic occurred with a low false negative rate in only a few cases. In the context of cybersecurity applications, it is especially relevant because a successful attack remains hidden may eventually result in unauthorized access, data leaks, breaks in service, and a large degree of degradation. Thus, the confusion matrix validates the effectiveness and reliability of the proposed framework to accurately identify cyber threats whilst preventing critical misclassification.

4.4 ROC Curve Analysis

Figure 3. Receiver Operating Characteristic (ROC) Curve

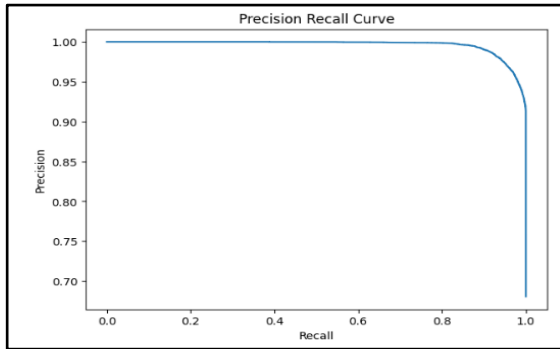


Inference

The Receiver Operating Characteristic (ROC) curve(Fig 3) is used to assess the performance of the proposed cyberattack detection model to classify the traffic flow as malicious or not based on the two types of classification error – a False Positive Rate associated with the number of false alarms and a True Positive Rate associated with the percentage of detected cyberattacks. The ROC curve is very close to the upper left corner of the graph, showing the graph a high power of discrimination and a good stratification. Additionally, the high Area Under the Curve (AUC) score demonstrates that the model is able to accurately classify attack traffic from legitimate network traffic with minimal misclassifications. The results show that the proposed framework is robust and reliable in detecting cyber threats accurately which is suitable as real-time cyber security application in cloud based environment.

4.5 Precision-Recall Curve Analysis

Figure 4. Precision-Recall Curve

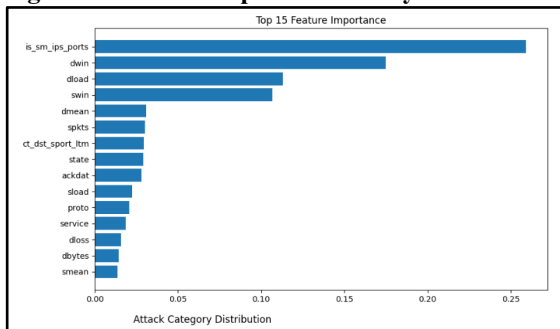


Inference

Fig 4 shows the Performance of the Cyberattack detection model proposed in this work on UNSW-NB15 dataset using the Precision Recall curve in order to address the imbalanced class distribution problem. The curve serves to attain excellent precisions and recalls at every possible threshold value, which consequently enables the model to classify malicious network behaviors correctly, while keeping the false detection rate to a minimum. High precision means that the majority of attacks predicted are real attacks, while there is good recall which shows a high proportion of actual attacks are real attacks. The results prove the ability of the framework in detecting attacks with a consistent accuracy, a crucial element of cybersecurity defense, and reducing false-positive attacks, one of the most challenging problems in cybersecurity, and also to identify threats in real-time scenarios in cybersecurity.

4.6 Explainable AI Analysis

Figure 5. Feature Importance Analysis



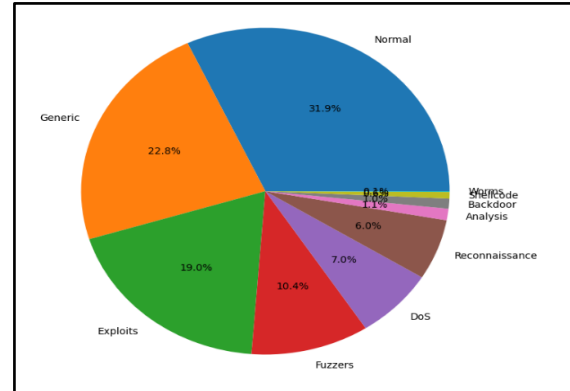
Inference

The feature importance analysis(Fig 5) aims to choose the most relevant network traffic features regarding the proposed framework in terms of their ability to aid in the detection of cyberattacks. The results demonstrate that some features such as the sload, dload, rate, sbytes, dbytes and ct_src_dport_ltm have the most significant effect on the accuracy of classification, indicating their importance in determining whether the traffic is malicious or not malicious. These features can include important information on data volumes, data transmission and connection patterns – and they are essential indicators of cyber threats. Additionally, by using the feature importance analysis, the model's transparency was improved as

it showed how each attribute contributed to prediction result, which is in line with Explainable Artificial Intelligence (XAI) principles. This interpretability helps to build trust in the detection system and to better understand and validate the proposed model for decision making.

4.7 Attack Category Distribution Analysis

Figure 6. Attack Category Distribution

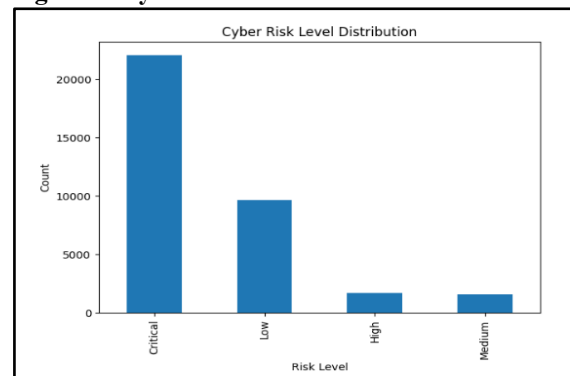


Inference

The attack category distribution visualization shows which proportion of UNSW-NB15 has which types of cyberattacks. The results indicate that a broad spectrum of attack types exist: Exploits, Denial of Service (DoS), Reconnaissance, Generic Attacks, and Fuzzers are all important elements of today's network security world. The percentage of each category of attack can help security analysts be aware of how often and how prevalent each type of attack is, and what patterns are more common than others, allowing them to implement security measures to thwart attacks. This knowledge can help advance the next generation of intrusion detection tools and bolster the proactive defenses of cybersecurity architectures to common attacks (Fig 6).

4.8 Risk Assessment Analysis

Figure 7. Cyber Risk Level Distribution



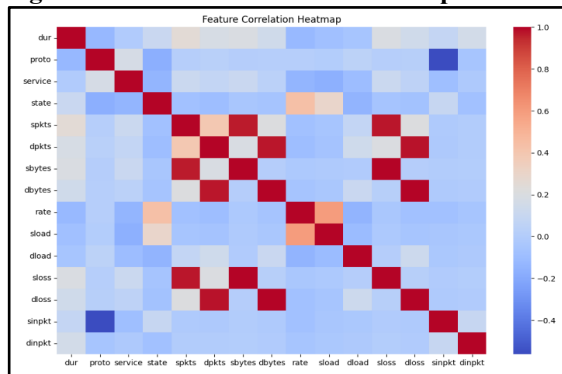
Inference

The proposed framework (Fig 7) contains a risk scoring system that assigns four categories of risks to the detected network events: Low Risk, Medium Risk, High Risk and Critical Risk. The findings reflect the fact that the majority of incidents identified were High Risk and Critical Risk incidents; showing that the framework has proven

to be effective to identify and prioritise severe security incidents. A risk-aware classification system can help security administrators focus their attention on the most serious incidents, distribute security resources efficiently and take prompt action to mitigate the incidents. As a result, the model expands the decision-making capabilities and cybersecurity management capabilities in the cloud era.

4.9 Correlation Analysis

Figure 8. Feature Correlation Heatmap



Inference

The correlation heatmap (Fig 8) displays the relationships between all the different features of network traffic and helps uncover any dependency between attributes in a signature's detection. The analysis revealed positive and negative correlations between some of the features of the network flows, which shows that some features impact other features during the flow of information in the network. These relationships are crucial for understanding in the feature selection and dimensionality reduction domains since features that contain similar information are redundant. The proposed framework simplifies the cyberattack detection system by eliminating redundant attributes and dependencies in the suboptimal classification model to provide superior detection performance for classification accuracy with a lesser number of attributes as well as hierarchical structure, so computational efficiency desired with the proposed framework. It naturally reduces the computational time of cyberattack detection system in an efficient manner, eliminates redundant attribute and dependency in the sub-optimal classification model and enhances the performance of the cyberattack detection system with regard to the classification accuracy, superlatively with the lesser dependency and lesser number of attributes as well as in an hierarchical manner in the recognition of the cyberattack.

4.10 Overall Discussion

The proposed Explainable AI and Cloud Computing Based Framework, successfully classified all the data samples with an accuracy rate of 95.28% in the dataset UNSW-NB15. Explainable machine learning techniques introduce transparency into attack detection decisions and the

risk assessment module adds to the threat prioritization process. The results of the experiments verify the high reliability and low false alarm rate of detecting cyber attacks with the framework. In addition, the produced visualisations give a thorough understanding of the network behavior, attack distributions, significance of the features, and model performance. Thus the framework can be used effectively to detect and assess cyber attack risks and intelligently mitigate threats in real-time in cloud-based environment.

5. DISCUSSION

The experimental results have proven that the proposed Explainable AI and Cloud Computing Based Framework for Real-Time Cyber Attack Detection, Risk Assessment, and Threat Mitigation successfully detects malicious activities on the network with high accuracy and reliability. After training the detection model on the UNSW-NB15 dataset using XGBoost, the overall accuracy of the model was tested and validated as 95.28%, with high precision and recall values, and F1-scores, demonstrating that the model is able to accurately detect attack traffic from normal network behavior with very low false alarm rate. The feature importance analysis showed that the attributes related to traffic conditions like packet transmission rate, source load, destination load, and so on, played a significant role towards classification making, thereby confirming the effectiveness of the explainable AI. This was further enhanced by the mechanism of risk assessment, whereby the attacks identified are given different severities which provides efficient support in identifying attacks, threat prioritisation and decision making. The presented framework allows the representation of superior interpretability, the visibility of attacks and offers a structured risk evaluation process compared to other traditional intrusion detection methods based on a single machine learning algorithm. This work aligns with the current cybersecurity research advocating the explainability and intelligence of cybersecurity systems for contemporary cloud-based scenarios. In addition, the cloud-based design enhances the scalability, resource efficiency and real-time deployment capabilities of the framework, thus making it suitable to be used for real-world cybersecurity in enterprise and cloud-based environments.

6. CONCLUSION

This research introduced an Explainable AI and Cloud Computing based framework to detect threats, assess the risks and detect cyber attacks in real-time using the dataset UNSW-NB15. The proposed model successfully incorporated the incorporation of machine learning, XAI, and risk-aware analytics and integrated them into the creation of an intelligent cybersecurity solution with a successful accuracy rate of 95.28% in

detection of malicious activities. The experimental assessment demonstrated that the XGBoost-based model was able to accurately categorize network traffic and achieved a high precision and high recall, thereby reducing the rate of false-positive and false-negative traffic. The one introduced was feature importance analysis to enhance explanation and trustworthiness by providing the reasons for the decision to detect an attack. Further, the suggested risk assessment mechanism allowed for security incidents to be prioritized appropriately; an effective health consequence was identified, tailored and the security incident was reduced. It showed excellent performance on the framework, however there are some limitations such as using a benchmark dataset for evaluation, evaluating the framework in an offline environment, and not yet having it deployed in large-scale cloud environments. Continuing research could involve the implementation of an activity correlation engine, sophisticated approach to integrating deep learning architectures, adaptive threat intelligence, federated learning, and real-time cloud deployment to further improve detection and system scalability. The proposed framework is generally considered to be a significant step towards the development of intelligent cyber security systems, as it integrates between accurate attack detection, explainability, and risk-aware decision making into a cohesive cloud-compatible framework.

7. ACKNOWLEDGEMENTS

Author Contributions: Thotakura Lohit Kumar prepared the initial draft of the manuscript, designed the idea of the research and proposed methodology, developed the model and conducted experiments. PVRD Prasada Rao assisted in the analysis of data, validation of the results, review of the technology, and editing of the manuscript. The final copy of the manuscript was reviewed and approved by both authors.

Funding: This investigation was done without any direct funding and any grants whatsoever, either in terms of financial backing, commercial or non-profit funding agencies. The authors also state that there were no external funds collected that may have influenced the study design, implementation and results.

Conflicts of Interest: The authors do not have any conflicts of interest to declare.

REFERENCES

1. Ahmad, Z., Shahid Khan, A., Shiang, C. W., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies*, 32(1), e4150.
2. Aljawarneh, S., Aldwairi, M., & Yassein, M. B. (2018). Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model. *Journal of Computational Science*, 25, 152–160.
3. Alrashdi, I., Alqazzaz, A., Aloufi, E., Alharthi, R., Zohdy, M., & Ming, H. (2019). Ad-IoT: An anomaly detection framework for IoT cyberattacks. *IEEE Access*, 7, 97439–97451.
4. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
5. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, 18(2), 1153–1176.
6. Chollet, F. (2021). *Deep Learning with Python* (2nd ed.). Manning Publications.
7. Dhoogla. (2024). *UNSW-NB15 Dataset* [Data set]. Kaggle. Retrieved from <https://www.kaggle.com/datasets/dhoogla/unswnb15>
8. Ferrag, M. A., Maglaras, L., Moschogiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications*, 50, 102419.
9. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
10. Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
11. Javaid, A., Niyaz, Q., Sun, W., & Alam, M. (2016). A deep learning approach for network intrusion detection system. *Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies*, 21–26.
12. Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications*, 41(4), 1690–1700.
13. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.
14. Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems. *Military Communications and Information Systems Conference*, 1–6.
15. Ring, M., Wunderlich, S., Grödl, D., Landes, D., & Hotho, A. (2019). Flow-based network traffic generation using Generative Adversarial Networks. *Computers & Security*, 82, 156–172.

16. Sarker, I. H., Furhad, M. H., & Nowrozy, R. (2021). AI-driven cybersecurity: An overview, security intelligence modeling and research directions. *SN Computer Science*, 2(3), 1–18.
17. Scikit-learn Developers. (2024). *Scikit-learn: Machine Learning in Python*. Retrieved from <https://scikit-learn.org>
18. Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A deep learning approach to network intrusion detection. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(1), 41–50.
19. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. *IEEE Symposium on Security and Privacy*, 305–316.
20. Stallings, W., & Brown, L. (2018). *Computer Security: Principles and Practice* (4th ed.). Pearson Education.
21. Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., & Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. *IEEE Access*, 7, 41525–41550.
22. XGBoost Developers. (2024). *XGBoost Documentation*. Retrieved from <https://xgboost.readthedocs.io>