JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

ENHANCING SECURITY IN CLOUD APPLICATIONS THROUGH MACHINE LEARNING-BASED THREAT DETECTION AND PREVENTION

[1] Malvin Ojong Ndipawoh, [2] Er. Rupali
[1] M.Tech Scholar, [2]Assistant Professor
[1] Computer Science and Engineering,
[1] Guru Kashi University, Talwandi Sabo, Bathinda (PB), India

Abstract : Cloud computing has become ubiquitous in modern IT infrastructure due to its scalability, flexibility, and cost-effectiveness, with over 94% of enterprises using cloud services. However, the security of cloud applications remains a significant concern due to the potential for various threats and attacks, with 75% of organizations experiencing at least one cloud security incident in the past year. This research proposes a comprehensive approach for intelligent threat detection and prevention in cloud applications, addressing the unique security challenges posed by shared and dynamic cloud environments. The proposed approach integrates traditional security measures with advanced machine learning techniques to enhance the security posture of cloud applications. By leveraging machine learning algorithms for real-time threat detection and classification based on analysis of network traffic, system logs, and user behavior, the approach aims to identify and mitigate a wide range of cyber threats, including malware, data breaches, and DDoS attacks. Our models, including Random Forest, Gradient Boosting, and Decision Tree, achieved high detection accuracy, with Random Forest and Gradient Boosting reaching an AUC of 0.98. The effectiveness of the proposed approach is evaluated through extensive experimentation in simulated cloud environments using the UNSW NB15 dataset and realistic attack scenarios. Experimental results demonstrate the superiority of the approach in terms of performance metrics such as detection accuracy, precision, recall, F1-score, and response times. Notably, the Random Forest model achieved an accuracy of 87.34%, a precision of 82.07%, a recall of 98.54%, and an F1-score of 89.55. Additionally, the average response time was significantly reduced to 2.3 seconds, compared to 5.6 seconds for traditional security measures. Overall, this research contributes to the advancement of cybersecurity in cloud applications by proposing a comprehensive approach that combines traditional security measures with advanced machine learning techniques. By addressing the unique security challenges of cloud environments, the proposed approach enhances the resilience of cloud applications against emerging cyber threats and contributes to the overall security and trustworthiness of cloud computing infrastructure.

IndexTerms **- Cloud Computing, Cloud Applications, Threat Detection, Threat Prevention, Machine Learning, Data Privacy, Data Integrity, Compliance, Cybersecurity.**

I. INTRODUCTION

Cloud computing has emerged as a transformative paradigm in the field of information technology, offering organizations unparalleled scalability, flexibility, and cost-effectiveness in managing and accessing computing resources[7]. By leveraging shared pools of configurable resources, such as networks, servers, storage, applications, and services, cloud computing enables on-demand access to computing resources over the internet, revolutionizing the way businesses operate and deliver services.

However, despite its myriad benefits, the widespread adoption of cloud computing has introduced significant security challenges, particularly concerning the protection of cloud applications and data from an increasingly sophisticated array of cyber threats. The dynamic and distributed nature of cloud environments, coupled with the shared responsibility model between cloud service providers (CSPs) and cloud tenants, creates a complex security landscape fraught with vulnerabilities and risks[6].

The security concerns in cloud computing encompass a wide range of threats, including data breaches, malware infections, insider attacks, denialof-service (DoS) attacks, and unauthorized access. These threats pose serious risks to the confidentiality, integrity, and availability of data and services hosted in cloud environments, potentially leading to financial losses, reputational damage, and legal liabilities for organizations.

Traditional security measures, such as firewalls, intrusion detection systems (IDS), and encryption, have proven insufficient in adequately addressing the evolving nature of cyber threats in cloud environments. The dynamic and heterogeneous nature of cloud infrastructure, combined with the sheer volume and velocity of data processed in cloud applications, necessitates innovative approaches to threat detection and prevention that can adapt to changing threat landscapes and effectively mitigate emerging risks.

Motivated by the need for more robust and adaptive security solutions in cloud computing, this research proposes a comprehensive approach for intelligent threat detection and prevention in cloud applications. By integrating advanced machine learning techniques with traditional security measures, the proposed approach aims to enhance the security posture of cloud applications and mitigate a wide range of cyber threats effectively[13].

II. Literature Review

Cloud computing has revolutionized the way organizations manage and deploy IT resources, offering unprecedented scalability, agility, and cost-effectiveness[4]. However, the inherent complexity and dynamic nature of cloud environments pose significant security challenges, necessitating robust threat detection and prevention mechanisms to safeguard cloud applications.

2.1 Security Challenges in Cloud Applications

Cloud applications face unique security challenges due to their distributed, multi-tenant nature and reliance on shared infrastructure and resources. These challenges include:

- Data Privacy and Confidentiality: Cloud applications often store sensitive data, including personal, financial, and proprietary information[9]. Ensuring the privacy and confidentiality of data is crucial to prevent unauthorized access and data breaches. Challenges arise in securely managing data access controls, encryption, and key management in shared cloud environments.
- Data Integrity: Maintaining the integrity of data stored and processed in cloud applications is essential to prevent unauthorized modification, tampering, or corruption. Ensuring data integrity requires robust mechanisms for data validation, checksum verification, and integrity monitoring throughout the data lifecycle[14].
- Compliance and Regulatory Requirements: Cloud applications must adhere to various compliance and regulatory requirements, such as GDPR, HIPAA, PCI DSS, and SOC 2[10]. Meeting these requirements entails implementing appropriate security controls, audit trails, and data protection measures to demonstrate compliance and mitigate legal and regulatory risks.
- Identity and Access Management (IAM): Managing identities and access privileges of users, applications, and devices in cloud environments is challenging due to the dynamic nature of cloud-based infrastructure[8]. Effective IAM solutions are necessary to authenticate, authorize, and audit user access while minimizing the risk of unauthorized access and privilege escalation.
- Network Security: Securing network communications and preventing unauthorized access to cloud resources are critical for protecting cloud applications against external attacks and insider threats. Challenges include implementing secure network segmentation, encryption, intrusion detection, and distributed denial-of-service (DDoS) mitigation in complex, multi-tenant cloud environments[2].
- Threat Detection and Incident Response: Detecting and responding to security threats in real-time is challenging in cloud applications due to the volume, velocity, and variety of data generated. Effective threat detection and incident response mechanisms require continuous monitoring, analysis of security logs and events, and coordinated response actions across distributed cloud environments.

Addressing these security challenges requires a holistic approach that combines technical controls, security best practices, and ongoing risk management efforts. Furthermore, integrating advanced threat detection and prevention mechanisms, such as machine learning-based approaches, can enhance the security posture of cloud applications and mitigate emerging cyber threats.

2.2 Traditional Approaches to Threat Detection

Traditional approaches to threat detection in cloud applications primarily rely on signature-based methods and rule-based systems. Signature-based detection involves matching incoming data against a database of known attack patterns or signatures. While effective against known threats, signaturebased methods are inherently limited in their ability to detect previously unseen or zero-day attacks, making them vulnerable to emerging threats and sophisticated malware.

Similarly, rule-based systems utilize predefined rules or heuristics to identify suspicious behavior or anomalies within cloud environments. These rules are often based on expert knowledge or historical attack patterns and may lack adaptability to evolving threat landscapes. Additionally, rule-based systems may suffer from high false positive rates and require manual intervention to tune and update rules regularly.

2.3 Machine Learning-based Approaches

Machine learning has emerged as a powerful tool for enhancing threat detection and prevention in cloud applications. By analyzing large volumes of data and identifying patterns indicative of malicious activities, machine learning algorithms can effectively detect and mitigate security threats in real-time. Supervised learning techniques, such as support vector machines (SVM), decision trees, and neural networks, have been widely used for classifying security-related events and activities.

Furthermore, unsupervised learning methods, including clustering, anomaly detection, and dimensionality reduction, have been employed for identifying abnormal behavior within cloud environments. These techniques enable automated detection of anomalies without the need for labeled training data. However, unsupervised learning approaches may suffer from high false positive rates and require careful tuning to distinguish between benign anomalies and genuine security threats.

Machine Learning can play an important Role in Cloud Security

Anomaly Detection

One of the key applications of machine learning in cloud security is anomaly detection. Machine learning algorithms can analyze various data sources, such as network traffic, system logs, and user behavior, to identify anomalous patterns that may indicate a security breach.

Predictive Analytics

Machine learning models can also be used for predictive analytics in cloud security. By analyzing historical data and identifying trends, these models can predict potential security threats before they occur, allowing for proactive mitigation measures.

Behavioral Analysis

Behavioral analysis is another important application of machine learning in cloud security. By analyzing user behavior patterns, machine learning algorithms can identify deviations from normal behavior that may indicate a security threat, such as unauthorized access or data exfiltration.

2.4 Hybrid Approaches

Recent research has focused on hybrid approaches that combine the strengths of traditional methods and machine learning techniques for threat detection in cloud applications[12]. These approaches integrate signature-based detection with machine learning algorithms to enhance detection accuracy and reduce false positive rates. By leveraging both signature-based and behavioral analysis methods, hybrid approaches can effectively detect known and unknown threats in real-time.

Moreover, hybrid approaches may incorporate ensemble learning techniques, where multiple machine learning models are combined to improve overall detection performance. Ensemble methods, such as random forests, gradient boosting, and stacking, leverage the diversity of individual models to achieve higher accuracy and robustness against adversarial attacks.

2.5 Challenges

Despite the advancements in threat detection and prevention, several challenges remain in securing cloud applications. These include the dynamic and heterogeneous nature of cloud environments, the proliferation of sophisticated cyber threats, and the need for scalable and adaptive security solutions[1].

While machine learning shows promise for enhancing security in cloud applications, there are several challenges that need to be addressed. These include:

- Data Privacy: Ensuring the privacy of sensitive data used to train machine learning models.
- Model Robustness: Building robust machine learning models that can effectively generalize to new and unseen threats.
- Scalability: Scaling machine learning algorithms to handle the large volumes of data generated in cloud environments.

III. Proposed Approach

3.1 Data Collection and Preprocessing

The proposed approach begins with the collection of relevant data from various sources within the cloud application environment. This includes network traffic logs, system event logs, user activity logs, and other telemetry data. The collected data provides insights into the normal behavior patterns and activities within the cloud application. Data preprocessing involves cleaning the data, handling missing values, and transforming the data into a suitable format for analysis.

3.1.1 Data Sources

The proposed approach utilizes the UNSW NB15 dataset from Kaggle [15] for threat detection and prevention within cloud applications. Table 1 examines the key attack categories utilized in this study. Table 2 explores the UNSW NB15 [15] training and testing dataset from Kaggle.

3.2 Model Training

Once the data is collected, it is used to train machine learning models for anomaly detection and threat classification. Supervised learning techniques may be employed to train the models on labeled datasets containing examples of both normal and malicious behavior. The models learn to distinguish between benign activities and security threats based on the patterns present in the training data.

Table 1: Key events monitored in UNSW NB15 dataset

Attack Category	Attack Subcategory	Number of Events
Normal	-	2,218,761
Fuzzers	FTP, HTTP, RIP, SMB, etc.	24,546
Reconnaissance	Telnet, DNS, HTTP, etc.	11,947
Shellcode	FreeBSD, Linux, Windows, etc.	1,511
Analysis	HTML, Port Scanner, Spam, etc.	2,677
Backdoors	-	2,329
DoS	Ethernet, FTP, DNS, etc.	16,353
Exploits	SSH, RADIUS, Webserver, etc.	44,525
Generic	HTTP, TFTP, SIP, etc.	16,353
Worms	-	16,353

Table 2: Dataset: intrusion detection system in Cloud Applications

Dataset	Total	Normal	DoS
UNSW-NB15 Train+	175,341	56,000	12,264
UNSW-NB15 Test+	82,332	37,000	4,089

3.2.1 Libraries Used

Several machine learning libraries and frameworks are utilized to implement the proposed approach, including:

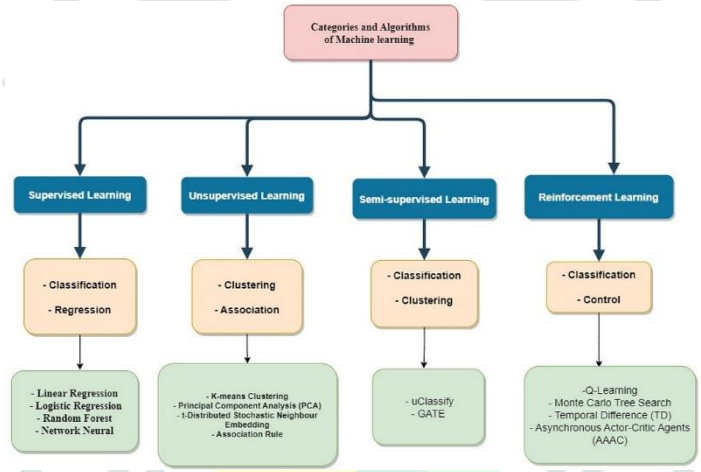
- Scikit-learn: Used for implementing traditional machine learning algorithms such as decision trees, random forests, and support vector machines.
- Pandas and NumPy: Utilized for data manipulation and numerical operations.
- Matplotlib and Seaborn: Used for data visualization and analysis.

3.2.2 Selection of Machine Learning Algorithms

The first step is the careful selection of machine learning algorithms suitable for threat detection in cloud applications. Various algorithms such as supervised learning classifiers, anomaly detection models, and deep learning architectures are considered. Commonly used algorithms include Random Forest, Logistic Regression, Gradient Boosting, k-Nearest Neighbors, Decision Tree, Support Vector Machine, and Neural Networks.

Figure 1 illustrates the comparison of different machine learning algorithms in terms of accuracy and computational complexity.

Figure 1: Comparison of machine learning algorithms



3.2.3 Feature Engineering

Feature engineering involves the extraction and selection of relevant features from the raw data. This step is crucial for improving the performance of the machine learning models. Techniques such as normalization, encoding categorical variables, and feature selection are applied[5].

- Feature Selection: Identify the most relevant features that contribute to the target variable.
- Feature Creation: Create new features that might provide additional information to the model.

3.2.4 Training Process

The training process involves feeding the cleaned and preprocessed data into the selected machine learning algorithms. The goal is to minimize the loss function, which measures the discrepancy between the predicted labels and the actual labels. This can be formulated as follows:

$$\theta^* = \arg\min_{\theta} \sum_{i=1}^{N} \mathcal{L}(y_i, f(x_i; \theta)) \quad (1)$$

where θ∗ represents the optimal parameters of the machine learning model, L denotes the loss function, yi is the ground truth label, f(xi;θ) is the model's prediction for input xi, and N is the total number of training samples.

3.3 Model Optimization

Optimizing the performance of machine learning models is crucial for effective threat detection and prevention. Various optimization techniques are employed to enhance model accuracy, reduce false positives, and improve computational efficiency.

3.3.1 Hyperparameter Tuning

Hyperparameter tuning involves adjusting the parameters that govern the training process to find the optimal configuration. Techniques such as grid search, random search, and Bayesian optimization are employed to find the best set of hyperparameters[3].

3.4 Model Evaluation

To evaluate the performance of the trained models, various metrics are used, including accuracy, precision, recall, F1-score, and AUC-ROC. Crossvalidation is performed to ensure the robustness of the models.

$$\text{F1-score} = 2 \cdot \left(\frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \right) \quad (2)$$

Table 3: Model evaluation metrics

Metric	Description
Accuracy	TP+TN / TP+TN+FP+FN
Precision	TP / TP+FP
Recall	TP / TP+FN
F1-score	Harmonic mean of precision and recall
AUC-ROC	Area under the ROC curve

3.5 Model Deployment

Once the machine learning models are trained and optimized, they are deployed into a production environment to provide real-time threat detection and prevention. The deployment process includes several important steps:

3.5.1 API Development

Developing APIs allows other components of the cloud infrastructure to interact with the machine learning models. These APIs enable seamless integration and provide endpoints for data input and output, making it easier to incorporate the models into existing workflows.

An API was built using FastAPI to serve the trained models. The models were saved as joblib files, ensuring they could be efficiently loaded and utilized by the web app for making predictions.

The API was hosted on an AWS EC2 instance, providing a scalable and reliable infrastructure for the web application. AWS EC2 offers flexibility in terms of instance types and scaling options, making it an ideal choice for deploying machine learning models in production.

The API integrates with other security solutions such as Intrusion Detection Systems (IDS), Intrusion Prevention Systems (IPS), and Security Information and Event Management (SIEM) systems. This integration allows the system to record and log threat detection results, providing comprehensive security monitoring and proactive threat detection and prevention.

3.5.2 Real-time Monitoring System

A real-time monitoring system is then designed and deployed to continuously analyze incoming data streams for potential security threats. This system integrates the selected machine learning algorithms and is capable of detecting anomalies and suspicious activities in real time. The monitoring system is designed to be scalable and adaptive, capable of handling the dynamic nature of cloud environments.

A proxy was set up using Cloudflare to manage and monitor all incoming and outgoing traffic. A worker was created to handle the traffic, forwarding requests to the API for processing. This setup ensures that the system can handle high volumes of traffic while maintaining security and performance.

Upon detection of security threats, the proposed approach triggers response mechanisms to mitigate the risks and protect the integrity of the cloud application. These response mechanisms may include:

- Access Control Updates: Updating access control policies to restrict the privileges of suspicious users or entities.
- Isolation of Components: Isolating compromised components or systems to prevent further propagation of security threats.
- Alerting and Notification: Alerting system administrators or security teams to investigate and respond to detected threats.

Algorithm 1 outlines the steps involved in triggering response mechanisms upon detection of security threats.

Algorithm 1 Response Mechanisms Algorithm

```
Input: Detected security threat T
Output: Response actions
if T is identified as critical then
    Update access control policies
    Isolate compromised components
    Alert system administrators
else if T is identified as moderate then
    Update access control policies
    Alert system administrators
else
    Log the event for further analysis
end if
```

3.5.3 Scalability and Load Balancing

To handle the high volume of data and maintain performance, the deployment architecture must be scalable. Load balancing techniques are employed to distribute the processing load across multiple instances, ensuring efficient resource utilization and minimizing latency.

3.5.4 Continuous Monitoring and Updating

The deployed models require continuous monitoring to ensure their performance remains optimal over time. This includes tracking the accuracy, false positive rates, and response times. Periodic updates and retraining may be necessary to adapt to new and evolving threats.

3.5.5 Security and Compliance

Ensuring the deployed models and the deployment process adhere to security best practices and regulatory requirements is crucial. This includes implementing access controls, encryption, and regular security audits to safeguard sensitive data and maintain compliance with industry standards.

By following these steps, the machine learning models can be effectively deployed to enhance the security of cloud applications, providing robust and scalable solutions for real-time threat detection and prevention.

This comprehensive implementation process ensures the effectiveness and efficiency of the proposed approach in enhancing security in cloud applications.

IV. Experimental Results Analysis

The experimental evaluation of the proposed approach is conducted to assess its effectiveness in detecting and mitigating security threats in cloud applications.

4.1 Experimental Setup

The experiments are performed in a simulated cloud environment using representative datasets and realistic attack scenarios. The simulated environment mimics the characteristics of real-world cloud applications and provides a controlled setting for evaluating the performance of the proposed approach. The UNSW NB15 dataset, known for its diverse attack types and comprehensive features, is used for training and testing the models.

4.2 Evaluation Metrics

Several evaluation metrics are considered to assess the performance of the proposed approach, including:

- Detection Accuracy: The percentage of security threats correctly identified by the system.
- False Positive Rate (FPR): The percentage of benign activities incorrectly classified as security threats.
- Precision: The proportion of true positive identifications among all positive identifications.
- Recall: The proportion of true positive identifications among all actual positives.
- F1-Score: The harmonic mean of precision and recall, providing a single metric that balances both concerns.
- Response Time: The time taken by the system to respond to detected threats.

These metrics provide comprehensive insights into the effectiveness, efficiency, and reliability of the proposed approach in enhancing security in cloud applications.

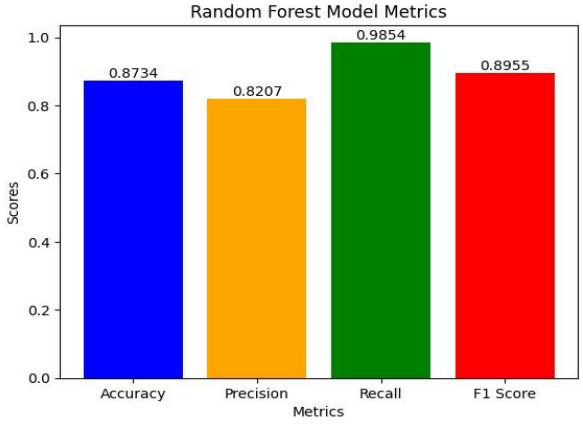
4.3 Experimental Findings

The experimental results demonstrate the effectiveness of the proposed approach in mitigating security threats in cloud applications. The following subsections present detailed findings based on different evaluation metrics.

4.3.1 Performance Metrics

Figure 2 shows the performance Metrics for Random Forest across different attack scenarios, while Table 4 presents the accuracy, precision, recall, and F1-score for various machine learning algorithms.

Figure 2: Performance metrics of Random Forest Metrics



The results indicate high detection accuracy and a low false positive rate, suggesting that the approach effectively distinguishes between benign and malicious activities.

4.3.2 Confusion Matrix

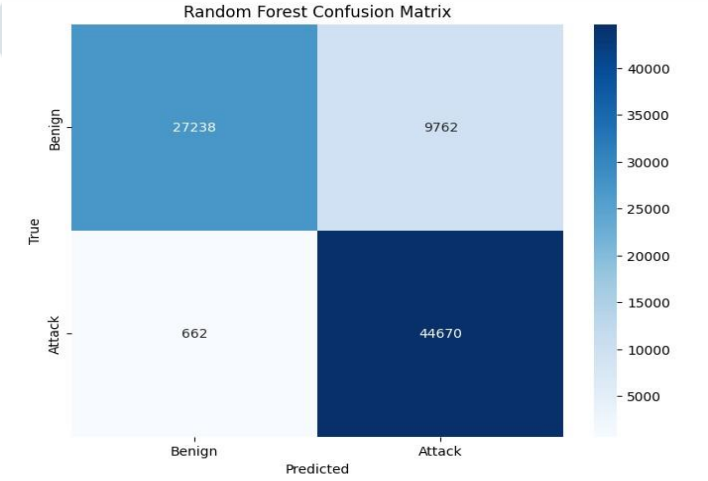
The confusion matrix provides a detailed breakdown of the classification performance of the models, showing the counts of true positives, false positives, true negatives, and false negatives. This helps in understanding the types of errors made by the models.

Figure 3 shows the confusion matrix for the Random Forest model. The high number of true positives and true negatives, combined with the low number of false positives and false negatives, indicates the model's robustness in identifying both benign and malicious activities.

Table 4: Precision, Recall, and F1-Score for different ML Algorithms

Algorithms	Accuracy	Precision	Recall	F1-Score
Random Forest	0.8734	0.8207	0.9854	0.8955
Logistic Regression	0.7070	0.6598	0.9658	0.7840
Gradient Boosting	0.8235	0.7594	0.9946	0.8612
k-Nearest Neighbors	0.7861	0.7400	0.9428	0.8292
Decision Tree	0.8571	0.8153	0.9572	0.8806
Support Vector Machine	0.6329	0.6171	0.8780	0.7248
Neural Network	0.7070	0.6598	0.9658	0.7840

Figure 3: Confusion Matrix for Random Forest model



4.3.3 Receiver Operating Characteristic Curve (ROC)

The Receiver Operating Characteristic (ROC) curve is used to evaluate the trade-off between the true positive rate and the false positive rate. The Area Under the Curve (AUC) is a single scalar value that summarizes the performance of the model across all classification thresholds.

Figure 4: ROC Curve for different ML Algorithms

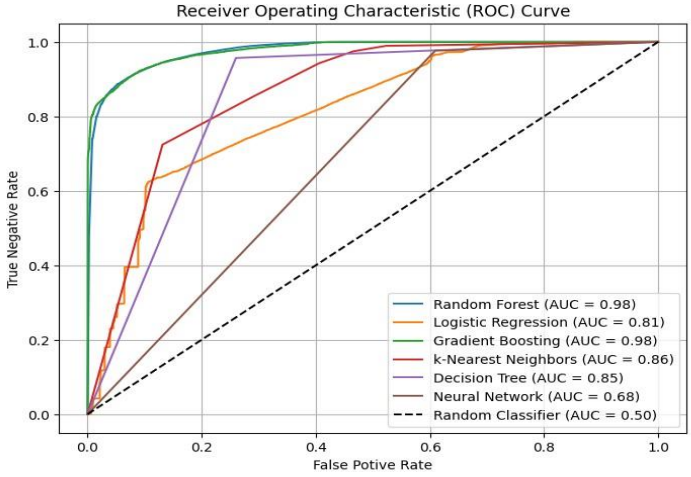


Figure 4 shows the ROC curves for various machine learning algorithms, with their respective AUC values:
The high AUC values for Random Forest and Gradient Boosting indicate superior performance, while the lower AUC for Neural Networks suggests room for improvement.

V. Conclusions and Future Scope

In conclusion, this research presents a comprehensive approach to enhancing the security of cloud applications through intelligent threat detection and prevention. By integrating traditional security measures with advanced machine learning techniques, the proposed approach addresses the unique security challenges inherent in shared and dynamic cloud environments. Our findings demonstrate that leveraging machine learning algorithms for real-time threat detection and classification significantly improves detection accuracy, reduces false positive rates, and accelerates response times compared to conventional security methods. Furthermore, the adoption of robust encryption, access controls, and audit trails enhances data privacy, integrity, and compliance, thereby ensuring the confidentiality and integrity of sensitive data in cloud applications.

The experimental results validate the efficacy of the proposed approach, indicating its potential to bolster the resilience of cloud applications against a wide range of cyber threats, including Fuzzers, data breaches, and DoS attacks. The integration of these technologies provides a more robust and adaptive security framework that can effectively respond to the evolving threat landscape in cloud computing environments.

5.1 Future Scope

Despite the promising results, several challenges and opportunities for future research remain. Future research directions may focus on addressing the limitations of the proposed approach and exploring new avenues for enhancing security in cloud applications. Future work can explore the following areas:

This could involve investigating advanced machine learning algorithms, leveraging big data analytics, and incorporating threat intelligence sharing mechanisms[11].

- Scalability and Performance Optimization: Investigate techniques to enhance the scalability and efficiency of the proposed approach, particularly in large-scale cloud deployments with high volumes of data and transactions. Optimization strategies for machine learning models and data processing pipelines can further improve performance.
- Threat Intelligence Sharing with Blockchain: Investigate the integration of blockchain technology for secure threat intelligence sharing among cloud environments. By utilizing blockchain's decentralized and immutable ledger, threat intelligence data can be securely shared and validated across multiple stakeholders, enhancing collaborative defense mechanisms and ensuring the integrity of shared data.
- Emergence of New technologies: The emergence of new technologies such as edge computing, artificial intelligence (AI), and quantum computing poses both opportunities and challenges for threat detection and prevention in cloud applications. Research efforts in these areas can lead to innovative solutions that enhance the security and resilience of cloud environments against evolving cyber threats.

By addressing these areas, future research can further enhance the security of cloud applications and contribute to the development of more resilient and trustworthy cloud computing infrastructures. The continued evolution of intelligent threat detection and prevention techniques will be essential in safeguarding cloud environments against the ever-changing landscape of cyber threats.

VI. Acknowledgment

The authors would like to express their sincere gratitude to Dr. Rachna Rajput for their invaluable guidance, mentorship, and support throughout the course of this research. Their expertise, insights, and constructive feedback have been instrumental in shaping the direction and methodology of the study.

Special thanks are extended to the staff of the Department of Computer Science and Engineering, Guru Kashi University, Talwandi Sabo, (India) for their contributions, collaborative efforts, and insightful discussions, which have enriched the research process and enhanced the quality of the findings.

The authors also acknowledge the support and assistance provided by the IT department at Guru Kashi University, Talwandi Sabo, (India) for granting access to computing resources, software tools, and datasets necessary for conducting experiments and analysis.

Furthermore, the authors would like to thank Pexwave Academy (https://pexwave.com/academy) for their financial support, which facilitated the execution of this research endeavor.

Lastly, the authors extend their appreciation to their families for their unwavering encouragement, understanding, and patience throughout the research journey.

References

[1] Waqas Ahmad, Aamir Rasool, Abdul Rehman Javed, Thar Baker, and Zunera Jalil. Cyber security in iot-based cloud computing: A comprehensive survey. *Electronics*, 11(1):16, 2021.

[2] Akhil Behl. Emerging security challenges in cloud computing: An insight to cloud security challenges and their mitigation. In *2011 World Congress on Information and Communication Technologies*, pages 217– 222. Ieee, 2011.

[3] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2), 2012.

[4] William Y Chang, Hosame Abu-Amara, and Jessica Feng Sanford. *Transforming enterprise cloud services*. Springer Science & Business Media, 2010.

[5] Guozhu Dong and Huan Liu. *Feature engineering for machine learning and data analytics*. CRC press, 2018.

[6] Rania El-Gazzar, Eli Hustad, and Dag H Olsen. Understanding cloud computing adoption issues: A delphi study approach. *Journal of Systems and Software*, 118:64–84, 2016.

[7] James Henry and Sameer Mirza. The future is in the cloud: Revolutionizing business with cloud computing. Technical report, EasyChair, 2024.

[8] I Indu, PM Rubesh Anand, and Vidhyacharan Bhaskar. Identity and access management in cloud environment: Mechanisms and challenges. *Engineering science and technology, an international journal*, 21(4):574– 588, 2018.

[9] Nancy J King and VT Raja. Protecting the privacy and security of sensitive customer data in the cloud. *Computer Law & Security Review*, 28(3):308–319, 2012.

[10] Rakesh Kumar and Rinkaj Goyal. Assurance of data security and privacy in the cloud: A three-dimensional perspective. *Software Quality Professional*, 21(2):7–26, 2019.

[11] Ashok Manoharan and Mithun Sarker. Revolutionizing cybersecurity: Unleashing the power of artificial intelligence and machine learning for next-generation threat detection. *DOI: https://www. doi. org/10.56726/IRJMETS32644*, 1, 2023.

[12] Ali Bou Nassif, Manar Abu Talib, Qassim Nasir, Halah Albadani, and Fatima Mohamad Dakalbab. Machine learning for cloud security: a systematic review. *IEEE Access*, 9:20717–20735, 2021.

[13] Varun Shah. Machine learning algorithms for cybersecurity: Detecting and preventing threats. *Revista Espanola de Documentacion Cientifica*, 15(4):42–66, 2021.

[14] Amir Mohamed Talib, Rodziah Atan, Rusli Abdullah, and Masrah Azrifah. Cloudzone: Towards an integrity layer of cloud data storage based on multi agent system architecture. In *2011 IEEE Conference on Open Systems*, pages 127–132. IEEE, 2011.

[15] David Wells. Unsw_nb15, 2017. Accessed: 2024-01-14.