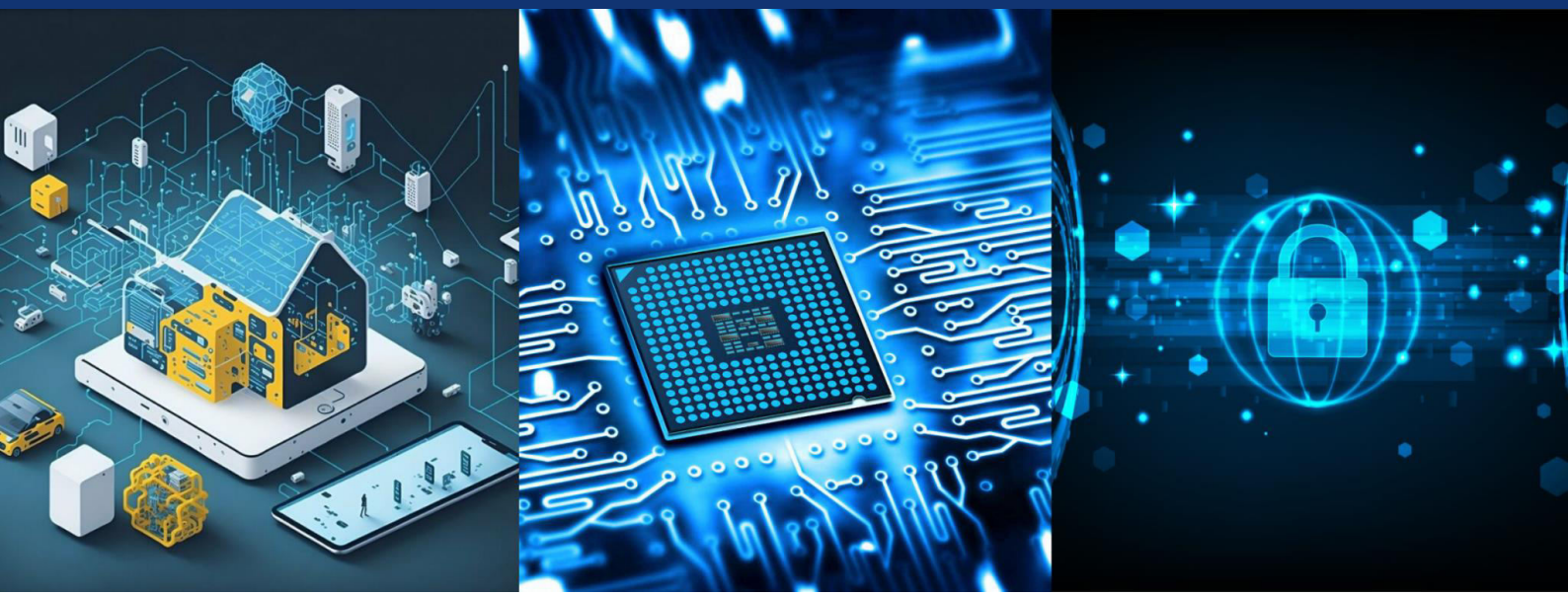
IJIRCCE

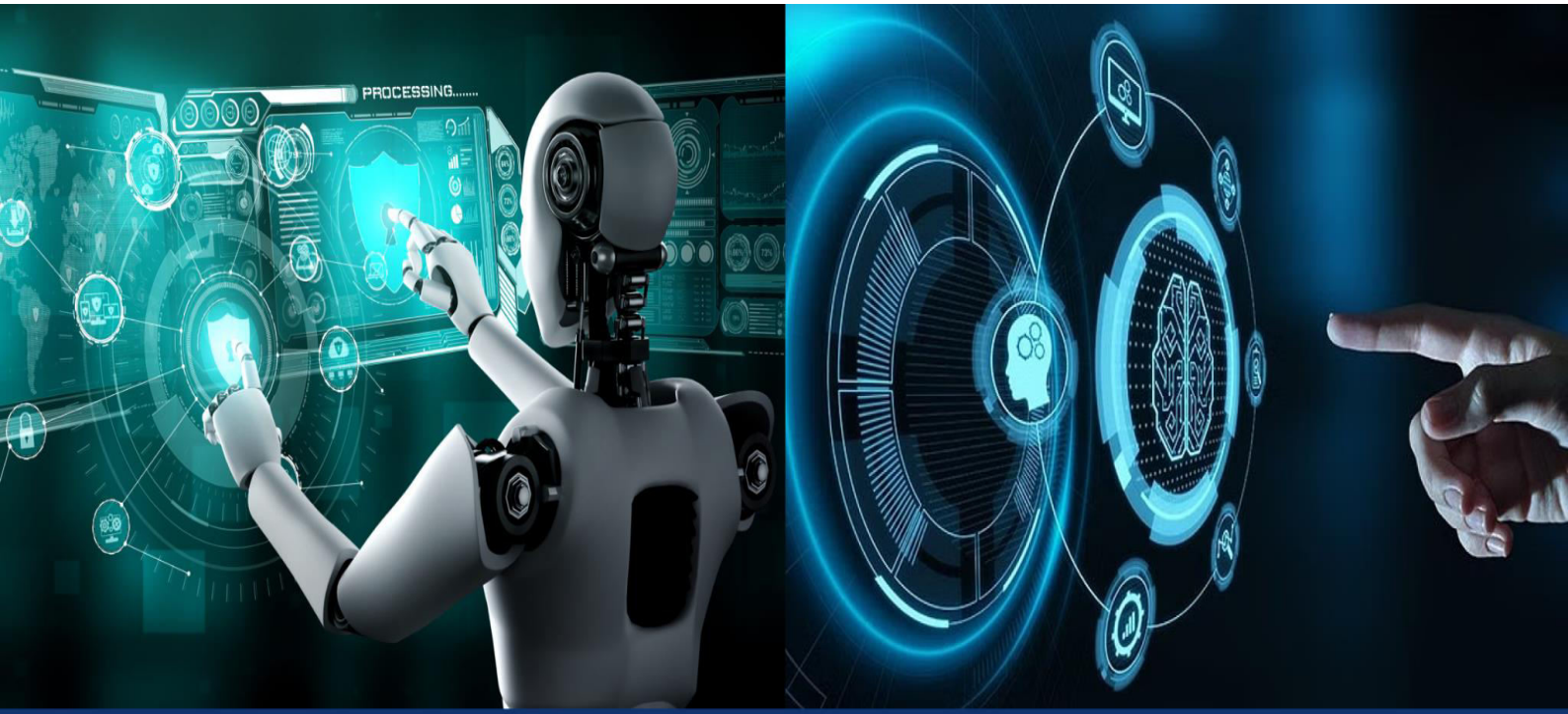
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.927

Volume 14, Issue 8, August 2026

© 2026 IJIRCCE | Volume 14, Issue 8, August 2026 DOI: 10.15680/IJIRCCE.2026.1408011

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Ransomware Early Detection System

Dharavath Anjali, Dr. M. Dhanalakshmi

Post Graduate Student, Department of Computer Science and Engineering, Jawaharlal Nehru Technological University, Hyderabad, India

Professor, Department of Computer Science and Engineering, Jawaharlal Nehru Technological University, Hyderabad, India

ABSTRACT: Ransomware has become one of the most destructive forms of cyberattacks, causing severe financial losses, data breaches, and disruption of critical services across organizations worldwide. Traditional signature-based antivirus solutions often fail to detect newly emerging and zero-day ransomware variants due to their dependence on previously known attack signatures. To overcome these limitations, this paper proposes the Ransomware Early Detection System (REDS), an intelligent machine learning-based framework designed to identify ransomware activities at an early stage before extensive file encryption occurs. The proposed system continuously monitors system behavior by analyzing file operations, process execution patterns, entropy values, CPU utilization, memory consumption, and other behavioral characteristics. The collected information is preprocessed and transformed into meaningful feature vectors that are used to train multiple machine learning classifiers, namely Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Support Vector Machine (SVM). The trained models classify system activities as either benign or ransomware in real time. Whenever suspicious behavior is identified, the system immediately generates security alerts and displays the detection results through an interactive Streamlit dashboard. Experimental evaluation demonstrates that the proposed framework provides excellent ransomware detection performance. Among the evaluated classifiers, Random Forest achieved an accuracy of 99.62%, precision of 99.62%, recall of 99.72%, and F1-score of 99.67%, outperforming the remaining models. The proposed REDS framework successfully combines real-time behavioral monitoring with machine learning to provide accurate, efficient, and proactive ransomware detection while minimizing false alarms. The system offers an effective solution for protecting organizational data and strengthening cybersecurity against evolving ransomware threats.

KEYWORDS: Ransomware Detection, Machine Learning, Cybersecurity, Behavioral Analysis, Random Forest, XGBoost, SVM, Threat Classification, Malware Detection, Real-Time Monitoring.

I. INTRODUCTION

Ransomware is one of the fastest-growing categories of cybercrime and has emerged as a major threat to individuals, organizations, healthcare institutions, financial services, and government agencies. Modern ransomware attacks encrypt valuable user data and demand financial payment in exchange for the decryption key. The rapid evolution of ransomware families and the emergence of zero-day variants have significantly reduced the effectiveness of traditional signature-based antivirus systems, making early detection an essential requirement for modern cybersecurity.

Machine learning has recently demonstrated significant potential in detecting ransomware based on behavioral analysis rather than predefined signatures. Instead of identifying known malware signatures, machine learning algorithms analyze system activities such as file modifications, process behavior, CPU utilization, memory usage, and entropy values to recognize malicious behavior. This behavioral approach enables the detection of both known and previously unseen ransomware attacks before extensive file encryption takes place.

The proposed Ransomware Early Detection System (REDS) utilizes machine learning techniques to identify ransomware attacks during their initial execution stages. The system continuously monitors system activities, performs data preprocessing, extracts behavioral features, trains multiple classification models, and predicts malicious behavior in real time. Random Forest, XGBoost, and Support Vector Machine (SVM) classifiers are employed to compare detection performance and identify the most effective model. Once ransomware activity is detected, REDS immediately generates warning notifications and displays comprehensive monitoring information through an interactive Streamlit dashboard.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Unlike traditional antivirus software that depends primarily on static malware signatures, REDS combines behavioral analysis with machine learning to provide proactive ransomware detection. The proposed framework improves detection accuracy, reduces false positives, supports continuous monitoring, and enables timely response before significant data loss occurs. The system therefore contributes to strengthening endpoint security and improving resilience against modern ransomware attacks.

II. RELATED WORK

The increasing sophistication of ransomware has motivated researchers to develop advanced detection mechanisms using machine learning, deep learning, and behavioral analysis techniques. Several studies have demonstrated that monitoring system activities rather than relying solely on signature databases significantly improves ransomware detection performance.

Numerous researchers have proposed machine learning-based ransomware detection systems using algorithms such as Decision Trees, Random Forest, Support Vector Machine, Naïve Bayes, K-Nearest Neighbors, and Logistic Regression. These methods classify applications based on behavioral features including file access frequency, registry modifications, CPU utilization, memory usage, API call sequences, and entropy measurements. Among these algorithms, ensemble learning techniques such as Random Forest and Gradient Boosting have consistently achieved higher detection accuracy due to their ability to learn complex decision boundaries while minimizing overfitting.

Recent research has also explored deep learning approaches including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Transformer-based architectures. These models automatically learn hierarchical feature representations from system events and demonstrate improved performance in detecting unknown ransomware variants. However, their computational complexity often limits deployment in real-time endpoint monitoring environments.

Several commercial security solutions integrate behavioral monitoring with Endpoint Detection and Response (EDR) platforms to enhance ransomware protection. While these solutions provide effective monitoring capabilities, they often require expensive infrastructure and proprietary security platforms. Furthermore, many existing systems focus primarily on detection after encryption has already begun, limiting their effectiveness in preventing data loss.

The proposed Ransomware Early Detection System (REDS) addresses these limitations by integrating behavioral feature extraction with multiple machine learning classifiers, including Random Forest, XGBoost, and Support Vector Machine. The framework continuously monitors system activities, performs real-time prediction, generates instant security alerts, and visualizes detection results through a Streamlit dashboard. Experimental evaluation demonstrates that the proposed system achieves high detection accuracy while maintaining efficient real-time performance, making it suitable for practical deployment in endpoint security environments.

III. PROPOSED ALGORITHM

A. System Architecture:

Fig. 1 The proposed Ransomware Early Detection System (REDS) follows a structured machine learning-based architecture to detect ransomware attacks in real time. The system begins by collecting ransomware and benign datasets from cybersecurity repositories and system activity logs. The collected data is preprocessed by removing duplicate records, handling missing values, and normalizing features to improve data quality. After preprocessing, the Feature Engineering module extracts important behavioral features such as file modification rate, entropy value, CPU usage, memory consumption, and process behavior. These features are used to train machine learning models including Random Forest, XGBoost, and Support Vector Machine (SVM). The trained models are evaluated using Accuracy, Precision, Recall, and F1-Score, and the best-performing model is selected for deployment. During real-time monitoring, the Prediction Engine analyzes new system activities, classifies them as benign or ransomware, and forwards suspicious results to the Threat Identification module. If ransomware is detected, the Alert Generation module immediately notifies the user, while the Dashboard Visualization module displays live alerts, detection statistics, graphs, and model performance for continuous monitoring.

© 2026 IJIRCCE | Volume 14, Issue 8, August 2026 DOI: 10.15680/IJIRCCE.2026.1408011

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

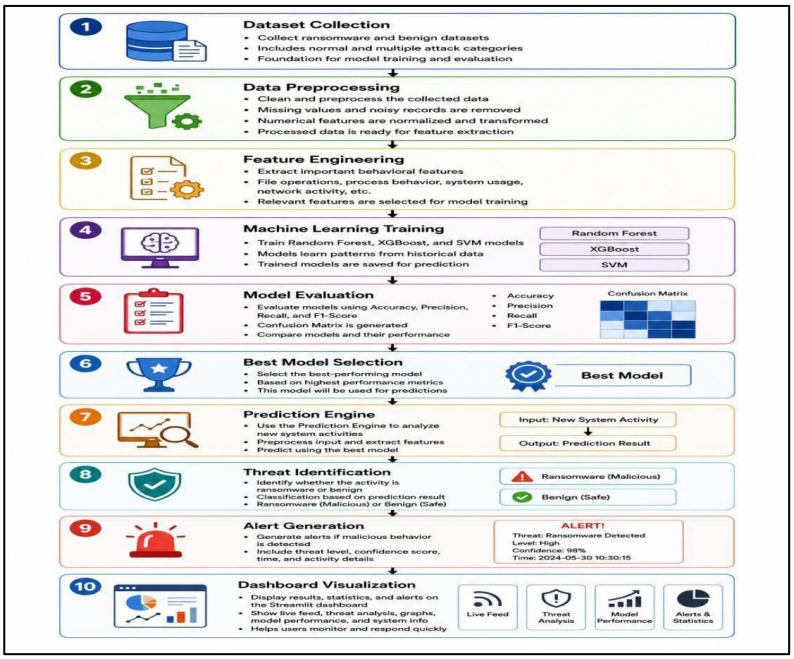


Fig. 1. Proposed System Architecture

B. Design Considerations:

1. Early Detection: Detects ransomware at an early stage before significant file encryption and data loss occur.
2. Behavior-Based Detection: Monitors file activities, process behavior, CPU usage, memory usage, and entropy instead of relying on signature-based detection.
3. Data Preprocessing: Cleans the dataset by removing duplicates, handling missing values, and normalizing data to improve model accuracy.
4. Feature Engineering: Extracts important features such as file modification rate, entropy, CPU usage, memory usage, and process behavior for effective detection.
5. Machine Learning Models: Uses Random Forest, XGBoost, and SVM to learn ransomware behavior and improve classification performance.
6. Best Model Selection: Evaluates models using Accuracy, Precision, Recall, and F1-Score to select the most accurate model for deployment.
7. Real-Time Monitoring: Continuously analyzes new system activities and predicts whether they are Benign or Ransomware.
8. Alert Generation: Generates instant alerts with threat details, confidence score, and detection time whenever ransomware is detected.
9. Dashboard Visualization: Provides a Streamlit dashboard to display live monitoring, alerts, threat statistics, graphs, and model performance.
10. Modular & Scalable Design: Supports easy integration of new datasets, machine learning algorithms, and future security enhancements.

C. Description of the Proposed Algorithm:

The proposed REDS algorithm begins by collecting ransomware and benign datasets containing system behaviors such as file operations, process execution, CPU usage, memory utilization, and network activities. The collected data is preprocessed by removing duplicate records, handling missing values, and normalizing numerical features. Next, the Feature Engineering module extracts important behavioral features, including entropy value, file modification rate, process behavior, CPU usage, and memory consumption, and converts them into feature vectors. These feature vectors are used to train Random Forest, XGBoost, and Support Vector Machine (SVM) models. The trained models are evaluated using Accuracy, Precision, Recall, and F1-Score, and the best-performing model is selected. During real-time operation, the Prediction Engine receives new system activities, extracts their features, and classifies them as Benign or Ransomware. If ransomware behavior is detected, the Threat Identification module confirms the attack and the Alert

© 2026 IJIRCCE | Volume 14, Issue 8, August 2026 DOI: 10.15680/IJIRCCE.2026.1408011

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Generation module immediately generates a warning notification. Finally, all prediction results, alerts, and performance statistics are displayed through the Streamlit dashboard, enabling continuous monitoring and early ransomware detection.

IV. PSEUDO CODE

Step 1: Load the feature-extraction-for-ransomware-detection-data ransomware and benign datasets collected from cybersecurity repositories and real-time system activity logs.
Step 2: Perform data preprocessing by:
· Removing missing and duplicate values.
· Eliminating irrelevant features.
· Handling inconsistent data records.
· Normalizing selected features for improved model performance.
Step 3: Separate the dataset into input features (X) and target labels (Y).
Step 4: Split the dataset into training and testing datasets.
Step 5: Perform feature engineering to identify the most significant behavioral features contributing to ransomware detection, including:
· File creation rate
· File modification rate
· File deletion rate
· Entropy value
· CPU utilization
· Memory usage
· Process execution behavior
· PE header features
Step 6: Initialize the machine learning classifiers: Random Forest (RF) , XGBoost and Support Vector Machine (SVM)
Step 7: Train the Random Forest, XGBoost, and SVM models using the training dataset.
Step 8: Test the trained models using the testing dataset.
Step 9: Calculate performance metrics such as: Accuracy,Precision,Recall,F1-Score and Confusion Matrix.
Step 10: Start real-time monitoring of the system using the Watchdog library.
Step 11: Monitor system activities including: File creation, File modification, File deletion,Process execution and CPU and memory utilization.
Step 12: Apply the same preprocessing and feature extraction steps to the monitored system activities.
Step 13: Pass the extracted feature vector to the trained machine learning model.
Step 14: Predict the system activity.
IF Prediction = Benign
Display "Normal System Activity"
Continue Monitoring
ELSE
· Display "Ransomware Detected"
· Proceed to Threat Identification
Step 15: Analyze the detected ransomware activity and identify:Threat type ,Confidence score, Detection time and Affected files.
Step 16: Generate appropriate response actions such as:
· Real-time alert generation
· Logging threat information
· Displaying warning notifications
· Updating the threat database
· Recording detection history
Step 17: Display real-time monitoring information, threat statistics, model performance metrics, confusion matrices, graphs, and alert history on the Streamlit Dashboard.
Step 18: Store prediction results, alerts, and monitoring logs for future analysis and model improvement.
Step 19: Continue monitoring incoming system activities for new ransomware events.
Step 20: End.

© 2026 IJIRCCE | Volume 14, Issue 8, August 2026 DOI: 10.15680/IJIRCCE.2026.1408011

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. EXPERIMENTAL RESULTS

A. Table and Graphs:

The proposed Ransomware Early Detection System (REDS) was implemented in Python and evaluated using ransomware and benign datasets. The Random Forest, XGBoost, and SVM models were evaluated using Accuracy, Precision, Recall, and F1-Score. Table 1 presents the performance comparison of all classifiers, while the graphs illustrate their ransomware detection performance.

ML Model	Accuracy	F1-Score	Recall	Precision
Random Forest	99.62%	99.67%	99.72%	99.62%
XGBoost	99.62%	99.66%	99.72%	99.60%
Support Vector Machine (SVM)	96.56%	96.99%	97.95%	96.05%

Table 1. Assessment of Performance

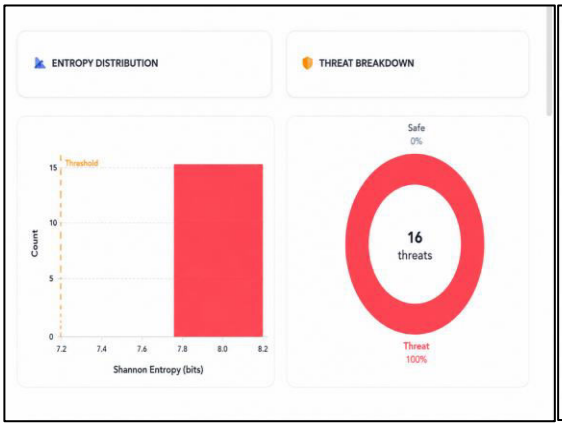


Fig.1.Entropy distribution and threat breakdown

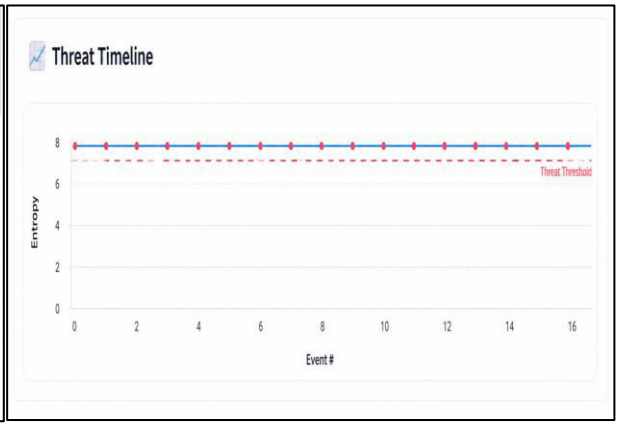


Fig.2.Entropy exceeds the predefined threat threshold

The proposed Ransomware Early Detection System (REDS) was implemented using the Python programming language and evaluated using ransomware and benign datasets collected from publicly available cybersecurity repositories and real-time system activity logs. The collected data was preprocessed to remove duplicate records, handle missing values, and normalize important features. Behavioral features such as file modification rate, entropy value, CPU utilization, memory usage, process execution behavior, and PE header attributes were extracted and used to train Random Forest, XGBoost, and Support Vector Machine (SVM) classifiers. The trained models were tested to accurately classify system activities as either Benign or Ransomware. Performance evaluation was carried out using standard machine learning metrics including Accuracy, Precision, Recall, F1-Score, and Confusion Matrix. Experimental results showed that the Random Forest model achieved the highest detection accuracy of 99.62%, outperforming XGBoost and SVM while maintaining low false positive and false negative rates, demonstrating its effectiveness in early ransomware detection.

In addition to ransomware detection, the proposed REDS framework provides real-time behavioral monitoring and automated alert generation to enhance system security. The system continuously monitors file operations, process activities, CPU usage, and entropy values using the Watchdog library. Whenever suspicious behavior is detected, the trained machine learning model classifies the activity and generates instant alerts containing the threat type, confidence score, detection time, and affected files. A Streamlit-based dashboard provides interactive visualization of live monitoring, threat analysis, model comparison graphs, confusion matrices, performance metrics, and alert history. The experimental results demonstrate that the proposed REDS framework effectively improves early ransomware detection, reduces potential data loss, enhances cybersecurity protection, and provides an efficient real-time monitoring solution for modern computing environments.

© 2026 IJIRCCE | Volume 14, Issue 8, August 2026 DOI: 10.15680/IJIRCCE.2026.1408011

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

B. *Evaluation and Interpretation of System Interfaces:*

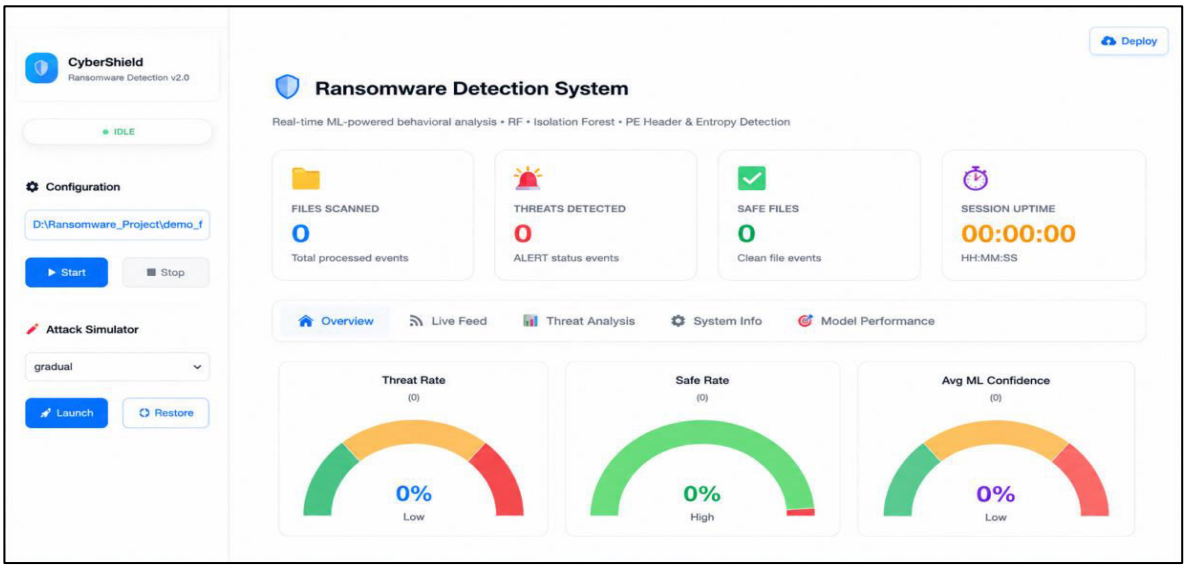


Fig. 3. Home Page

A real-time ransomware detection dashboard that monitors file activities, identifies potential threats using machine learning models, and provides instant alerts with system performance insights.

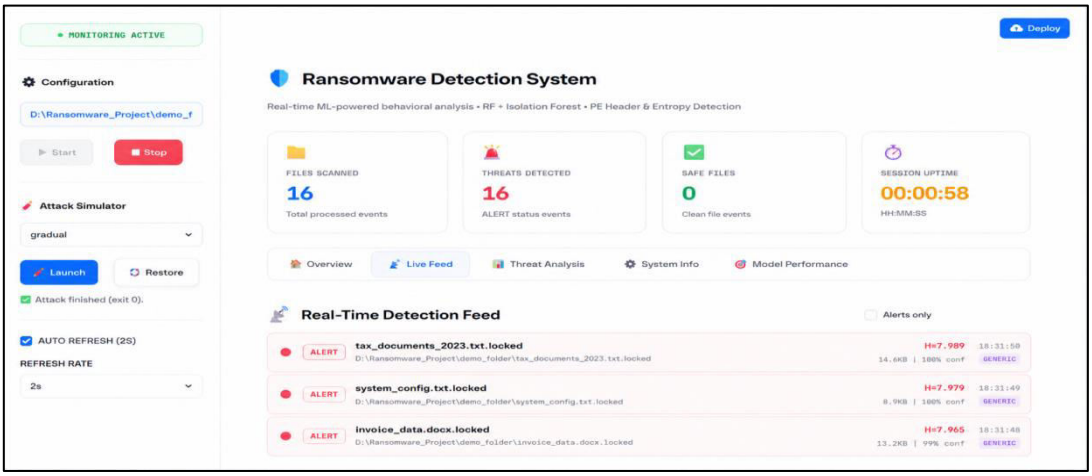


Fig.4. Dashboard After Dataset Upload

© 2026 IJIRCCE | Volume 14, Issue 8, August 2026 DOI: 10.15680/IJIRCCE.2026.1408011

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

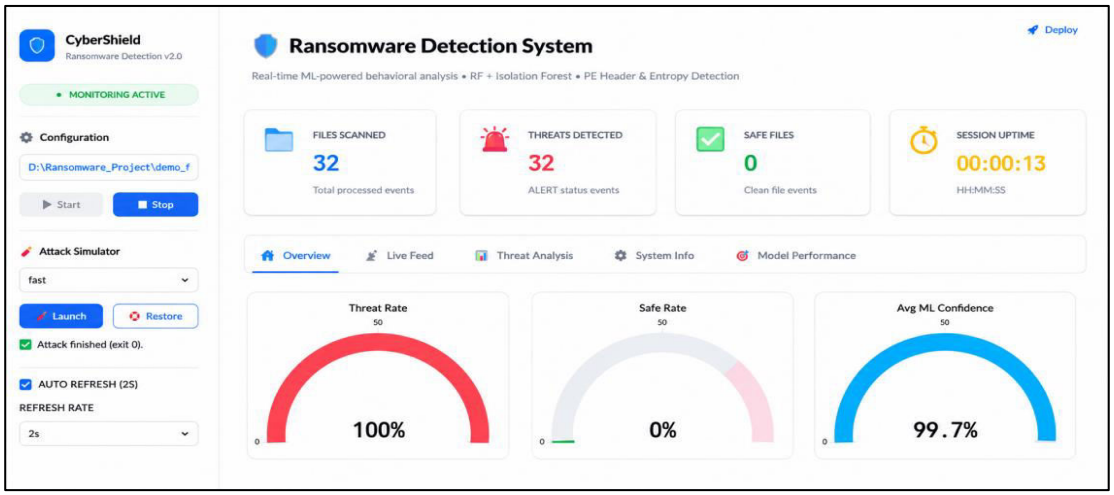
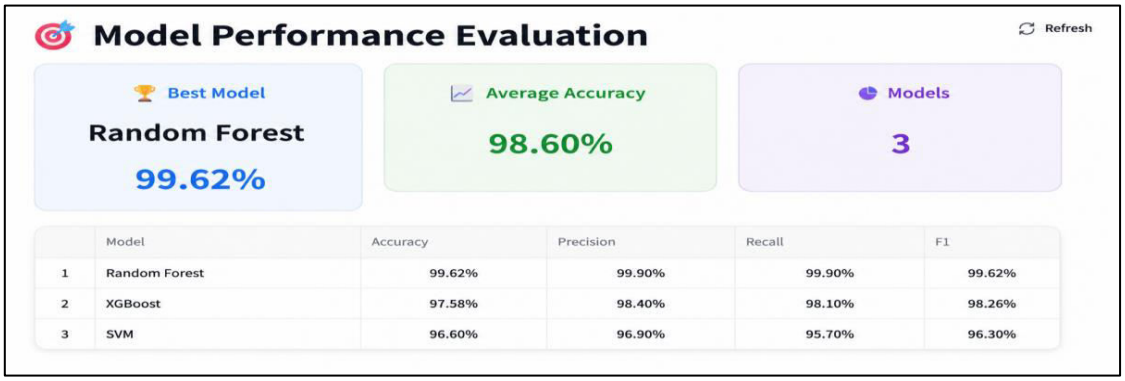


Fig.5. Instant security alerts fast by the machine learning-based monitoring system.



	Model	Accuracy	Precision	Recall	F1
1	Random Forest	99.62%	99.90%	99.90%	99.62%
2	XGBoost	97.58%	98.40%	98.10%	98.26%
3	SVM	96.60%	96.90%	95.70%	96.30%

Fig.6.Model Performance Evaluation

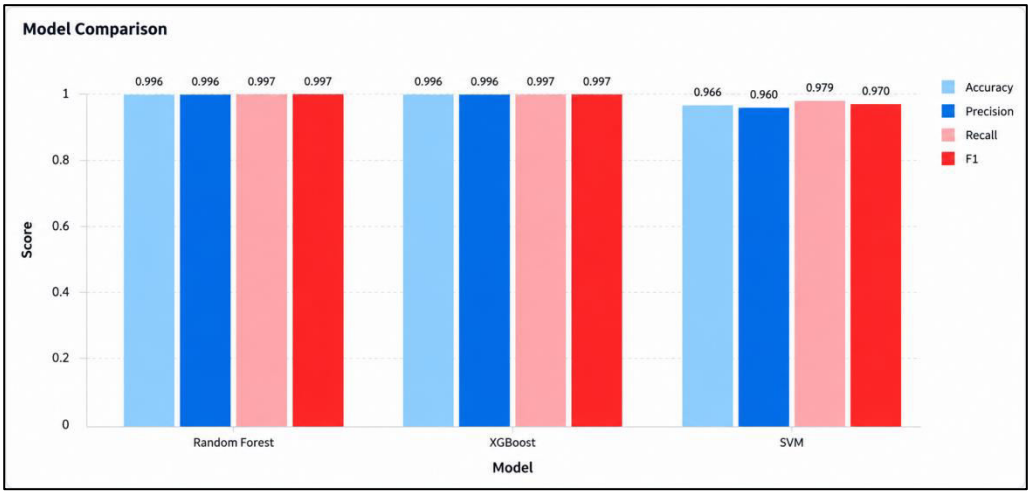


Fig.7.Model Comparison Graph

© 2026 IJIRCCE | Volume 14, Issue 8, August 2026 DOI: 10.15680/IJIRCCE.2026.1408011

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

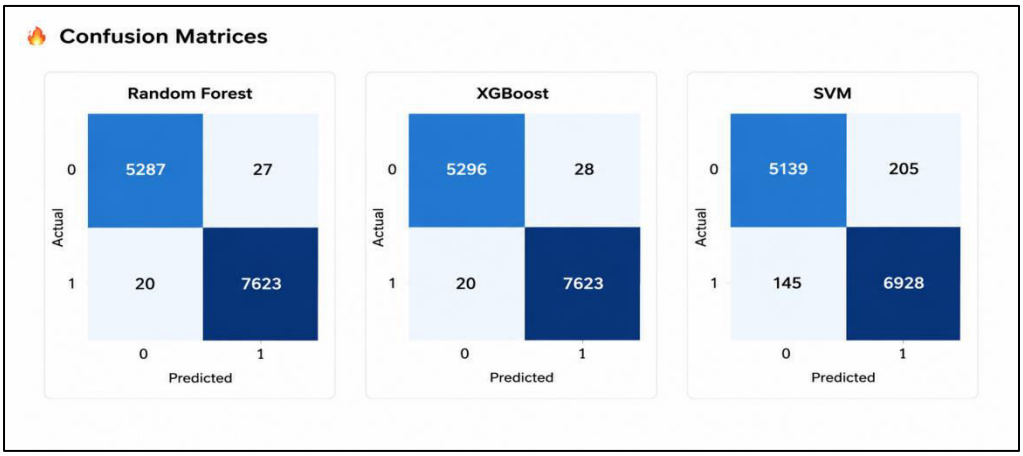


Fig.8. Model Performance Metrics

C. *Metrics for Evaluation:*

The performance of the proposed Ransomware Early Detection System (REDS) is evaluated using standard machine learning classification metrics, including Accuracy, Precision, Recall, and F1-Score. These metrics measure the effectiveness of the Random Forest, XGBoost, and Support Vector Machine (SVM) models in accurately distinguishing between benign system activities and ransomware attacks. The evaluation helps identify the best-performing model for real-time ransomware detection while minimizing false alarms and missed detections.

Let:

- TP (True Positives): Correctly identified ransomware activities.
- TN (True Negatives): Correctly identified benign (normal) activities.
- FP (False Positives): Benign activities incorrectly classified as ransomware.
- FN (False Negatives): Ransomware activities incorrectly classified as benign.

Accuracy:

Accuracy represents the proportion of correctly classified system activities to the total number of activities analyzed. It measures the overall effectiveness of the machine learning model in identifying both benign and ransomware activities.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{....Eq. (5)}$$

Precision:

Precision is the ratio of correctly predicted ransomware activities to the total number of activities predicted as ransomware. A higher precision indicates fewer false alarms.

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{....Eq.(6)}$$

Recall (Detection Rate):

Recall measures the proportion of actual ransomware attacks that are correctly detected by the machine learning model. A higher recall indicates the system can identify most ransomware attacks with fewer missed detections.

$$\text{Recall} = \frac{TP}{TP + FN} \quad \text{....Eq.(7)}$$

F1-Score

The F1-Score is the harmonic mean of Precision and Recall. It provides a balanced evaluation of the classifier by considering both false positives and false negatives, making it an important metric for ransomware detection.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

$$\text{F1Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad \text{....Eq.(8)}$$

VI. CONCLUSION AND FUTURE SCOPE

The proposed Ransomware Early Detection System (REDS) effectively detects ransomware attacks at an early stage using behavioral analysis and machine learning techniques. The system monitors file operations, process behavior, entropy values, CPU utilization, memory usage, and executable characteristics to accurately classify system activities as either benign or ransomware. By implementing Random Forest, XGBoost, and Support Vector Machine (SVM) algorithms, the framework achieves high detection performance, with the Random Forest classifier providing the highest accuracy of 99.62%. The integration of Watchdog for real-time monitoring and a Streamlit-based dashboard enables continuous threat detection, instant alert generation, and effective visualization of security events. Experimental results demonstrate that the proposed REDS framework provides accurate, reliable, and efficient ransomware detection while minimizing false positives and false negatives. Future work includes integrating deep learning models, cloud and IoT support, automated threat mitigation, SIEM/EDR integration, and Explainable AI (XAI) techniques to further improve detection accuracy, scalability, and real-time cybersecurity protection.

REFERENCES

[1] S. Saleh, "A Survey of Ransomware Detection Methods," 2025.
[2] J. Ispahany, "Ransomware Detection Using Machine Learning: A Review, Research Limitations and Future Directions," 2024.
[3] D. Smith, "Machine Learning Algorithms and Frameworks in Ransomware Detection," 2022.
[4] A. Almashhadani, M. Kaiiali, S. Sezer, and P. O'Kane, "A Multi-Classifier Network-Based Crypto Ransomware Detection System: A Case Study of Locky Ransomware," IEEE Access, vol. 7, pp. 47053–47067, 2019.
[5] N. Sgandurra, L. Muñoz-González, R. Mohsen, and E. C. Lupu, "Automated Dynamic Analysis of Ransomware: Benefits, Limitations and Use for Detection," arXiv preprint arXiv:1609.03020, 2016
[6] S. Scaife, H. Carter, P. Traynor, and K. R. B. Butler, "Cryptolock (and Drop It): Stopping Ransomware Attacks on User Data," Proceedings of the IEEE International Conference on Distributed Computing Systems (ICDCS), pp. 303–312, 2016.
[7] K. Cabaj and W. Mazurczyk, "Using Software-Defined Networking for Ransomware Mitigation: The Case of CryptoWall," IEEE Network, vol. 30, no. 6, pp. 14–20, 2016.
[8] X.. Ding and W. Feng, ‘‘Ananomaly detection method based on feature mining for wireless sensornetworks,’’ Int. J. Sens. Netw., vol. 36, no. 3,pp. 167–173, 2021
[9] N.Scaife, H. Carter, P. Traynor, and K. R. B. Butler, ‘‘CryptoLock (and drop It): Stopping ransomware attacks on user data,’’ in Proc. IEEE 36th Int. Conf. Distrib. Comput. Syst. (ICDCS), Jun. 2016, pp. 303–312
[10] K.Wang, M. Tong, J. Pang, J. Wang, and W. Han, ‘‘XRAD: Ransomware address detection method based on Bitcoin transaction relationships,’’ACM Trans. Web, vol. 18, no. 4, pp. 1–33, Oct. 2024.
[11] M. M. Ahmadian and H. R. Shahriari, ‘‘2entFOX: A framework for high survivable ransomwares detection,’’ in Proc. 13th Int. Iranian Soc. Cryptol. Conf. Inf. Secur. Cryptol. (ISCISC), Sep. 2016, pp. 79–84.
[12] J.Song, R. Paul, J. Yun, H. Kim, and Y. Choi, ‘‘CNN-based anomaly detection for packet payloads of industrial control system,’’ Int. J. Sens. Netw., vol. 36, no. 1, pp. 36–49, 2021.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details