iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** II **Month of publication:** February 2026

DOI: https://doi.org/10.22214/ijraset.2026.77437

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue II Feb 2026- Available at www.ijraset.com

CHRIS: Cyber Security Hub for Responsible Intelligence Scanning

Alby Ponnachan[1], Devi L R[2], Darsan S S[3], Karthik Lal S[4], Aron Colins Pereira[5]
Department of Computer Science and Engineering St. Thomas Institute for Science & Technology, Thiruvananthapuram, India

Abstract: ***The rapid growth of digital services and interconnected systems has led to an increase in sophisticated cyber threats such as phishing, malware, network intrusions, ransomware, and deepfake-based attacks. Conventional security solutions are often limited to single-domain detection and lack adaptability, explainability, and user-centric intelligence. This paper presents CHRIS (Cyber Security Hub for Responsible Intelligence System), a unified, web-based cybersecurity platform that integrates Machine Learning (ML), Deep Learning (DL), and Generative AI to provide comprehensive and explainable threat detection. CHRIS incorporates six security modules: phishing detection, malware detection, network intrusion detection, password strength evaluation, deepfake detection, and ransomware detection, all accessible through a single interface. Random Forest, XGBoost, and Xception models are employed for predictive analysis, while Google Gemini is integrated to generate natural-language explanations, recommendations, and interactive assistance via an AI-powered chatbot. Experimental analysis demonstrates that the proposed system achieves high detection accuracy while significantly improving interpretability and usability. The results highlight the effectiveness of combining predictive security analytics with Generative AI, making CHRIS a practical and scalable solution for next-generation cybersecurity applications.***

Keywords: Cybersecurity, Phishing Detection, Malware Detection, Network Intrusion Detection System, Deepfake Detection, Ransomware Detection, Machine Learning, Deep Learning, Generative AI, Google Gemini.

I. INTRODUCTION

The rapid digital transformation of modern society has significantly increased the complexity and frequency of cyber threats. Attacks such as phishing, malware, network intrusions, ransomware, and deepfake-based social engineering pose serious risks to individuals and organizations. Traditional security solutions often operate in silos, addressing only a single class of attack and relying heavily on rule-based mechanisms that fail to adapt to evolving threats. Recent advances in Machine Learning (ML) and Deep Learning (DL) have enabled intelligent threat detection by learning complex patterns from data, while Generative AI (GenAI) has opened new possibilities for explainability and user interaction.

This paper presents CHRIS (Cybersecurity Hub for Responsible Intelligence Scanning), a unified web-based cybersecurity platform that integrates multiple ML and DL models with Generative AI to provide comprehensive threat detection and user-centric risk intelligence. CHRIS consolidates six major cybersecurity modules—phishing detection, malware detection, network intrusion detection, password strength evaluation, deepfake detection, and ransomware detection—into a single platform, complemented by an AI-powered chatbot. The novelty of CHRIS lies in its holistic design and the 1 integration of Google Gemini to generate natural-language explanations, recommendations, and interactive cybersecurity guidance to clearly highlight the contributions of this work, the main contributions of CHRIS are summarized as follows

1) A unified, web-based cybersecurity platform that integrates multiple ML, DL, and Generative AI techniques to address diverse cyber threats within a single system.
2) Design and implementation of six independent yet interoperable security modules covering phishing, malware, network intrusion, password security, deepfake detection, and ransomware detection.
3) Adoption of explainable Machine Learning models combined with Google Gemini to provide human-readable risk explanations, recommendations, and interactive assistance.
4) A proactive ransomware detection mechanism based on real-time behavioural monitoring, automated quarantine, and AI-assisted incident analysis.
5) Comprehensive experimental evaluation demonstrating high detection accuracy and practical usability across multiple cybersecurity domains

II. RELATED WORK

Extensive research has explored ML and DL techniques for individual cybersecurity problems. For clarity, existing works can be broadly grouped into thematic categories covering phishing detection, malware detection, network intrusion detection, deepfake detection, and password security.

1) *Phishing and Malware Detection:* Deep learning-based malware detection systems have demonstrated strong performance by learning discriminative features from static and dynamic malware representations [1], [11]. Phishing detection has been addressed using traditional ML algorithms such as Random Forest (RF), Support Vector Machines (SVM), and ensemble approaches, as well as DL-based models that analyse URLs and webpage content [2], [9], [10]. These studies highlight the effectiveness of feature-driven and ensemble learning methods in identifying deceptive and malicious artifacts.
2) Network Intrusion Detection Systems: Network Intrusion Detection Systems (NIDS) have widely adopted XGBoost due to its high accuracy, scalability, and ability to handle imbalanced datasets [5], [6], [7]. Recent studies have also proposed hybrid and ensemble IDS models combining deep neural networks with XGBoost to improve detection of zero-day attacks [8], [18], [19], emphasizing the importance of robust learning techniques for complex network environments.
3) *Deepfake Detection:* Deepfake detection research has shown that CNN architectures such as Xception outperform conventional models in identifying manipulated images, particularly on benchmark datasets such as FaceForensics++ [4], [13], [14], [15]. These works demonstrate that deep convolutional features are highly effective in capturing subtle visual artifacts introduced during image manipulation.
4) *Password Security and Explainability:* Password strength evaluation has traditionally relied on rule-based metrics, but human-centric approaches such as zxcvbn provide more realistic strength estimation by modelling user behaviour [16]. While prior work focuses primarily on individual security domains, very limited research addresses the integration of multiple threat detection mechanisms with Generative AI driven explainability in a single platform. CHRIS addresses this gap by unifying heterogeneous ML/DL models with an AI assistant for enhanced usability and situational awareness.

III. PROPOSED METHADOLOGY

CHRIS– Cybersecurity Hub for Responsible Intelligence Scanning is designed to provide a unified, modular, and intelligent approach to detecting and mitigating multiple cybersecurity threats. The system follows a hybrid methodology by integrating machine learning–based detection with rule-based and heuristic techniques, ensuring both accuracy and real-time responsiveness.

The methodology begins with data acquisition, where inputs such as URLs, executable files, system activities, passwords, network traffic, or images are submitted by the user through a centralized interface. These inputs are pre-processed and routed internally by the CHRIS core engine to the appropriate detection module without direct user interaction with individual modules.

Each module then performs threat analysis using its respective approach. Machine learning models are employed for phishing URL detection and static malware detection, while heuristic and behaviour-based methods are used for ransomware detection and password strength evaluation. Planned modules such as Network Intrusion Detection and Deepfake Detection rely on deep learning architectures for advanced anomaly and pattern recognition.

Finally, the system generates detection results, alerts, and recommendations, which are displayed on a unified dashboard. All activities and outcomes are logged securely for auditing and future improvements. This modular and scalable methodology allows CHRIS to adapt to emerging threats, support phased development, and ensure effective real-time cybersecurity protection.

IV. SYSTEM ARCHITECTURE

CHRIS is designed using a layered and modular system architecture that enables seamless integration of multiple cybersecurity services while maintaining scalability, flexibility, and ease of deployment. The architecture is organized into four primary layers: the User Interface Layer, Application and API Layer, Intelligence and Analytics Layer, and Data and Monitoring Layer. This separation of concerns ensures that each component can evolve independently without affecting the overall system stability.

At the *User Interface Layer*, CHRIS provides a unified web-based dashboard developed using React.js. This layer serves as the primary interaction point for users, allowing them to submit URLs for phishing analysis, upload files for malware and deepfake detection, monitor ransomware alerts, evaluate password strength, and view network intrusion reports. The interface is designed to be intuitive and responsive, presenting model predictions alongside probability scores, alerts, and AI-generated explanations in a clear and user-friendly manner.

The *Application and API Layer* is implemented using FastAPI and acts as the central communication backbone of the system. Each cybersecurity module is exposed as an independent RESTful API endpoint, enabling concurrent processing of requests.

This layer handles request validation, routing, authentication, and response formatting. The modular API design allows individual services, such as phishing detection or ransomware monitoring, to be updated or scaled independently, supporting future extensions of the platform.

The Intelligence and Analytics Layer form the core of CHRIS and hosts all Machine Learning, Deep Learning, and Generative AI components. Random Forest models are used for phishing and malware detection, XGBoost is employed for network intrusion detection, and the Xception deep learning architecture is utilized for deepfake detection. These predictive models generate classification results and confidence scores, which are further enriched through integration with Google Gemini. The Generative AI component transforms technical outputs into natural-language explanations, risk summaries, and actionable recommendations, enhancing interpretability and user trust.

The *Data and Monitoring Layer* is responsible for handling datasets, feature extraction pipelines, model artifacts, and real-time system monitoring. Static datasets are used for training and evaluation, while dynamic data streams—such as file system events for ransomware detection—are captured in real time using the watchdog library. This layer also manages secure storage for quarantined files, logs, and alert histories, enabling auditing and forensic analysis. Overall, the system architecture of CHRIS emphasizes modularity, explainability, and real- time responsiveness. By combining service-oriented design with intelligent analytics and Generative AI driven insights, the architecture supports a holistic and scalable cybersecurity platform capable of addressing diverse and evolving threat landscapes

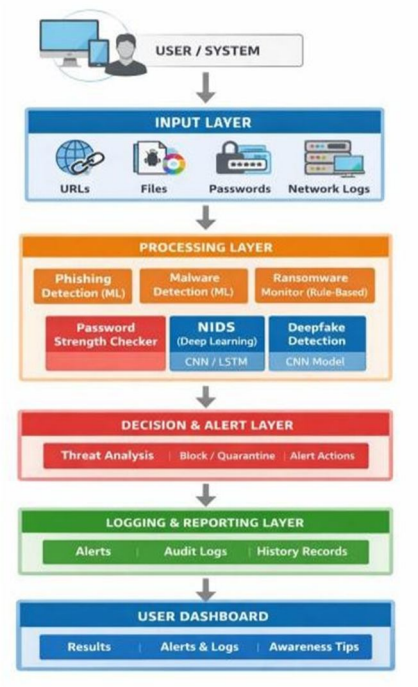


Fig.1 System Architecture

V. USER INTERFACE LAYER

The User Interface (UI) Layer is the final layer of the CHRIS framework, responsible for providing an interactive and user-friendly dashboard that presents all detection results in a single centralized view. Since CHRIS integrates multiple threat detection modules, the UI layer plays a critical role in ensuring that users can easily understand security status without needing technical expertise.

This layer mainly focuses on visualization, alert communication, and user awareness. Whenever a threat is detected (phishing URL, malware file, ransomware activity, network intrusion, weak password, or deepfake media), the UI displays the output in real time along with severity level or confidence score. It also provides clear warnings and recommended actions such as avoiding a URL, quarantining a file, changing a password, or verifying suspicious media. Another important function of this layer is security awareness and transparency. Instead of showing only "malicious/benign," the UI explains why the input was flagged, helping users understand attack patterns and improve their cyber hygiene. The UI also maintains a summary of logs, module status, and recent detections, enabling administrators to track system security posture and respond quickly to incidents.

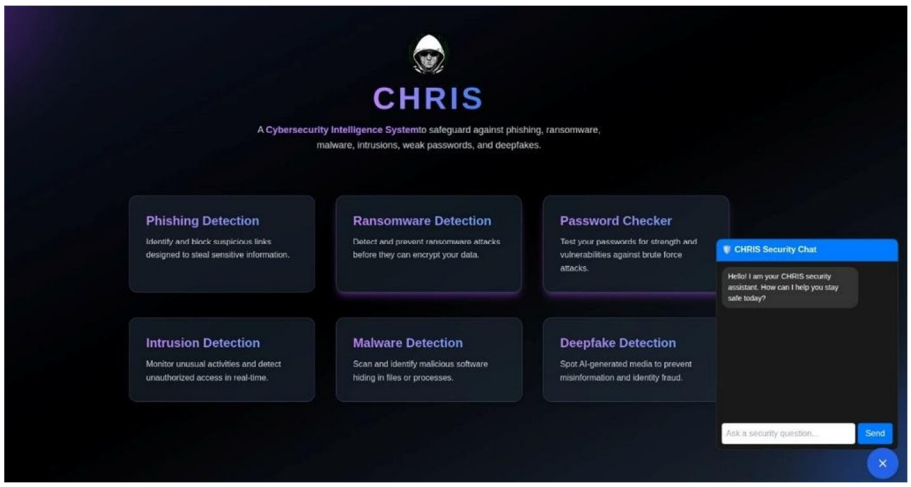


Fig 2. UI/UX

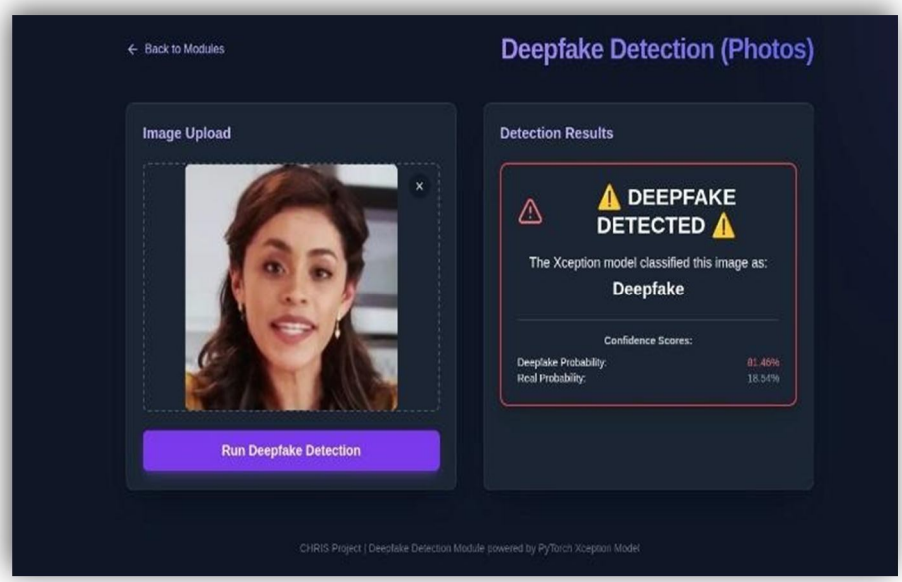


Fig 3. Deepfake Detector

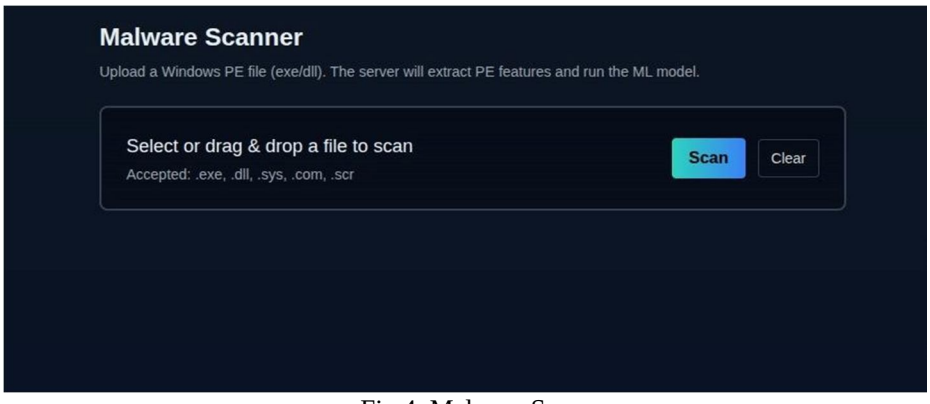


Fig 4. Malware Scanner

Fig 5. Passwords Cheeker and suggestion

Fig 6. Ransomware Detector

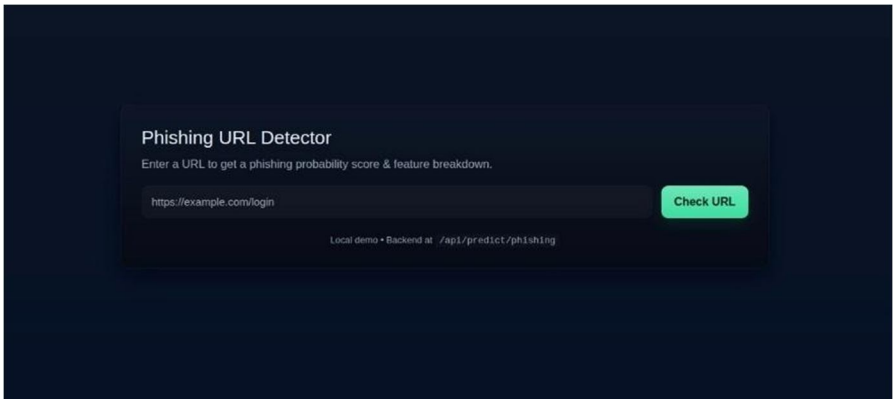


Fig 7. Phishing URL Detection

Fig 8. NIDS

VI. SYSTEM MODULES

This section describes the overall methodological pipeline adopted in CHRIS, followed by a detailed explanation of each security module. The system follows a common workflow consisting of data collection, feature extraction or representation, model training, and real-time inference. Depending on the nature of the threat, Machine Learning or Deep Learning models are employed to analyze structured data, executable features, network traffic, images, or system behaviour. The outputs of these models are further enriched using Generative AI to provide explainable, user- friendly insights. The following subsections present the methodology of each module in detail.

A. Phishing Detection

The phishing detection module in CHRIS is designed to identify malicious URLs that attempt to deceive users into revealing sensitive information such as login credentials or financial data. The system employs a Random Forest (RF) classifier, chosen for its robustness, ability to handle heterogeneous feature sets, and strong performance in phishing detection tasks reported in prior studies [2], [9], [10]. The model is trained on a dataset consisting of both phishing and legitimate URLs. A comprehensive set of lexical, structural, and host-based features is extracted from each URL.

Lexical features include URL length, number of special characters, presence of suspicious keywords, use of IP addresses instead of domain names, and abnormal token patterns. Structural features capture properties such as subdomain depth, URL redirection behaviour, and encoding techniques, while host-based features include domain age, DNS records, and HTTPS usage when available. During inference, the Random Forest classifier outputs a probability score representing the likelihood that a given URL is phishing. Instead of providing a binary decision alone, CHRIS emphasizes risk awareness by presenting this probability to the user, enabling informed decision-making. To further enhance interpretability, the extracted features contributing to the prediction are summarized and passed to Google Gemini, which generates a detailed natural-language explanation. This explanation highlights suspicious characteristics of the URL, such as misleading domain names or excessive use of obfuscation techniques, and contextualizes the risk in a user-friendly manner. The integration of Generative AI transforms traditional phishing detection into an explainable and interactive process, bridging the gap between technical ML outputs and end-user understanding. This approach not only improves detection accuracy but also increases user trust and cybersecurity awareness, making the phishing detection module a key component of the CHRIS platform.

B. Malware Detection

The malware detection module in CHRIS focuses on identifying malicious executable files before execution, thereby reducing the risk of system compromise. A static analysis-based approach is adopted, leveraging a Random Forest (RF) classifier due to its ability to handle high- dimensional feature spaces, robustness against overfitting, and strong performance in malware classification tasks reported in existing literature [1], [11], [17]. In this module, Portable Executable (PE) files are analysed to extract a rich set of discriminative features without running the file in a live environment. The extracted features include PE header attributes, section-based statistics, entropy measures, imported and exported functions, resource information, and f ile metadata. Entropy-related features are particularly important, as malware often employs packing or encryption techniques that result in unusually high entropy values. Similarly, abnormal import tables and suspicious API usage patterns serve as strong indicators of malicious intent. Prior to model training, feature selection is performed using ensemble-based techniques to eliminate redundant and irrelevant attributes, thereby reducing computational overhead and improving generalization. The selected features are then used to train the Random Forest classifier on labeled datasets consisting of benign and malicious samples. During inference, the trained model predicts whether a given file is benign or malicious, along with an implicit confidence derived from ensemble voting. The choice of Random Forest enables CHRIS to capture complex, non-linear relationships between file features while maintaining interpretability through feature importance analysis. This makes the 4-malware detection module both efficient and explainable, allowing it to be seamlessly integrated into a real-time web-based environment. By relying on static analysis, the module avoids the risks and resource costs associated with dynamic execution, making it suitable for rapid malware screening in practical cybersecurity deployments.

C. Network Intrusion Detection System

The Network Intrusion Detection System (NIDS) module in CHRIS is designed to identify malicious activities and anomalous patterns within network traffic generated by a personal computer or local network environment. Unlike traditional signature-based IDS solutions, this module adopts a data driven approach using XGBoost, a powerful gradient boosting algorithm well-suited for structured and tabular network traffic data

[5]–[7]. XGBoost is selected due to its high detection accuracy, resistance to overfitting, and ability to handle imbalanced datasets, which are common characteristics of real-world intrusion detection scenarios.

The methodology begins with the collection of network traffic features derived from packet-level or f low-level statistics. These features typically include protocol types, connection duration, packet counts, byte transfer statistics, flag indicators, and temporal behavior metrics. Such attributes capture both short-term anomalies, such as port scanning and denial-of-service attempts, and long-term behavioural deviations indicative of stealthy intrusions. The extracted features are pre-processed through normalization and encoding techniques before being fed into the XGBoost classifier. During training, the model learns complex, non-linear relationships between benign and malicious traffic patterns by iteratively optimizing decision trees using gradient boosting. The trained model is capable of detecting known attack categories as well as previously unseen (zero-day) intrusions by identifying deviations from learned normal traffic behaviour.

During deployment, the NIDS module performs near real-time analysis of incoming network data and assigns each instance a classification label along with a confidence score. These outputs enable timely alerts and situational awareness for users. The use of XGBoost also allows limited interpretability through feature importance analysis, helping security analysts understand which traffic attributes contributed most to intrusion detection decisions. This makes the NIDS module both effective and practically deployable within the CHRIS platform.

D. Password Strength Checker

The password strength checker module in CHRIS is designed to evaluate and improve user passwords using a human-centric security approach. Instead of relying solely on rigid rule-based policies, the system employs the widely adopted zxcvbn library, which estimates password strength based on the number of guesses required to crack a password [16]. This approach accounts for real-world password creation habits, including the use of common words, keyboard patterns, dates, and personal information.

When a user submits a password for evaluation, zxcvbn analyses its structural composition and assigns a strength score along with estimated cracking times under various attack models. This feedback provides a more realistic assessment of password security compared to traditional metrics such as length or character variety alone. Weak passwords are flagged, and detailed explanations are generated to inform users why a particular password is vulnerable. To further enhance usability and security awareness, Google Gemini is integrated into this module.

When a weak password is detected, Gemini generates context-aware suggestions to improve the 5 existing passwords while preserving memorability. Additionally, users can request entirely new strong passwords, which are generated based on best practices such as sufficient length, randomness, and resistance to dictionary and brute-force attacks. By combining zxcvbn's quantitative strength estimation with Generative AI-driven recommendations, this module not only evaluates password security but actively assists users in adopting stronger authentication practices. This interactive design significantly improves user engagement and reduces the likelihood of password-related security breaches.

E. Deepfake Detection

The deepfake detection module in CHRIS addresses the growing threat of AI-generated and manipulated media used in misinformation, fraud, and social engineering attacks. This module employs the Xception deep learning architecture, a convolutional neural network known for its effectiveness in image forensics and manipulation detection [4], [13]–[15]. Xception leverages depth wise separable convolutions, enabling efficient learning of fine-grained visual artifacts introduced during image synthesis and facial manipulation.

Unlike many prior works that rely on the FaceForensics++ benchmark, the proposed system is trained using publicly available deepfake image datasets obtained from Hugging Face. These datasets consist of a diverse collection of real and synthetically manipulated facial images generated using multiple deepfake and generative techniques. The use of Hugging Face datasets provides flexibility, ease of access, and diversity in manipulation styles, making the model more adaptable to real-world deepfake scenarios. During preprocessing, facial regions are detected and cropped where applicable, resized to a fixed resolution, and normalized to ensure consistent input representation. Data augmentation techniques such as horizontal flipping, brightness variation, and compression artifacts are applied to improve generalization and robustness.

The Xception network is then fine-tuned on the Hugging Face dataset to learn discriminative spatial features that capture subtle inconsistencies in texture, illumination, and facial boundaries that are often imperceptible to human observers. During inference, the trained model classifies an input image as either genuine or deepfake, producing a confidence score that reflects the likelihood of manipulation. This probabilistic output enables CHRIS to provide nuanced risk assessments rather than strict binary decisions.

The deepfake detection module plays a crucial role in mitigating identity impersonation, fake profile creation, and AI driven social engineering attacks. By leveraging Xception and diverse datasets from Hugging Face, the proposed approach achieves a strong balance between detection accuracy, adaptability, and computational efficiency, making it well suited for deployment in a real-time, web-based cybersecurity platform such as CHRIS.

F. Ransomware Detection

The ransomware detection module in CHRIS is designed for early-stage and real-time identification of ransomware behaviour on a personal computer. Unlike traditional approaches that rely solely on signature-based detection, this module adopts a behaviour-driven monitoring strategy combined with automated response and Generative AI-assisted analysis. 6 The system continuously monitors a user-defined directory using the watchdog library, which enables real-time observation of file system events such as file creation, modification, deletion, and renaming.

This real-time monitoring forms the first line of defence against ransomware attacks. When a file is created or modified, the system immediately inspects its extension and metadata. If a file is detected with a known or suspicious ransomware-associated extension (e.g.,. locked, .wannacry), the system triggers an alert. To prevent further damage, CHRIS employs an automatic quarantine mechanism. Once a file is f lagged as potentially malicious, it is immediately isolated by moving it to a secure quarantine directory. This rapid containment strategy ensures that the suspected ransomware cannot propagate or encrypt additional user data, thereby limiting the overall impact of the attack. In addition to passive monitoring, the ransomware module includes a controlled behaviour simulation framework to validate detection effectiveness. The built-in run test suite simulates common ransomware behaviours in a safe environment.

These simulations include rapid file creation and modification to mimic encryption speed, generation of high-entropy files that resemble encrypted content, and mass renaming of files to simulate extension locking. By reproducing realistic ransomware activity patterns, the system can be tested and tuned without exposing the host system to actual malware. When ransomware-like behaviour is detected, CHRIS enhances situational awareness through its AI powered security assistant. The React-based user interface opens a chatbot sidebar, and relevant information such as file paths, timestamps, and behavioural metadata is forwarded to Google Gemini. Gemini generates a detailed natural-language report that includes an initial threat assessment, potential risks, and recommended mitigation actions. This explanation-driven response bridges the gap between low-level system events and user understanding. By combining real-time file system monitoring, automated quarantine, behavioural simulation, and Generative AI-based analysis, the ransomware detection module provides proactive, explainable, and user-centric protection. This approach is particularly effective against both known and zero-day ransomware variants, making it a critical component of the CHRIS platform

VII. IMPLEMENTATION DETAILS

The implementation of CHRIS follows a modular and scalable design, enabling seamless integration of multiple security services within a unified web-based platform. The backend is developed using FastAPI, chosen for its high performance, asynchronous request handling, and ease of integration with machine learning pipelines. Each security module is exposed as an independent RESTful API endpoint, allowing concurrent processing of user requests such as URL analysis, file scanning, network traffic evaluation, and image-based deepfake detection.

All Machine Learning and Deep Learning models are implemented in Python using standard libraries such as scikit-learn, XGBoost, and TensorFlow/Keras. Trained models are serialized using joblib or model checkpointing techniques and loaded into memory at runtime to ensure low-latency inference. Feature extraction pipelines for phishing, malware, and intrusion detection are optimized to minimize preprocessing overhead while preserving detection accuracy. For example, static malware features are extracted directly from executable metadata, and network intrusion features are derived from flow-level statistics. 7

The frontend is built using React.js, providing an interactive dashboard for users to upload files, submit URLs, monitor alerts, and visualize detection outcomes. Prediction results are displayed along with probability scores, performance indicators, and AI-generated explanations. The Google Gemini API is securely integrated into the backend and invoked selectively for tasks requiring natural-language reasoning, such as phishing risk summaries, password improvement suggestions, ransomware incident explanations, and chatbot-based assistance. This separation ensures that core detection remains efficient while Generative AI enhances interpretability and user engagement. The entire system is designed to operate as a web-based application, making it platform-independent and suitable for real-world deployment in personal and organizational environments

VIII. RESULTS AND ANALYSIS

The performance of CHRIS was evaluated across its major detection modules using standard classification metrics, including accuracy, precision, recall, and F1-score. These metrics provide a comprehensive understanding of each model's effectiveness in identifying cyber threats while minimizing false positives and false negatives.

The *phishing detection module*, based on the Random Forest algorithm, achieved an accuracy of 96%, with a precision of 96%, recall of 97%, and an F1-score of 96%. These results indicate strong discriminative capability in identifying malicious URLs while maintaining a low false- negative rate, which is critical for protecting users from deceptive attacks.

The *malware detection module* demonstrated the highest overall performance among the evaluated components. Using a Random Forest classifier trained on static executable features, the system achieved an accuracy of 99.41%, precision of 98.76%, recall of 99.31%, and an F1- score of 99.03%. The high recall value highlights the model's effectiveness in detecting malicious files, while the strong precision indicates a low rate of benign file misclassification.

For the *Network Intrusion Detection System*, the XGBoost-based model achieved an accuracy of 88.75%, precision of 91%, recall of 89%, and an F1-score of 89%. Although the accuracy is comparatively lower than file-based detection tasks, the results are consistent with prior studies on network intrusion detection, where data imbalance and traffic variability pose significant challenges. The high precision demonstrates the model's ability to reduce false alerts, which is essential for practical deployment.

The *deepfake detection module*, implemented using the Xception architecture and trained on datasets sourced from Hugging Face, achieved an accuracy of 95.3%, precision of approximately 91.6%, recall of 100%, and an F1-score of 95.63%. The perfect recall indicates that the model successfully identified all manipulated images in the evaluation set, while the high F1-score reflects a balanced performance between detection sensitivity and precision.

Overall, the experimental results validate the effectiveness of integrating multiple ML and DL models within a single platform. In addition to strong quantitative performance, the inclusion of *Generative AI explanations via Google Gemini* significantly enhances qualitative aspects such as interpretability, user trust, and situational awareness. This combination of high detection accuracy and explainable intelligence distinguishes CHRIS from traditional single-purpose cybersecurity solutions response to threats.

Table 1. Result Analyses

S. No	Module	Algorithm Used	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
1	Phishing Detection	Random Forest	96.00	96.00	97.00	96.00
2	Malware Detection	Random Forest	99.41	98.76	99.31	99.03
3	Network Intrusion Detection	XGBoost	88.75	91.00	89.00	89.00
4	Deepfake Detection	XceptionNet	95.30	91.60	100.00	95.63

IX. CONCLUSIONS

This paper presented CHRIS (Cyber Security Hub for Responsible Intelligence System), a unified and intelligent cybersecurity platform that integrates Machine Learning, Deep Learning, and Generative AI to address a wide spectrum of modern cyber threats. By consolidating phishing detection, malware detection, network intrusion detection, password strength evaluation, deepfake detection, and ransomware detection into a single web-based system, CHRIS overcomes the limitations of traditional siloed security solutions. The modular architecture ensures scalability and flexibility, enabling each detection component to operate independently while contributing to a holistic security view. Experimental evaluations demonstrate that the proposed models achieve high detection accuracy across multiple threat domains. Random Forest models provide reliable and interpretable performance for phishing and malware detection, while the XGBoost-based intrusion detection system effectively identifies anomalous network behaviour with reduced false alerts. The Xception-based deepfake detection module successfully identifies manipulated media using datasets sourced from Hugging Face, highlighting the adaptability of the system to diverse and evolving data sources.

In addition, the ransomware detection module emphasizes early-stage behavioural monitoring and automated containment, offering proactive protection against one of the most destructive forms of cyber-attacks. A key contribution of CHRIS lies in the integration of Generative AI through Google Gemini, which significantly enhances explainability, user awareness, and interaction. Instead of presenting raw model predictions, the system provides natural-language explanations, contextual risk summaries, and actionable recommendations. This human- centric design bridges the gap between complex security analytics and end-user understanding, thereby improving trust and usability.

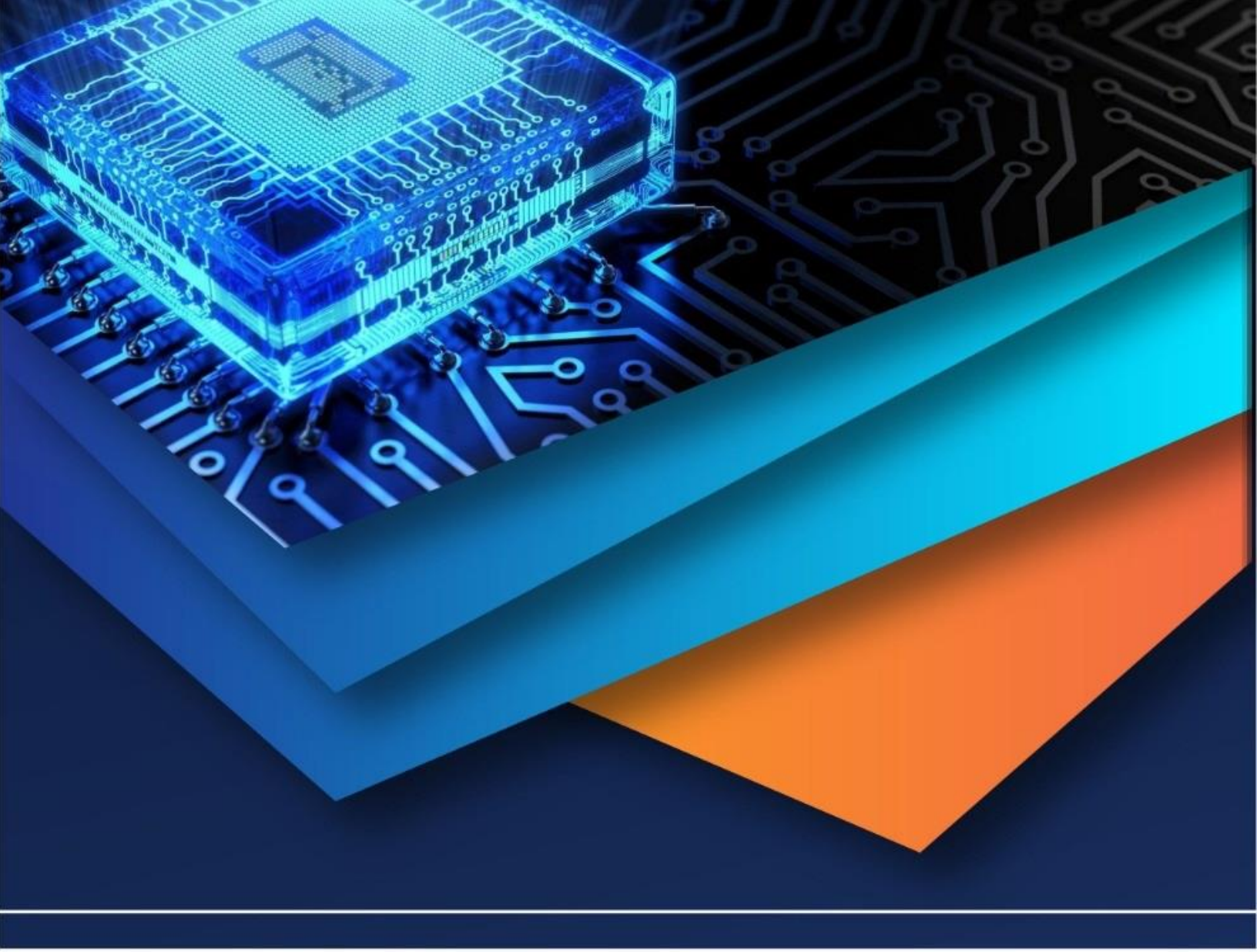
Overall, CHRIS demonstrates that combining predictive security analytics with Generative AI-driven intelligence can lead to more effective, explainable, and user-friendly cybersecurity solutions. The proposed platform is well suited for real-world deployment in personal and organizational environments. Future work will focus on extending CHRIS with dynamic malware analysis, real-time network traffic capture, multimodal deepfake detection (image and video), and continuous learning mechanisms to further improve resilience against emerging and zero-day threats.

X. ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Department of Computer Science and Engineering, St. Thomas Institute for Science and Technology, Kerala, for providing the necessary infrastructure, resources, and academic support to successfully carry out this work. We extend our heartfelt thanks to our project guide and faculty members for their continuous guidance, valuable feedback, and encouragement throughout the design and development of the CHRIS framework. We also acknowledge the support and cooperation of our classmates and peers who contributed through discussions, suggestions, and testing during various stages of implementation. Finally, we express our gratitude to everyone who directly or indirectly helped in the successful completion of this project.

REFERENCES

[1] Z. Li, Q. Yan, R. Deng, W. Liu, and D. Wang, “A survey on phishing attacks and countermeasures,” IEEE Communications Surveys & Tuto rials, vol. 20, no. 1, pp. 1–28, First Quarter 2018.
[2] S. Marchal, J. Francois, R. State, and T. Engel, “PhishStorm: Detecting phishing with streaming analytics,” IEEE Transactions on Network and Service Management, vol. 11, no. 4, pp. 458–471, Dec. 2014.
[3] Y. Ye, T. Li, Q. Jiang, and Y. Wang, “CIMDS: Adapting postprocessing techniques of associative classification for malware detection,” IEEE Transactions on Systems, Man, and Cybernetics, vol. 40, no. 3, pp. 298 307, May 2010.
[4] M. Egele, T. Scholte, E. Kirda, and C. Kruegel, “A survey on automated dynamic malware-analysis techniques and tools,” ACM Computing Sur veys, vol. 44, no. 2, pp. 1–42, 2012.
[5] N. Scaife, H. Carter, P. Traynor, and K. Butler, “CryptoLock (and Drop It): Stopping ransomware attacks on user data,” in Proc. IEEE International Conference on Distributed Computing Systems (ICDCS), 2016, pp. 303 312.
[6] A. Sgandurra, L. Mu˜noz-Gonz´alez, R. Mohsen, and E. C. Lupu, “Au tomated dynamic analysis of ransomware: Benefits, limitations, and use for detection,” arXiv preprint arXiv:1609.03020, 2016.
[7] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
[8] N. M. Karie, V. R. Kebande, H. S. Venter, and N. I. Choo, “On the importance of forensic readiness in digital investigations,” Digital Investigation, vol. 32, pp. 200–214, 2020.
[9] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, “Toward generating a new intrusion detection dataset and intrusion traffic characterization,” in Proc. International Conference on Information Systems Security and Privacy (ICISSP), 2018, pp. 108–116.
[10] A. H. Lashkari, G. Draper-Gil, M. S. I. Mamun, and A. A. Ghorbani, “Characterization of Tor traffic using time-based features,” in Proc. International Conference on Information Systems Security and Privacy (ICISSP), 2017, pp. 253–262.
[11] W. Wang, M. Zhu, J. Wang, X. Zeng, and Z. Yang, “End-to-end encrypted traffic classification with one-dimensional CNN,” in Proc. IEEE International Conference on Intelligence and Security Informatics (ISI), 2017, pp. 43–48.
[12] R. Sommer and V. Paxson, “Outside the closed world: On using machine learning for network intrusion detection,” in Proc. IEEE Symposium on Security and Privacy, 2010, pp. 305–316.
[13] M. Masood, M. Nawaz, K. M. Malik, A. Javed, A. Irtaza, and H. Malik, “Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward,” Applied Intelligence, vol. 53, no. 4, pp. 3974–4026, 2023.
[14] H. Zhao, W. Zhou, D. Chen, W. Zhang, and N. Yu, “Multi-attentional deepfake detection,” in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 2185–2194.
[15] R. Lanzino, F. Fontana, A. Diko, M. R. Marini, and L. Cinque, “Faster than lies: Real-time deepfake detection using binary neural networks,” in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024, pp. 3771–3780.
[16] J. Guan, H. Zhou, Z. Hong, E. Ding, J. Wang, C. Quan, and Y. Zhao, “Delving into sequential patches for deepfake detection,” Advances in Neural Information Processing Systems, vol. 35, pp. 4517–4530, 2022.
[17] B. Liu, M. Ding, T. Zhu, and X. Yu, “TI2Net: Temporal identity inconsistency network for deepfake detection,” in Proc. IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2023, pp. 4691 4700.
[18] H. Zhao, W. Zhou, D. Chen, W. Zhang, and N. Yu, “Self-supervised transformer for deepfake detection,” arXiv preprint arXiv:2203.01265, 2022.
[19] M. S. M. Altaei, “Detection of deep fake in face images using deep learning,” Wasit Journal of Computer and Mathematics Science, vol. 1, no. 4, pp. 60–71, 2022.
[20] D. Battista, “Political communication in the age of artificial intelligence: An overview of deepfakes and their implications,” Society Register, vol. 8, no. 2, pp. 7–24, 2024.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY