IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, December 2025
Impact Factor: 7.67

Android Based Ransomware Detection using Explainable AI (XAI) Integrated Machine learning

Shravani M. R, Y Srushti, Poornima R M, Varun, Preetham D. R
Dept. Information Science and Engineering
Global Academy of Technology, Bengaluru, India
Shravanimr7@gmail.com, Ysrushti2004@gmail.com, Poornima.rm@gat.ac.in
Varunbg030@gmail.com, preethamrally555@gmail.com

Abstract: *The widespread adoption of Android as the leading mobile operating system has made it a primary target for various forms of cyberattacks, particularly ransomware. Android's open-source architecture and extensive app ecosystem, while enabling innovation, also expose it to significant security risks. Ransomware attacks on Android devices encrypt user data or lock access, demanding payment for restoration. These attacks not only cause financial loss but also compromise user privacy and system integrity.*

Keywords: *Android*

I. INTRODUCTION

The widespread adoption of Android as the leading mobile operating system has made it a primary target for various forms of cyberattacks, particularly ransomware. Android's open-source architecture and extensive app ecosystem, while enabling innovation, also expose it to significant security risks. Ransomware attacks on Android devices encrypt user data or lock access, demanding payment for restoration. These attacks not only cause financial loss but also compromise user privacy and system integrity.

Consequently, the need for robust, intelligent, and adaptive ransomware detection mechanisms has become more critical than ever.

Traditional signature-based antivirus systems and heuristic approaches often fall short when dealing with rapidly evolving ransomware variants. Attackers continuously modify code structures, behaviors, and permissions to evade detection, rendering conventional techniques inadequate. Machine learning (ML) and deep learning (DL) models have shown promising results in identifying previously unseen malware by learning complex behavioral and feature-based patterns. However, despite their impressive detection accuracy, these models often function as "black boxes," providing little to no insight into their decision-making processes. This lack of transparency hinders trust, especially in critical security applications.

Explainable Artificial Intelligence (XAI) emerges as a solution to address this limitation by introducing interpretability and transparency into AI-driven systems. XAI enables researchers and practitioners to understand, visualize, and justify model predictions, thus bridging the gap between accuracy and explainability. When integrated into Android ransomware detection, XAI techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) can provide detailed reasoning behind why an application is classified as malicious or benign. This capability not only supports better decision-making but also aids in enhancing user confidence and regulatory compliance.

The integration of XAI within Android ransomware detection frameworks also facilitates security analysts in identifying key features contributing to malicious behaviors, such as abnormal permission requests, encryption routines, or suspicious network activities. This interpretability helps in refining detection models, improving dataset quality, and designing more secure app development practices. Furthermore, it allows nontechnical stakeholders, including endusers and policymakers, to gain a clearer understanding of how AI systems make security-related decisions, fostering

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, December 2025 Impact Factor: 7.67

accountability and trust in automated cybersecurity solutions.

This research focuses on developing an Android ransomware detection model that combines the predictive power of AI with the transparency of XAI. By leveraging both static and dynamic analysis of Android applications, the proposed framework aims to achieve high detection accuracy while maintaining explainability. The ultimate goal is to build an intelligent, transparent, and trustworthy ransomware detection system that not only identifies threats effectively but also provides interpretable insights into its reasoning process. Such integration is expected to play a vital role in advancing the future of secure, explainable, and responsible mobile cybersecurity.

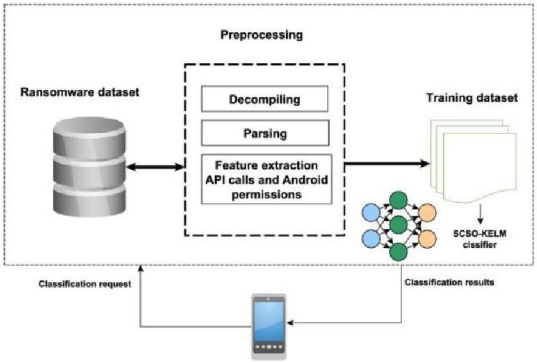

Figure 1: Planned Android Ransomware Detection

The proposed system introduces an intelligent Android ransomware detection framework that integrates Explainable Artificial Intelligence (XAI) to enhance both detection accuracy and interpretability. The framework operates in multiple stages, beginning with the collection of Android application packages (APKs) from both benign and malicious sources. These applications undergo static and dynamic analysis to extract relevant features such as permissions, API calls, system behaviors, network activities, and encryption patterns. The extracted feature set is then preprocessed and optimized using feature selection techniques to reduce redundancy and improve model efficiency.

A machine learning or deep learning classifier—such as Random Forest, XGBoost, or a neural network—is trained to distinguish between ransomware and legitimate applications. To overcome the black-box limitation of traditional AI models, the proposed system incorporates XAI methods such as SHAP and LIME, which generate human-understandable explanations for each classification decision. These explanations highlight the most influential features that led to a particular detection outcome, enabling analysts to interpret and validate the model's reasoning. The proposed approach not only achieves high detection accuracy but also ensures transparency, trustworthiness, and accountability in AIbased Android ransomware detection.

The increasing dependency on Android devices for communication, banking, and data storage has made them prime targets for ransomware attacks. Android ransomware encrypts user data or restricts device access, demanding payment for recovery and causing severe financial and privacy losses. Traditional detection systems, such as signaturebased and heuristic approaches, struggle to identify newly emerging or obfuscated ransomware variants due to their reliance on known patterns. Although machine learning and deep learning models have improved detection capabilities, they often operate as opaque "black-box" systems that provide no insight into their decision- making processes. This lack of transparency limits user trust, makes it difficult for analysts to validate model predictions, and hinders regulatory compliance in security-critical environments.

Therefore, there is a pressing need for an intelligent ransomware detection framework that not only provides accurate and early threat identification but also integrates Explainable Artificial Intelligence (XAI) to offer clear, interpretable, and justifiable explanations for each detection outcome. Such a solution would enhance both the reliability and accountability of AI-driven Android security systems.

II. RELATED WORKS

Research on Android ransomware and broader mobile malware detection has expanded rapidly as attackers target the dominant mobile platform with increasingly sophisticated payloads. Several recent surveys and literature reviews

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, December 2025 Impact Factor: 7.67

synthesize threat taxonomies, common ransomware behaviors, and detection challenges specific to Android, highlighting issues such as code obfuscation, repackaging, and the diversity of delivery channels. These reviews emphasize the urgent need for adaptive detection techniques that go beyond signature matching to address novel and polymorphic ransomware families.

A large body of work compares and combines static and dynamic analysis techniques to extract discriminative features from Android applications. Static analysis studies focus on manifest permissions, API call patterns, and opcode or bytecode features, while dynamic approaches capture runtime behaviors such as file system access, network communications, and cryptographic operations inside sandboxes or emulators. Hybrid pipelines that fuse static and dynamic feature sets have shown improved robustness against obfuscation and unpacking, but they also introduce higher collection overhead and complexity.

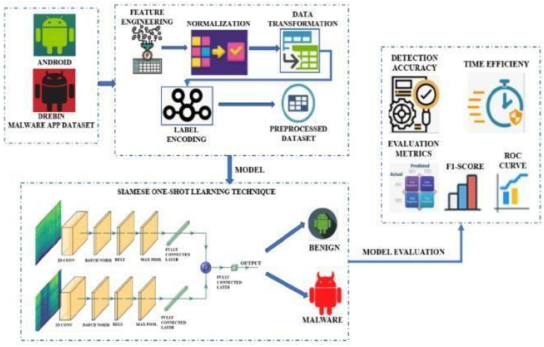

Figure 2: Android Malware detection

Machine learning and deep learning methods have been widely applied to the ransomware detection problem, producing strong classification performance in many experimental settings. Researchers have used classical classifiers (Random Forest, SVM, XGBoost) and neural architectures to model behavioral and traffic-based features, and ensemble techniques have recently been proposed to boost accuracy and resilience. Still, many high-performing models are developed and tested on constrained datasets, and differences in feature engineering, dataset composition, and evaluation protocols make cross-paper comparison difficult.

More recently, Explainable AI (XAI) techniques have been introduced to make malware and ransomware classifiers interpretable. Work applying modelagnostic explainers (LIME, SHAP) and other post-hoc methods to Android malware models demonstrates how feature attribution can reveal which permissions, API calls, or dynamic behaviors drove a given decision. XAI studies report benefits for analyst validation, false-positive triage, and model debugging, while also noting challenges such as explanation stability, scalability to large feature spaces, and the potential for adversarial manipulation of explanations.

While XAI-enhanced detection frameworks show promise, current literature reveals important gaps: few studies focus specifically on ransomware as opposed to general malware, many explanations remain post-hoc and local without offering global model insight, and practical deployment constraints (performance, data collection cost, privacy) are often underexamined. These gaps motivate integrated solutions that combine robust static/dynamic feature extraction, scalable ML/DL classifiers, and principled XAI methods to produce accurate, interpretable, and operationally viable Android ransomware detectors.

Module:

1. Data Collection and Preprocessing Module

This module focuses on gathering Android application samples from trusted and malicious sources to create a balanced dataset for model training and evaluation. Benign applications are collected from official repositories such as Google Play Store, while ransomware samples are obtained from reputable malware databases and research platforms. The collected APK files are then subjected to preprocessing steps including unpacking, decompilation, and feature extraction. Noise, redundant features, and irrelevant data are removed to enhance dataset quality. Proper labeling of

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, December 2025 Impact Factor: 7.67

ransomware and benign samples ensures that the system learns to differentiate malicious behaviors effectively during model training.

2. Feature Extraction and Selection Module

The next stage involves extracting both static and dynamic features that characterize Android ransomware behavior. Static features such as permissions, API calls, intent filters, and manifest configurations are obtained through code analysis, whereas dynamic features such as system calls, file access patterns, encryption routines, and network activities are captured during controlled execution in an emulator or sandbox.

Feature selection algorithms like Information Gain, Chi-square, or Principal Component Analysis (PCA) are applied to identify the most relevant attributes, reducing computational complexity while retaining essential behavioral indicators for accurate classification.

3. Machine Learning-Based Detection Module

This module is responsible for training and deploying an intelligent detection model using the optimized feature set. Various supervised learning algorithms— such as Random Forest, XGBoost, and Deep Neural Networks—are evaluated to identify the most effective classifier for ransomware detection. The model learns to differentiate between benign and malicious behaviors by recognizing patterns and anomalies within the extracted features. Performance metrics such as accuracy, precision, recall, and F1score are used to assess model effectiveness. This ensures the detection system can identify both known and previously unseen ransomware variants with high reliability.

4. Explainable AI (XAI) Integration Module

To overcome the limitations of traditional "black-box" models, this module incorporates Explainable AI techniques such as LIME (Local Interpretable ModelAgnostic Explanations) and SHAP (SHapley Additive exPlanations). These methods provide human-understandable explanations for the model's predictions by identifying which features contributed most to a classification decision. For example, the system can explain that certain suspicious permissions, encryption patterns, or API calls influenced the detection of ransomware. This transparency helps security analysts and users build trust in the AI system, enables better debugging of false positives, and ensures compliance with explainability requirements in cybersecurity research.

5. Evaluation and Visualization Module

The final module evaluates the overall performance of the proposed framework and visualizes both detection outcomes and interpretability results. Confusion matrices, ROC curves, and precision-recall graphs are used to assess detection performance, while XAI visualization tools illustrate feature importance and model decision paths. These visual outputs help analysts understand how the detection system arrives at its conclusions and identify areas for further improvement. This module ensures the system not only achieves high detection accuracy but also provides meaningful insights that enhance decision-making and trustworthiness in AI-based ransomware defense mechanisms.

The primary contribution of this research lies in the development of an intelligent and transparent Android ransomware detection framework that integrates

Explainable Artificial Intelligence (XAI) to enhance both accuracy and interpretability. Unlike traditional blackbox models that focus solely on detection performance, the proposed system emphasizes transparency by providing clear explanations for its decisions through methods such as SHAP and LIME. The study contributes a comprehensive dataset of benign and malicious Android applications, incorporating both static and dynamic behavioral features to improve robustness against code obfuscation and evasion techniques. Furthermore, the research introduces a hybrid machine learning approach that combines feature optimization with model interpretability, ensuring efficient and explainable threat detection. The integration of visualization tools for XAI-based explanations allows security analysts to identify critical factors influencing model predictions, thereby improving trust, accountability, and decision-making in cybersecurity environments. Overall, this work contributes to the growing field of explainable mobile security by bridging the gap between high-performing AI models and human-understandable analysis in Android ransomware

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, December 2025 Impact Factor: 7.67

detection.

The motivation behind this research stems from the growing prevalence and sophistication of ransomware attacks targeting Android devices, which have become integral to daily life for communication, banking, and data management. As mobile platforms store sensitive personal and financial information, the consequences of ransomware infections such as data encryption, extortion, and privacy violations—pose serious risks to individuals and organizations alike. Traditional malware detection systems, which rely heavily on signature-based and heuristic approaches, fail to keep pace with the rapidly evolving and obfuscated nature of modern ransomware. Although artificial intelligence and machine learning techniques have significantly improved detection accuracy, they often operate as black-box systems, providing limited insight into how and why specific predictions are made. This lack of interpretability undermines user trust and makes it challenging for cybersecurity professionals to validate model outcomes. Therefore, integrating Explainable AI (XAI) into Android ransomware detection serves as a crucial step toward building transparent, trustworthy, and accountable security systems. By combining high detection accuracy with interpretability, this research aims to empower analysts with meaningful explanations, enhance user confidence, and advance the development of responsible AI solutions in mobile cybersecurity.

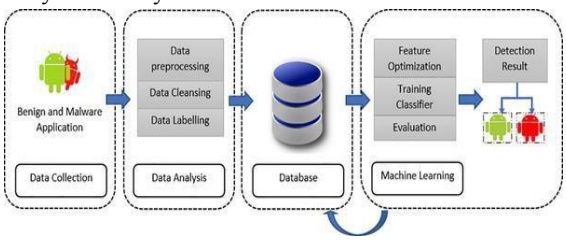

Figure: 3 Static analysis approach for Android permission

The proposed algorithm for Android ransomware detection begins with data acquisition and preprocessing, where Android application packages (APKs) are collected from both benign and malicious sources. Each APK undergoes static and dynamic analysis to extract critical features such as permissions, API calls, system behaviors, network activities, and encryption routines. Feature selection techniques, such as Principal Component Analysis (PCA) or Information Gain, are applied to identify the most relevant attributes, reducing dimensionality and improving the model's learning efficiency. The preprocessed feature set is then used to train a machine learning or deep learning classifier, such as Random Forest, XGBoost, or a neural network, to distinguish between ransomware and legitimate applications based on learned behavioral patterns.

To enhance transparency and interpretability, the algorithm integrates Explainable AI (XAI) techniques, such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model- Agnostic Explanations). Once the classifier makes a prediction, XAI methods generate humanunderstandable explanations, identifying which features contributed most to labeling an application as malicious or benign. The output includes both the detection result and a ranked list of influential features, allowing analysts to validate decisions, understand ransomware behaviors, and improve trust in the AI system. Finally, performance metrics, such as accuracy, precision, recall, and F1-score, are computed, and visualization tools are used to present both detection results and explanation insights, creating

an effective, interpretable, and actionable ransomware detection framework for Android devices. III Literature Survey:

1. "XRan: Explainable Deep LearningBased Ransomware Detection Using Dynamic Analysis" Authors: S. Gulmez, et al. Published in: Computers & Security, 2024

This paper introduces XRan, an explainable deep learning model for ransomware detection that leverages dynamic analysis to identify behavioral patterns during ransomware execution. The model employs a Convolutional Neural Network (CNN) architecture and integrates Interpretable Model-Agnostic Explanations (LIME) and SHAP to provide transparent decision-making processes. Experimental results demonstrate the model's effectiveness in detecting ransomware with high accuracy while offering insights into the features influencing its predictions.

2. "Explainable AI for Android Malware Detection: Towards Understanding

Why the Models Perform So Well?" Authors: Yue Liu, Chakkrit Tantithamthavorn, Li Li, Yepang Liu Published in: arXiv, 2022

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, December 2025 Impact Factor: 7.67

This study applies XAI techniques to machine learning-based Android malware detection systems to explore the reasons behind their high performance. By training classic ML models and employing XAI methods, the research identifies that models often rely on temporal inconsistencies in the dataset rather than actual malicious behaviors, leading to over-optimistic performance metrics. The findings highlight the importance of realistic experimental designs and the need for interpretability in evaluating model reliability.

3. "Explainable Artificial Intelligence for Cybersecurity: A Survey" Authors: F. Charmet, et al. Published in: Journal of Network and Systems Management, 2022
This comprehensive literature review examines the intersection of XAI and cybersecurity, focusing on applications such as intrusion detection and malware classification. The paper discusses the challenges of applying XAI in cybersecurity, including the trade-off between model complexity and interpretability, and the potential for adversarial attacks on XAI systems. It provides a structured overview of existing approaches and suggests directions for future research to enhance the transparency and trustworthiness of cybersecurity systems.

4. "Explainable Artificial Intelligence for Malware Analysis: A Survey of Techniques, Applications, and Open Challenges" Authors: Harikha Manthena, Shaghayegh Shajarian, et al. Published in: arXiv, 2024
This survey presents a comprehensive review of state-of-the-art machine learning techniques for malware analysis, with a specific focus on explainability methods. The paper examines existing XAI frameworks, their application in malware classification and detection, and the challenges associated with making malware detection models more interpretable. It also explores recent advancements and highlights open research challenges in the field of explainable malware analysis, aiming to bridge the gap between high detection accuracy and model interpretability.

5. "Android Malware Classification Using Explainable AI" Authors: J. Sugunan Nair Published in: DIVA Portal, 2025
This research explores the use of XAI techniques in Android malware classification, addressing the topic of interpretability in machine learning models. By applying XAI methods to various ML models, the study aims to provide insights into the decision-making processes of these models, enhancing their transparency and trustworthiness. The findings underscore the importance of incorporating explainability into malware detection systems to improve their reliability and acceptance in real-world applications.Dataset Collection and Visualization

1. Dataset Collection Overview
The dataset for this study is constructed from a combination of benign and ransomware-infected Android applications (APKs) to provide a comprehensive foundation for model training and evaluation. Benign samples are collected from official repositories such as the Google Play Store, ensuring that they represent a variety of categories and functionalities. Malicious samples, specifically ransomware variants, are sourced from reputable malware repositories such as VirusTotal, Drebin, and AndroZoo, which provide well-labeled and verified malware datasets. For each APK, both static features (permissions, API calls, manifest data) and dynamic behaviors (system calls, file access patterns, encryption routines, network traffic) are extracted. Additionally, for visual and explainability analysis, certain dynamic behaviors are captured as images or graphs, such as API call sequences, permission heatmaps, and system call traces. These images serve as input for explainable AI techniques, helping to visually represent the features contributing to ransomware detection. This hybrid dataset of raw APK data, extracted features, and behavioral images ensures that the proposed system can learn effectively while providing interpretable insights through XAI visualizations.

IV. PROPOSED TECHNIQUES:

The future scope of Android ransomware detection with Explainable AI (XAI) integration is highly promising, as mobile ransomware continues to evolve in complexity and sophistication. One potential direction is the integration of real-time detection systems on mobile devices, allowing AI models to monitor application behaviors continuously and prevent ransomware attacks before significant damage occurs. By combining edge computing with XAI, these systems could offer both high-performance detection and instant explainability, enabling users and security analysts to make

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, December 2025 Impact Factor: 7.67

informed decisions immediately. Additionally, incorporating multi-modal data sources, such as network traffic patterns, user interaction logs, and cloudbased app behaviors, could further enhance the robustness and accuracy of ransomware detection frameworks.

Another significant avenue for future research lies in improving the interpretability and trustworthiness of AI models. Advanced XAI methods could be developed to provide more granular and context-aware explanations, helping analysts understand complex ransomware strategies and facilitating regulatory compliance in sensitive environments. Furthermore, the adoption of federated learning techniques could allow multiple devices to collaboratively improve detection models without sharing sensitive user data, ensuring privacy while maintaining model performance. Overall, expanding the application of XAI in Android ransomware detection not only strengthens cybersecurity defenses but also promotes the responsible and transparent use of artificial intelligence in protecting mobile ecosystems.

V. FUTURE WORK

The future scope of Android ransomware detection with Explainable AI (XAI) integration is highly promising, as mobile ransomware continues to evolve in complexity and sophistication. One potential direction is the integration of real-time detection systems on mobile devices, allowing AI models to monitor application behaviors continuously and prevent ransomware attacks before significant damage occurs. By combining edge computing with XAI, these systems could offer both high-performance detection and instant explainability, enabling users and security analysts to make informed decisions immediately. Additionally, incorporating multi-modal data sources, such as network traffic patterns, user interaction logs, and cloudbased app behaviors, could further enhance the robustness and accuracy of ransomware detection frameworks.

Another significant avenue for future research lies in improving the interpretability and trustworthiness of AI models. Advanced XAI methods could be developed to provide more granular and context-aware explanations, helping analysts understand complex ransomware strategies and facilitating regulatory compliance in sensitive environments. Furthermore, the adoption of federated learning techniques could allow multiple devices to collaboratively improve detection models without sharing sensitive user data, ensuring privacy while maintaining model performance. Overall, expanding the application of XAI in Android ransomware detection not only strengthens cybersecurity defenses but also promotes the responsible and transparent use of artificial intelligence in protecting mobile ecosystems.

VI. CONCLUSION

In conclusion, the integration of Explainable Artificial Intelligence (XAI) into Android ransomware detection presents a significant advancement in mobile cybersecurity. By combining machine learning and deep learning techniques with XAI, the proposed framework not only achieves high accuracy in identifying ransomware threats but also provides transparent and interpretable insights into the model's decision-making process. This interpretability enables security analysts and end-users to understand the underlying reasons for classifying an application as malicious or benign, thereby building trust and enhancing the effectiveness of response strategies. The hybrid approach of using static and dynamic feature analysis ensures that the system remains robust against evolving ransomware variants and code obfuscation techniques, addressing the limitations of traditional detection methods.

Moreover, the study highlights the importance of explainability in AI-driven security systems, particularly in mobile environments where user trust and data privacy are critical. The visualization of feature contributions through XAI techniques empowers analysts to refine models, reduce false positives, and gain deeper insights into ransomware behavior. This research lays the groundwork for future advancements, such as real-time detection, federated learning, and multi- modal feature integration, which can further strengthen Android security. Overall, the proposed framework not only enhances ransomware detection capabilities but also fosters responsible and transparent AI deployment in cybersecurity, contributing to safer and more resilient mobile ecosystems.

REFERENCES

[1] Liu, Y., Tantithamthavorn, C., Li, L., & Liu, Y. (2022). Explainable AI for Android Malware Detection: Towards Understanding Why the Models Perform So Well? arXiv. This study investigates the overoptimistic performance of

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, December 2025 Impact Factor: 7.67

machine learning models in Android malware detection, attributing it to temporal inconsistencies in training datasets. arXiv

[2] Kulkarni, M., & Stamp, M. (2024). XAI and Android Malware Models. arXiv. The paper applies various XAI techniques to machine learning and deep learning models trained on Android malware datasets, enhancing model interpretability. arXiv

[3] Gulmez, S., et al. (2024). XRan: Explainable Deep Learning-Based Ransomware Detection Using Dynamic Analysis. Computers & Security. Introduces XRan, a deep learning model for ransomware detection that incorporates explainability through dynamic analysis and XAI techniques.
ScienceDirect

[4] Rabby, M. F., & Sultana, R. (2025). Explainable AI-Driven Hybrid Feature Fusion for Robust Android Malware Detection. Research Square. Presents a hybrid feature fusion approach combining static, dynamic, and image- based features with XAI for enhanced Android malware detection. Sciety

[5] Vanjire, S. S., et al. (2024). A Novel Method of Detecting Malware on Android Mobile Devices Using Deep Learning and Explainable AI. Bulletin of Electrical Engineering and Informatics. Proposes a deep learning framework integrated with XAI techniques for effective Android malware detection. beei.org

[6] Palma, C., et al. (2024). Explainable Machine Learning for Malware Detection on Android Applications. Information. Explores machine learning techniques for Android malware detection, emphasizing the importance of feature selection and explainability. MDPI

[7] Meti, S., Sidramayyanmath, V., & Patil, S. (2025). Ransomware Detection Using Machine Learning and Explainable AI. In Artificial Intelligence: Theory and Applications. Discusses the integration of machine learning and XAI for effective ransomware detection, highlighting the importance of model transparency. ResearchGate

[8] Kinkead, M., et al. (2021). Towards Explainable CNNs for Android Malware Detection. Procedia Computer Science. Investigates the application of convolutional neural networks for Android malware detection with a focus on interpretability. ScienceDirect

[9] Ullah, S., et al. (2024). The Revolution and Vision of Explainable AI for Android Malware Detection and Protection. Computers & Security. Provides a comprehensive overview of XAI's role in enhancing Android malware detection and protection mechanisms. ScienceDirect

[10] Najibi, M., et al. (2025). Towards a Robust Android Malware Detection Model Using Explainable AI. Journal of Computer Security. Proposes a robust model for Android malware detection that incorporates XAI to improve model reliability and interpretability. ScienceDirect

[11] Youssofi, C. E., et al. (2025). YoloMal-XAI: Interpretable Android Malware Classification Using RGB Images and YOLO11. Journal of Cybersecurity and Privacy. Introduces YoloMal-XAI, a framework that transforms Android application files into RGB images for malware classification, enhancing interpretability. MDPI

[12] Meti, S., Sidramayyanmath, V., & Patil, S. (2025). Ransomware Detection Using Machine Learning and Explainable AI. In Artificial Intelligence: Theory and Applications. Explores the integration of machine learning and XAI for ransomware detection, emphasizing the need for transparent AI systems. ResearchGate

[13] Kulkarni, M. (2023). Explainable AI for Android Malware Detection. San Jose State University. Applies XAI techniques to machine learning models for Android malware detection, enhancing model interpretability. SJSU ScholarWorks

[14] Nair, J. S. (2025). Android Malware Classification Using Explainable AI. DIVA Portal. Investigates the use of XAI in Android malware classification, addressing the interpretability of AI-based detection mechanisms. DIVA Portal

[15] Rabby, M. F., & Sultana, R. (2025). Explainable AI-Driven Hybrid Feature Fusion for Robust Android Malware Detection. Research Square. Presents a hybrid feature fusion approach combining static, dynamic, and image- based features with XAI for enhanced Android malware detection.