Undergraduate Conference on Intelligent Computing and Systems (UCICS 2025)
26-27 February, 2025; Varendra University, Rajshahi, Bangladesh

Securing Android Ecosystem: Refining Ransomware Detection with Stacking Models and Explainable AI

Nafis Fuad
Dept of CSE
Bangladesh Army University of Science and Technology
Saidpur, Bangladesh
nafisfuad888pen@gmail.com

Ananna Hoque Shathi
Dept of CSE
Bangladesh Army University of Science and Technology
Saidpur, Bangladesh
anannahoque.cse10@gmail.com

Hasan Muhammad Kafi
Dept of CSE
Bangladesh Army University of Science and Technology
Saidpur, Bangladesh
engr.kafi@gmail.com

Taher Muhammad Mahdee
Dept of CSE
Bangladesh Army University of Science and Technology
Saidpur, Bangladesh
t.mahdee@gmail.com

Masroor Mahmud
Dept of CSE
Bangladesh Army University of Science and Technology
Saidpur, Bangladesh
herritagemahib@gmail.com

Dip Karmakar
Dept of CSE
Bangladesh Army University of Science and Technology
Saidpur, Bangladesh
rdip42810@gmail.com

***Abstract*—As Android devices confront an escalating ransomware threat, this study introduces a refined defense mechanism using ensemble stacking models. Utilizing a comprehensive Android network dataset rich in diverse network features, we explore the nuances of ransomware detection. While traditional machine learning and deep learning models exhibit commendable performance, our pursuit of heightened accuracy and true positive rates leads us to craft a bespoke ensemble model. Demonstrating remarkable achievements, particularly a high accuracy rate, this approach is fortified by explainable AI, specifically LIME (Local Interpretable Model-agnostic Explanations), offering transparency into the models' decision-making processes. This study not only elevates the standards of Android ransomware detection but also highlights the imperative for precision and resilience in contemporary cybersecurity.**

Keywords*— *Android Ransomware Detection, Ensemble Stacking Models, Explainable AI, LIME, and Cybersecurity

I. Introduction

Within the intricate realm of cybersecurity, Android devices emerge not only as pioneers of technological innovation but also as primary defenders in the ever-changing fight against ransomware threats. The escalating frequency of ransomware attacks underscores the imperative for a nuanced and resilient defense strategy. This study responds to this imperative by crafting a sophisticated solution tailored explicitly to the unique challenges posed by the Android ecosystem [1].

Our investigative journey commences with a meticulous analysis of an expansive Android network dataset, purposefully curated to encompass a diverse array of network features. These features collectively present a comprehensive panorama of ransomware behavior, forming the bedrock for the subsequent development of a robust defense mechanism. While traditional machine learning and deep learning models demonstrate commendable proficiency in deciphering ransomware patterns, our pursuit of precision and heightened accuracy propels us to engineer a bespoke ensemble stacking model [2]–[5].

In our dedication to transparency and interpretability, we seamlessly integrate explainable AI [6], with a specific emphasis on LIME, to illuminate the decision-making processes inherent in our models. Beyond the technical intricacies, this research narrative is imbued with a human touch, aspiring to instill trust and confidence in the security of the Android ecosystem.

The remainder of the paper is structured as follows: The literature is reviewed in Section II, our methodology is addressed in Section III, the results and analysis of our experiments are presented in Section IV, the discussion is presented in Section V, and the study's conclusion is presented in Section VI.

II. Literature Review

The collected literature delves into the landscape of mobile cybersecurity, with a specific focus on Android ransomware. In study [7], the vulnerability of Android smartphones, constituting 73% of the market share, is underscored. The paper provides a comprehensive review of Android ransomware threat scenarios, spanning from 2015 to 2020, and suggests directions for future research, emphasizing the need for specialized defenses against evolving cyber threats.

The paper [8] contributes significantly by outlining the basics of Android malware, its evolution, and the tools for analysis. The review critically evaluates existing approaches, particularly those leveraging machine learning and deep learning for classification. It identifies research gaps and emphasizes the imperative to develop robust techniques to combat the increasing sophistication of malware.

Addressing the critical threat of ransomware on smartphones, [9] introduces an innovative methodology. This approach employs evolutionary-based machine learning,

utilizing the binary particle swarm optimization algorithm and the synthetic minority oversampling technique (SMOTE) for classification. The paper highlights the importance of non-classical, intelligent techniques in safeguarding against evolving threats. In the realm of forensic analysis and detection of Android ransomware, [10] focuses on overcoming the limitations of supervised machine learning. The proposed RansomDroid framework utilizes clustering-based unsupervised machine learning with a Gaussian Mixture Model, aiming to enhance detection, accuracy and robustness by addressing issues such as mislabeling of historical targets and unforeseen Android ransomware. Extending beyond Android ransomware, [11] explores challenges in securing IoT systems in industrial settings.

The paper [12] provides a holistic survey of cognitive computing applications. It emphasizes the role of cognitive computing in addressing challenges related to data analysis and knowledge-rich automation. It emphasizes the role of cognitive computing in addressing challenges related to data analysis and knowledge-rich automation. It also emphasizes its significance in addressing intricate challenges in mobile cybersecurity.

The study conducted in [13] presents a comprehensive survey of current literature on Explainable AI (XAI) methods for cybersecurity applications. It highlights the limitations of existing AI techniques in terms of transparency and interpretability, underlining the importance of XAI in building more explainable models, a valuable resource for understanding the state-of-the-art in XAI and its potential applications in addressing cyber threats.

III. METHODOLOGY

In our experiments, we followed a structured methodology to ensure robust results. Initially, we collected data from relevant sources, ensuring a diverse and representative dataset. After collection, we performed data preprocessing. Once the data was prepared, we selected a set of machine learning models suitable for ensemble stacking. The chosen models were then trained. After training, the models were stacked in an ensemble framework, where their predictions were combined to enhance overall accuracy and robustness.

To further interpret the predictions and enhance model transparency, we utilized explainable AI techniques called LIME. Fig. 1 shows the flowchart of our methodology.

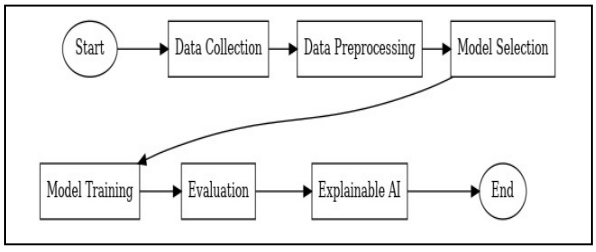

Fig. 1. Flowchart of Our Methodology

A. Dataset Description

The dataset used in this study was sourced from Kaggle. The dataset comprises network monitoring records from

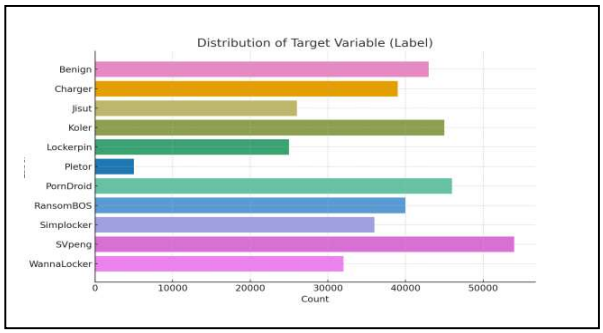

Fig. 2. Class Wise Distribution of Dataset

android devices, encompassing both benign and ransomware traffic types. With 203,556 rows and 85 columns, it includes ten types of Android ransomware: SVpeng, PornDroid, Koler, RansomBOS, Charger, Simplocker, WannaLocker, Jisut, Lockerpin, and Pletor. The distribution of data labels shows varying counts across ransomware types, with SVpeng having the highest at 54,161 records, visualized in Fig. 2.

B. Data Preprocessing

1) *Cleaning:* We have cleaned the dataset by removing duplicates and handling missing values.
2) *Feature Selection:* The dataset is refined by removing redundant or uncorrelated features using a heatmap-guided method.
3) *Label Encoding:* In order to make categorical variables compatible with machine learning models, we converted them into numerical representation using label encoding.
4) *Standardization:* To provide uniform scales and handle outliers, we used Standard Scaler to standardize numerical features.
5) *Train-Test Split:* To make evaluating the model easier, we separated the dataset into 80% training and 20% testing sets.
6) *Exploratory Data Analysis (EDA):* To address class imbalances and comprehend the distribution of the target variable (the "Label"), we performed visual analysis.

These pivotal preprocessing stages formed the foundation for developing effective Android ransomware detection models.

C. Ensemble Stacking Architecture

Ensemble stacking model consists of diverse algorithms such as Logistic Regression, Random Forest, Decision Tree, MLP, K-nearest Neighbors using a meta-classifier (Logistic Regression) to record multiple aspects of data. By integrating the predictions of several models, ensemble stacking reduces the shortcomings of each one. Ensemble stacking enhances overall prediction accuracy and lowers the possibility of false positives and negatives by combining many classifiers and leveraging their complimentary capabilities. The equation (1) represents stacking classifier-

$$P(Y) = Logistic(\sum_{i=1}^{N} w_i . Model_i\,(X)) \qquad (1)$$

where P(Y) is the predicted probability of the class, w_i are weights, and $Model_i(X)$ is the prediction from each base model. Fig. 3. Shows the ensemble stacking architecture used in our experiment.

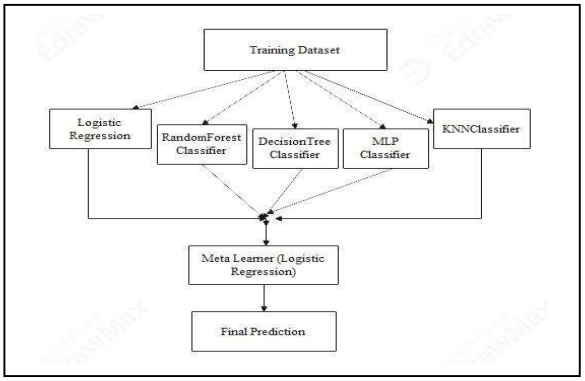

Fig. 3. Ensemble Stacking Architecture

IV. RESULT AND ANALYSIS

A. Model Performance Metrics

The Table I compares our ensemble stacking model and other traditional machine learning algorithms including Random Forest, K-nearest Neighbors (KNN), MLP Classifier, Logistic Regression and Decision Tree. When the outcomes are summarized, it is clear that the Ensemble Stacking model stands out as the best performer, earning a perfect 100% across all metrics- Accuracy, Precision, Recall, and F1-score. This indicates its extraordinary ability to appropriately classify instances. The Decision Tree model also excels, boasting high accuracy and precision at 99.8%. In contrast, the K-nearest Neighbors model displays comparatively lower scores, indicating moderate effectiveness with an accuracy of 75.2%. Logistic Regression performs less optimally, securing the lowest scores among the listed models, with an accuracy of 58.9%.

TABLE 1. MODEL PERFORMANCE METRICS

Model Name	Accuracy	Precision	Recall	F1-Score
Random Forest	0.967	0.967	0.967	0.966
K-nearest Neighbors	0.752	0.754	0.752	0.752
MLP Classifier	0.885	0.887	0.885	0.885
Logistic Regression	0.589	0.585	0.589	0.581
Decision Tree	0.998	0.998	0.998	0.998
Ensemble Stacking	1.00	1.00	1.00	1.00

B. Explaining Model Predictions with LIME

LIME is a machine learning explainability technique that interprets predictions from sophisticated black-box models. LIME is model-agnostic, meaning it works with any classifier or regressor (e.g., neural networks, decision trees, ensembles, etc.).

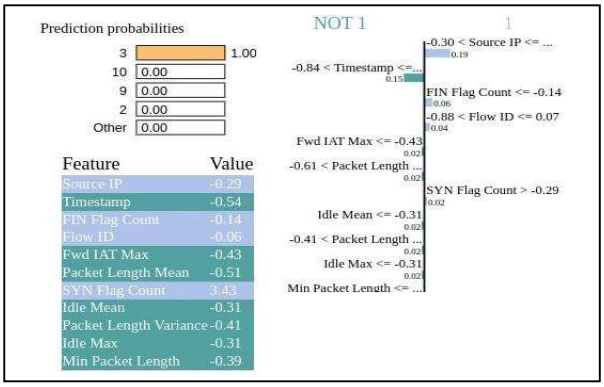

Fig. 4. LIME Explanation for 9000th Instance

LIME also helps to explain predictions from the meta-model, understand individual base models and identify important features. Fig. 4. explains the prediction made by our proposed model for 9000th instance. From this Fig, we can see our model predicts class 3 (Jisut) with certainty (1), while other classes are assigned a probability of 0. The figure also shows the positive and negative influence of various features on the model's prediction.

C. Research Goal Clarification

The primary goal of this research is to surpass the performance of existing Android ransomware detection models. Traditional machine learning and deep learning models demonstrate commendable results. However, the introduction of an ensemble model, particularly the Stacking Classifier, aims for unparalleled accuracy and precision. The emphasis is on safeguarding all Android-related devices against ransomware attacks.

D. Significance of Explainable AI (LIME)

The utilization of Explainable AI, exemplified by LIME, provides transparency into the decision-making process of complex models. By visualizing feature importance, as demonstrated in the LIME output for instance 9000, stakeholders gain insights into how the model arrives at its predictions. In applications like Android ransomware detection, where precision and interpretability are critical, this openness is essential for maintaining confidence in the model's findings.

V. DISCUSSION

A. Model Performance

The models exhibit varying degrees of performance. The Stacking Classifier outshines others, achieving a perfect score across all metrics, showcasing its effectiveness in Android ransomware detection. Random Forest and Decision Tree models demonstrate high accuracy and precision, underlining their robustness in handling the dataset's complexity. Logistic Regression and K-nearest Neighbors show comparatively lower performance, emphasizing the importance of selecting appropriate models for the task.

B. Explainable AI Insights

The LIME analysis provides valuable insights into the decision-making process of the models, especially for instances like 9000. For the Stacking Model, the certainty in predicting class 3 aligns with the overall model accuracy, reinforcing the model's reliability. Logistic Regression assigns a high probability to class 3, showcasing its role in contributing to the ensemble's decision.

C. Research Goal Achievement

The research introduces a powerful ensemble model to meet its goal of outperforming existing algorithms for Android ransomware detection. This new method improves detection effectiveness and is a major step forward in combating threats from Android ransomware. Moreover, the study uses explainable AI methods more precisely, LIME to shed light on the inner workings of the model. This transparency makes the model's decision-making process easier to understand.

VI. Conclusions

In summary, this study offers a very accurate and precise ensemble model and advances the field of Android ransomware detection. Our proposed Stacking Classifier orchestrates the merging of deep learning with traditional machine learning models, producing remarkable results. Transparency in decision-making is crucial, and the incorporation of Explainable AI more especially, LIME improves the model's interpretability. The model's predictions may be trusted by stakeholders, and the knowledge gathered by LIME offers a better comprehension of how the model differentiates between ransomware incidents. This study establishes a standard for Android ransomware detection going forward and highlights the value of Explainable AI and ensemble models in improving interpretability and performance. Moreover, it can serve as a strong protection system for android devices.

References

[1] Q. Hou et al., 'Large-scale security measurements on the android firmware ecosystem', in Proceedings of the 44th International Conference on Software Engineering, 2022, pp. 1257–1268.

[2] M. Rashid, J. Kamruzzaman, T. Imam, S. Wibowo, and S. Gordon, 'A tree-based stacking ensemble technique with feature selection for network intrusion detection', Applied Intelligence, vol. 52, no. 9, pp. 9768–9781, 2022.

[3] A. K. Mishra and S. Paliwal, 'Mitigating cyber threats through integration of feature selection and stacking ensemble learning: the LGBM and random forest intrusion detection perspective', Cluster Computing, vol. 26, no. 4, pp. 2339–2350, 2023.

[4] H. Mohanty, A. H. Roudsari, and A. H. Lashkari, "Robust stacking ensemble model for darknet traffic classification under adversarial settings," Computers & Security, vol. 120, p. 102830, 2022.

[5] R. Soleymanzadeh, M. Aljasim, M. W. Qadeer, and R. Kashef, 'Cyberattack and fraud detection using ensemble stacking', AI, vol. 3, no. 1, pp. 22–36, 2022.

[6] D. Minh, H. X. Wang, Y. F. Li, and T. N. Nguyen, 'Explainable artificial intelligence: a comprehensive review', Artificial Intelligence Review, pp. 1–66, 2022.

[7] S. Sharma, R. Kumar, and C. Rama Krishna, 'A survey on analysis and detection of Android ransomware', Concurrency and Computation: Practice and Experience, vol. 33, no. 16, p. e6272, 2021.

[8] M. Dhalaria and E. Gandotra, 'Android malware detection techniques: A literature review', Recent Patents on Engineering, vol. 15, no. 2, pp. 225–245, 2021.

[9] I. Almomani et al., 'Android ransomware detection based on a hybrid evolutionary approach in the context of highly imbalanced data', IEEE Access, vol. 9, pp. 57674–57691, 2021.

[10] S. Sharma, C. R. Krishna, and R. Kumar, 'RansomDroid: Forensic analysis and detection of Android Ransomware using unsupervised machine learning technique', Forensic Science International: Digital Investigation, vol. 37, p. 301168, 2021.

[11] R. J. Raimundo and A. T. Rosário, 'Cybersecurity in the internet of things in industrial management', Applied Sciences, vol. 12, no. 3, p. 1598, 2022.

[12] A. G. Sreedevi, T. N. Harshitha, V. Sugumaran, and P. Shankar, 'Application of cognitive computing in healthcare, cybersecurity, big data and IoT: A literature review', Information Processing & Management, vol. 59, no. 2, p. 102888, 2022.

[13] Z. Zhang, H. Al Hamadi, E. Damiani, C. Y. Yeun, and F. Taher, 'Explainable artificial intelligence applications in cyber security: State-of-the-art in research', IEEE Access, vol. 10, pp. 93104–93139, 2022.