JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 12(1), MARCH, 2024
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

An Interactive Dashboard for Visualizing and Evaluating Mobile Malware Detection Performance Using Machine Learning and Deep Learning Algorithms

Charles Ikpeama, Azita Pourghasem
Department of Cybersecurity,
University of Hertfordshire (UH), UK.

ABSTRACT

In recent years, mobile devices' ubiquitous presence has been paralleled by an escalating surge in mobile malware threats, highlighting unprecedented security concerns. This research endeavor delineates the development of an advanced interactive dashboard, tailored to visualise and critically assess an array of machine and deep learning algorithms for mobile malware detection. Designed as a specialized interface, the dashboard equips cybersecurity experts, data analysts, and researchers with the means to engage dynamically with pertinent data, offering exploratory avenues for crucial performance metrics—namely, accuracy, precision, recall, and the F1 score. Built on a dataset that integrates mobile malware classifications, the dashboard extends instantaneous perceptiveness into algorithmic efficacy. A rigorous analytical process underscores the algorithmic selection, including Random Forest, Logistic Regression, Naive Bayes, Neural Networks (NN), and Deep Neural Networks (DNN). Each algorithm is meticulously evaluated to ascertain its merit and relevance for detecting mobile malware. The research magnifies the significance of state-of-the-art deep learning modalities, primarily focusing on NNs and DNNs, aiming to bolster the precision and resilience of detection mechanisms. The dashboard's incorporation of these paradigms stands instrumental in enhancing defenses against progressive malware challenges. Supplemented by diverse graphical representations, the dashboard elucidates intricate facets of algorithmic output, while diagrams such as the Entity-Relationship (ERD), Sequence, Deployment, State, Use Case, and Activity cohesively offer an exhaustive insight into the system's structural blueprint and operative capacities. Conclusively, this undertaking epitomises a pivotal evolution in mobile malware detection paradigms. By orchestrating cutting-edge algorithms, all-encompassing datasets, and lucid visual aids, it forges a pathway towards effectively navigating and mitigating the complexities intrinsic to mobile malware threats. This novel interface stands as a vanguard, priming stakeholders with an astute instrument for strategic decision-making and fostering progressive strides in mobile security domains.

ARTICLE INFO
Article History
Received: November, 2023
Received in revised form: January, 2023
Accepted: March, 2024
Published online: March, 2024

KEYWORDS
Interactive Dashboard, Mobile Malware, Machine Learning, Deep learning Algorithms

INTRODUCTION

Smart Mobile devices are now essential tools for communication, work, and data access. With their rapid growth, there's been a corresponding increase in cyber threats, particularly mobile malware targeting mobile software and operating systems. (Cinar and Kara, 2023). According to Ibrahim et al., this poses a significant risk to individual, corporate, and governmental digital security, leading to potential data breaches and compromised infrastructures.

To counter these threats, advanced mobile malware detection mechanisms are required (Ciaramella et al., 2020). While traditional signature-based methods remain

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

important, they struggle to keep up with the evolving nature of mobile malware, highlighting the need for more advanced detection solutions (Akintola et al., 2020).

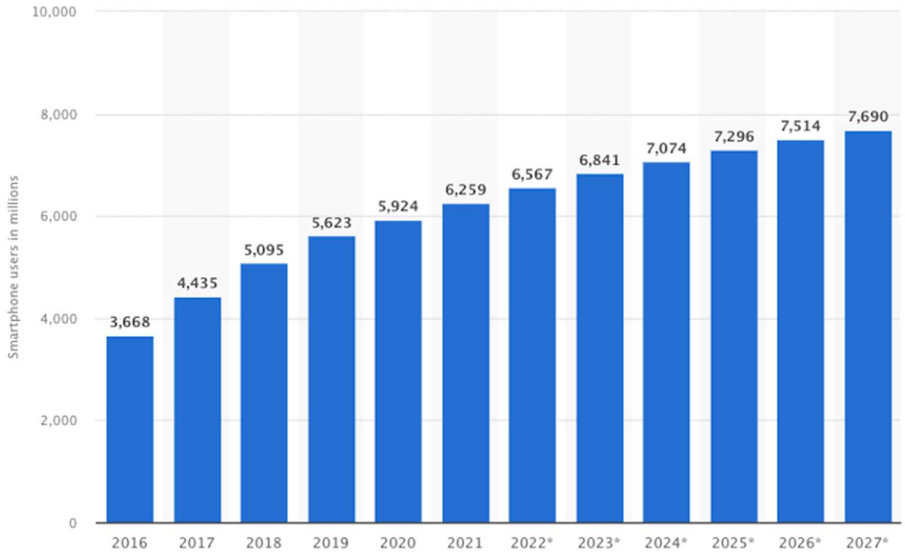

Figure 1: Smart phone user base growth smartphone user base growth *(Patel,2023)*

This research introduces an interactive dashboard designed to visualise and evaluate the performance of various machine learning and deep learning algorithms for mobile malware detection. The dashboard combines a user-friendly interface with real-time analytics, aiming to assist cybersecurity professionals in malware detection. The following chapters detail the project's approach, design, implementation, and results, emphasizing the dashboard's importance in enhancing mobile security against a growing range of malware threats

Statement of Problem

According to Krishnappa, in today's digital environment, the widespread use of mobile devices has led to an increase in mobile malware threats, affecting individuals, businesses, and the broader digital ecosystem. Traditional signature-based detection methods struggle to address these threats due to the constant evolution of malware techniques (Tayyab et al.,2022). Evaluating the performance of detection algorithms is complex and requires specialized expertise. (Alsanad and Altuwaijri,2022).

The primary challenge is to develop a solution with a clear interface that allows users to assess and compare the performance of various machine learning and deep learning algorithms for mobile malware detection. This solution should provide real-time visualization, interactive result analysis, and the flexibility to adapt to the changing nature of mobile malware. The ultimate goal is to provide cybersecurity professionals, researchers, and decision-makers with a tool for choosing the best algorithms, thereby enhancing strategies to combat mobile malware.

Objectives of the Study

This project is anchored on the following core objectives:

1. Dashboard Development: Design a dashboard for user-driven visualization and comparison of machine and deep learning algorithms for mobile malware detection.
2. Performance Evaluation: Include a range of machine and deep learning algorithms and evaluate key performance metrics such as accuracy, precision, recall, and F1 score.
3. Real-time Data Display: Provide real-time graphical displays of algorithm performance using charts and graphs for quick insights.
4. Interface Design: Develop a user-friendly UI that allows easy switching between

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

algorithms, performance metrics, and visualization styles.
5. Validation: Ensure the accuracy of the dashboard's results by comparing them with known benchmarks and industry standards.
6. Documentation: Provide detailed documentation and a user manual to help users navigate the dashboard and understand metrics.
7. Ethical Considerations: Address ethical issues related to malware dataset use and ensure adherence to ethical standards throughout the project.
8. Knowledge Sharing: Share the project's findings and methods through academic papers, presentations, and open-source contributions.

By achieving these objectives, the project aims to provide the cybersecurity field with a practical tool to improve decision-making for mobile malware detection and advance cybersecurity research and practice.

LITERATURE REVIEW

With the proliferation of mobile devices, there's a significant increase in malware threats, particularly targeting OS and app vulnerabilities. Traditional defense mechanisms often lag in tackling these threats. Notably:

1. Trojans can disguise as genuine apps, jeopardizing user data and device operations (Alzubaidi 2022).
2. Adware and spyware introduce privacy threats by overloading users with unwanted ads and spying on user activities (Srivastava 2021).
3. The rise in mobile transactions has seen a spike in ransomware attacks targeting encrypted data (Alotaibi and Fawad 2022).

Given these dynamic threats, there's a growing need for innovative cybersecurity measures. The application of AI, especially machine learning (ML) and deep learning (DL), is proving valuable in this realm. ML techniques fall into three primary categories: supervised, unsupervised, and semi-supervised. Notably, Xie et al. (2020) explored semi-supervised learning, introducing the MUSCLE approach, which combines both labeled and unlabeled data, showcasing its potential in cybersecurity.

Deep learning, especially neural networks, is becoming increasingly relevant in cybersecurity. Deldar and Abadi (2023) provided an in-depth analysis of deep learning for detecting novel malware. Ramos et al. (2022) also demonstrated deep learning's versatility, applying deep semi-supervised and self-supervised methods to medical imaging—hinting at their potential applicability in malware detection.

In essence, as malware techniques evolve, there's an urgent need for the ongoing development of AI-enhanced solutions to ensure robust cyber defense.

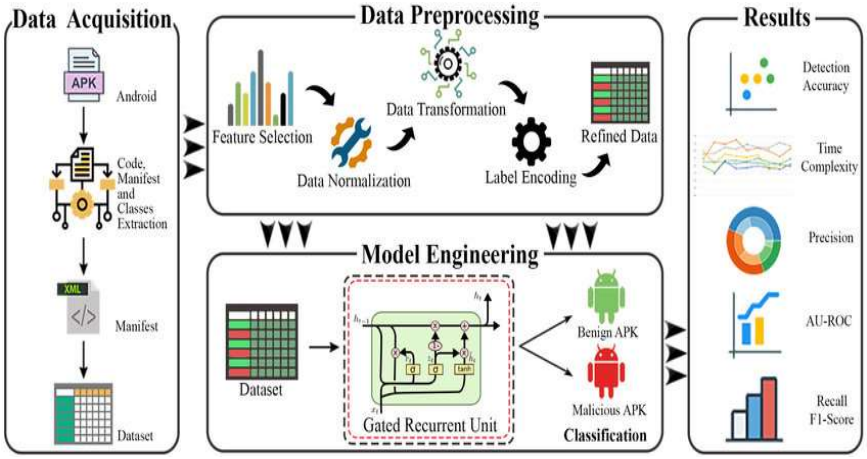

Figure 2: The Simplified Overview of Proposed GRU-based Android Malware Detection scheme (Iram et al., 2020)

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

Evaluating malware detection algorithms using metrics such as accuracy, precision, and recall is vital for assessing their real-world efficiency. Kisambu and Mjahidi (2022) highlighted the significance of accuracy in evaluating ML algorithms for detecting malware phishing attacks. Meanwhile, Bibi et al. (2020) used precision and recall to examine a deep learning model's capability against advanced Android malware, emphasizing the metrics' role in understanding false positives and negatives. Alazzam et al. (2022) underscored the F1-score's value, especially when dealing with imbalanced data, as it offers a harmonized measure of an algorithm's performance. These metrics are essential not just for quantitative assessment but also for enhancing the algorithm's adaptability to changing cyber threats.

Comparative studies, backed by standardized datasets, are crucial in malware detection, providing a means to measure different algorithmic effectiveness. Alkahtani and Aldhyani (2022) showcased this by comparing several machine and deep learning algorithms on benchmark datasets for Android mobile devices, emphasizing the need for empirical, data driven evaluations as follows:

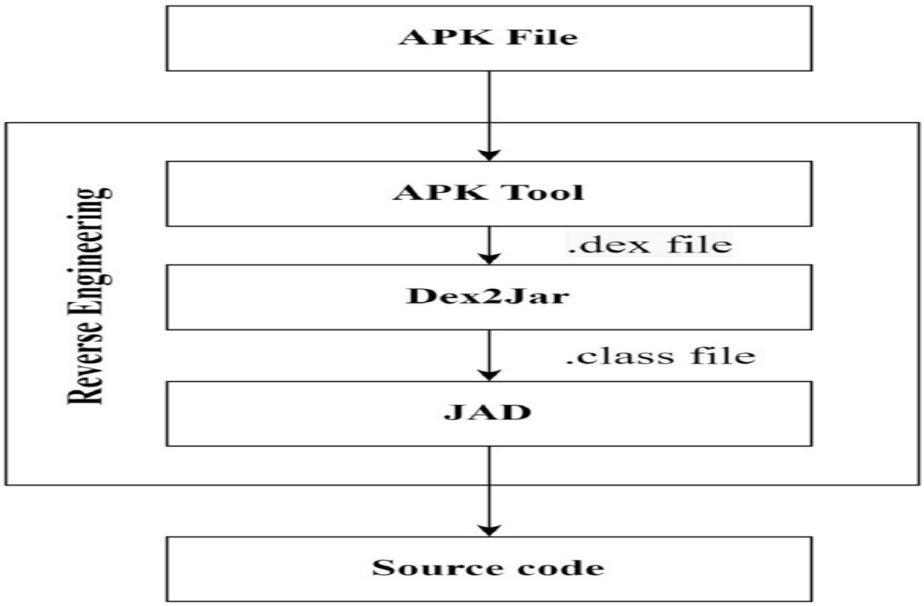

Figure 3: Android application package (APK) reverse engineering (Rafiq, Husnain, et al., 2022)

1. Machine Learning in IIoT: BhupalNaikD et al. (2022) compared various machine learning algorithms for anomaly detection in IIoT data, emphasizing the need for domain-specific datasets.
2. Repacked Malware in Android: Rafiq et al. (2022) tackled repacked malware with AndroMalPack, stressing the removal of repacked apps from datasets for improved accuracy.
3. Feature Extraction and Selection: This process is central to mobile malware detection. It turns raw app data into actionable insights, focusing on app behaviors and structures. Catal et al. (2021) identified predominant features, while Tokmak et al. (2021) explored both static and dynamic analysis methods for detection. Duraisamy and Subbiah (2022) highlighted refining extracted features for efficiency and accuracy.
4. Dataset Considerations: Datasets are crucial for algorithm development and evaluation. Quality, representativeness, and pre-processing of datasets are essential. Riaz et al. (2022) and Lavanya Bharathi and Chandrabose (2022) highlighted pre-processing's importance, like scaling and

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

de-noising. Rao and Babu (2023) addressed imbalanced datasets with IGAN, emphasizing proper data pre-processing for deep learning.

5. Visualization and User Interaction: Visualization translates complex data into understandable visuals. Menin et al. (2021) introduced an interactive tool, ARViz, for exploring associations. Mayer and colleagues (2022) developed a web-based application for analyzing surgical data, emphasizing interactive visual exploration. Srinivasan et al. (2020) presented DataBreeze, which offers a multimodal data exploration experience. Zong and team (2022) emphasized non-visual affordances for accessible data visualization. The emphasis is on converting intricate data into actionable insights through interactive dashboards and tools.

Machine Learning and Deep Learning in Mobile Malware Detection

The rapid increase in mobile device usage has resulted in a corresponding rise in mobile malware, driving the need for advanced detection methods. Deep learning, a subset of machine learning, has become essential for identifying patterns in extensive datasets to address this challenge. Mercaldo et al. (2022) studied deep learning's role in mobile malware detection, testing various Convolutional Neural Network models on a significant dataset of Android applications. Their research emphasized the importance of understanding how models make decisions and showed deep learning's effectiveness in distinguishing between benign and malicious apps. Meanwhile, Ramos et al. (2022) utilized a semi-supervised learning approach for diabetic retinopathy detection, demonstrating its potential adaptability for mobile malware detection, even in contexts with limited labeled data.

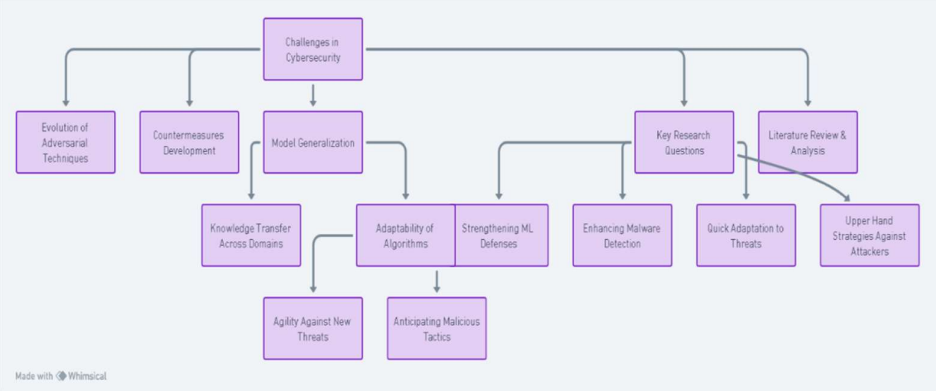

Figure 4: Challenges and open research questions in cybersecurity

Performance Metrics for Algorithm Evaluation

Performance metrics are essential in mobile malware detection as they quantitatively assess the efficiency and reliability of detection algorithms. Kisambu and Mjahidi (2022) highlighted the value of metrics like accuracy, precision, recall, and F1-score in evaluating machine learning algorithms for malware-based phishing attacks. Alazzam et al. (2022) also underscored the importance of these metrics, using them to assess a wrapper-based approach for Android malware detection. Meanwhile, Kim et al. (2023) introduced a new metric, the time-series aware precision and recall (TaPR), specifically designed for time-series data in industrial control systems. Collectively, these studies emphasize the critical role of performance metrics in refining and enhancing malware detection methods as cyber threats continue to evolve.

Existing Interactive Dashboard Solutions

Various interactive dashboard solutions exist for data analytics and visualization, including Tableau, Microsoft

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

Power BI, Metabase, and Plotly. While Tableau and Power BI provide dynamic charts and user-friendly interfaces, Metabase offers features tailored for non-technical users, and Plotly integrates Python and JavaScript for in-depth visualisations, especially when paired with Dash for interactive web applications. For the project, Dash/Plotly was selected due to its flexibility, integration potential, and efficiency in delivering interactive visuals (Michael et al., 2019).

Research Design

The following chapter explains the methodology employed to fulfil the objectives of this project. This systematic approach underpins the creation of an interactive dashboard designed for visualizing and examining the efficacy of machine learning and deep learning algorithms in mobile malware detection. The subsequent sections delve into the detailed methodologies that determine the project's direction.

Table 1: Dataset Processing Methodology

STEP NUMBER	DESCRIPTION
1. Data Acquisition	Collect data from various trusted sources, e.g., VirusShare.
2. Data Inspection	Examine the dataset for initial anomalies, missing values, and understand the data distribution.
3. Data Cleaning	Remove or impute missing values, handle duplicates, and remove unnecessary columns | such as 'Name' and 'md5'
4. Data Transformation	Normalize or standardize features to bring them to a similar scale, especially for algorithms sensitive to feature scales.
5. Feature Extraction	Extract new features from existing ones, e.g., using domain knowledge or automated feature extraction techniques.
6. Data Splitting	Divide the dataset into training and test sets to ensure proper evaluation.
7. Oversampling/Undersampling	Handle class imbalance by either oversampling the minority class or under sampling the majority.
8. Feature Selection	Use techniques like correlation analysis, recursive feature elimination, or others to select the most relevant features.
9. Final Dataset Review	Inspect the pre-processed dataset before Feeding it to machine learning models.

RESULT DISCUSSION, PROJECT MANAGEMENT AND RELEVANCE

Dataset

Data integrity and relevance are paramount in any data science endeavour. Recognising this, the dataset procurement and selection phase was approached with meticulous care and adherence to best practices.

For the purpose of this project, we utilized a comprehensive dataset sourced from VirusShare (*https://virusshare.com/*). This platform functions as a reservoir of malware samples, continually updated to serve security researchers, incident responders, and other professionals in the infosec community. As of **15th September 2023**, VirusShare's system reported a total of **68,889,089** malware samples.

For the specific requirements of our research, we extracted **96,724** malware files from VirusShare. In addition, to create a comprehensive dataset, we supplemented this malware data with **41,323** legitimate or benign binaries (comprising .exe and .dll files). By amalgamating these data sources, we

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

constructed a CSV file titled "malData.csv", which serves as the foundation for our project's analysis.

Given the sensitive nature of the malware samples used, extreme caution was taken during the data extraction and handling processes. The overarching goal of this exercise is to foster a deeper understanding of mobile malware characteristics, leveraging both malicious and benign datasets to draw meaningful insights.

The culmination of rigorous research methodology, data collection, and algorithmic processing is the generation of results that offer insights into the landscape of mobile malware detection. This section delves into the outcomes derived from the implemented algorithms, the interactive dashboard visualizations, and the subsequent analysis. Through a combination of quantitative data, visual representations, and interpretative analysis, we aim to shed light on the effectiveness of the chosen methodologies and their implications in the broader context of mobile security. As we navigate through the results, it's essential to view them in light of the project's objectives, the inherent limitations, and the overarching goal of enhancing mobile cybersecurity.

Table 2: Comparative Analysis of Algorithm Performance

ALGORITHMS OBSERVATIONS	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Logistic Regression	98.04	97.15	96.35	|96.75
Random Forest	99.53	99.12	99.33	99.22
Naive Bayes	47.14	36.40	99.76	53.33
Neural Network	99.23	98.40	99.07	98.73
DeepNeural Network	99.41	98.81	99.23	99.02

Dashboard Interface Overview

The heart of our project's data visualization lies in the interactive dashboard, meticulously designed to offer users an intuitive and immersive experience. The dashboard serves as a bridge between raw data and actionable insights, translating complex algorithmic outcomes into comprehensible visual narratives.

Table 3: Dashboard Interface Components

COMPONENT	DESCRIPTION
Algorithm Dropdown Selector	Allows the user to select and view the performance metrics for a specific machine learning or deep learning algorithm. Options include algorithms like Logistic Regression, Random Forest, etc.
Metric Dropdown Selector	Enables the user to pick a specific performance metric (like Accuracy, Precision, etc.) or 'All' to view. This controls the visual representation in the graphs below.
Performance Chart	A dynamic bar or pie chart (based on the metric selection) that displays the performance metrics of the chosen algorithm.
Algorithm Distribution Graph	A bar chart that visually represents the distribution of performance metrics for the selected algorithm.
Metric Distribution Graph	Another bar chart showing how the chosen performance metric distributes across different algorithms.

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

Layout Design

The dashboard's layout is structured to guide users seamlessly through the data. A clean, organized interface ensures that users can quickly locate and interact with the desired metrics without feeling overwhelmed. The design prioritizes clarity and simplicity, ensuring that even individuals without a technical background can navigate and understand the presented data.

Dynamic Components

One of the dashboard's standout features is its dynamic components. Dropdown menus allow users to filter and select specific data subsets, tailoring the visualizations to their unique queries. Checkboxes provide options to toggle between different data points, offering a comparative view that can reveal hidden patterns or trends.

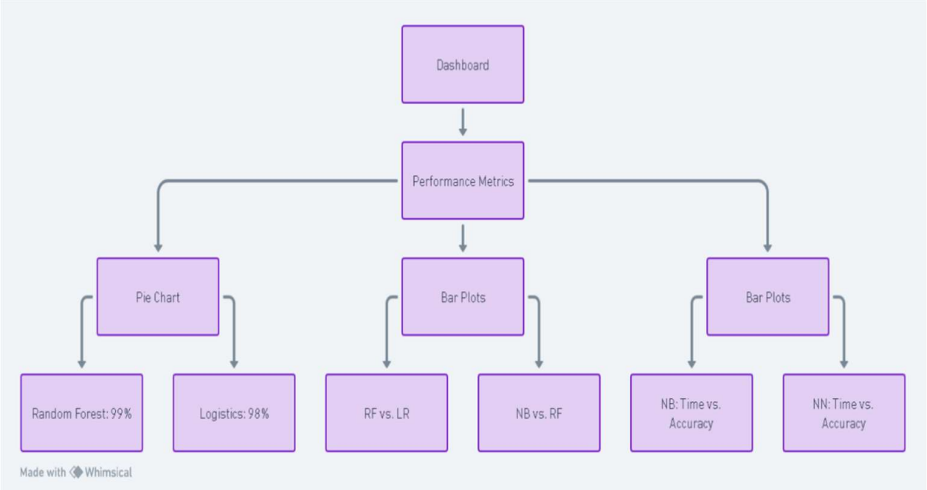

Figure 5: Dashboard performance Visualisation Spectrum

Interactive Plots

Central to the dashboard are the interactive plots, primarily pie charts, which offer a visual representation of the "performance_metrics" dataset. These plots are not just static images; they respond to user interactions. Hovering over a segment of the pie chart, for instance, might reveal additional information or percentages, enhancing the depth of analysis. The choice of pie charts, in particular, allows for an immediate understanding of proportions and distributions, making them ideal for representing the diverse metrics in our dataset.

User-Centric Design

The overarching philosophy behind the dashboard's design is user-centricity. Recognizing that the best tool is one that can be used effectively, every element of the dashboard, from its color palette to its interactive features, has been chosen with the user in mind. The goal is to empower users, regardless of their expertise level, to explore, analyze, and derive meaningful insights from the data.

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 12(1), MARCH, 2024
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

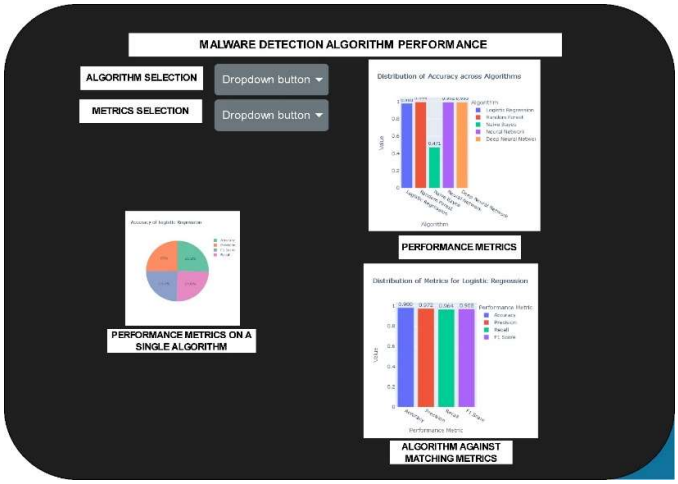

Figure 62: Dashboard UI/UX

The dashboard interface is more than just a tool; it's a gateway to understanding the intricacies of mobile malware detection. Through its thoughtful design and interactive features, it invites users to engage with the data, fostering a deeper appreciation of the project's findings and their implications.

Data Visualization and Analysis

Data visualisation is a powerful tool, transforming raw numbers into visual narratives that can be easily interpreted and understood. In the context of mobile malware detection, visual representations provide a clear picture of algorithmic performance, highlighting areas of strength and potential improvement.

Visualisation Types:

The dashboard incorporates a variety of visualization tools:

Plots: These are used to represent data points in a two-dimensional space, allowing for a clear comparison between different algorithms or metrics. For instance, scatter plots might be used to visualise the relationship between precision and recall for different algorithms.

Charts: Charts, such as bar and pie charts, offer a visual breakdown of categorical data. A bar chart might display the accuracy of each algorithm, while a pie chart could represent the distribution of true positives, false positives, true negatives, and false negatives for a particular algorithm.

Graphs: Graphs, especially line graphs, are instrumental in showcasing trends over time or across different parameters. They might be used to track the performance of an algorithm as it processes different batches of data.

User-Driven Analysis:

One of the standout features of the dashboard is its interactivity. Users can:

1. Switch between algorithms: This allows for a side-by-side comparison, highlighting the strengths and weaknesses of each algorithm in real-time.
2. Select metrics: Depending on the focus, users can choose to view results based on accuracy, precision, recall, or the F1 score, ensuring a tailored analysis experience.
3. Choose visualization types: Depending on the data's nature and the user's preference, different visualization tools can be selected, ensuring the most effective representation of the data.

Interpretation and Insights:

Beyond mere visualization, the dashboard facilitates data interpretation. The visual tools are complemented by descriptive statistics and insights, guiding users through the data's nuances. For instance, while a high precision might be celebrated, the dashboard might also highlight a corresponding drop in recall, prompting a more in-depth analysis.

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

Table 4: Key Findings Overview

KEY FINDINGS	DESCRIPTION
Best performing algorithm	The "Random Forest" algorithm displayed the highest accuracy, outperforming other models with an accuracy score of 0.9953.
Least effective algorithm	"Naive Bayes" showed significantly lower performance in terms of accuracy, with a score of 0.4714, indicating it may not be the best choice for this dataset.
Neural Network Variance	The deep neural network slightly outperformed the basic neural network in terms of F1 Score, suggesting the additional complexity could provide benefit in some scenarios.
Recall Importance	Despite its low accuracy, the Naive Bayes model had the highest recall of 0.9976, indicating its potential ability to catch the majority of positive cases, albeit with a higher false positive rate.
Dashboard Utility	The interactive dashboard proved invaluable in swiftly comparing and visualizing the performance of various models, streamlining decision-making processes.

CONCLUSION

The findings presented here are not just a testament to the rigorous methodologies employed but also highlight the nuances and intricacies of mobile malware detection. From the initial data collection to the final visualization on the interactive dashboard, every step has contributed to a richer understanding of the domain.

The project's contributions to the field of mobile malware detection are outlined. It highlights the endeavor's dual focus on deepening understanding and pioneering new methodologies. Innovations and Insights, the project's multifaceted approach to malware detection is discussed. By incorporating diverse algorithms such as Random Forest and Deep Neural Networks, the project has developed a comprehensive detection mechanism capable of addressing the intricate nature of mobile malware. Additionally, the project emphasizes the importance of feature engineering in optimizing algorithmic performance, shedding light on its pivotal role in the detection process.

RECOMMENDATIONS

The rapidly evolving landscape of cybersecurity and the increasing sophistication of mobile malware threats necessitate a forward-thinking approach. Section 5.4 delves into the potential avenues for future enhancements, ensuring that the project remains at the forefront of innovation and effectiveness in malware detection.

1. A significant enhancement planned for the dashboard is the capability for users to upload datasets in real-time. This feature will empower users to not only analyze predefined datasets but also to evaluate their own data on-the-fly. Upon uploading, the dashboard will process the dataset, run the algorithms, and instantly visualize the respective metrics. This real-time interaction will provide users with immediate insights, enhancing the dashboard's utility and interactivity.
2. As mobile malware types continue to evolve, there's a pressing need to continuously update and expand the dataset. Future iterations could involve integrating real-time threat intelligence feeds, collaborating with cybersecurity firms for live data, and sourcing from emerging repositories.
3. The field of machine learning and AI is in constant flux. While the project currently employs algorithms like Random Forest and Deep Neural Networks, there's potential to integrate more advanced models in the future, such as Generative Adversarial Networks or Transformer-based models.
4. Beyond the real-time dataset upload feature, the dashboard offers numerous opportunities for enhancement. Advanced visualization techniques, predictive analytics features, and compatibility with emerging technologies will be pivotal for the dashboard's future versions.

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

5. As the project grows, revisiting and reinforcing ethical considerations, especially around data privacy, will be crucial. Incorporating advanced anonymization techniques and ensuring GDPR compliance are potential areas of focus.
6. The domain of mobile malware detection offers vast potential for collaborative research. Partnerships with academic institutions, cybersecurity firms, and governmental agencies can lead to richer datasets, advanced algorithms, and more robust detection mechanisms.

REFERENCES

Akhtar, Muhammad Shoaib, and Tao Feng. "Malware Analysis and Detection Using Machine Learning Algorithms." *Symmetry*, vol. 14, no. 11, 1 Nov. 2022, p. 2304, www.mdpi.com/2073-8994/14/11/2304, https://doi.org/10.3390/sym14112304.

Akintola, Abimbola G., et al. "Empirical Analysis of Forest Penalizing Attribute and Its Enhanced Variations for Android Malware Detection." *Applied Sciences*, vol. 12, no. 9, 6 May 2022, p. 4664, https://doi.org/10.3390/app12094664. Accessed 25 Sept. 2022.

AlAbdulaali, Abeer, et al. "Designing Multimodal Interactive Dashboard of Disaster Management Systems." *Sensors*, vol. 22, no. 11, 5 June 2022, p. 4292, https://doi.org/10.3390/s22114292. Accessed 17 July 2022.

Alazzam, Hadeel, et al. "An Improved Binary Owl Feature Selection in the Context of Android Malware Detection." *Computers*, vol. 11, no. 12, 30 Nov. 2022, p. 173, https://doi.org/10.3390/computers11120173. Accessed 21 Jan. 2023.

Alkahtani, Hasan, and Theyazn H. H. Aldhyani. "Artificial Intelligence Algorithms for Malware Detection in Android-Operated Mobile Devices." *Sensors*, vol. 22, no. 6, 15 Mar. 2022, p. 2268, https://doi.org/10.3390/s22062268. Accessed 14 May 2022.

Almomani, Iman, et al. "Android Malware Analysis in a Nutshell." *PLoS ONE*, vol. 17, no. 7, 5 July 2022, p. e0270647, www.ncbi.nlm.nih.gov/pmc/articles/PMC9255778/#:~:text=There%20are%20a%20lot%20of, https://doi.org/10.1371/journal.pone.0270647.

Alomari, Esraa Saleh, et al. "Malware Detection Using Deep Learning and Correlation-Based Feature Selection." *Symmetry*, vol. 15, no. 1, 1 Jan. 2023, p. 123, www.mdpi.com/2073-8994/15/1/123, https://doi.org/10.3390/sym15010123. Accessed 18 Feb. 2023.

Alsanad, Ahmed, and Sara Altuwaijri. "Advanced Persistent Threat Attack Detection Using Clustering Algorithms." *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 9, 2022, https://doi.org/10.14569/ijacsa.2022.0130976. Accessed 17 Nov. 2022.

Alzubaidi, Abdulaziz. "Recent Advances in Android Mobile Malware Detection: A Systematic Literature Review." *IEEE Access*, vol. 9, 2021, pp. 146318–146349, https://doi.org/10.1109/access.2021.3123187.

Anandhi, V., et al. "Performance Evaluation of Deep Neural Network on Malware Detection: Visual Feature Approach." *Cluster Computing*, vol. 25, no. 6, 18 Aug. 2022, pp. 4601–4615, https://doi.org/10.1007/s10586-022-03702-3. Accessed 20 Nov. 2022.

Anderson, Janna, and Lee Rainie. "The Positives of Digital Life." *Pew Research Center: Internet, Science & Tech*, Pew Research Center: Internet, Science & Tech, 3 July 2018, www.pewresearch.org/internet/2018/07/03/the-positives-of-digital-life/.

Arrieta, Jose, et al. "Deep Semi-Supervised and Self-Supervised Learning for Diabetic Retinopathy Detection." *ArXiv (Cornell University)*, 6 Mar. 2023, https://doi.org/10.1117/12.2669723.

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

Accessed 10 Sept. 2023.

B, Harikrishnan N. "Confusion Matrix, Accuracy, Precision, Recall, F1 Score." *Medium*, 1 June 2020, medium.com/analytics-vidhya/confusion-matrix-accuracy-precision-recall-f1-score-ade299cf63cd.

Baniecki, Hubert, and Przemyslaw Biecek. "The Grammar of Interactive Explanatory Model Analysis." *The Grammar of Interactive Explanatory Model Analysis*, 14 Feb. 2023, https://doi.org/10.1007/s10618-023-00924-w. Accessed 13 July 2023.

BCS. *BCS, the CHARTERED INSTITUTE for IT CODE of CONDUCT for BCS MEMBERS*. 3 June 2015.

Bharathi, Lavanya, and Shanthi Chandrabose. "Machine Learning-Based Malware Software Detection Based on Adaptive Gradient Support Vector Regression." *International Journal of Safety and Security Engineering*, vol. 12, no. 1, 28 Feb. 2022, pp. 39–45, https://doi.org/10.18280/ijsse.120105 . Accessed 20 May 2022.

Bibi, Iram, et al. "A Dynamic DL-Driven Architecture to Combat Sophisticated Android Malware." *IEEE Access*, vol. 8, 2020, pp. 129600–129612, https://doi.org/10.1109/access.2020.3009819. Accessed 10 Oct. 2021.

Brownlee, Jason. "Logistic Regression for Machine Learning." *Machine Learning Mastery*, 22 Sept. 2016, machinelearningmastery.com/logistic-regression-for-machine-learning/.

"Call for Information: Unauthorised Access to Online Accounts and Personal Data." *GOV.UK*, www.gov.uk/government/consultations/unauthorised-access-to-online-accounts-and-personal-data/call-for-information-unauthorised-access-to-online-accounts-and-personal-data#:~:text=1. Accessed 13 Mar. 2023.

Catal, Cagatay, et al. "Applications of Deep Learning for Mobile Malware Detection: A Systematic Literature Review." *Neural Computing and Applications*, 23 Oct. 2021, https://doi.org/10.1007/s00521-021-06597-0. Accessed 8 Nov. 2021.

Chapman, Cameron . "A Complete Overview of the Best Data Visualization Tools." *Toptal Design Blog*, 2019, www.toptal.com/designers/data-visualization/data-visualization-tools.

Chaw, Jun Kit, et al. "Interactive Dashboard with Visual Sensing and Fast Reactivity." *The Journal of the Institution of Engineers, Malaysia*, vol. 82, no. 3, 24 Nov. 2022, https://doi.org/10.54552/v82i3.103. Accessed 16 Feb. 2023.

Chowdhury, Naseef-Ur-Rahman, et al. *Android Malware Detection Using Machine Learning: A Review*.

Ciaramella, Giovanni, et al. "Introducing Quantum Computing in Mobile Malware Detection." *Proceedings of the 17th International Conference on Availability, Reliability and Security*, 23 Aug. 2022, https://doi.org/10.1145/3538969.3543816. Accessed 10 Sept. 2023.

Cinar, Ahmet Cevahir, and Turkan Beyza Kara. "The Current State and Future of Mobile Security in the Light of the Recent Mobile Security Threat Reports." *Multimedia Tools and Applications*, 30 Jan. 2023, https://doi.org/10.1007/s11042-023-14400-6. Accessed 17 Feb. 2023.

"Do You Know What Deep Learning Is?" *Oracle.com*, 2022, www.oracle.com/artificial-intelligence/machine-learning/what-is-deep-learning/.

Donges, Niklas . "Recurrent Neural Networks 101: Understanding the Basics of RNNs and LSTM." *Built In*, 2019, builtin.com/data-science/recurrent-neural-networks-and-lstm.

Duraisamy Soundrapandian, Pradeepkumar, and Geetha Subbiah. "MULBER: Effective Android Malware Clustering Using Evolutionary Feature Selection and Mahalanobis Distance Metric." *Symmetry*, vol. 14, no. 10, 21 Oct. 2022, p. 2221, https://doi.org/10.3390/sym14102221 . Accessed 4 Dec. 2022.

Fang, Yong, et al. "DeepDetectNet vs

Corresponding author: Charles Ikpeama
✉ *charlyikpeama@yahoo.com*
Department of Cybersecurity, University of Hertfordshire (UH), UK.

RLAttackNet: An Adversarial Method to Improve Deep Learning-Based Static Malware Detection Model." *PLOS ONE*, vol. 15, no. 4, 23 Apr. 2020, p. e0231626, https://doi.org/10.1371/journal.pone.0231626. Accessed 18 Aug. 2022.

Fatemeh Deldar, and Mahdi Abadi. *Deep Learning for Zero-Day Malware Detection and Classification: A Survey*. 24 June 2023, https://doi.org/10.1145/3605775. Accessed 9 July 2023.

Forcepoint. "What Is Mobile Malware?" *Forcepoint*, 4 Dec. 2018, www.forcepoint.com/cyber-edu/mobile-malware.

Gavrishchaka, Valeriy, et al. "Synergy of Physics-Based Reasoning and Machine Learning in Biomedical Applications: Towards Unlimited Deep Learning with Limited Data." *Advances in Physics: X*, vol. 4, no. 1, 1 Jan. 2019, p. 1582361, https://doi.org/10.1080/23746149.2019.1582361. Accessed 16 Nov. 2020.

Guendouz, Mohamed, and Abdelmalek Amine. "A New Feature Selection Method Based on Dragonfly Algorithm for Android Malware Detection Using Machine Learning Techniques." *International Journal of Information Security and Privacy*, vol. 17, no. 1, 10 Mar. 2023, pp. 1–18, https://doi.org/10.4018/ijisp.319018. Accessed 30 Mar. 2023.

Ibrahim, A., et al. "The Challenges of Leveraging Threat Intelligence to Stop Data Breaches." *Semantic Scholar*, 2020, pdfs.semanticscholar.org/d20c/96a5047860dab5517af4189542bc06bf796c.pdf. Accessed 24 Mar. 2021.

Ijcsis, Journal of Computer Science. "Journal of Computer Science IJCSIS February 2017 Part I.pdf." *Www.academia.edu*, www.academia.edu/36011040/Journal_of_Computer_Science_IJCSIS_February_2017_Part_I_pdf. Accessed 10 Sept. 2023.

Jayaswal, Vaibhav. "Performance Metrics: Confusion Matrix, Precision, Recall, and F1 Score." *Medium*, 15 Sept. 2020, towardsdatascience.com/performance-metrics-confusion-matrix-precision-recall-and-f1-score-a8fe076a2262.

Kim, Ga-Yeong, et al. "A Study on Performance Metrics for Anomaly Detection Based on Industrial Control System Operation Data." *Electronics*, vol. 11, no. 8, 12 Apr. 2022, p. 1213, https://doi.org/10.3390/electronics11081213. Accessed 7 Aug. 2022.

Kim, Yu-kyung, et al. "A Systematic Overview of the Machine Learning Methods for Mobile Malware Detection." *Security and Communication Networks*, vol. 2022, 22 July 2022, p. e8621083, www.hindawi.com/journals/scn/2022/8621083/, https://doi.org/10.1155/2022/8621083.

Krishnappa, Triveni. "MOBILE SECURITY THREATS and a SHORT SURVEY on MOBILE AWARENESS: A REVIEW." *International Journal of Engineering Applied Sciences and Technology*, vol. 7, no. 8, 1 Dec. 2022, pp. 83–88, https://doi.org/10.33564/ijeast.2022.v07i08.007. Accessed 17 Apr. 2023.

L. Kisambu, Kambey, and Mohamedi Mjahidi. "Evaluation of Machines Learning Algorithms in Detection of Malware-Based Phishing Attacks for Securing E-Mail Communication." *Artificial Intelligence and Machine Learning*, 23 July 2022, https://doi.org/10.5121/csit.2022.121202. Accessed 19 Sept. 2022.

Lee, Hyunjong, et al. "Robust IoT Malware Detection and Classification Using Opcode Category Features on Machine Learning." *IEEE Access*, vol. 11, 2023, pp. 18855–18867, https://doi.org/10.1109/access.2023.3247344. Accessed 12 Apr. 2023.

Lu, Kai, et al. "Malware Detection Based on the Feature Selection of a Correlation Information Decision Matrix." *Mathematics*, vol. 11, no. 4, 13 Feb. 2023, p. 961, www.mdpi.com/2227-7390/11/4/961/pdf,

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

https://doi.org/10.3390/math11040961. Accessed 10 Sept. 2023.

Mayer, Benedikt, et al. "Interactive Visual Exploration of Surgical Process Data." *International Journal of Computer Assisted Radiology and Surgery*, 21 Oct. 2022, https://doi.org/10.1007/s11548-022-02758-1. Accessed 26 Dec. 2022.

Mazaed Alotaibi, Fahad, and Fawad. "A Multifaceted Deep Generative Adversarial Networks Model for Mobile Malware Detection." *Applied Sciences*, vol. 12, no. 19, 20 Sept. 2022, p. 9403, https://doi.org/10.3390/app12199403. Accessed 16 Oct. 2022.

Menin, Aline, et al. "ARViz: Interactive Visualization of Association Rules for RDF Data Exploration." *HAL (Le Centre Pour La Communication Scientifique Directe)*, 1 July 2021, https://doi.org/10.1109/iv53921.2021.00013. Accessed 10 Sept. 2023.

Mercaldo, Francesco, et al. "Towards Explainable Quantum Machine Learning for Mobile Malware Detection and Classification." *Applied Sciences*, vol. 12, no. 23, 24 Nov. 2022, p. 12025, https://doi.org/10.3390/app122312025. Accessed 23 Feb. 2023.

Mian, Tariq Saeed. "An Unsupervised Neural Network Feature Selection and 1D Convolution Neural Network Classification for Screening of Parkinsonism." *Diagnostics*, vol. 12, no. 8, 25 July 2022, p. 1796, https://doi.org/10.3390/diagnostics12081796. Accessed 23 Aug. 2022.

Miao, Y., et al. "RESEARCH on EVALUATION of the SPALLING DISEASE of NICHE for BUDDHA in GROTTOES." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLVI-M-1-2021, 28 Aug. 2021, pp. 441–447, isprs-archives.copernicus.org/articles/XLVI-M-1-2021/441/2021/isprs-archives-XLVI-M-1-2021-441-2021-relations.html, https://doi.org/10.5194/isprs-archives-XLVI-M-1-2021-441-2021.

Michael Burch et al. "Finding the outliers in scanpath data." Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications (2019). https://doi.org/10.1145/3317958.3318225.

Naik, Bhupal. "Comparative Analysis of Machine Learning-Based Algorithms for Detection of Anomalies in IIoT." *International Journal of Information Retrieval Research (IJIRR)*, vol. 12, no. 1, 2022, pp. 1–55, ideas.repec.org/a/igg/jirr00/v12y2022i1p1-55.html.

Office, International Commissioner's. "Data Protection Impact Assessments." *Ico.org.uk*, 19 May 2023, ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/.

Okikiola, Folasade Mercy, et al. "A DIABETES PREDICTION CLASSIFIER MODEL USING NAIVE BAYES ALGORITHM." *FUDMA JOURNAL of SCIENCES*, vol. 7, no. 1, 28 Feb. 2023, pp. 253–260, https://doi.org/10.33003/fjs-2023-0701-1301.

Orlovskyi, Dmytro Leonidovych, et al. "DEVELOPMENT of a MODEL and a SOFTWARE SOLUTION to SUPPORT the ANALYTICAL DASHBOARDS DESIGN PROBLEM." *Bulletin of National Technical University "KhPI". Series: System Analysis, Control and Information Technologies*, vol. 0, no. 1 (3), 9 July 2020, pp. 58–67, https://doi.org/10.20998/2079-0023.2020.01.11. Accessed 17 Mar. 2023.

Patel, Akash, et al. "Prediction of Stock Market Using Artificial Intelligence." *Papers.ssrn.com*, 7 May 2021, papers.ssrn.com/sol3/papers.cfm?abstract_id=3871022.

Patel, Samar. "Mobile App Market Growth

Corresponding author: Charles Ikpeama
✉ *charlyikpeama@yahoo.com*
Department of Cybersecurity, University of Hertfordshire (UH), UK.

Statistics & Trends 2023 and Beyond." *Mind Inventory*, 5 Aug. 2022, www.mindinventory.com/blog/mobile-app-usage-growth-statistics/.

Rafiq, Husnain, et al. "AndroMalPack: Enhancing the ML-Based Malware Classification by Detection and Removal of Repacked Apps for Android Systems." *Scientific Reports*, vol. 12, no. 1, 14 Nov. 2022, https://doi.org/10.1038/s41598-022-23766-w. Accessed 30 Nov. 2022.

Ragab, Mahmoud, et al. "Enhancement of Predicting Students Performance Model Using Ensemble Approaches and Educational Data Mining Techniques." *Wireless Communications and Mobile Computing*, vol. 2021, 7 Dec. 2021, pp. 1–9, https://doi.org/10.1155/2021/6241676. Accessed 6 July 2022.

Rao, Yamarthi Narasimha, and Kunda Suresh Babu. "An Imbalanced Generative Adversarial Network-Based Approach for Network Intrusion Detection in an Imbalanced Dataset." *Sensors*, vol. 23, no. 1, 3 Jan. 2023, p. 550, https://doi.org/10.3390/s23010550. Accessed 12 Jan. 2023.

Riaz, Sharjeel, et al. "Malware Detection in Internet of Things (IoT) Devices Using Deep Learning." *Sensors*, vol. 22, no. 23, 1 Jan. 2022, p. 9305, www.mdpi.com/1424-8220/22/23/9305, https://doi.org/10.3390/s22239305. Accessed 5 Dec. 2022.

Sharma, Ravi Mohan, et al. "CFSBFDroid: Android Malware Detection Using CFS + Best First Search-Based Feature Selection." *Mobile Information Systems*, vol. 2022, 7 July 2022, pp. 1–15, https://doi.org/10.1155/2022/6425583. Accessed 25 Aug. 2022.

Srinivasan, Arjun, et al. "Interweaving Multimodal Interaction with Flexible Unit Visualizations for Data Exploration." *IEEE Transactions on Visualization and Computer Graphics*, 2020, pp. 1–1, https://doi.org/10.1109/tvcg.2020.2978050. Accessed 18 Dec. 2020.

Srivastava, Aviral, et al. "IMPERCEPTIBLE MALWARE: BYPASSING MODERN AV - ENGINES by AI-ASSISTED CODE." *International Journal of Engineering Applied Sciences and Technology*, vol. 6, no. 6, 1 Oct. 2021, pp. 146–150, https://doi.org/10.33564/ijeast.2021.v06i06.022. Accessed 9 Jan. 2023.

Sriyanto, et al. "MiMaLo: Advanced Normalization Method for Mobile Malware Detection." *International Journal of Modern Education and Computer Science*, vol. 14, no. 5, 8 Oct. 2022, pp. 24–33, www.mecs-press.org/ijmecs/ijmecs-v14-n5/IJMECS-V14-N5-3.pdf, https://doi.org/10.5815/ijmecs.2022.05.03. Accessed 10 Sept. 2023.

Tayyab, Umm-e-Hani, et al. "A Survey of the Recent Trends in Deep Learning Based Malware Detection." *Journal of Cybersecurity and Privacy*, vol. 2, no. 4, 28 Sept. 2022, pp. 800–829, https://doi.org/10.3390/jcp2040041. Accessed 29 Sept. 2022.

Thoidis, Iordanis, et al. "Semi-Supervised Machine Condition Monitoring by Learning Deep Discriminative Audio Features." *Electronics*, vol. 10, no. 20, 11 Oct. 2021, p. 2471, https://doi.org/10.3390/electronics10202471. Accessed 1 Nov. 2021.

Tokmak, Mahmut, et al. "Deep Learning Based Malware Detection Tool Development for Android Operating System." *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, vol. 12, no. 4, 20 Dec. 2021, pp. 28–56, https://doi.org/10.18662/brain/12.4/237. Accessed 9 Jan. 2022.

Xie, Hanchen, et al. "MUSCLE: Strengthening Semi-Supervised Learning via Concurrent Unsupervised Learning Using Mutual Information Maximization." *ArXiv.org*, 30 Nov. 2020, arxiv.org/abs/2012.00150. Accessed 10 Sept. 2023.

Yacouby, Reda, and Dustin Axman.

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.

"Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models." *ACLWeb*, Association for Computational Linguistics, 1 Nov. 2020, aclanthology.org/2020.eval4nlp-1.9/.

Yew Jie, Loo, et al. "Metrics and Benchmarks for Empirical and Comprehension Focused Visualization Research in the Sales Domain." *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 12, no. 3, 1 Dec. 2018, p. 1340, https://doi.org/10.11591/ijeecs.v12.i3.pp1340-1348. Accessed 22 June 2022.

Yuxin, Ding, and Zhu Siyi. "Malware Detection Based on Deep Learning Algorithm." *Neural Computing and Applications*, vol. 31, no. 2, 25 July 2017, pp. 461–472, https://doi.org/10.1007/s00521-017-3077-6.

Yuxin, D., Siyi, Z. "Malware Detection Based on Deep Learning Algorithm." *Neural Computing and Applications*, vol. 31, no. 2, 25 July 2017, pp. 461–472, https://doi.org/10.1007/s00521-017-3077-6. Accessed 29 July 2022.

Zong, Jonathan, et al. "Rich Screen Reader Experiences for Accessible Data Visualization." *Computer Graphics Forum*, vol. 41, no. 3, 1 June 2022, pp. 15–27, https://doi.org/10.1111/cgf.14519. Accessed 10 Sept. 2023.

Corresponding author: Charles Ikpeama
charlyikpeama@yahoo.com
Department of Cybersecurity, University of Hertfordshire (UH), UK.