Applying Machine Learning to Identify Malicious Behavior

[1]Ch.Yashwanth, [2]D. Vikram Kumar, [3]Divya S,[4]V Swathi

[1,2,3,] UG Scholars,[4] Assistant Professor

[1,2,3,4] Department of Computer Science and Engineering,

[1,2,3,4] Guru Nanak Institutions Technical Campus, Hyderabad, Telangana, India.

Abstract: Malicious assaults, with an emphasis on URLs, are detected using a new technique that makes use of machine learning techniques. We use hybrid machine learning models in conjunction with ensemble approaches for Natural Language Processing (NLP). To extract pertinent information, we preprocess a dataset that includes both malicious and genuine URLs. We improve our models' accuracy and efficiency by using strategies like Grid Search Hyper Parameter Optimisation and Canopy feature selection. Evaluation measures that show the effectiveness of our method include precision, accuracy, recall, F1-score, and specificity. Our hybrid machine learning system, which incorporates natural language processing (NLP), performs better than current models, providing strong protection against malevolent threats and improving cyber security, according to comparative analysis.

.

Keywords: Machine learning, Natural Language Processing (NLP), Cybersecurity, Canopy feature selection, Grid Search Hyperparameter Optimization, evaluation metrics

INTRODUCTION

Malicious cyber-attacks, often using social engineering tactics, aim to deceive users into revealing personal or financial information. Attackers impersonate legitimate sources, sending fraudulent messages via email or social media to trick victims into sharing sensitive data or downloading malicious attachments. Social media attacks have increased in recent years due to the ease of reaching a large number of users globally. Reports from the Anti-Malicious Working Group (APWG) highlight a surge in malicious attacks, with a 250,000 increase in one month in January 2021. Financial institutions, social media platforms, and email services were the most targeted sectors in 2021, as attackers primarily seek to steal financial information and identities.

To combat these attacks, organizations typically rely on human expertise for detection. However, the similarity between legitimate and fake messages makes manual identification difficult. Attackers continue to refine their techniques, such as creating malicious URLs resembling trusted websites (e.g., "https://www.faceb00k.com/"). Detecting these malicious URLs is critical, as traditional methods, such as blacklists, fail to catch new or unknown threats. Additionally, machine learning models require time-consuming manual feature extraction, which is inefficient when attackers continuously create new malicious URLs.

To improve detection, we're integrating Natural Language Processing (NLP) for better understanding of URLs, using ensemble methods for more accurate predictions, and employing hybrid machine learning models (e.g., decision trees, support vector machines, neural networks). We also implement feature selection and grid search for optimization, aiming to enhance performance, evaluated through precision, recall, and accuracy metrics**.**

OBJECTIVE

The objective of this project is to conduct a comprehensive survey and analysis of HTML and URL-based malicious attacks, focusing specifically on the development and evaluation of machine learning models for automated detection. The project aims to address the increasing sophistication of malicious techniques by investigating current state-of-theart methodologies. The scope includes a detailed exploration of data preprocessing techniques, feature extraction methods, model design considerations, and performance metrics employed in existing machine learning models for malicious detection. By comparing these models based on various criteria, the objective is to provide a comprehensive understanding of their strengths and limitations. Furthermore, the project seeks to contribute to the field by identifying gaps in current research and proposing potential areas for future improvement and innovation. Overall, the objective is to enhance the knowledge base surrounding URL malicious attacks, paving the way for more effective and advanced approaches to malicious threat detection.

PROBLEM STATEMENT

In the rapidly evolving digital landscape, cyberattacks targeting web-based services have become increasingly sophisticated, with malicious URLs serving as a primary vector for delivering threats such as phishing, malware distribution, and data breaches. Existing detection methods often struggle with accuracy, scalability, and adaptability to emerging attack patterns. There is a critical need for an advanced detection framework that leverages cutting-edge machine learning techniques and Natural Language Processing (NLP) to accurately distinguish between malicious and legitimate URLs. This research aims to develop a hybrid machine learning model that integrates NLP-based ensemble methods, optimized through Canopy feature selection and Grid Search hyperparameter tuning, to enhance detection accuracy, efficiency, and robustness. The proposed solution will address current limitations in URL-based threat detection systems, providing a more effective defense mechanism against evolving cyber threats.

LITERATURE SURVEY

Title: A Comprehensive Survey of Machine Learning-Based Approaches for Malicious Website Detection

Authors: L. Tang and Q. H. Mahmoud

Year: 2021

Description:

With the expansion of the Internet, ensuring network security has become a critical concern. A secure online environment forms the foundation for the Internet’s continued development. Malicious websites pose a significant cyber threat, using deceptive URLs to trick users into revealing sensitive information such as login credentials and financial data. Cybersecurity is an ongoing battle between attackers and defenders, with both malicious tactics and detection methods evolving over time. Traditional detection techniques, such as blacklists and whitelists, struggle to identify newly emerging malicious links, emphasizing the need for more advanced prediction models.

The rise of machine learning (ML) has enabled more accurate and proactive malicious link detection. This survey provides a comprehensive overview of cutting-edge techniques for detecting malicious websites. It covers the lifecycle of phishing attacks, explores common anti-phishing strategies, and focuses on ML-based detection methods. Key aspects discussed include data collection, feature extraction, model development, and performance evaluation. The paper also presents a detailed comparison of various ML-based detection methods, offering valuable insights into the strengths and limitations of different approaches.

METHODOLOGY – ALGORITHMS USED

Data Set:

A structured set of data is called a dataset. It is often arranged in rows and columns, with each row denoting a distinct observation or instance and each column denoting a particular attribute or feature of that instance. Spreadsheets, databases, text files, and unique formats for purposes are just a few of the different formats that datasets can take.

Data Cleaning:

Finding and fixing mistakes or inconsistencies in a dataset to increase its quality and dependability for analysis is known as data cleaning. It entails duties include dealing with outliers, resolving missing values, eliminating duplication, fixing errors, and standardising formats.

NLP Feature Extraction:

The process of turning text input into numerical or categorical features that may be applied to machine learning tasks is known as feature extraction in natural language processing (NLP).

ML Model:

Computational algorithms known as machine learning models use data to identify patterns and relationships that can be used to inform predictions or choices. They use a range of methods, including deep learning, clustering, regression, and classification, and are trained on labelled or unlabelled datasets to solve problems and generalise patterns, allowing automated decision-making across a range of fields.

Train Model:

A model is trained by feeding it a dataset in order to identify patterns and relationships. To reduce prediction errors, optimisation procedures like gradient descent are used to modify the model's parameters. To determine how well the trained model performs in producing precise predictions or classifications, its performance is next assessed using a different validation dataset**.**

Test and Deployment:

Testing a model entail evaluating how well it performs on unknown data to make sure it satisfies target accuracy levels and generalises well. Real-time prediction, integration into production systems, and performance monitoring for continuous improvement and maintenance are all part of deployment.

Proposed Technique:

Natural Language Processing (NLP) with Linear Regression

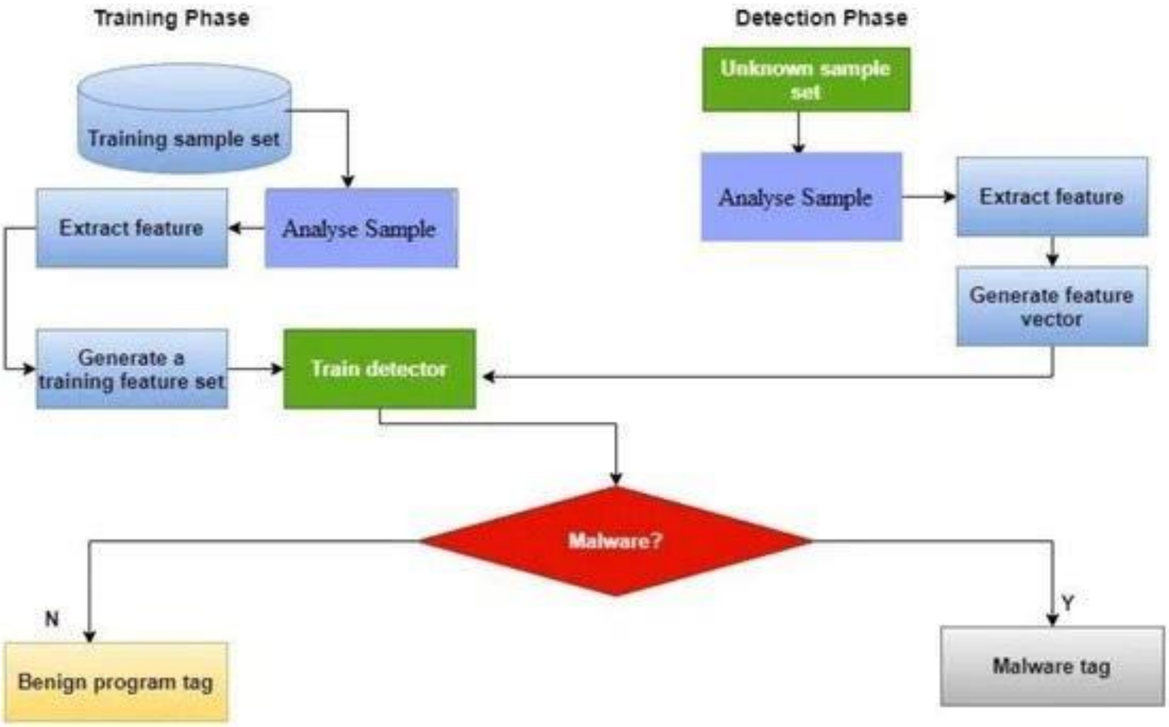
Technical Definition:

The proposed system combines Natural Language Processing (NLP) and Linear Regression to detect and classify malicious URLs. NLP techniques are applied to analyze the textual structure of URLs, extracting relevant features that indicate potentially harmful behavior. These features are then used as input for a Linear Regression model, which predicts the likelihood of a URL being malicious. The system is trained on a dataset containing labeled examples of both legitimate and malicious URLs, enabling it to learn distinguishing characteristics. By integrating NLP-based feature extraction with predictive modeling, the system provides an efficient, real-time solution for identifying and mitigating online threats**.**

Advantages:

- Effective in high dimensional spaces.
- This ensemble approach improves generalization and reduces overfitting, resulting in a more robust and accurate model.
- It is very easy to understand.

SYSTEM ARCHITECTURE



RESULT

The proposed system effectively detects malicious URLs using a hybrid machine learning model integrated with Natural Language Processing (NLP) techniques. By applying feature extraction, Canopy feature selection, and Grid Search hyperparameter optimization, the system achieves high accuracy and efficiency. Evaluation metrics such as precision, recall, and F1-score validate its performance, while comparative analysis shows that it outperforms existing models, offering a reliable defense against cyber threats.

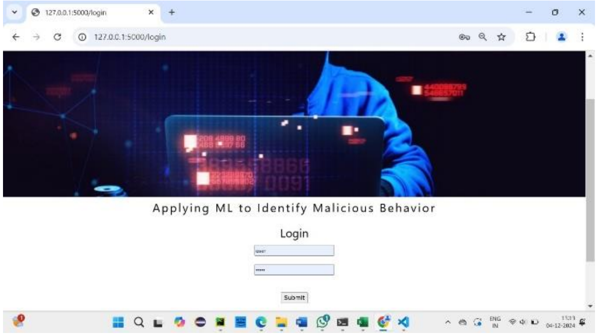
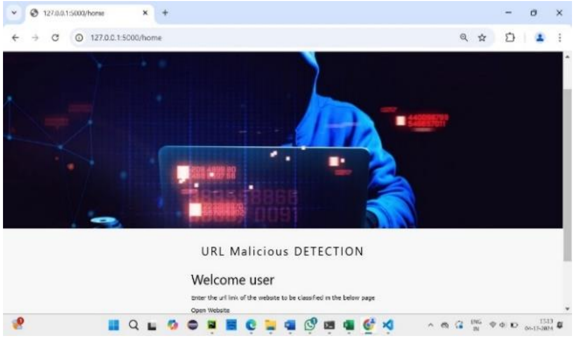


Fig.1: Login page

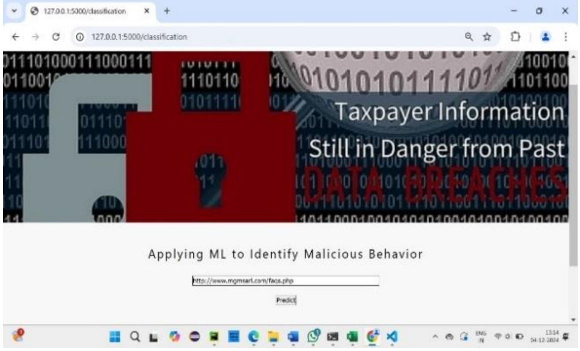


Welcome user

Enter the url link of the website to be classified in the below page

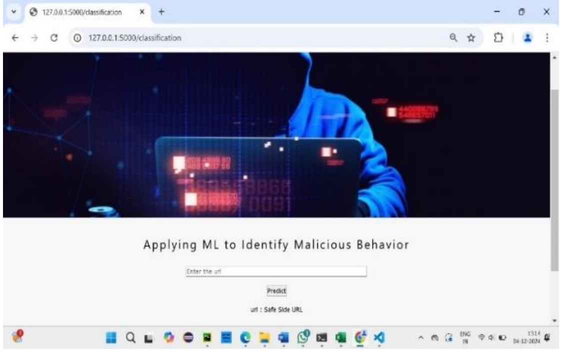
Open Website

Fig.2: Home page


127.0.0.1:5000/classification
127.0.0.1:5000/classification
Taxpayer Information
Still in Danger from Past
DATA BREACHES
Applying ML to Identify Malicious Behavior
Predict

http://www.mgmsarl.com/facs.php

Fig.3: Input URL


127.0.0.1:5000/classification
127.0.0.1:5000/classification
Applying ML to Identify Malicious Behavior
Enter the url
Predict

url : Safe Side URL

Fig.4: Output

CONCLUSION

In conclusion, our innovative method of identifying malicious attacks—which focusses on URLs and uses machine learning techniques—represents a substantial breakthrough in cyber security. We have created a system that performs better at differentiating between dangerous and valid URLs by combining machine learning models with Natural Language Processing (NLP) ensemble techniques. We have improved the precision and effectiveness of our models by carefully preprocessing the dataset, extracting pertinent characteristics, and using strategies like parameter optimisation and selection. Measures of evaluation such as precision, accuracy, recall, F1-score, and specificity continuously demonstrate how successful our strategy is. The superiority of our hybrid machine learning system, which provides strong defence against malevolent threats and improves cyber security posture, is highlighted by a comparison with current models.

REFERENCE

[1] Y. Zhang, Y. Xiao, K. Ghaboosi, J. Zhang, and H. Deng, ‘‘A survey of cyber crimes,’’ Secur. Commun. Netw., vol. 5, no. 4, pp. 422–437, 2012.

[2] APWG Developers. (2021). Malicious Activity Trends Report.

[3] M. Lei, Y. Xiao, S. V. Vrbsky, and C.-C. Li, ‘‘Virtual password using random linear functions for on-line services, ATM machines, and pervasive computing,’’ Comput. Commun., vol. 31, no. 18, pp. 4367–4375, Dec. 2008.

[4] P. Burda, L. Allodi, and N. Zannone, ‘‘Don’t forget the human: A crowdsourced approach to automate response and containment against spear malicious attacks,’’ in Proc. IEEE Eur. Symp. Secur. Privacy Workshops (EuroS PW), Sep. 2020, pp. 471–476.

[5] P. Prakash, M. Kumar, R. R. Kompella, and M. Gupta, ‘‘PhishNet: Predictive blacklisting to detect malicious attacks,’’ in Proc. IEEE INFOCOM, Mar. 2010, pp. 1–5.

[6] W. Zhang, Y.-X. Ding, Y. Tang, and B. Zhao, ‘‘URL web page detection based on on-line learning algorithm,’’ in Proc. Int. Conf. Mach. Learn. Cybern., vol. 4, Jul. 2011, pp. 1914–1919.

[7] A. C. Bahnsen, E. C. Bohorquez, S. Villegas, J. Vargas, and F. A. González, ‘‘Classifying malicious URLs using recurrent neural networks,’’ in Proc. APWG Symp. Electron. Crime Res. (eCrime), 2017, pp. 1–8.

[8] B. Cui, S. He, X. Yao, and P. Shi, ‘‘URL URL detection with feature extraction based on machine learning,’’ Int. J. High Perform. Comput. Netw., vol. 12, no. 2, pp. 166–178, 2018.

[9] Y. Fang, C. Zhang, C. Huang, L. Liu, and Y. Yang, ‘‘Malicious email detection using improved RCNN model with multilevel vectors and attention mechanism,’’ IEEE Access, vol. 7, pp. 56329–56340, 2019. [10] J. Feng, L. Zou, O. Ye, and J. Han, ‘‘Web2Vec: Malicious webpage detection method based on multidimensional features driven by deep learning,’’ IEEE Access, vol. 8, pp. 221214–221224, 2020.

[11] H. Cheng, J. Liu, T. Xu, B. Ren, J. Mao, and W. Zhang, ‘‘Machine learning based low-rate DDoS attack detection for SDN enabled IoT networks,’’ Int. J. Sens. Netw., vol. 34, no. 1, pp. 56–69, 2020.

[12] S. Christin, É. Hervet, and N. Lecomte, ‘‘Applications for deep learning in ecology,’’ Methods Ecol. Evol., vol. 10, no. 10, pp. 1632–1644, Oct. 2019.

[13] A. Aggarwal, A. Rajadesingan, and P. Kumaraguru, ‘‘PhishAri: Automatic realtime malicious detection on Twitter,’’ in Proc. eCrime Res. Summit, Oct. 2012, pp. 1–12.

[14] H. Ma, Y. Zuo, and T. Li, ‘‘Vessel navigation behavior analysis and multiple-trajectory prediction model based on AIS data,’’ J. Adv. Transp., vol. 2022, pp. 1–10, Jan. 2022.

[15] J. Fang, B. Li, and M. Gao, ‘‘Collaborative filtering recommendation algorithm based on deep neural network fusion,’’ Int. J. Sens. Netw., vol. 34, no. 2, pp. 71–80, 2020.

. [16] E. S. Gualberto, R. T. De Sousa, T. P. De Brito Vieira, J. P. C. L. Da Costa, and C. G. Duque, ‘‘The answer is in the text: Multi-stage methods for malicious detection based on feature engineering,’’ IEEE Access, vol. 8, pp. 223529–223547, 2020.