International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 03 | March - 2024 SJIF Rating: 8.176 ISSN: 2582- 3930

A Systematic Review of Evaluating the Efficiency of Hybrid Machine Learning Techniques in Unmasking Phishing URLs

Dr. Pradip Paithane[1], Siddheshwar Patil[2], Abhishek Kubde[3],

Rohit Mehetre[4], Shubham Gaikwad[5]

[1]Department of Artificial Intelligence and Data Science, VPKBIET, Baramati
[2]Department of Artificial Intelligence and Data Science, VPKBIET, Baramati
[3]Department of Artificial Intelligence and Data Science, VPKBIET, Baramati
[4]Department of Artificial Intelligence and Data Science, VPKBIET, Baramati
[5]Department of Artificial Intelligence and Data Science, VPKBIET, Baramati

---

Abstract - Phishing URLs pose significant threats to individuals and organizations, exploiting vulnerabilities to perpetrate scams and fraud. The Indian Cybercrime Coordination Center (I4C) faces challenges in effectively addressing these threats, necessitating a systematic examination of phishing URL detection methods. This review article focuses on evaluating the efficacy of hybrid machine learning algorithms, particularly URL-based techniques, in combating phishing. Leveraging the unparalleled accuracy and performance of hybrid machine learning models, this research represents a groundbreaking approach to early detection and mitigation of phishing URLs, which are a prevalent cause of fraud and hacking globally. Recent advancements in hybrid machine learning have facilitated the integration of multiple methods to enhance accuracy and reliability in identifying and thwarting phishing attempts. This study contributes to the ongoing efforts in cybersecurity by shedding light on the potential of hybrid machine learning techniques in unmasking phishing URLs, thereby bolstering defenses against cyber threats.

***Key Words*:** Phishing URLs, Hybrid Machine Learning, accuracy, URL-based techniques, Indian cybercrime coordination center (I4C)

1. INTRODUCTION

One of the core areas of computer science, computer security, is significantly impacted by criminal activities directed towards Internet users. Attacks and security problems of many kinds began to surface as the Internet and information technology applications developed. Beginning in the early 1990s, as the Internet gained worldwide popularity and accessibility, security risks additionally started to change. Attackers have targeted these sensitive data, and a particular type of assault known as phishing first surfaced in the mid-1990s.

One of the main topics covered in this thesis is phishing, which is a type of socially engineered online identity theft. In an ever-evolving cybersecurity landscape, the proliferation of phishing attacks poses a serious threat to both individuals and organizations. Multiple approaches that can identify phishing at various stages have been put forth by researchers looking at the issues surrounding phishing attacks. Some identify phishing attempts right down to the site page. Furthermore, certain models are designed to identify phishing attempts at an earlier email level, when the attacker is still attempting to persuade the recipient to visit the bogus website. This is a superior approach since, in the event that a phishing assault is discovered at the webpage level, it will first examine each Uniform Resource Locator (URL) the user tries to open before granting access, which will slow down website navigation.

Second, since the attack is discovered sooner, users are safer when phishing attacks are identified at the email level. For example, certain codes may be downloaded onto the user's device to infect it when the user views a web page. Furthermore, databases of phishing emails are always accessible, while a phony webpage only lasts for roughly 46 hours on average.

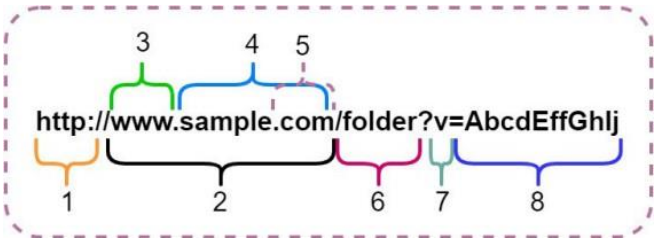

Fig -1: URL presentation based on HTTP

In fig1, label1 represents HTTP (Hypertext transfer protocol), which is used to acquire resources based on client requests. Label 2 denotes a hostname, and the hostname is further subdivided into three subdomains: top-level domain (also known as web address), and domain labeled 6 relates to a web server's directory. A label 7 "v" character with the value "AbcdEffGhIJ" and a label 6 "?" character in a URL initialize the parameter x. URLs are often used to represent website addresses.

To effectively counter the growing threat of phishing attacks on the Internet, comprehensive URL-based phishing detection system leverages various algorithms of machine learning like Logistic Regression, Random Forest, Decision Tree, XGBoost, KNN, etc. Phishing detection using hybrid machine learning models highlights the growing threat of phishing attacks in the digital age and highlights the need for advanced and accurate detection methods. It introduces the concept of hybrid machine learning and highlights its potential to increase recognition accuracy by combining the strengths of multiple algorithms. Phishing, often carried out via fraudulent URLs, exploits a human vulnerability to extract sensitive information or deliver malicious payloads. This report details the development and capabilities of a phishing URL detection system and highlights its importance in mitigating cyber threats and strengthening defenses against malicious actors. By examining the complexity of such a system, we discover the essential role it plays in providing a secure digital environment for users, contributing to the ongoing fight against cyber threats.

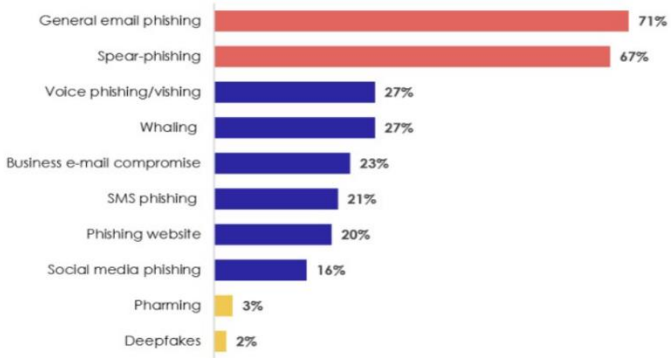

Fig -2: Types of phishing involved in most significant security incident

Figure 2 above illustrates the widespread problem of phishing assaults. Email is still the most popular technique, with hackers using bulk emails (generic phishing) or customized emails addressed to particular targets (spear phishing). Additional techniques including phone calls (vishing), text messages (SMS phishing), and phony websites are becoming more and more popular. Whaling, which targets CEOs, and Business Email Compromise (BEC) are two examples of highly focused attacks that highlight the importance of being vigilant. Phishing poses a variety of risks since its perpetrators are always changing their strategies.

1.1 The major contributions of this study are as follows.

1. Phishing URL-based cyberattack detection is proposed in this study to prevent crime and protect people's privacy. Model's decision-making for specific cases. On a global scale, it unveils overarching patterns within the entire model, enabling a holistic comprehension of its behavior.

2. The dataset consists of 11000+ phishing URL attributes that help classify phishing URLs based on these attributes.

3. Machine learning models have been applied, such as decision tree (DT), linear regression (LR), naive Bayes (NB), random forest (RF), gradient boosting machine(GBM), support vector classifier (SVC), K- Neighbors classifier (KNN), and the proposed hybrid model (LR+SVC+DT) LSD with soft and hard voting, which can accurately classify the threats of phishing URLs.

4. Cross-fold validation with a grid search parameter based on the canopy feature selection technique was used with the proposed LSD hybrid model to improve prediction results.

5.The proposed methodology must be evaluated using evaluation parameters, such as accuracy, precision, recall, specificity, and F1-score.

2. RELATED WORK

2.1 Phishing detection using random forest with NLP-features

Ozgur Koray Sahingoz, Ebubekir Buber, Onder Demir, Banu Diri Machine Learning Models: Random Forest with NLP-Based Features. The study presents a high accuracy real-time anti-phishing system that uses a variety of machine learning (ML) algorithms and features for URL-based phishing detection. Nevertheless, issues pertaining to NLP-based features are not examined, and the performances of the seven classification algorithms are not thoroughly analyzed. The Naïve Bayes classification is a probabilistic machine learning method, which is not only straightforward but also powerful.

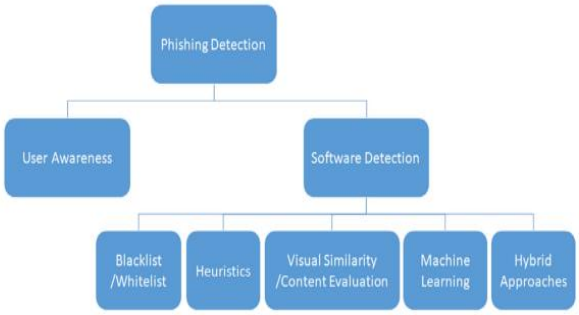

Fig -3: Classification of phishing URLs proposed based on designed methodological structure

Due to its simplicity, efficiency and good performance, it is preferred in lots of application areas such as classification of texts, detection of spam emails/intrusions, etc. It is based on the Bayes theorem, which describes the relationship of conditional probabilities of statistical quantities.

2.2 Phishing detection using hybrid machine learning model

S. Raman Kumar Jog, ABDUL KARIM, MOBEEN SHAHROZ, KHABIB MUSTOFA, AND SAMIR BRAHIM BELHAOUAR. Machine Learning Models: Logistic Regression (LR), Support Vector Classifier (SVC), and Decision Trees (DT) combined into a hybrid model. In

summary, the hybrid model (LSD) that has been suggested seeks to improve phishing detection efficiency and accuracy.

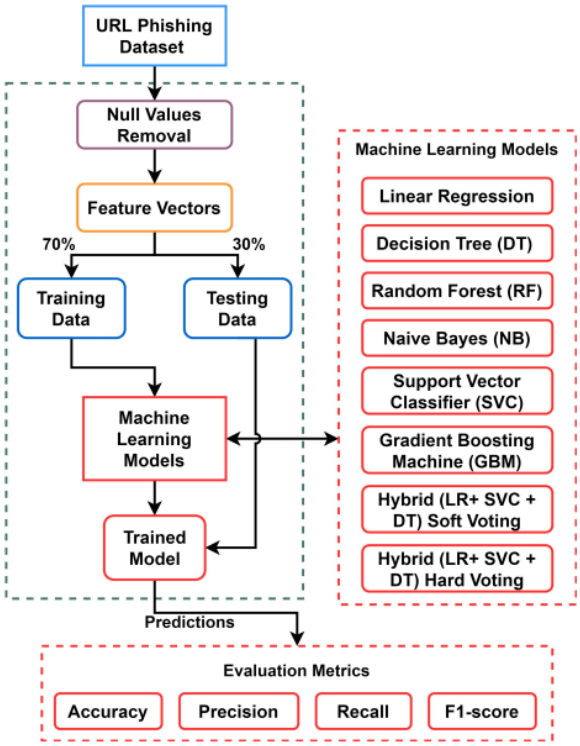

Fig -4: Classification of phishing URLs proposed based on designed methodological structure

It makes use of grid search, hyper parameter optimization, and canopy feature selection. On the other hand, false positive or negative rates are not discussed, and the study does not provide any information on scalability for huge datasets. Phishing Detection Using Machine Learning Techniques: Logistic Regression, Ada Booster, Random Forest, K-Nearest Neighbors (KNN), Neural Networks, Support Vector Machines (SVM), Gradient Boosting, Naive Bayes, and XGBoost. The study focuses on phishing prevention techniques, including email filtering, user education, and real-time machine learning detection. It also gives us a detailed explanation of the workings of different machine learning models. However, it does not explore potential vulnerabilities or weaknesses in the ma- chine learning models, and dataset details are omitted, impacting generalizability.

2.3 Phishing Email Detection Using Improved RCNN Model With Multilevel Vectors and Attention Mechanism

YONG FANG, CHENG ZHANG, CHENG HUANG,LIANG LIU, AND YUE YANG[4]Machine Learning Models : Deep learning model named THEMIS (advanced version of RCNN). THEMIS enhances embedding for improved performance detection. This approach allows for modeling emails at multiple levels, including the email header, email body, character level, and word level simultaneously, enhancing the model's ability to capture relevant features. However, the study notes that some phishing emails without an email header may reduce efficiency and accuracy. It is more accurate than previous methods, and it is also more efficient and robust.

1. The Bidirectional Long Short-Term Memory (Bi-LSTM) enhances the RCNN model. Next, an enhanced RCNN model is used to model the email at several levels. When as little noise as possible is introduced, the email's context can be better understood.

2. The email header and body are subjected to the attention mechanism, with varying weights allocated to each section. This allows the model to concentrate on more distinct and valuable data found in the email header and body.

3. On an unbalanced dataset, the THEMIS model presented in this work performs admirably. The accuracy is 99.848, and THEMIS's evaluation metrics are better than current detecting technology.

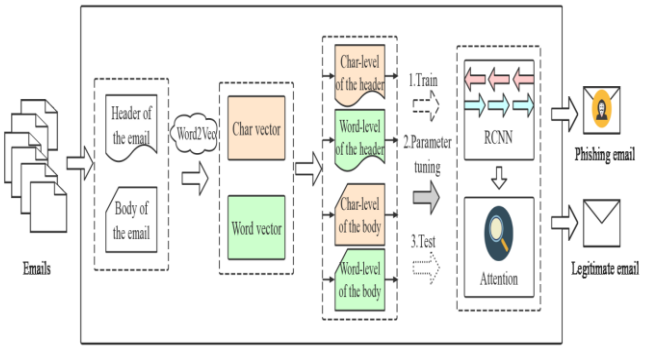

Fig -5: The framework for classifying phishing emails and legitimate emails using RCNN

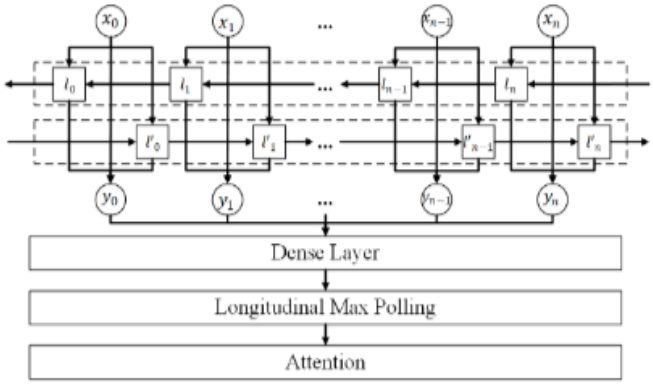

Fig -6: Improved RCNN-Attention model

$$c_l(w_i) = f(W^{(l)}c_l(w_{i-1}) + W^{(sl)}e(w_{i-1}))$$
$$c_r(w_i) = f(W^{(r)}c_r(w_{i+1}) + W^{(sr)}e(w_{i+1}))$$
$$x_i = [c_l(w_i); e(w_i); c_r(w_i)]$$
$$y_i^{(2)} = tanh(W^{(2)}x_i + b^{(2)})$$
$$y^{(3)} = \max_{i=1}^{n} y_i^{(2)}$$

Fig -7: RECURRENTCONVOLUTIONAL NEURAL NETWORKS AND ATTENTION MECHANISM

2.4 PhiDMA– A phishing detection model with multi-filter approach

Gunikhan Sonowal, K.S. Kuppusamy. Phishing is a major cyber security problem that results in financial loss. A single detection model cannot detect all types of phishing threats. The PhiDMA model detects phishing sites using many filters. The model gives suggestions about the actual website that the visitor is seeking to view. The model provides an interface that is accessible to people with visual impairments. The model detected phishing sites with an accuracy of 92.72. The model outperformed prior models in terms of accuracy. Future improvements will be made to boost the model's performance. The interface will be more accessible to those with multiple disabilities. Phishing detection model called PhiDMA is proposed, which incorporates five layers: Auto upgrade whitelist layer, URL features layer, Lexical signature layer, String matching layer, and Accessibility Score comparison layer.

1. To develop a model titled PhiDMA which attempts to handle the phishing problem with five different filters.

2. To incorporate the accessibility score of a web page as a phishing indicator. The proposed model has attempted this by building an accessibility score filter.

3. To develop a browser plugin based prototype of PhiDMA model with accessibility features in the interface.

4. To perform the comparative analysis of the proposed model on various metrics with the existing phishing detection approaches.

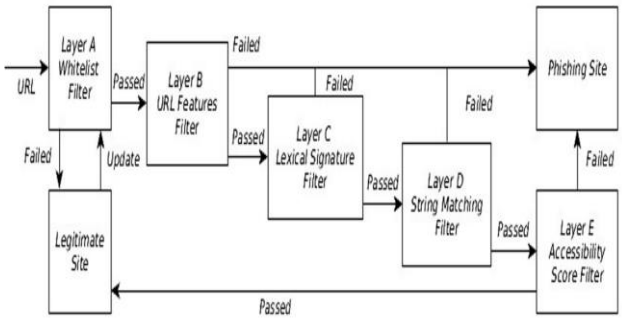

Fig -9: PhiDMA model Architecture

Algorithm 2: PhiDMA

Input : PhiDMA takes the URL from the user
Output: PhiDMA verify the URL

```
PhiDMA(url)
if url ∈ wlist then
    return True
else
    fea_{1..n} ← eFea(url)
    fSco ← cSco(fea_{1..n})
    if fSco ≥ Thr_f then
        return False
    else
        ls ← gISig(url)
        U_Link ← sEng(ls)
        if U_Link = Null then
            return False
        else
        end
        if LCS(url, uLink) <
          Thr_lcs && edis(url, uLink) < Thr_ed) then
            return False
        else
        end
        if accSco(uLink, url) ≥ Thr_a then
            return False
        else
            wlist ← update(uLink)
            display(uLink)
            return True
        end
    end
end
```

Algorithm 1: URL Filter

```
Function W_ip(url)
    if IP Address in URL then
        return 1
    end
    return 0
Function W_age(url)
    if Domain Age < T_DA then
        return 1
    end
    return 0
Function W_length(url)
    if length(Domain Length) > T_DL then
        return 1
    end
    return 0
Function W_dot(url)
    if count(Dots in Host Name) > T_ND then
        return 1
    end
    return 0
Function W_sus(url)
    if Suspicious symbols in host then
        return 1
    end
    return 0
Function urlFilter(url)
    Score = W_ip(url) + W_age(url) + W_length(url) +
      W_dot(url) + W_sus(url)
    if Score > Threshold then
        return True
    end
    return False
```

Fig -8: PhiDMA Algorithm

2.5 Hybrid ensemble feature selection framework for enhancing machine learning-based phishing detection systems

Kang Leng Chiew, Choon Lin Tan, Kok Sheik Wong, Kelvin S.C. Yong, Wei King Tiong Machine Learning Models: Hybrid Ensemble Feature Selection (HEFS), Random Forest Classifier. The study uses a real browser for feature extraction, improving robustness. However, detailed accuracy results for baseline features with classifiers other than Random Forest are provided. The phishing dataset used for benchmarking HEFS is not specified. According to experimental findings, HEFS works best when combined with the Random Forest classifier, identifying phishing and authentic websites with 94.6% accuracy while utilizing just 20.8% of the original characteristics. It consists of two phases: the first phase uses the CDF-g algorithm to generate primary feature subsets, which are then used in a data perturbation ensemble to produce secondary feature subsets.The second phase derives a set of baseline features from the secondary feature subsets using a function perturbation ensemble.

2.6 Detection of Online Phishing Email using Dynamic Evolving Neural Network Based on Reinforcement Learning

Sami Smadi, Nauman Aslam, and Li Zhang. The paper proposes a novel framework that combines a neural network with reinforcement learning. The proposed model adapts itself to produce a new phishing email detection system. The model solves the problem of a limited dataset by automatically adding more emails to the offline dataset in the online mode. A novel

algorithm is proposed to explore any new phishing behaviors in the new dataset. The pro- posed technique achieves high accuracy, TPR, and TNR at 98.63, 99.07, and 98.19, respectively. The model has acceptable error rates, with low false positive and false negative rates. The model can adapt and dynamically enhance the phishing email detection system. The model can explore and detect new phishing behaviors in real-time. The proposed algorithm improves the performance of existing detection methods.The Dynamic Evolving Neural Network (DENNuRL) algorithm, which the paper develops using Reinforcement Learning, enables the neural network (NN) to evolve dynamically and construct the optimal NN for problem solving. The suggested methodology, which includes DENNuRL, RL-Agent, experimental design, system model, and pre-processing.

Table -1: Survey of relevant research papers

Author	Brief Description	Result Summary
O.K.sahingoz, E.Buber, O.Demir, B.Diri (2019)	The study suggests a real-time anti-phishing system that makes use of features based on natural language processing (NLP) and seven distinct categorization algorithms.	It achieved 97.98% accuracy
Karim, M.Shahroz, k.Mustofa, S.B.Belhaouari (2023)	Machine learning models such as decision tree, linear regression, random forest, naive Bayes, gradient boosting classifier, K-neighbors classifier, support vector classifier, and a hybrid LSD model are used to defend against phishing attacks with high accuracy and efficiency .	The use of machine learning models such as decision tree, linear regression, random forest, naive Bayes, gradient boosting classifier, K-neighbors classifier, support vector classifier, and a hybrid LSD model contributes to achieving high accuracy in detecting phishing attacks
Y.Fang, C.Zhang, C.Huang,L.Liu, Y.Yang(2019)	A novel phishing email detection model named THEMIS is presented in the paper. It makes use of an enhanced recurrent convolutional neural network (RCNN) model that includes multilevel vectors and an attention mechanism.	THEMIS achieves an overall accuracy of 99.848%
Gunikhan Sonowal, K.S.Kuppusamy ,(2020)	A five-layer phishing detection model called PhiDMA is suggested, comprising the following layers: URL features, auto upgrade whitelist, lexical signature, string matching, and accessibility score comparison.	The accuracy of the model in identifying phishing sites is 92.72%
K.L.Chiew, C.L.Tan, K.Won (2019)	HEFS is considered a highly desirable and practical feature selection technique for machine learning-based phishing detection systems	HEFS paired with Random Forest achieves 94.6 \% accuracy in distinguishing phishing from authentic websites, utilizing just 20.8% of original characteristics.
Sami Smadi Nauman, Li Zhang (2018)	An innovative framework that combines reinforcement learning and a neural network is put forth to identify phishing attempts in the online mode.	The suggested method achieves good performance levels with 98.63\% accuracy, 99.07\% true positive rate (TPR), and 98.19\% true negative rate (TNR), according to thorough testing on well-known datasets.

3. ANALYSIS AND DISCUSSION

3.1 Dataset

The dataset was saved as a CSV file. The collection contained 11054 items and 33 attributes taken from over 11000 websites. Some common attributes shared by phishing and legitimate website URLs include UsingIP, LongURL, ShortURL, Symbol@, Redirecting//, PrefixSuffix, SubDomains, HTTPS, DomainRegLen, Favicon, NonStdPort, HTTPSDomainURL, RequestURL, AnchorURL, LinksInScriptTags, and ServerFormHandler. The dataset was divided into two categories: phishing and legitimate. The dataset consisted of vectors that had to be adjusted.

Dataset Link:
https://www.kaggle.com/code/eswarchandt/website phishing/input?select=phishing.csv

3.2 Evaluation parameters

Several evaluation factors must be used to assess machine learning performance. The machine-learning algorithm generates predictions as its output. The evaluation parameters measure the number of correct and incorrect predictions produced by the model across both the legitimate and phishing classes. Accuracy, precision, recall, specificity, and the F1-score were all used.

1. Accuracy

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2. Precision

$$Precision = \frac{TP}{TP + FN}$$

3. Recall

$$accuracy = \frac{TP}{TP + N}$$

4. F1-Score

$$F1 - score = \frac{2 * Precision * recall}{TP + TN + FP + FN}$$

3.3 Modeling

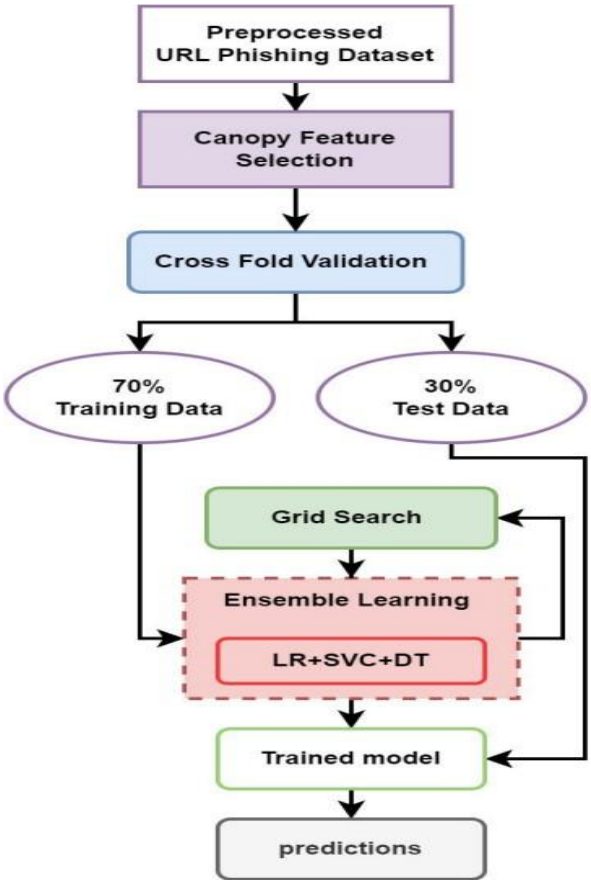

Fig -10: System Architecture

Using a multilayered strategy, this technique seeks to identify suspicious websites. It begins with a dataset of labeled URLs (phishing or real) and selects the most relevant attributes for distinguishing them. Consider it like sifting through sand for gold nuggets. Following that, the data is divided into two parts: 70 for training the model, similar to how a detective learns to identify dishonesty, and 30 for testing its abilities. The ”detective” in this case is a conglomeration of three machine learning algorithms: logistic regression, support vector machine, and decision tree. Each provides distinct advantages in cracking the case of fake website. The model then sharpens its skills on the training data, tweaking its internal parameters like a detective honing their observation skills. Finally, it is tested using unseen data, demonstrating its accuracy in detecting phishing attempts.

3.4 Observation

According to the preceding reference articles, the two most accurate models are the deep learning model dubbed THEMIS and the hybrid machine learning model (LR+SVC+DT), with accuracy values of 99.848 and 95.23, respectively. THEMIS (Throughput-oriented Efficient Multi-stage Inception for Scalable Object Detection) is an improved object detection architecture that improves on the capabilities of R-CNN. It seeks to attain high accuracy while being efficient, especially in real-time or resource-constrained applications. This approach enables for the simultaneous modeling of emails at many levels, including the email header, email body, character level, and word level, thereby improving the model’s capacity to capture essential information. A hybrid machine learning model that combines Logistic Regression (LR), Decision Trees (DT), and Support Vector Classification (SVC) makes use of the advantages of each approach. LR handles linear relationships, DT captures non-linear patterns, and SVC ensures robust classification. During training, features are processed by LR, DT, and SVC independently. The individual predictions are then combined, often using a weighted average, to generate a final output. This ensemble approach enhances overall model performance by exploiting the complementary capabilities of each algorithm.

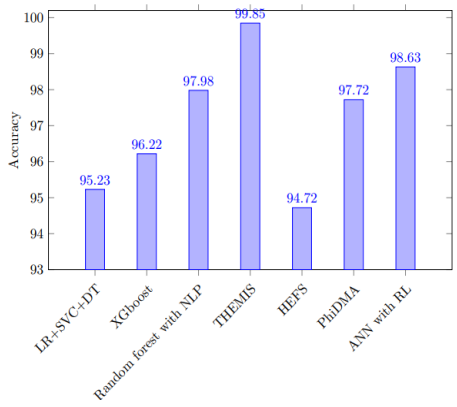

Fig -11: Accuracy of ML-DL Models

3.3 Contribution

1. Our review study makes major contributions to the field of machine learning by thoroughly assessing and synthesizing the landscape of hybrid models and diverse individual models.

2. First, we provide a comprehensive examination of several hybrid techniques, highlighting their benefits and drawbacks. This contains a thorough discussion of combinations of Decision Trees, Support Vector Machines, and Neural Networks.

3. Second, we conduct a comparative analysis with traditional standalone models, providing insights into situations in which hybrid models outperform or complement single models.

4. Further- more, our research reveals growing patterns in ensemble approaches and model stacking, laying the

groundwork for future advances in model combination strategies.

5. By combining these findings, this study provides a helpful resource for scholars and practitioners seeking an informed viewpoint on the efficacy and uses of hybrid machine learning models in diverse domains.

4. CONCLUSION

The review paper systematically evaluates diverse ma- chine learning approaches for phishing detection based on URL analysis. The studies showcased hybrid models, each with unique strengths and weaknesses. While these models demonstrate enhanced accuracy and efficiency, gaps exist, such as a lack of in-depth analysis on false positive/negative rates, scalability challenges, and omitted dataset details. Despite these limitations, the review illuminates valuable insights, offering a foundation for future research to refine and advance URL-based phishing detection systems using hybrid machine learning approaches.

Supplementary information

If your article has accompanying supplementary file/s please state so here. Authors reporting data from electrophoretic gels and blots should supply the full unprocessed scans for key as part of their Supplementary information. This may be requested by the editorial team/s if it is missing. Please refer to Journal-level guidance for any specific requirements.

Acknowledgements

Acknowledgements are not compulsory. Where included they should be brief. Grant or contribution numbers may be acknowledged. Please refer to Journal-level guidance for any specific requirements.

Declarations

- Funding
  Not applicable
- Conflict of interest/Competing interests (check journal-specific guidelines for which heading to use
  Not applicable
- Ethics approval and consent to participate
  Not applicable
- Consent for publication
  Not applicable
- Data availability
  Not applicable
- Materials availability
  Not applicable
- Code availability
  Not applicable
- Author contribution
  Not applicable

If any of the sections are not relevant to your manuscript, please include the heading and write `Not applicable' for that section.

REFERENCES

1. Zouina, M., Outtaj, B.: A novel lightweight url phishing detection system using svm and similarity index. Human-centric Computing and Information Sciences 7(1), 1–13 (2017)
2. Wang, S., Khan, S., Xu, C., Nazir, S., Hafeez, A.: Deep learning-based efficient model development for phishing detection using random forest and blstm classifiers. Complexity 2020, 1–7 (2020)
3. Abdelhamid, N., Ayesh, A., Thabtah, F.: Phishing detection based associative classification data mining. Expert Systems with Applications 41(13), 5948–5959 (2014)
4. Karim, A., Shahroz, M., Mustofa, K., Belhaouari, S.B., Joga, S.R.K.: Phishing detection system through hybrid machine learning based on url. IEEE Access 11, 36805–36822 (2023)
5. Sahingoz, O.K., Buber, E., Demir, O., Diri, B.: Machine learning based phishing detection from urls. Expert Systems with Applications 117, 345–357 (2019)
6. Shahrivari, V., Darabi, M.M., Izadi, M.: Phishing detection using machine learning techniques. arXiv preprint arXiv:2009.11116 (2020)
7. Buber, E., Diri, B., Sahingoz, O.K.: Nlp based phishing attack detection from urls. In: Intelligent Systems Design and Applications: 17th International Conference on Intelligent Systems Design and Applications (ISDA 2017) Held in Delhi, India, December 14-16, 2017, pp. 608–618 (2018). Springer
8. Paithane, P.M.: Random forest algorithm use for crop recommendation. ITEGAM-JETIA 9(43), 34–41 (2023)
9. Paithane, P.M.: Yoga posture detection using machine learning. Artificial Intelligence in Information and Communication Technologies, Healthcare and Education: A Roadmap Ahead 27 (2022).
10. Chiew, K.L., Tan, C.L., Wong, K., Yong, K.S., Tiong, W.K.: A new hybrid ensemble feature selection framework for machine learning-based phishing detection system. Information Sciences 484, 153–166 (2019)
11. Paithane, P., Kakarwal, S.: Lmns-net: Lightweight multiscale novel semantic-net deep learning approach used for automatic pancreas image segmentation in ct scan images. Expert Systems with Applications 234, 121064 (2023)
12. Fang, Y., Zhang, C., Huang, C., Liu, L., Yang, Y.: Phishing email detection using improved rcnn model with multilevel vectors and attention mechanism. IEEE Access 7, 56329–56340 (2019)
13. Rao, R.S., Pais, A.R.: Detection of phishing websites using an efficient featurebased machine learning framework. Neural Computing and applications 31, 3851– 3873 (2019)
14. Sonowal, G., Kuppusamy, K.: Phidma–a phishing detection model with multifilter approach. Journal of King Saud University-Computer and Information Sciences 32(1), 99–112 (2020)

15. Alam, M.S., Vuong, S.T.: Random forest classification for detecting android malware. In: 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, pp. 663–669 (2013). IEEE
16. Smadi, S., Aslam, N., Zhang, L.: Detection of online phishing email using dynamic evolving neural network based on reinforcement learning. Decision Support Systems 107, 88–102 (2018)
17. Paithane, P.M., Kakarwal, S.: Automatic pancreas segmentation using a novel modified semantic deep learning bottom-up approach. International Journal of Intelligent Systems and Applications in Engineering 10(1), 98–104 (2022)
18. Buber, E., Dırı, B., Sahingoz, O.K.: Detecting phishing attacks from url by using nlp techniques. In: 2017 International Conference on Computer Science and Engineering (UBMK), pp. 337–342 (2017). IEEE
19. Feng, F., Zhou, Q., Shen, Z., Yang, X., Han, L., Wang, J.: The application of a novel neural network in the detection of phishing websites. Journal of Ambient Intelligence and Humanized Computing, 1–15 (2018)
20. Wagh, S.J., Paithane, P.M., Patil, S.: Applications of fuzzy logic in assessment of groundwater quality index from jafrabad taluka of marathawada region of maharashtra state: A gis based approach. In: International Conference on Hybrid Intelligent Systems, pp. 354–364 (2021). Springer
21. Shirazi, H., Hayne, K.: Towards performance of nlp transformers on url-based phishing detection for mobile devices. International journal of ubiquitous systems and pervasive networks (2022)