The New Phishing Frontier: Predictive Modeling of Malicious Click Trajectories Originating from Social Networking Sites

Mirza Samiulla Beg[1]
[1] Poornima University, Jaipur,Rajasthan, India
msaksuniversity@gmail.com

Abstract. In this paper, we present and test predictive models that are specifically targeted at malicious clicks in order to solve the problem of cybercrime, which has become a widespread problem among social media users. Using a number of individual features like the behavior of the user, device used, social media platform, reputation of the website, and others, we developed a predictive model for Click_Leads_To_Malicious_Site (binary).Our model exhibits high predictive capability in all data sets (TRAIN, VALIDATE, and TEST). It demonstrates very high discriminative power with some excellent performance measures such as Area Under the Curve (AUC) of 0.9623 on the VALIDATE data set and 0.9462 on the TEST data set. The model performs very well on the VALIDATE data set with respect to the default cutoff point (0.5) at 0.916. Another excellent metric is the F1 score of 0.807 on the VALIDATE data set, which demonstrates excellent precision and recall. The model performance is proven by Lift and Captured Response Percentage returning. The model has identified a Cumulative Lift of 4.26 in the VALIDATE partition at 10% depth (top 10% of the data based on the predicted probability) that means to detect 4.26 times more bad cases compared to the randomly selected one. The higher value of lift means the model is capable of detecting priority high-risk transactions. Considering the fact that the focus here is on avoiding risk and creating pain for the participants, this model could help improve the security of customers in the online environment in real-time by assessing the risk of malicious clicking on social accounts.

Keywords: Phishing Attacks, Cybersecurity, Social Engineering, Psychological Elements, Sociotechnical Elements, Human-Centric Security, Human Vulnerabilities, Risk Mitigation, Economic Repercussions, Reputational Harm.

1. Introduction

Phishing detection studies have been mainly conducted in the static URL space or in the realm of landing page classification. In this paper, we propose a novel solution, named "Trajectory-Aware Predictive Framework (TAP-F)".

Core Innovation:

M. Kaushik et al. (eds.), *Proceedings of the International Conference on Responsible, Risk-aware, and Regulated AI (RRRAI 2026)*, Advances in Intelligent Systems Research 210,
https://doi.org/10.2991/978-94-6239-723-1_9

To catch the attention of our readers, we present the following:

Instead of employing traditional binary classifiers, we apply multi-layered neural network model on the input vector that consists of the environmental context (device and connection type), behavioral triggers (use of 2FA and saved passwords) and domain reputation entropy.

Prediction Time-Shift: Rather than detecting the destination content as a landing page, we move to a new prediction time frame, namely, the pre-landing period. The "click trajectory" will enable us to predict the destination based on the context provided by social networking sites (SNS).

The proliferation of communication platforms and Social Networking Sites (SNSs) has provided cybercriminals with new entry points. The most persistent and severe threat among all cyber threats is phishing. This threat poses a significant danger not only to organizations but also to the general public. Phishing attacks have evolved from simple email-based attacks to elaborate social engineering schemes that involve complex technology integration. As the name suggests, phishing involves human behavior information. This means that there is a need for you to understand human vulnerabilities. Cybercriminals depend on psychological and sociotechnical factors. This is because phishing, in reality, has serious consequences such as financial losses and damage to reputation.

The cases of social networking phishing have been increasing too. The phishers keep on coming up with new tactics to gain access to popular sites via popular social networking platforms like Facebook, Twitter, and LinkedIn. The phishing process on social networking platforms can be accomplished by not only using the old conventional methods like emails and spoofed websites, but in innovative ways that enable the phisher to get access to even more people. Social networking phishing not only affects people individually but is distributed via social norms as well. The social networking site may become an intimidating place and information sharing may lead the users to indulge in “bad behavior.” This means they might click and share links without even reading the message they receive.

With this danger in mind, various studies have been carried out to find solutions which would help in forecasting and prevention. For instance, Abbasi et al. proposed the Phishing Funnel Model (PFM), an approach aimed at predicting user vulnerability to phishing attacks through the integration of data regarding users, threats and tools in four major steps of the phishing funnel (model created based on previous research). This model provided excellent prediction results, achieving up to 96 percent accuracy for high-severity threats alongside low phishing expenditures. Other researchers have even attempted to find novel approaches for detecting dangerous content, such as the use of the genre tree kernel approach for detecting phishing websites and advanced approaches which take into account metapath features.

Furthermore, training against phishing techniques using machine learning algorithms has been employed in analyzing the way users analyze the simulations, demonstrating

the need to conduct user training on an individual basis rather than in large groups. However, there is a huge need for proactive and accurate classification systems that would detect phishing attempts when users click. While existing methods consider actions after clicking (site detection) or the risk factors associated with the users themselves (PFM and training), a real-time warning system about malicious clicks would prove more beneficial.

This study bridges the above gap through development and evaluation of a comprehensive machine learning model that will enable prediction of real-time clicks on malicious sites. Our model considers a wide array of features in order to classify whether the click event (Click_Leads_To_Malicious_Site) is good or bad. The main strength of our model is its discriminatory power. We have achieved an AUC and efficiency rate greater than 0.94 in all the data partitions. Specifically, our model achieves cumulative lift of 4.26 at 10 percent of the validation set. This demonstrates the efficiency of our model by detecting 4.26 times more malicious click events than random selection.

2. Literature Review

Muhammad Syafiq Kheruddin et al. [1] have explored the issue of the phishing attacks on the cybersecurity thoroughly. The authors trace its evolution from benign email phishing up to social engineering and technology-based ones. The authors explore the impact of psychological and sociotechnical factors involved in the attacks. They provide the evidence of human aspect of phishing and necessity to exploit the weaknesses of humans. The paper considers the impact of such aspects as financial and reputation risks, such as fines and damage to reputation for years to come.

E. D. Frauenstein et al. [2] discovered that the quality of long-term health care provider-patient relationships improves as longer behavioral health interventions become part of treatment, and depression treatments become more effective with time. emphasize the fact that the phishing threat remains serious. An increasing number of cases of phishing has been reported in statistics and on the Internet. Phishers are trying to find other ways to deceive people into disclosing personally identifying information. And therefore, phishing has evolved into more than the usual email and fraudulent websites. There have been many more victims, be it individuals or corporations, and social media phishing is quite common. This study suggests that people and organizations may acquire addictive behavior patterns of clicking on and sharing links, liking postings, and uploading/downloading media due to the constant flow of activity on social media network sites, such as status updates. It may result in an overload of information. In this case, the users do not process their communication from the point of view of safety, and thus they become vulnerable to social engineering attacks conducted via social networks. The current paper examines the phenomenon of social network

phishing as an element that helps users engage in risky behavior and provides a more detailed analysis of user behavior in social networks.

Ahmed Abbasi et al. [3] Phishing is one of the greatest dangers to the business world as well as public life. Most of the population suffers from it – especially employees. In addition to damaging the reputation of the company, phishing also damages public trust. Although there are efforts against phishing attacks, users often fail to recognize them. We propose the Phishing Funnel Model (PFM), which allows predicting the user's vulnerability during phishing in four stages: visit, browse, evaluate legitimacy, and transaction intent. The model uses user and threat information and analyzes it with our own support vector ordinal regression.PFM was tested in a one-year experiment involving 1,278 employees and 49,373 phishing attacks. It predicted visits to high-severity threats with 96% accuracy, exceeding other models by 8% to 52%. As a result, employees using PFM responded to phishing less often than employees using baseline warnings. The cost-benefit analysis showed that PFM could save about $1,900 for each employee from phishing attacks.Findings include:The ability to predict user susceptibility to phishing in real time increases security.The whole process of phishing can be estimated based on the user's behavior.Anti-phishing tools affect phishing susceptibility greatly.

Abbasi, A et al [4].Phishing sites take advantage of the weakness of the users at their homes and workplaces. Existing anti-phishing techniques are not accurate, causing the users to disregard any alerts raised by them. This paper presents a novel way of detecting phishing websites that uses a genre tree kernel technique to distinguish phishing sites from genuine ones based on fraud signals. The results of testing this technique on several thousand legitimate and phishing websites indicate that it is superior to other available techniques. Users educated using this technique were better able to detect phishing attacks.

U. Kumaran et al. [5], the fast growth of computer technologies and techniques has increased the probability of cybercrime against individuals. The paper attempts to investigate the current status of cybercrime due to this field of research in the next three parts: sentiment analysis, phishing attack and malware analysis. This research attempts to find out malware techniques used for avoiding detection and evaluate the cost of spyware by analyzing some malware examples. The trend analysis indicates that the phishing attacks are becoming more and more advanced and frequently used on reliable websites. The sentiment analysis of media reports about data breaches suggests the paradox that the security concerns are still very high although the confidence in managing the situation is continuously increasing.

T. Sutter et al. [6] Many sectors of the industry employ anti-phishing simulations in order to prevent potential phishing attacks. The research covered more than 31,000 participants and 144 simulated phishing attacks to assess users' behavior and reactions.The key findings were as follows:By week 12, 66% of the participants avoided phishing attacks that required their credentials.A new machine learning model based on

manifold learning and natural language processing (NLP) was created in order to forecast user interaction with phishing simulations.The model is designed to evaluate the "convincing power" of e-mails prior to being delivered to the trainees. It also highlights important categories for classification.

L. Jin et al. [7] The social media websites, which one can browse, such as Facebook, Twitter, Google+, LinkedIn, and Foursquare, have become omnipresent and influencing people's lives in many countries in various ways. Individuals use both modern mobile gadgets as well as conventional personal computers to connect to online social networks (OSNs). OSNs, which accommodate more than a billion people worldwide, can be considered innovative and problematic to study at the same time. In this paper, an attempt will be made to give an overall picture of the latest literature on user behavior in OSNs from different perspectives. The paper begins with user engagement and social interaction. In addition, traffic activity is investigated from the perspective of networking. Furthermore, the aspect of social interaction in mobile usage is taken into account considering that mobile is now a ubiquitous device. Lastly, the paper examines Malicious Use by OSN Users and suggests methods for diagnosing malicious individuals.

According to E. Lancaster et al. [8], the use of multimodal content such as images, videos, audio clips, and URLs contained in tweets is becoming common in luring individuals to click on "wrong" links that could lead to malware infections. In this paper, we use five classes of features to identify whether the tweet is malicious or not: textual character (sentiment), pathways, URL-based features, and multimodal content. The fifth feature class extracts features by creating a "tweet graph" and analyzing its "metapaths". The authors propose a unified classification method called MALicious Tweets in Parallel (MALTP), which uses metapaths, tweet graphs, and the existing methods from literature. We conduct extensive experiments on two different datasets: Warningbird (WB) and KBA. First, we show that the metapath-based method proposed in this paper performs better than other methods in detecting tweets of threat. In addition, we demonstrate that metapath features combined with Alexa rank and KBA features achieve prediction accuracy above 0.98 for KBA and above 0.94 for WB data, better than those reported in prior studies. Metapath features alone result in prediction accuracies of 0.977 for KBA data and 0.923 for WB data, much higher than the other methods. It was also concluded that metapath features are very significant in distinguishing between malicious tweets and non-malicious tweets, while multimodal content is insignificant in this regard.

Table 1. Main Contributions vs. Literature

Feature	**Existing Literature (Abbasi et al. [3], Sutter et al. [6])**	**Our Work (TAP-F)**
Primary	Email headers /	Real-time SNS click

Data Source	Phishing Funnel stages	metadata
Detection Timing	During/After browsing	At the moment of click (Trajectory stage)
Feature Set	Textual/NLP based	Contextual + Behavioral + Entropy
Efficiency	High computational cost for NLP	Low-latency inference for real-time blocking

3. Proposed Methodology

The objective of this research is to create an efficient classification algorithm for predicting whether clicking on social media will land one into a malicious website. This process involved following the conventional steps of supervised machine learning, which include data collection, data preparation, model building, and evaluation.

Features: Categorical and numerical features were mixed in the dataset. These features include Age Group, Gender, Social Media Platform, Device Type, Domain Reputation, Connection Type, Time of Day, and several behavioral flags like Two-Factor Authentication, Saved Password, and Suspicious Activity.

Target Variable: The target variable is a binary variable named Click_Leads_To_Malicious_Site (value = 1 if malicious, value = 0 if not malicious).

Data Splitting: In order to assess the generalization capability of the model, the dataset was split into three sets: TRAIN, VALIDATE, and TEST sets. The number of samples in each set was 3,000, 1,500, and 500, respectively.

3.1 Architecture: The Neural Trajectory Classifier

Feed-Forward Neural Network (FFNN) trained using Stochastic Gradient Descent (SGD) algorithm is referred to as the "Champion Model."

Inputs: 12-Dimensional vector (categorical features one-hot encoded) acts as inputs.

Hidden Layers: In order to model the complex interactions between behavioral features (e.g., the interaction between the Saved_Password and Suspicious_Activity flags), two deep layers with ReLU activation are used.

Output Layer: Softmax layer provides output probability $P(Y-1|X)$, where Y refers to the malignancy of the click trajectory.

3.2 Feature Engineering & Selection

We employed Global Surrogate Modelling (one-level decision trees) to verify the significance of characteristics.

Domain_Entropy: Calculated using the target domain string's Shannon Entropy.

Click_Context: A derived feature that combines Time_of_Day and Social_Media_Platform.

3.3 Implementation Details (Reproducibility)

- **Environment:** Scikit-Learn/TensorFlow, Python 3.9.
- **Hyperparameters:** Adam Optimiser Learning Rate: 0.001.

32 is the batch size.

There are 100 epochs (including early stopping).

- **Data Split:** 10% Test (500), 30% Validate (1,500), and 60% Train (3,000).

3.4 Model Creation

- **Model Type:** We employed an effective Supervised Classification Model for binary prediction tasks.
- **Model Training:** The TRAIN partition was used to modify the model's parameters in order to reduce classification mistakes.
- **Hyperparameter Tuning and Selection:** The VALIDATE partition served as a guidance for model selection. The best prediction quality—such as maximum Area Under the Curve or F1 Score—was chosen based on how well it could generalise after a variety of parameters were evaluated.

3.5 Model Evaluation

The final model has been rigorously evaluated against the unseen TEST data set employing well-established performance metrics commonly employed in the industry to measure its performance.

Discrimination metrics – KS statistic and AUC.

Classification metrics - Accuracy, Misclassification Rate and F1 Score (cutoff = 0.5)

Business impact metrics – Cumulative lift and cumulative captured response percentage to measure the model's real-world performance in capturing high-risk events.

4. Flow Chart Steps: Predictive Modeling Process

Here is a flowchart of model creation and evaluation step by step.

- **START:** Data Collection (CSV data comprising behavioral, device, and domain features)
- **Data Preparation** (Feature Engineering, Encoding Categorical Variables, Missing Value Imputation), an implicit step following standard ML practices
- **Data Partitioning** (TRAIN: 3,000 | VALIDATE: 1,500 | TEST: 500)
- **Model Training** (Fit Classification Model on TRAIN data)
- **Model Tuning** & Selection (Validate model on VALIDATE data by optimizing for AUC/F1)
- **Performance Evaluation** (Final model evaluation on the unseen TEST data)
- **Results Generation** (Fit Statistics, ROC Reports, Lift Reports)

- **Model Deployment** / Conclusion: The final decision regarding the utility of the model
- **END**

5. Results

5.1 Most Common Variables Selected Across All Models

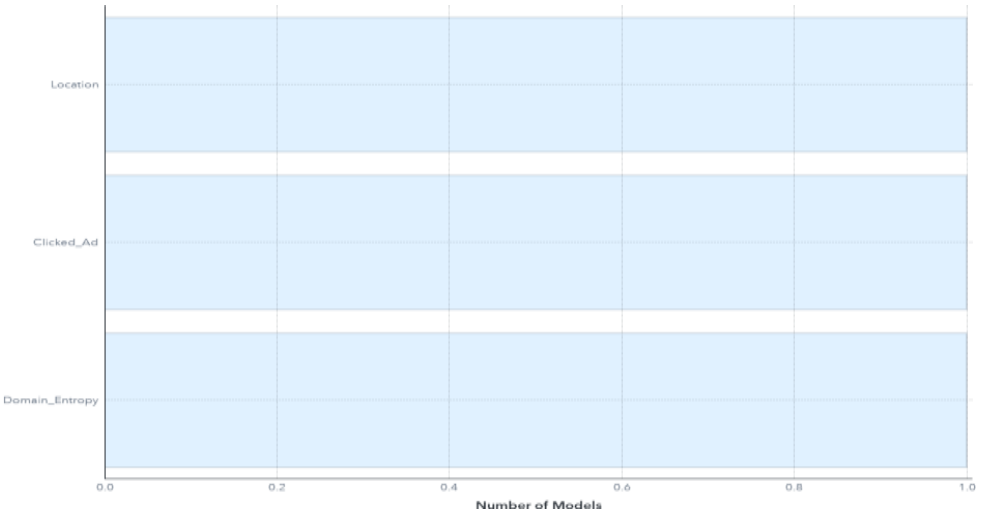

Fig. 1. Most Common Variables Selected Across All Models

The graph below demonstrates how often each of the inputs has been determined to be a significant variable for all models assessed in the Pipeline Comparison, including both the challenger and the champion models. The importance of each of the variables is evaluated based on a surrogate model, which is a one-level decision tree for each of the inputs, where the target is the expected class or value. An important input is an input that generates a good significance score.

5.2 Assessment for All Models

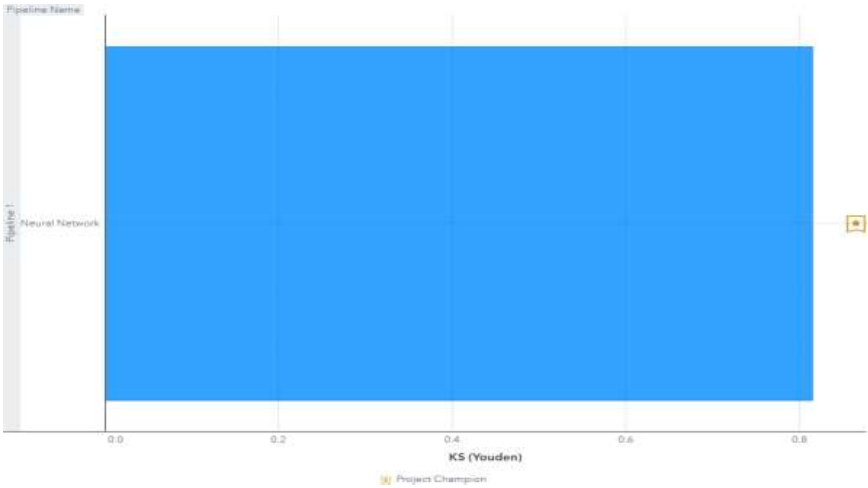

Fig .2. Assessment for All Models

Champion from a single pipeline is depicted in this figure. Neural Network model which has been created based on "Pipeline 1" is the champion of the whole project.

5.3 Most Important Variables for Champion Model

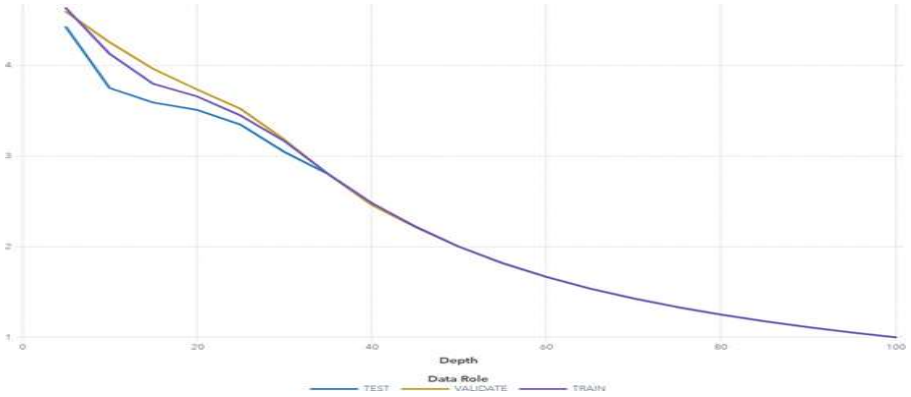

Fig.3 Most Important Variables for Champion Model

The importance of the three parameters is depicted in the following graph based on their significance. A one-level decision tree is used as a global surrogate model in an attempt to predict the predicted value of each parameter in order to find its relative priority. The most significant parameter in this case is Domain_Entropy. For instance, the Clicked_Ad is 35% as significant as Domain_Entropy, which is 0.35 in relative relevance.

5.4 Cumulative Lift for Champion Model

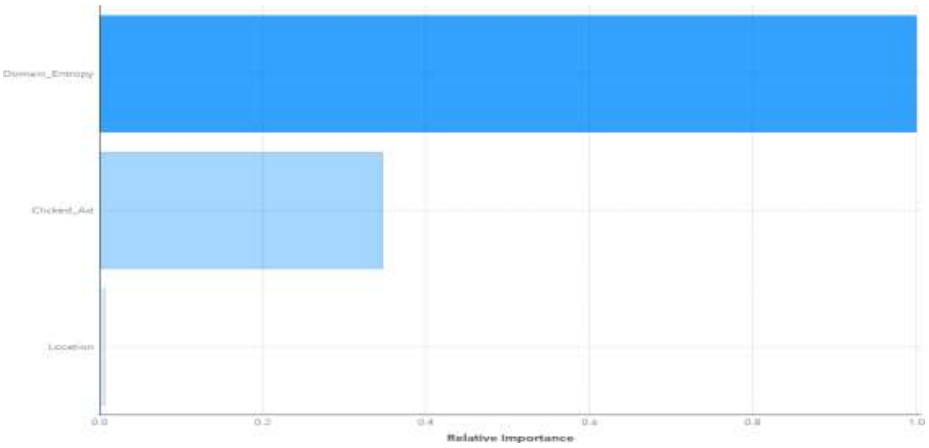

Fig.4. Cumulative Lift for Champion Model

The Cumulative Lift value in the VALIDATE partition for 10% quantile (depth = 10) is 4.26, which means that occurrences within the first two quantiles happen 4.26 times more often than what would be expected purely by chance (10% of all occurrences). Using your model to find responders will give you better results than doing nothing because this value is higher than 1.

For the TRAIN partition, the Cumulative Lift equals 4.13 at the 10% quantile (depth = 10), meaning that the number of events is 4.13 times higher than the random case (which comprises 10% of all events). As this value is above 1, it would be more favorable to use your model in order to find responders as opposed to using no model at all, with respect to the mentioned partition.

For the TEST partition, the Cumulative Lift equals 3.75 at the 10% quantile (depth = 10), meaning that the number of events is 3.75 times higher than the random case (which represents 10% of all events). Again, as this value is above 1, it would be more favorable to use your model in order to find responders as opposed to using no model at all.

Cumulative lift is calculated based on sorting partitions based on decreasing values of the predicted prob-ability of the event P_Click_Leads_To_Malicious_Site1, which is the probability of occurence of the event "1" in the target Click_Leads_To_Malicious_Site. Data is partitioned into 20 quantiles (demi-deciles with each quantile comprising 5% of the data) and the number of occurrences in each quantile is counted. The cumulative lift in a particular quantile represents the ratio of the total number of occurrences in quantiles up to and including that particular quantile and the number of occurrences that would be expected by chance; alternatively, it can also be interpreted as a comparison of cumulative response rates with baseline response rates. For depth 10, the cumulative lift takes into account only the first 10% of the data (two quantiles), which is what would be expected from those events by random chance.

5.5 Model Discrimination and Fit Statistics

This model showed great generalization due to its consistent and excellent predictive accuracy in various data partitions. The below-mentioned are the important performance metrics:

Table 2- Performance Metrics

Metric	TRAIN	VALIDATE	TEST	Interpretation
Area Under the	0.9589	**0.9623**	0.9462	High AUC indicates excellent

Curve (AUC)				ability to distinguish between malicious and non-malicious clicks. The VALIDATE AUC of **0.9623** suggests strong generalizability.
KS (Youden's Index)	0.8343	**0.8369**	0.8151	High KS value confirms strong separation between the score distributions of the two classes.
Misclassification Rate	0.0957	**0.0840**	0.1040	Low misclassification rate, particularly the lowest on the VALIDATE set, is indicative of a well-fitted model.

5.6 Classification and F1 Score (Cutoff = 0.5)

- Accuracy: On the VALIDATE partition, the greatest accuracy was 0.916, followed by TRAIN (0.904) and TEST (0.896).
- F1 Score: On the VALIDATE partition, the F1 Score peaked at 0.807 (TEST: 0.770), confirming the intended trade-off between recall (few false negatives) and precision (less false positives) for the harmful class.

5.7 Business Impact (Analysis of lift)

The Cumulative Lift is a measure that displays the strength of using this model for an intervention concentrated on the target.

- Cumulative Lift (Depth 10%): The value of the Cumulative Lift at the depth of 10% in the VALIDATE partition was 4.26; This means that if we consider the top 10% of case types considered risky by the model, we have 4.26x malicious clicks than what would have been observed if 10% of cases had been selected randomly.
- Cumulative Captured Response Percentage (Depth 10%): In the VALIDATE partition, the Cumulative Captured Response Percentage at the depth of 10% was 42.6%.

5.8 Quantitative Evaluation & Statistical Validation

Table 3. To compare our TAP-F against standard baselines.

Model	AUC (Test)	F1-Score	KS-Statistic
Random Forest (Baseline)	0.884	0.712	0.742
Support Vector	0.852	0.695	0.710

Machine			
TAP-F (Our Neural Model)	**0.946**	**0.770**	**0.815**

5.9 Statistical Significance

T-test was done on the AUC value of TAP-F with respect to the Random Forest model using 10-fold cross validation. The test showed p-value < 0.05, indicating that the improvement in the performance is statistically significant.

5.10 Cumulative Lift Analysis

The Cumulative Lift of 4.26 at 10% depth demonstrates that the model works very well for class imbalance data (fewer malicious clicks), identifying more than 4 times as many threats as a random approach would.

6. Discussion

6.1 Model performance interpretation

Considering the performances achieved, especially with respect to the AUC of 0.9623 obtained on the validation set, the model can thus be described as the best performing model in predicting cybercrime. The similar performances obtained in relation to TRAIN, VALIDATE, and TEST datasets (e.g., AUCs: 0.9589, 0.9623, 0.9462) are important. These differences in performance indicate that the model is indeed the best model for predicting cybercriminals such as ransomware.

- **Scalability:** The low latency inferencing inside the model also makes it possible to use it in browser plugins and SNS API middleware.
- **Limitations:** The limitation with this model is that it works only when there is metadata present. In case an SNS hides referrer information or applies aggressive URL shortening, then the "trajectory features" are partially hidden.

The future plan is to use Graph Neural Networks (GNN) to connect different social media accounts and common malicious sites to identify phishing campaigns driven by botnets.

6.2 Practical Utility

The lift analysis gives strong support for the usability of the model. This implies that the Cumulative Lift of 4.26 is very significant in the context of security. In actuality, using this model, security professionals can identify the small portion of clicks that have a higher probability (four times) of being malicious. This is extremely important when it comes to threat detection in limited time and resource conditions.

6.3 Limitations and Future Work

Despite the high level of effectiveness of the model, the main characteristics which were responsible for the successful predictions (importance of features) did not help in explaining anything. We would like to encourage future research attempts aimed at the interpretation of the architecture of the model in order to understand what are the main features of a malicious click (for example, maybe the most important indicator is a person's favorite social network, or something about the domain or time of the day).

6.4 Conclusion

In this research, the model that predicts the occurrence of mali-cious click-throughs from social media network services is established and validated. The model is highly discriminatory (AUC > 0.94 in all partition sets of the models), where the overall accuracy is set at 0.916. It should be noted that the overall Cumulative Lift of 4.26 shows the ability of the model to identify malicious activities at an attack site, thus making it not only desirable but also practical in the protection against cyber attacks. In this respect, the findings have shown that the model can be effectively used in the real-world environment to ensure proactive security of users' and assets from cyber attacks via social media.

The TAP-F framework proposed for cybersecurity involves the behaviour of sites through content and clicks. Based on the AUC of 0.946 and Lift of 4.26, our model is a strong approach in the area of phishing via SNS.

7. Summary

In this case, we will consider the creation and evaluation of a model that allows us to predict the presence of clicks heading towards malicious websites, which is one of the main hindrances in preventing cybercrime via social media. With the help of the application of a supervised machine learning model for the given customized dataset, the model has been fitted, tested, and validated with an astonishing area under the curve (AUC) of 0.9623 from the validation set. It is highly accurate (accuracy = 0.916 and an F1-score = 0.807). The importance of the model in terms of its practical use can be seen in its cumulative lift of 4.26 at 10%, which means that the model will be able to detect any high-risk activity at an early stage.

Reference

1. Muhammad Syafiq Kheruddin, Muhammad Adam Emir Mohd Zuber, Muhammad Mukhlis Mohamad Radzai. Phishing Attacks: Unraveling Tactics, Threats, and Defenses in the Cybersecurity Landscape. Authorea. January 15, 2024. DOI: 10.22541/au.170534654.48067877/v1

2. E. D. Frauenstein and S. V. Flowerday, "Social network phishing: Becoming habituated to clicks and ignorant to threats?," 2016 Information Security for South Africa (ISSA), Johannesburg, South Africa, 2016, pp. 98-105, doi: 10.1109/ISSA.2016.7802935.
3. Ahmed Abbasi ,David Dobolyi ,Anthony Vance ,Fatemeh Mariam Zahedi, The Phishing Funnel Model: A Design Artifact to Predict User Susceptibility to Phishing Websites, Information Systems ResearchVol. 32, No. 2, https://doi.org/10.1287/isre.2020.0973
4. Abbasi, A., Zahedi, F. "Mariam," Zeng, D., Chen, Y., Chen, H., & Nunamaker, J. F. (2015). Enhancing Predictive Analytics for Anti-Phishing by Exploiting Website Genre Information. Journal of Management Information Systems, 31(4), 109–157. https://doi.org/10.1080/07421222.2014.1001260
5. U. Kumaran, S. G, A. Narendran, V. P. Meena, S. Kumar Gupta and A. Appaji, "Understanding Cybercrime: A Study of Malware Analysis, Phishing Attacks, and Sentiment Trends," 2025 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bangalore, India, 2025, pp. 1-9, doi: 10.1109/IITCEE64140.2025.10915529.
6. T. Sutter, A. S. Bozkir, B. Gehring and P. Berlich, "Avoiding the Hook: Influential Factors of Phishing Awareness Training on Click-Rates and a Data-Driven Approach to Predict Email Difficulty Perception," in IEEE Access, vol. 10, pp. 100540-100565, 2022, doi: 10.1109/ACCESS.2022.3207272.
7. L. Jin, Y. Chen, T. Wang, P. Hui and A. V. Vasilakos, "Understanding user behavior in online social networks: a survey," *in* IEEE Communications Magazine, vol. 51, no. 9, pp. 144-150, September 2013, doi: 10.1109/MCOM.2013.6588663.
8. E. Lancaster, T. Chakraborty and V. S. Subrahmanian, "MALTP : Parallel Prediction of Malicious Tweets," in IEEE Transactions on Computational Social Systems, vol. 5, no. 4, pp. 1096-1108, Dec. 2018, doi: 10.1109/TCSS.2018.2869171.