







# Email Reading Behavior-Informed Machine Learning Model to Predict Phishing Susceptibility

Ning Xu<sup>1,2</sup> , Jiluan Fan<sup>1,2</sup> , and Zikai Wen<sup>3</sup>  

<sup>1</sup> Institute of Artificial Intelligence, Guangzhou University, Guangzhou, China  
{xuning, fanjiluan}@e.gzhu.edu.cn

<sup>2</sup> Guangdong Provincial Key Laboratory of Blockchain Security, Guangzhou, China

<sup>3</sup> Computational Media and Arts Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China  
zikaiwen@ust.hk

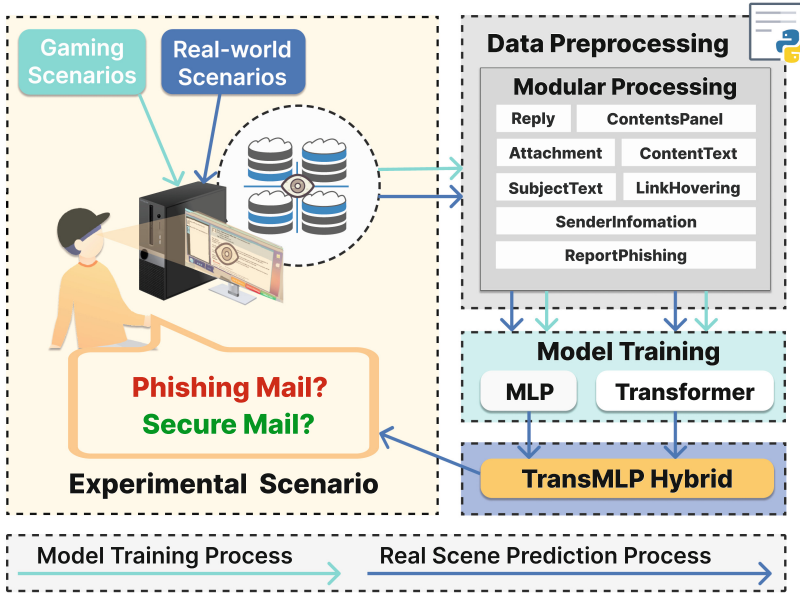
**Abstract.** As phishing threats intensify, incidents like the “COVID-19 vaccination form” phishing website underscore the limitations of relying solely on traditional firewall-based defenses. Consequently, there is a growing inclination towards user-centered anti-phishing solutions, exemplified by training games such as *What.Hack*. But could we proactively notify users in real time when they are on the brink of a scam or when their attention wanes? Our research explores machine learning and eye-tracking to identify email-reading weak spots and gauge a user’s risk of succumbing to phishing lures. We put forth innovative hybrid models, *TransMLP Link* and *TransMLP Hybrid*, melding the strengths of both Transformer and MLP. Our method also facilitates consistent interpretation of eye-tracking data across varied email interfaces and displays. Our *TransMLP Hybrid* model boasts an 88.75% accuracy rate, outperforming the standard Transformer model. Our research points to the future of anti-phishing tools that elegantly combine technological advancements with insights into human behavior.

**Keywords:** Anti-Phishing · User Modeling · Machine Learning

## 1 Introduction

Phishing is a growing concern in the digital age. It involves seemingly genuine emails, messages, and links that trick users into revealing personal data or downloading harmful software. The rise of such attacks, especially those leveraging pandemic themes, has been alarming [2,9]. An infamous example is the fake “COVID-19 Vaccination Form” site that falsely posed as an official NHS platform, leading users into fraudulent vaccine registrations [19].

Traditionally, firewalls have been used to combat phishing by maintaining updated blocklists and allowlists [7,13,14]. However, they struggled to



**Fig. 1.** Overview of the Model Training and Real-World Prediction Process for Phishing Email Detection.

counter new domains that are not yet listed [21]. As a solution, recent methods stressed the importance of educating users [3, 24, 30]. Training platforms like *What.Hack* [30] have sprung up to strengthen this first line of defense. Nonetheless, an unsettling 95% of phishing breaches result from human oversights [1]. This brings forth a question: Can we alert users in real time if they are about to fall for a scam or if their attention drifts?

To tackle this problem, we employed machine learning and eye-tracking techniques to analyze how users engage with emails, aiming to predict their vulnerability to phishing. Our research delved into the Transformer model, assessing its potential to gauge user focus; the Multilayer Perceptron (MLP) model, fine-tuned for eye-tracking data; and innovative hybrid models, *TransMLP Link* and *TransMLP Hybrid*, blending the best of both Transformer and MLP.

Moreover, we developed a technique to consistently interpret eye-tracking data across various devices and email applications, associating specific gaze points with their meaning in the email’s layout. This approach translated the raw eye-tracking data into eight key areas reflecting the main regions of an email interface. A detailed overview of our approach, from data collection to applying our hybrid model in real scenarios, can be found in Fig. 1.

In our experiment with 25 participants, we gathered eye-gazing patterns and user interactions while they interacted with genuine and phishing emails and played the *What.Hack* anti-phishing game. We utilized the in-game data to train

our phishing prediction models and the real-world data for testing. While the Transformer model delivered an 80.63% accuracy, the *TransMLP Hybrid* model stood out by achieving an impressive 88.75% accuracy rate.

In essence, our contributions are threefold:

1. The new design of hybrid models, *TransMLP Link* and *TransMLP Hybrid*, synergizing Transformer and MLP.
2. The new approach to uniformly interpret eye-tracking data across diverse email reading environments.
3. The experiment showcased the outstanding performance of *TransMLP Hybrid* with an 88.75% accuracy.

In the evolving landscape of anti-phishing, the dual challenges of innovative phishing tactics and human vulnerabilities necessitate more comprehensive defense strategies. This paper studies the intricate relationship between email reading behaviors, eye-tracking data, and their potential to inform machine learning models that predict phishing susceptibility.

We begin by examining the historical context of phishing attacks and the defense mechanisms in place, laying the groundwork for our innovative approach. Subsequently, we elucidate our machine learning models, emphasizing the novel integration of Transformer and MLP architectures. Following this, we detail our designed experiment, setting the stage for a thorough analysis of our results and their broader implications. By evaluating the effectiveness of our models and examining the underlying factors, we present a feasible strategy that combines advanced technological methods with deep insights into human behavior, paving the way for a significantly enhanced anti-phishing defense.

## 2 Related Work

The related work section explores phishing tactics, human vulnerabilities, and defense strategies designed to counteract these threats. The limitations of existing anti-phishing strategies led us to study the ability to leverage email reading eye-tracking data to train machine learning models to predict phishing susceptibility more effectively.

### 2.1 Phishing Email Attacks and Defense

Phishing is a cyber-attack where attackers pose as trustworthy entities to steal credentials or introduce malware. Research has identified three primary human vulnerabilities in defending against phishing attacks: a lack of system and security knowledge [4], challenges in detecting visual deception [10], and inattention [20]. For example, phishing emails often employ deceptive hyperlinks and subtle cues, such as spelling mistakes, to mislead users [12].

To address these vulnerabilities, a range of strategies has been developed to counteract phishing due to user negligence. These include anti-phishing training, active warning systems, and detection techniques using machine learning [8, 16, 24, 25, 27, 30]. Role-playing phishing simulation games [24, 30] aim to increase

users' security knowledge and awareness, and alert mechanisms were designed to notify users of potential threats [16]. Therefore, modeling how humans recognize phishing emails and implementing protective measures are crucial in preventing successful breaches.

Machine learning models are a mainstay in the detection of phishing emails [23]. For instance, Shie et al. [25] utilized deep learning and feature extraction to identify phishing emails. Additionally, Subasi et al. [27] assessed Adaboost and other boosting algorithms for detecting phishing websites, leveraging a dataset from the UCI repository to improve classifier accuracy. Despite these advances, even the most sophisticated machine learning model occasionally misses phishing threats. Thus, creating automated detection methods for phishing risks when users access their emails could provide an added layer of protection, significantly reducing the chances of successful attacks.

## 2.2 Eye-Tracking for User Intention Prediction

The Eye-Mind Hypothesis (EMH) suggests that during a task, an individual's focal point and cognitive thought are intrinsically linked — what they see often mirrors what they think [18]. In this context, eye-tracking data becomes pivotal in decoding visual attention and cognitive operations. With this premise, we postulate that specific eye-tracking patterns might be indicative of an individual's vulnerability to phishing emails.

Recent research in intent recognition through eye-tracking [5, 15, 17, 29] predominantly revolves around predicting the location or object of a user's attention. A research direction in this area aims to forecast subsequent attentional shifts of users [17, 29]. For instance, leveraging eye movement patterns from VR goggles, Nicolas et al. [26] developed a model to predict users' upcoming focal points. Deng et al. [11] utilized logistic regression to project user menu selections. Bhattacharya et al. [6] took a step further to investigate if readers' eye movements alone could gauge the authenticity of news headlines. Despite these advancements, such models remain unable to assess user susceptibility to phishing endeavors.

In a parallel development, Huang et al. [16] designed an array of visual cues to deter phishing, aiding users in distinguishing malicious emails from legitimate ones. However, the trigger for these alerts rests upon conclusive firewall detections. If a firewall deems an email safe, no alert is generated. This underscores an opportunity: if we can determine a user's lack of attentiveness while reading a phishing email, a timely alert could also be triggered.

## 3 Prediction Models Design

The section explores machine learning models for analyzing email reading behaviors using eye-tracking data. We start with the Transformer model, detailing its mechanics and applications in understanding user attentiveness. We then discuss the Multilayer Perceptron (MLP) model and how to make it process eye-tracking statistics. Finally, we introduce two new variant models, combining the best of both Transformer and MLP, to better predict phishing email susceptibility.

### 3.1 Transformer Model

**Model Background:** The Transformer model, proposed by Vaswani et al. [28], addresses the performance bottlenecks of recurrent neural networks in processing long data sequences. It comprises an encoder and a decoder, both stackable with multiple layers that comprise the self-attention layer and the feed-forward layer.

While the self-attention mechanism of the Transformer model processes data, it does not inherently consider the order of the input sequence. To enable sequence processing, positional encoding (PE) is necessary. The formula for positional encoding is:

$$PE_{(pos,2i)} = \sin(pos/10000^{2i/d}), PE_{(pos,2i+1)} = \cos(pos/10000^{2i/d}), \quad (1)$$

where  $d$  denotes the embedding vector's dimension,  $pos$  signifies the position in the data processing sequence, and  $i \in [0, d]$  represents the dimensions of the positional encoding vector.  $2i$  and  $2i + 1$  designate the even and odd dimensions of the positional embedding vector respectively.

The Transformer model may employ an  $h$  multi-head attention mechanism to capture richer feature information, which is essential for our application's purpose. Within the multi-head self-attention layer, the input vector undergoes three linear transformations to obtain the query vector  $Q$ , key vector  $K$ , and value vector  $V$ . The formula for multi-head attention computation is:

$$\text{MultiHead}(Q, K, V) = \text{Concatenation}(\text{head}_1, \dots, \text{head}_h)W^O, \quad (2)$$

where each  $\text{head}_i$  represents the output vector of the  $i$ -th attention head, and  $W^O$  is a linear transformation matrix. The formula for each head is:

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V), \quad (3)$$

with the matrices  $W_i^Q$ ,  $W_i^K$ , and  $W_i^V$  being linear transformations. The dimension of each head helps define the scaled dot-product attention:

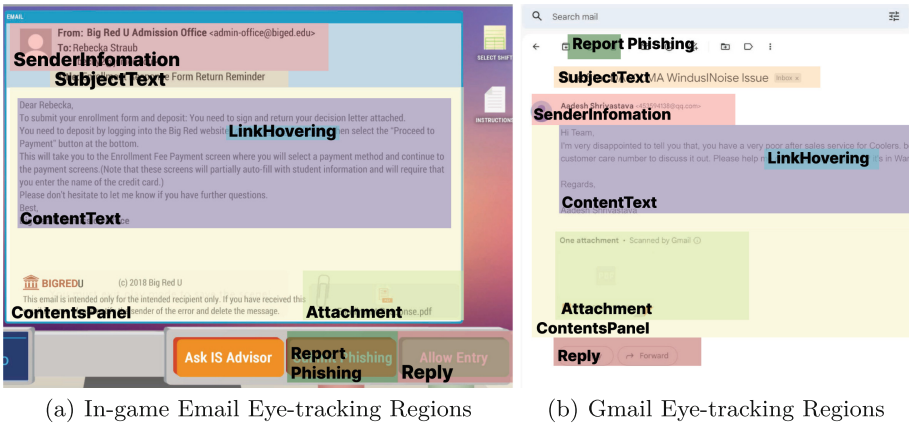
$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d/h}}\right)V. \quad (4)$$

The feed-forward network within the Transformer model is a two-layer neural network, which employs residual connections [17] or layer normalization [5] to facilitate model convergence and prevent gradient disappearance or explosion.

**Model Implementation:** We trained a series of Transformer-based models using temporal features to perform binary classification on email reading behavior. The aim is to determine whether users are careless about verifying the authenticity of emails.

The eye-tracking data for each user and email serves as a sequence input to the Transformer model. The eye-tracking data are chronologically organized into sequences according to the user’s history of processing emails.

To provide consistency in interpreting eye-tracking data, regardless of the screen size or email application in use, we developed a method to map location points from the eye-tracking data to their respective semantic meanings. The transformed data comprises eight spatial attributes, specifically: *SenderInformation*, *SubjectText*, *Reply*, *ReportPhishing*, *ContentText*, *ContentsPanel*, *Attachment*, and *LinkHovering*. These attributes align with the core email functions’ UI regions, as depicted in Fig. 2. Furthermore, our model incorporates a temporal feature. Each feature vector captures the needed eye-tracking information during each time step.



**Fig. 2.** Eye-tracking Mapping for Email Interaction Zones across Two Different Email Application Interfaces.

Our Transformer encoder consists of two blocks, each containing one multi-head self-attention layer and one feed-forward layer. Within the self-attention layer, the input vector is divided into three segments, each of which undergoes a linear transformation. Subsequently, these transformed segments are subjected to scaled dot-product attention calculations. The resulting output vectors from each head are combined and processed through a linear transformation matrix to produce the final output of the self-attention layer as the input of the feed-forward layer. The feed-forward layer includes two linear layers with a *ReLU* activation function in between them. After the input undergoes transformation by a fully connected layer, the activation function provides a nonlinear transformation. A subsequent fully connected layer further modifies the output, producing a tensor that maintains the input’s dimensions. Within each encoder block, the input is processed by both the self-attention mechanism and the feed-forward network, with the output being reintegrated with the original input through a residual connection.

### 3.2 Multilayer Perceptron Model

**Model Background:** The Multilayer Perceptron (MLP) model [22] employs multiple layers of neurons to enact nonlinear transformations, facilitating the extraction of higher-level features from input data. An MLP is composed of input, hidden, and output layers. Each layer houses multiple neurons, and each neuron processes the output from the preceding layer. The calculations performed by a neuron involve both a linear transformation that weights the output of the previous layer by the neuron’s own weights and a subsequent nonlinear transformation via an activation function. This combination generates the neuron’s final output. The computational formulation for MLP is given by:

$$r = f(W^{(L)} f(W^{(L-1)} f(W^{(L-2)} \dots f(W^{(1)} x + B^{(1)}) \dots + B^{(L-2)}) + B^{(L-1)}) + B^{(L)}, \quad (5)$$

in this equation,  $f$  denotes the activation function,  $x$  is the input data, and  $W^{(i)}$  and  $B^{(i)}$  symbolize the weights and biases for the  $i$ -th layer, respectively. Generally, the terminal layer of the MLP model uses the *sigmoid* function to transform the previous network’s output into two probability values, and the model picks the higher probability value as the final output.

**Model Implementation:** We employed an MLP model comprising six fully connected layers, using *ReLU* as the activation function and incorporating a dropout method to combat overfitting. The input to this model is derived from eye-tracking data, which we processed into 16 statistical features. These features come from eight previously identified spatial features related to the UI areas of core email functions. For each spatial feature, we calculated two values: the count and the total duration of user fixations. This data was then flattened into a one-dimensional vector. The model produces an output in the form of a probability value, representing the likelihood of a sample being a phishing email that successfully deceives the recipient.

For the training phase, we opted for the Adam optimizer over the stochastic gradient descent algorithm, enabling faster convergence and allowing distinct learning rates for individual parameters. Furthermore, we set the learning rate of each parameter group using a cosine annealing schedule to dynamically modify the learning rate, progressively decreasing it throughout training for better control and stability.

For the loss function, we used the binary cross-entropy for better training stability. This function is mathematically represented as:

$$Loss(y, p) = -(y \log(p) + (1 - y) \log(1 - p)), \quad (6)$$

in this equation,  $y$  represents the ground truth, indicating if the user failed to recognize the deceptive phishing email. Meanwhile,  $p \in [0, 1]$  denotes the model’s predicted probability that we aim to align with the  $y$  value.

### 3.3 TransMLP Model Variants Design

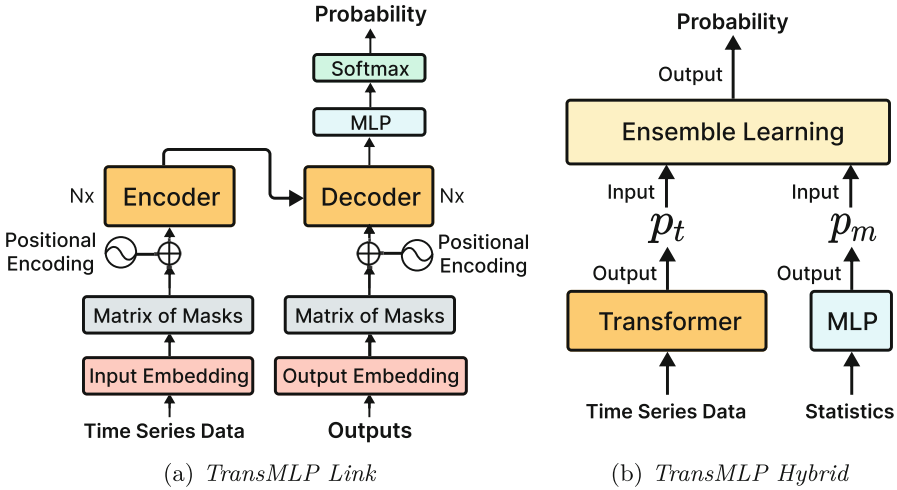


Fig. 3. Architectures of Two Proposed TransMLP Variants.

**TransMLP Link:** The *TransMLP Link* model is a new variant that diverges from the traditional Transformer model. The model integrates a multi-layer transformer encoder to enhance input data feature extraction, as shown in Fig. 3(a). In addition to this integration, the model employs an MLP model in lieu of the standard single fully connected layer to facilitate nonlinear transformations on the Transformer output. This design choice enables the *TransMLP Link* to achieve better nonlinear modeling capabilities relative to ML models that rely solely on a single fully connected layer.

**TransMLP Hybrid:** Rather than merely linking the output of the Transformer directly to the MLP’s input, we designed the *TransMLP Hybrid* model to harmoniously integrate the strengths of both Transformer and MLP paradigms, as shown in Fig. 3(b). This model harnesses eye-tracking statistical features derived from time series data to train the MLP component. To produce the final output, an ensemble learning strategy is employed, judiciously weighing the predictions from both the Transformer and MLP models to optimize performance.

## 4 Experiment Design

Our experiment centered on leveraging eye-tracking data to enhance the ability of machine learning models to assess the phishing risk of an email as read by a user. We also aimed to evaluate the two Transformer model variants that we proposed, *TransMLP Link* and *TransMLP Hybrid*, comparing their performance to the basic Transformer model.



## 4.1 Participant Recruitment

We collected eye-tracking and user interaction data from 25 participants (10 females and 15 males) for model training and testing. Participants were recruited via social media and snowball sampling. The participants need to be over 18 years old, have no prior training in anti-phishing, and be affiliated with the first author’s institution, which was targeted by the collected real phishing emails.

## 4.2 Experimental Method

First, we gathered data from participants in real-world application settings to assess their reactions to phishing and legitimate emails while using their everyday email applications. We designed simulated interfaces mimicking Gmail and NetEase Mail, which contained 6 phishing emails and 5 safe emails. These email addresses and contents were sourced from actual reported phishing cases. During the exercise, participants chose to reply or report the emails while we recorded their gaze data using the 7invensun A3 eye tracking device.

Then, we captured their gaze behavior while they engaged with the anti-phishing training game, *What.Hack*. This game comprises 5 levels, each emphasizing different email attributes to identify phishing attempts. In Level 1, volunteers inspected the sender’s email address. By Level 3, they were also evaluating potential phishing links, and by Level 5, they assessed attachments alongside previous checks. We observed that all participants had completed the game.

We recorded gaze positions, mouse movements, and link-hovering events throughout these two activities.

## 4.3 Data Post-processing

For the dataset obtained from participants playing *What.Hack*, we preserved the time series data for the Transformer model and computed the accumulated statistical data for the MLP model. We collected a total of 18,720 eye-tracking fixation events. We processed them into 1,019 events of user reaction to emails. We also computed the overall fixation duration and the number of event occurrences. All data has been anonymized. We will release the database<sup>1</sup> after implementing differential privacy measures to enhance user data protection.

## 5 Findings and Discussions

In our comprehensive analysis of phishing email susceptibility prediction models, the *TransMLP Hybrid* model distinctly stood out for its accuracy and adaptability to various phishing email challenges. Furthermore, our findings highlighted that eye movement patterns offer valuable insights into factors that influence prediction accuracy.

---

<sup>1</sup> <https://github.com/zikaiwen/EmailEye-PhishPredict>.

In the subsequent discussions, we introduced our novel technique for organizing eye-gazing data in email interfaces, ensuring consistent data collection across diverse devices and email applications. Additionally, we discussed the potential of merging malicious link detection with behavior-driven alerts, providing a robust defense against phishing attacks.

## 5.1 Findings

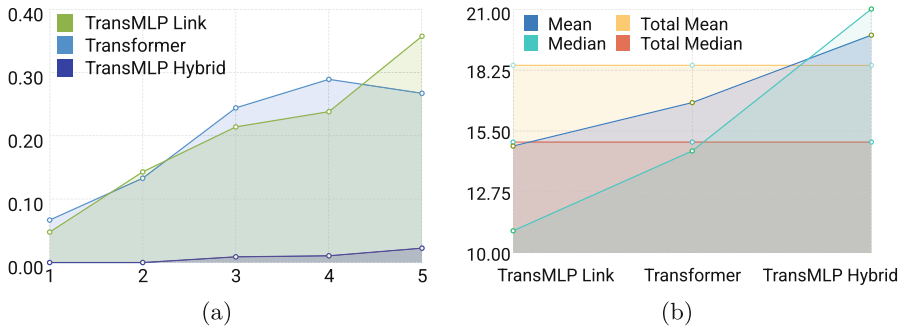
Our three primary findings are the results from the model accuracy comparison, the relationship between phishing email complexities and model error rates, and the correlation between saccade counts and model error rates.

**Model Accuracy Comparison:** We evaluated the performance of three models: Transformer, *TransMLP Link*, and *TransMLP Hybrid*. The Transformer achieved an accuracy of 80.63% in predicting the phishing email’s susceptibility. This accuracy saw a slight increase to 80.75% when augmented with MLP using *TransMLP Link*. However, the *TransMLP Hybrid*, which was trained on both game statistical and time series data, outperformed the others, achieving an accuracy of 88.75%. This suggests the *TransMLP Hybrid* is the most effective model when considering both statistical and sequential data. Detailed outcomes are provided in Table 1.

**Table 1.** Accuracy Rates of Transformer, *TransMLP Link*, and *TransMLP Hybrid* Models in Real-World and In-game Scenarios

Model Name	Real-World Accuracy(%)	In-game Accuracy	
		Testing(%)	Training(%)
Transformer	80.63	80.88	82.58
TransMLP Link	80.75	79.90	82.21
<b>TransMLP Hybrid</b>	<b>88.75</b>	<b>89.82</b>	<b>90.16</b>

**Phishing Email Difficulty and Prediction Error Rates:** We delved into the performance of the models as they predicted user intent during the game *What.Hack*, which is designed with escalating complexities across its 5 levels to simulate varying phishing email attributes. Starting at Level 1, participants primarily focused on scrutinizing the sender’s email address. By Level 3, their evaluation expanded to include potential phishing links. By the time they reached Level 5, they were also assessing email attachments in addition to their previous tasks. Of all the models, *TransMLP Hybrid* stood out by consistently registering the lowest error rate across every level of difficulty. Conversely, the other two models struggled more with discerning user intent in the face of complex phishing emails, as depicted in Fig. 4(a).



**Fig. 4.** Model Performance Analysis. (a) represents the prediction error rate of different models at different difficulty levels, and the vertical ordinate represents the prediction error rate. (b) represents the statistical data of different saccades when different models predict incorrectly, and the vertical ordinate represents the number of saccades.

**Saccade Counts and Prediction Errors Rates:** The average number of saccades (rapid eye movements) observed when models made inaccurate predictions was 18.37, with a median of 15.00. Notably, *TransMLP Link* exhibited more errors when there were fewer saccades. Conversely, as the number of saccades increased, the accuracy of *TransMLP Hybrid* predictions appeared to decline. These trends are illustrated in Fig. 4(b).

## 5.2 Discussions

**Modularizing Eye-Gazing Points in Email UI for Enhanced Feasibility and Effectiveness:** To ensure broader applicability and improved feature learning for classifying phishing email susceptibility, we developed a modularization technique for eye-gazing location data. This method divides email application interfaces into eight specific modules: *SenderInformation*, *SubjectText*, *Reply*, *ReportPhishing*, *ContentText*, *ContentsPanel*, *Attachment*, and *LinkHovering*. This structure enables the formation of a consistent dataset that is not tied to absolute coordinate positions. It thus overcomes the challenges posed by differing screen resolutions and email applications, ensuring the collected data from mouse and eye-gazing events remains relevant and usable.

**Integrating Malicious Link Detection and Behavior Intervention for Comprehensive Anti-Phishing:** Building on our research, there is potential to merge malicious link detection and behavioral intervention alerts. This holistic approach, fusing user intent recognition, machine learning classification, and effective UI warnings, can substantially lower the risk of phishing incidents.

## 6 Conclusion and Future Work

In our explorative research into machine learning’s capabilities, we honed in on the Transformer model and its variants, particularly in the context of predicting

phishing email susceptibility using eye-tracking data. Our ambition was to discern and understand the nuances of user attentiveness during email interactions, with the goal of leveraging this information to optimize phishing risk evaluations.

Among the models we evaluated, the *TransMLP Hybrid* emerged as a clear frontrunner. Its precision, coupled with its adaptability to diverse phishing scenarios, set it apart. Moreover, our study underscored the pivotal role that eye movement patterns play in determining prediction accuracy. Even though the *TransMLP Hybrid* model was exemplary in its performance. There lies an exciting challenge in enhancing this model further by augmenting its model architecture that marries eye-tracking data with other related behavioral indicators.

Looking ahead, our research has paved the way for several promising trajectories. The innovative technique we introduced for standardizing eye-gazing data in email interfaces marks a substantial advancement in ensuring consistent and reliable data collection across varying platforms. Furthermore, our discussions around merging malicious link detection with behaviorally-driven alerts have underscored a pressing need and significant opportunity for creating comprehensive defense mechanisms against phishing attacks. This multi-faceted approach, blending technology with human behavioral insights, could form the cornerstone of next-generation anti-phishing solutions.

**Acknowledgments.** The authors gratefully acknowledge support from the China Postdoctoral Science Foundation under grant number 2022M720889. The authors would like to thank the anonymous reviewers for their valuable comments and helpful suggestions.

## References

1. Alkhalil, Z., Hewage, C., Nawaf, L., Khan, I.: Phishing attacks: a recent comprehensive study and a new anatomy. *Front. Comput. Sci.* **3**, 563060 (2021)
2. Aonzo, S., Merlo, A., Tavella, G., Fratantonio, Y.: Phishing attacks on modern android. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, pp. 1788–1801, 2018
3. Arachchilage, N.A.G., Love, S.: A game design framework for avoiding phishing attacks. *Comput. Hum. Behav.* **29**(3), 706–714 (2013)
4. Arachchilage, N.A.G., Love, S.: Security awareness of computer users: a phishing threat avoidance perspective. *Comput. Hum. Behav.* **38**, 304–312 (2014)
5. Bednarik, R., Eivazi, S., Vrzakova, H.: A computational approach for prediction of problem-solving behavior using support vector machines and eye-tracking data. In: Nakano, Y.I., Conati, C., Bader, T. (eds.) *Eye Gaze in Intelligent User Interfaces: Gaze-based Analyses, Models and Applications*, pp. 111–134. Springer London, London (2013). [https://doi.org/10.1007/978-1-4471-4784-8\\_7](https://doi.org/10.1007/978-1-4471-4784-8_7)
6. Bhattacharya, N., Rakshit, S., Gwizdka, J., Kogut, P.: Relevance prediction from eye-movements using semi-interpretable convolutional neural networks. In: *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*, pp. 223–233, 2020
7. Caputo, D.D., Pflieger, S.L., Freeman, J.D., Johnson, M.E.: Going spear phishing: Exploring embedded training and awareness. *IEEE Secur. Priv.* **12**(1), 28–38, 2014

8. Chanti, S., Chithralekha, T.: Classification of anti-phishing solutions. *SN Comput. Sci.* **1**(1), 11 (2020)
9. Cui, Q., Jourdan, G-V., Bochmann, G V., Couturier, R., Onut, I-V.: Tracking phishing attacks over time. In: Proceedings of the 26th International Conference on World Wide Web, pp. 667–676, 2017
10. Das, S., Christena, N-E., Camp, L.J.: Evaluating user susceptibility to phishing attacks. *Inf. Comput. Secur.* **30**(1), 1–18, 2022
11. John, B.D., Peacock, C., Zhang, T., Murdison, T.S., Benko, H., Jonker, T.R.: Towards gaze-based prediction of the intent to interact in virtual reality. In: ACM Symposium on Eye Tracking Research and Applications, pp. 1–7, 2021
12. Dhamija, R., Tygar, J.D., Hearst, M. :Why phishing works. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 581–590, 2006
13. Jr, R.C.D., Carver, C., Ferguson, A.J.:Phishing for user security awareness. *Comput. Secur.* **26**(1):73–80, 2007
14. Han, X., Kheir, N., Balzarotti, D. Phisheye: live monitoring of sandboxed phishing kits. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, pp. 1402–1413, 2016
15. Huang, C.-M., Andrist, S., Sauppé, A., Mutlu, B.: Using gaze patterns to predict task intent in collaboration. *Front. Psychol.* **6**, 1049 (2015)
16. Huang, L., Jia, S., Balcetis, E., Zhu, Q.: Advert: an adaptive and data-driven attention enhancement mechanism for phishing prevention. *IEEE Trans. Inf. Forensics Secur.* **17**, 2585–2597 (2022)
17. Ishii, R., Ooko, R., Nakano, Y.I., Nishida, T. Effectiveness of gaze-based engagement estimation in conversational agents. In: Eye Gaze in Intelligent User Interfaces: Gaze-Based Analyses, Models and Applications, pp. 85–110, 2013
18. Just, M.A., Carpenter, P.A.: A theory of reading: from eye fixations to comprehension. *Psychol. Rev.* **87**(4):329, 1980
19. Kay, R., phish, F.: Fake mandatory Covid-19 vaccine form, 2023. <https://www.inky.com/en/blog/fake-mandatory-Covid-19-vaccine-form>
20. Koggalahewa, D., Yue, X., Foo, E.: An unsupervised method for social network spammer detection based on user information interests. *J. Big Data* **9**(1), 1–35 (2022)
21. Miyamoto, Daisuke, Hazeyama, Hiroaki, Kadobayashi, Youki: An Evaluation of Machine Learning-Based Methods for Detection of Phishing Sites. In: Köppen, Mario, Kasabov, Nikola, Coghill, George (eds.) *ICONIP 2008*. LNCS, vol. 5506, pp. 539–546. Springer, Heidelberg (2009). [https://doi.org/10.1007/978-3-642-02490-0\\_66](https://doi.org/10.1007/978-3-642-02490-0_66)
22. Murtagh, F.: Multilayer perceptrons for classification and regression. *Neurocomputing* **2**(5–6), 183–197 (1991)
23. Sharma, P., Dash, B., Ansari, M F.: Anti-phishing techniques-a review of cyber defense mechanisms. *Int. J. Adv. Res. Comput. Commun. Eng. ISO*, 3297:2007, 2022
24. Sheng, S., et al.: Anti-Phishing Phil: the design and evaluation of a game that teaches people not to fall for phish. In: Proceedings of the 3rd Symposium on Usable Privacy and Security, pp 88–99, 2007
25. Shie, E.W.S.: Critical analysis of current research aimed at improving detection of phishing attacks. *Sel. Comput. Res. pap.* **45**, 2020
26. Stein, N., Bremer, G., Lappe, M.: Eye tracking-based LSTM for locomotion prediction in VR. In: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 493–503. IEEE, 2022

27. Subasi, A., Molah, E., Almkallawi, F., Chaudhery, T.J.: Intelligent phishing website detection using random forest classifier. In: 2017 International Conference on Electrical and Computing Technologies and Applications (ICECTA), pp. 1–5. IEEE, 2017
28. Vaswani, A., et al.: Attention is all you need. *Advances in neural information processing systems*, **30**, 2017
29. Wei, P., Liu, Y., Shu, T., Zheng, N., Zhu, S-C.: Where and why are they looking? jointly inferring human attention and intentions in complex tasks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6801–6809, 2018
30. Wen, Z.A., Lin, Z., Chen, R., Andersen, E.: What. hack: engaging anti-phishing training through a role-playing phishing simulation game. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–12, 2019