RESEARCH ARTICLE OPEN ACCESS

Multidimentional Phishing Susceptibility Pridiction Model for Effective Mitigation of Potential Phishing Attack

Adamu Muhammad Tukur[1], Hamza Audi Giade[2], Isah Muhammad Lamir[3]; Yusuf Chindo[4]

[1]Department of Computer science. AbubakarTatari Ali Polytechnic, Bauchi,Bauchi State Nigeria.
E-mail: babaadamu6644@gmail.com

[2]Department of Computer science.AbubakarTatari Ali Polytechnic, Bauchi,Bauchi State Nigeria.
E-mail: gidadohenz@gmail.com

[3]Department of Computer science.AbubakarTatari Ali Polytechnic, Bauchi,Bauchi State Nigeria.
E-mail: muhammadlamir75@gmail.com

[4]Department of Computer science.AbubakarTatari Ali Polytechnic, Bauchi,Bauchi State Nigeria.
E-mail: yusufmusachindo@gmail.com

--************************----------------------------------

Abstract:

In recent years, most of our daily activities, such as online banking and online shopping, are increasingly linked to the Internet, which makes our lives more manageable and convenient wherever we are. However, these services come with some serious security risk that can cost Internet users dearly. Experts agree that phishing is one of the critical issues that have yet to be resolved, especially on its prediction accuracy. This is due to the fact that phishing is directed toward people instead of machines. The MPSPM phishing prediction model developed by Rundong Yang et al. in 2022 has achieved prediction accuracy of 89.04% on the phishing dataset as their research did not capture cognitive processes, which is one of the five factors that influence phishing susceptibility in the MPSPM model. The objective of this research is to use MPSPM prediction model and make improvement on the prediction accuracy of the model by including cognitive processes as an input into the MPSPM model. The results from the study indicate that identifiable models can accurately predict potential phishing victims, with the MPSPM model achieving a correct detection rate of 93.68% on the phishing dataset.

Keywords **—Phishing, Prediction, Susceptibility, Attack, Model.**

--************************----------------------------------

I. INTRODUCTION

important role in today's modern business activities and social activities [1], providing a lot of online services that tend to make our lives easier. With the aid of these services, we can access information anywhere. For instance, as more people have grown accustomed to it, online banking has become very popular [2]. There is no doubt that various types of attacks have become more common due to the pervasiveness of internet technology for information sharing. Replay and phishing are prominent examples, as are pharming, masquerading, and denial of service [3]. However, it has been established that the inadequate nature of security measures has greatly contributed to the network's vulnerability [4], as well as identity hiding, fame, and notoriety [5].

Phishing involves sending or disseminating various electronic mails that appear to come from a reputable person or organization in order to trick a target into disclosing sensitive personal information like a credit card number, a password, or other biometric details that could be used on the owner's behalf to carry out illegal activities. Phishing is a medium for numerous computer attacks that allow for the distribution of socially inspired messages to various internet users by requesting valuable and confidential information that can then be used against them to conduct illegal transactions on their behalf [5]. Phishers send messages via SMS, computer games, VoIP, and websites [6]. According to experts, phishing is still a serious issue that hasn't been resolved, because phishing attacks target people rather than machines [7].

The Multidimensional Phishing Susceptibility Prediction Model (MPSPM) is a phishing prediction model developed by Rundong Yang et al in 2022. It has achieved prediction accuracy of 89.04% on the phishing dataset that they used in their research. In their research, they did not captured cognitive processes, which is one of the five factors that influence phishing susceptibility in the MPSPM model. This research focused on finding factors that contribute to user phishing vulnerability and improving the accuracy of MPSPM susceptibility prediction model developed by (Rundong Yang et al, 2022). Our objective is to use MPSPM prediction model for user phishing susceptibility developed by [7] and make improvement on the prediction accuracy of the model by including cognitive processes as an input into the MPSPM model.

In this work, we use the current hypotheses of the existing work and improved our understanding of the variables that influence phishing susceptibility. We improved the prediction accuracy by examining the existing variables that have the biggest effects on susceptibility to phishing, and our work increase MPSPM prediction accuracy of 89.04%.

II. LITERATURE REVIEW

This section will first discuss the susceptibility to phishing in general, with a focus on phishing attacks. Then, it will compare various phishing prediction models. Finally, it will discuss similar or related works on susceptibility to phishing models based on multidimensional features that have been developed by previous researchers.

A. Susceptibility to Phishing Attack

Researchers in [8] identified the factors that influence users' susceptibility to phishing attacks on social networking sites, which are often targeted by phishers due to the large number of potential victims and behavioral vulnerabilities. They developed a theoretical framework to investigate phishing susceptibility on social networking sites. They used Structural Equation Modeling (SEM) to analyze the data collected during the study. The SEM results revealed that individuals who have a high level of conscientiousness (a personality trait associated with being organized, responsible, and careful) were less susceptible to phishing attacks. Thus, they suggested that being conscientious can play a protective role in reducing the likelihood of falling for phishing scams. However, their studies have certain limitations. Firstly, the sample used in the research consisted exclusively of students. As a result, the findings may not be representative of the general public. Their small sample size further restricts the generalizability of the results, and their conclusions may not apply to a broader population.

In order to improve the understanding of phishing susceptibility, they recommended expanding the model to include additional elements such as knowledge, social norms, perceived risk, and self-efficacy. These factors may provide more insights into how susceptible people are to phishing attacks in various contexts

Researchers in [9] investigated the factors influencing individuals' susceptibility to phishing emails using the Signal Detection Theory (SDT) framework. While the phishing susceptibility measures have been validated, the cognitive processes underlying individual differences in these measures are still not fully understood. Thus, they proposed and tested a theoretical path model that explored the influence of several factors on users' susceptibility to phishing emails. These factors include the Big Five Personality traits (i.e., openness, conscientiousness, extraversion, agreeableness, and emotional stability), knowledge and experience, and the cognitive processing of emails, specifically mail elaboration. However, their studies have certain limitations. Firstly, they acknowledges that the phishing susceptibility measures used in the research (sensitivity, judgment criterion, and correct rate in the phishing email detection task) only reflect a portion of an individual's vulnerability to phishing in real-life situations. In the real world, the likelihood of falling victim to phishing attacks is influenced by numerous other important variables such as personal interests, specific contexts, and the tactics employed by phishers.

B. Prediction Models of User Susceptibility to Phishing Attacks

Researchers in [10] highlighted the fact that phishing attacks can still succeed even when anti-phishing tools are in place, primarily because people struggle to recognize phishing attempts when they encounter them. Most existing research focuses on examining the static aspects of phishing behavior. To improve phishing susceptibility prediction, the researchers proposed a model called the "Dynamic-Static Model" (DSM). This model combines both dynamic and static features to enhance accuracy in identifying individuals who are more susceptible to phishing attacks.

In their studies, the test subjects mainly consisted of college students. However, the researchers highlighted that the sample data was not equally distributed, which could be a potential reason why they achieved less than 90% correctness in their results. Unequal distribution in the sample data may introduce biases and impact the model's performance.

Researchers in [11] highlighted the importance of understanding why some individuals are more susceptible to phishing attacks than others. The primary objective of the study is to understand why certain individuals are more susceptible to phishing attacks compared to others. Phishing attacks often rely on social influence and persuasion strategies to deceive users into revealing sensitive information. To explore this susceptibility to social influence, the researchers developed a scale called the "Susceptibility to Persuasion Strategies Scale." This scale is based on a dual-process model of persuasion and a framework that considers various social influence factors. However, the study has certain limitations. They did not assessed the impact of multiple persuasion principles in a single email, meaning that the influence of combining different persuasive tactics within a single phishing email was not tested. Additionally, the study design involved using only one genuine and one phishing email for each of the persuasion principles. This restricted approach might not fully capture the complexity of real-world phishing scenarios, where attackers often employ multiple strategies and variations.

Another researchers in [12] addressed the issue of phishing attacks, where attackers use social engineering techniques to exploit human carelessness and obtain sensitive information. Most current anti-phishing mechanisms often fail to detect malicious pages that lack visual and textual details, making them ineffective against certain phishing attacks. They proposed an approach called "piracema.io," which is a rule-based model for predicting phishing attempts. The model was designed to improve prediction accuracy, especially against phishing pages that have more sophisticated and rich content based on the page reputation. Their experiments achieved accuracy above 90% and demonstrated the ability to overcome some of the flaws present in native anti-phishing solutions. However, to optimize the understanding of terms and keywords, they used Natural Language Processing (NLP), which at times led to certain issues during the study.

C. Similar/Related Work

Researchers in [13] highlighted phishing is considered as one of the most dangerous attacks on digital security as it targets one of the weakest links in a network system that is the human factor. They developed a multidimensional phishing susceptibility prediction model (MPSPM) to identify individuals who are more likely to fall victim to phishing attempts. The MPSPM model achieved a prediction accuracy of 89.04%. However, they highlighted that they did not capture cognitive processes, which are one of the five factors influencing phishing susceptibility in the MPSPM model.

III. METHODOLOGY

In this study, 2,121 volunteers were recruited to participate. The phishing susceptibility model (MPSPM) designed by Rundong Yang et al (2022) is shown in Figure 1. It is divided into three parts: feature extraction part, classification part, and the prediction part. The MPSPM model is used to predict user susceptibility to phishing attacks, and considers five categories of decision factors that affect susceptibility: demographics, personality, cognitive processes, knowledge and experience, and security behavior. These factors are used as features for prediction using multidimensional features and multiple supervised machine learning methods such as LR, SVM, RF, GBDT, AdaBoost, and XGBoost. The rest of this section will detail the susceptibility factors and model specifics.

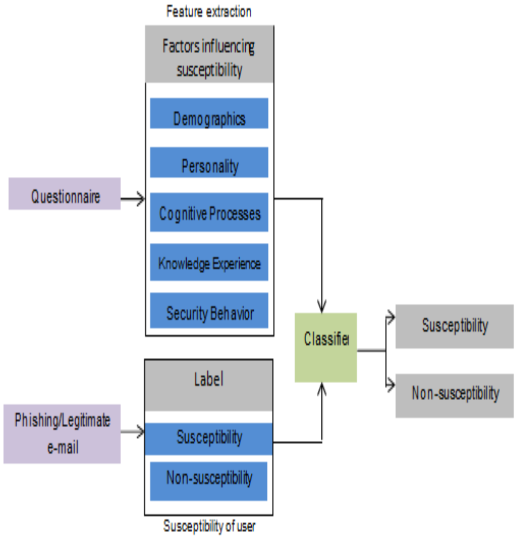


Figure 1: MPSPM Model

Demographic Factors:

Researchers in [14, 15] have studied demographic characteristics as one of the important factors that influence phishing susceptibility. According to [16], age has a significant impact on phishing susceptibility.

Personality Factor:

Researchers in [17] highlighted the Big Five's analysis about the relationship between the principle of persuasive and Five-Factor Model (FFM) of personality. The "Social Engineering Personality Framework" theory was proposed, and the impact of the Big Five personality traits (openness, extraversion, neuroticism, agreeableness and conscientiousness) on phishing susceptibility was explained.

Cognitive Processes Factor:

Researchers in [18] highlighted the mental operations through which people learn and comprehend. Human cognition, which consists of recognition, thinking, judgment, and memory, is created through these processes. Researchers in [19] proposed a theoretical framework to investigate factors for phishing emails and described how malicious emails impact people susceptibility to phishing.

Knowledge and Experience Factor:

The most significant risk factors for phishing are knowledge and experience, and numerous studies have found that these two factors have significant impact on success or failure of phishing attacks. Computer, network security, network usage, and knowledge related to social engineering make up the majority of the knowledge about phishing. According to the literature [20], people with high level of knowledge are unlikely to fall victim for phishing attacks and are also unlikely to click on phishing emails. Another authors [21] concluded that knowledge and experiences help people in distinguishing legitimate from phishing emails, knowledgeable people responded to fewer than unknowledgeable.

Behaviors Security Factor:

A security behavior scale (SeBIS) was developed by [22] to analyze the intention related to assessing user's security behavior; there are 16 points in the scale which are categories into the following:

Creating passwords

System updates

System updates, and

A proactive awareness

They highlighted the relationship between Self-Beliefs Inventory Short Form (SeBIS) and other psychometric tools in the literature. Domain-Specific Procrastination Scale (DoSpeRT) and SeBIS had a positive correlation, according to their results, whereas General Decision-Making Style (GDMS) and SeBIS had both negative and positive correlations between procrastination and rational decision-making, respectively.

Model Prediction:

A thorough analysis of those individuals to identify the most effective intervention strategies for phishing prevention is conducted. Also, the prediction model is used to identify whether individuals with high phishing susceptibility share any common traits. Finding an accurate method that can reliably identify people who are highly susceptible to phishing is compulsory in this study.

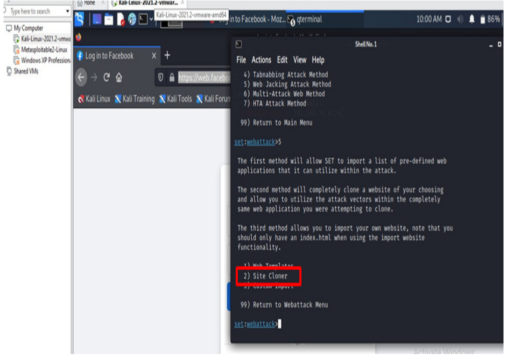
IV. RESULT AND DISCUSSION

This chapter provides in details, the step by step methods followed to analyze the collected phishing data and makes phishing susceptibility prediction using the following supervised machine learning algorithms: Adaboost, XGboost, GBDT, SVM, LR, and RF.

The phishing experiment involved sending two types of emails; legitimate and phishing emails, to a total of 2,121 volunteers.

Using SET Toolskit to Create Phishing Site

The below screenshots show how a social engineering toolskit (SET) in kali linux was used to cloned a legitimate website and used it to test the respondents.



```
  4) Tabnabbing Attack Method
  5) Web Jacking Attack Method
  6) Multi-Attack Web Method
  7) HTA Attack Method

 99) Return to Main Menu

set:webattack>5

 The first method will allow SET to import a list of pre-defined web
 applications that it can utilize within the attack.

 The second method will completely clone a website of your choosing
 and allow you to utilize the attack vectors within the completely
 same web application you were attempting to clone.

 The third method allows you to import your own website, note that you
 should only have an index.html when using the import website
 functionality.

  1) Web Templates
  2) Site Cloner
  3) Custom Import

 99) Return to Webattack Menu

set:webattack>
```

Cloning W

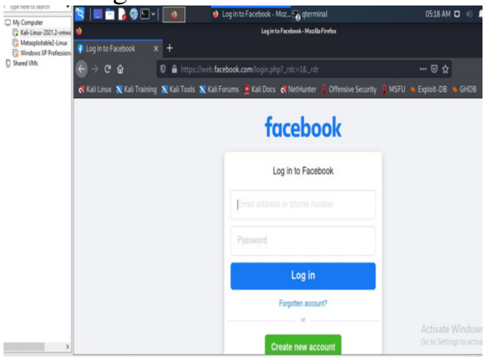
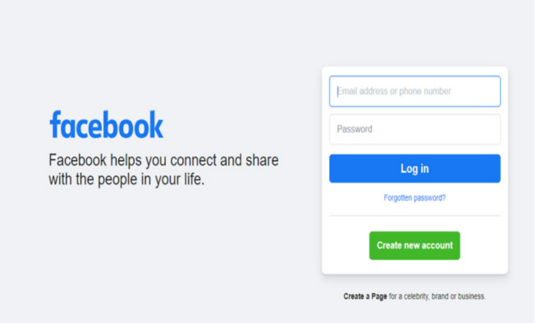


Figure 3: Cloned Facebook Loging Page



facebook

Facebook helps you connect and share with the people in your life.

Email address or phone number
Password
Log in
Forgotten password?
Create new account
Create a Page for a celebrity, brand or business.

Figure 4: Legitimate Facebook Loging Page

Figure 3 and 4 shows how the phishing and the legitimate sites almost look the same and the addresses to these (phishing and legitimate) sites are included in the questionnaire for testing the respondents on whether they can recognized the phishing site.

Frequency Distribution of Collected Phishing Dataset

Table 1 presents a list of individuals who fell victim to the attacks, totaling 1,647 people. This group represents individuals who exhibited a high susceptibility to phishing, with approximately 77.7% of them being successfully phished. This percentage significantly exceeds the average expected rate. To assess participants' ability to distinguish between genuine and fake websites, a questionnaire was designed to captured information on respondents' demographic, personality, cognitive processes, knowledge experiences and their phishing susceptibility.

Table 1: Multiple features frequency distribution

Attribute	Features	Category	Frequency	Percentage
Demographics	Age	<=20	148	7
		21-35	1441	67.9
		36-50	368	17.4
		51-65	148	7
		66++	16	0.8
	Education Level	High School	27	1.3
		Undergraduates	1542	72.7
		Graduates	552	26
	Gender	Female	1016	47.9
		Male	1105	52.1
Personality	Personality	Agreeableness	796	37.5
		Conscientiousness	409	19.3
		Extraversion	132	6.2
		Neuroticism	90	4.2
		Openness	694	32.7
Cognitive Processes	Cognitive Processes	Low	1651	77.8
		Middle	470	22.2
Knowledge Experience	Computer Knowledge	Low	321	15.1
		Middle	1481	69.8
		High	319	15
	Network Security Knowledge	Low	1530	72.1
		Middle	527	24.8
		High	64	3
	Social Engineering Knowledge	Low	1649	77.7
		Middle	409	19.3
		High	63	3
Susceptibility	Phished	Yes	1647	77.7
		No	474	22.3

Demographics:

Age: The majority of respondents fall within the age range of 21-35 (67.9%), followed by 36-50 (17.4%) and <=20 (7%).

Education Level: The majority of respondents are undergraduates (72.7%), followed by graduates (26%) and high school-educated individuals (1.3%).

Gender: The dataset consists of a nearly equal distribution between female (47.9%) and male (52.1%) respondents.

Personality:

Personality Traits: The most common personality trait among the respondents is agreeableness (37.5%), followed by openness (32.7%), conscientiousness (19.3%), extraversion (6.2%), and neuroticism (4.2%).

Cognitive Processes:

Cognitive Processes: The majority of respondents have a low level of cognitive processes (77.8%), while a smaller portion falls into the middle category (22.2%).

Knowledge and Experience:

Computer Knowledge: The majority of respondents have a middle level of computer knowledge (69.8%), followed by low knowledge (15.1%) and high knowledge (15%).

Network Security Knowledge: Most respondents have a low level of network security knowledge (72.1%), while some have a middle level (24.8%) and a few have a high level (3%).

Social Engineering Knowledge: A significant number of respondents have a low level of social engineering knowledge (77.7%), followed by middle knowledge (19.3%) and high knowledge (3%).

Susceptibility:

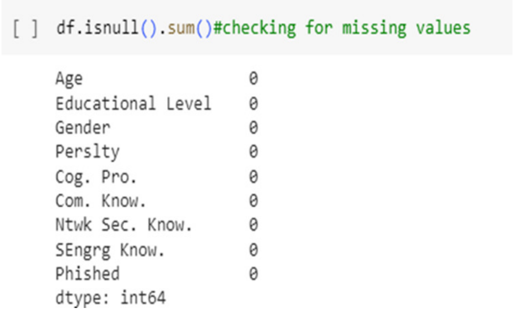
Phished: The dataset includes a large portion of respondents who have been phished (77.7%) compared to those who haven't (22.3%).

Data preprocessing

Data preprocessing is a crucial step in the data analysis pipeline. It involves transforming the raw data into a format suitable for modeling.

Data Cleaning

This stage involves handling missing data as can be seen in the below screenshot



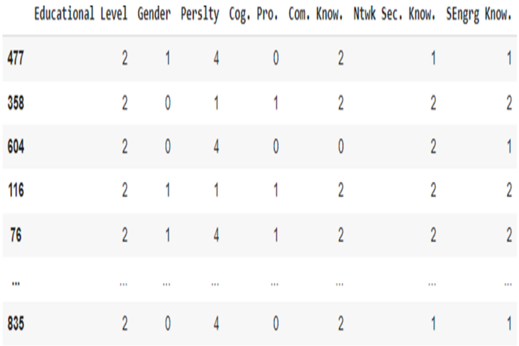
```
[ ]  df.isnull().sum()#checking for missing values

     Age                  0
     Educational Level    0
     Gender               0
     Perslty              0
     Cog. Pro.            0
     Com. Know.           0
     Ntwk Sec. Know.      0
     SEngrg Know.         0
     Phished              0
     dtype: int64
```

Figure 5: Checking for Missing Values

The above image shows how pandas is imported as pd into my colab and df.isnull().sum() function was used to checked for any missing value in the dataset.

Encoding Categorical Variables

Categorical variables cannot be directly used as inputs for many machine learning algorithms. Encoding these variables enables algorithms to understand and process them effectively.



	Educational Level	Gender	Perslty	Cog. Pro.	Com. Know.	Ntwk Sec. Know.	SEngrg Know.
477	2	1	4	0	2	1	1
358	2	0	1	1	2	2	2
604	2	0	4	0	0	2	1
116	2	1	1	1	2	2	2
76	2	1	4	1	2	2	2
...	...	...	...	...	...	...	...
835	2	0	4	0	2	1	1

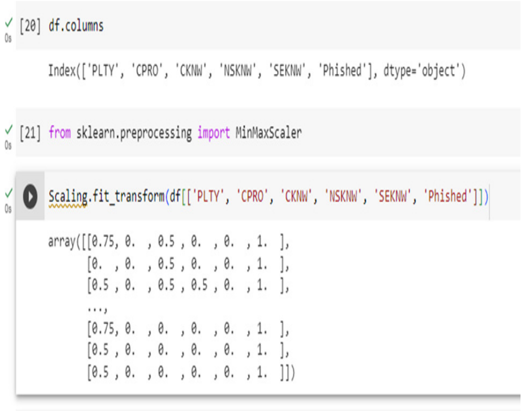
Figure 6: Variables Encoding

Above image shows the new look of my phishing dataset after its encoding to numerical form for my model training, validation and testing.

Variable encoding can help determine the importance of different variables in a predictive model. By encoding categorical variables and analyzing the impact of these encoded variables on the model's performance, one can identify which variables are more influential in making accurate predictions.

Data transformation

Data transformation involves converting the data into a suitable format for analysis. This may include feature scaling, such as standardization and normalization to ensure that all features are on a similar scale, and Feature selection techniques are employed to identify the most relevant and informative features for analysis as shown in the following figure

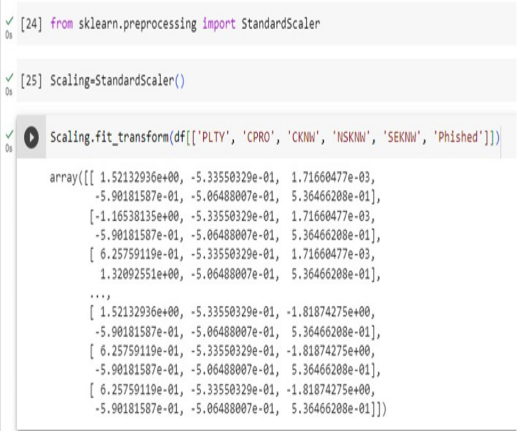


```
✓ [20] df.columns
0s
       Index(['PLTY', 'CPRO', 'CKNW', 'NSKNW', 'SEKNW', 'Phished'], dtype='object')

✓ [21] from sklearn.preprocessing import MinMaxScaler
0s

✓      Scaling.fit_transform(df[['PLTY', 'CPRO', 'CKNW', 'NSKNW', 'SEKNW', 'Phished']])
0s
       array([[0.75, 0.  , 0.5 , 0.  , 0.  , 1.  ],
              [0.  , 0.  , 0.5 , 0.  , 0.  , 1.  ],
              [0.5 , 0.  , 0.5 , 0.5 , 0.  , 1.  ],
              ...,
              [0.75, 0.  , 0.  , 0.  , 0.  , 1.  ],
              [0.5 , 0.  , 0.  , 0.  , 0.  , 1.  ],
              [0.5 , 0.  , 0.  , 0.  , 0.  , 1.  ]])
```

Figure 7: Feature Scaling

Above image shows how to use MinMaxScaler from sklearn.preprocessing to normalize dataset.

Data normalization ensures that all features in the dataset have a similar scale. When features have significantly different scales, it can lead to biased results in many machine learning algorithms. Normalizing the data eliminates this issue by bringing all features to a common scale, allowing algorithms to make fair and accurate comparisons.



```
[24] from sklearn.preprocessing import StandardScaler

[25] Scaling=StandardScaler()

Scaling.fit_transform(df[['PLTY', 'CPRO', 'CKNW', 'NSKNW', 'SEKNW', 'Phished']])

array([[ 1.52132936e+00, -5.33550329e-01,  1.71660477e-03,
        -5.90181587e-01, -5.06488007e-01,  5.36466208e-01],
       [-1.16538135e+00, -5.33550329e-01,  1.71660477e-03,
        -5.90181587e-01, -5.06488007e-01,  5.36466208e-01],
       [ 6.25759119e-01, -5.33550329e-01,  1.71660477e-03,
         1.32092551e+00, -5.06488007e-01,  5.36466208e-01],
       ...,
       [ 1.52132936e+00, -5.33550329e-01, -1.81874275e+00,
        -5.90181587e-01, -5.06488007e-01,  5.36466208e-01],
       [ 6.25759119e-01, -5.33550329e-01, -1.81874275e+00,
        -5.90181587e-01, -5.06488007e-01,  5.36466208e-01],
       [ 6.25759119e-01, -5.33550329e-01, -1.81874275e+00,
        -5.90181587e-01, -5.06488007e-01,  5.36466208e-01]])
```

Figure 8: Data Standardization

Above image shows how StandardScaler from sklearn.processing was used in google colab to standardized the phishing dataset.

Many machine learning algorithms assume that the features are normally distributed and have equal variances. Standardizing the data helps meet these assumptions, which can improve the performance of certain algorithms. Algorithms such as logistic regression and support vector machine that are used in this research can benefit from standardized data as it can help them converge faster and make better predictions.

Handling imbalanced dataset

An imbalanced dataset refers to a dataset in which one or more classes are significantly underrepresented compared to other classes. Below is a pie chart representation of my imbalanced phishing dataset

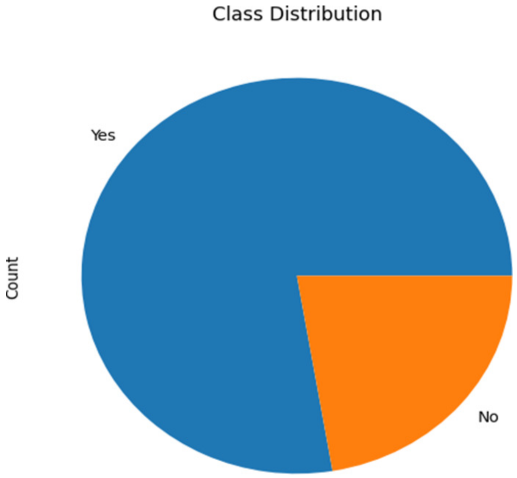
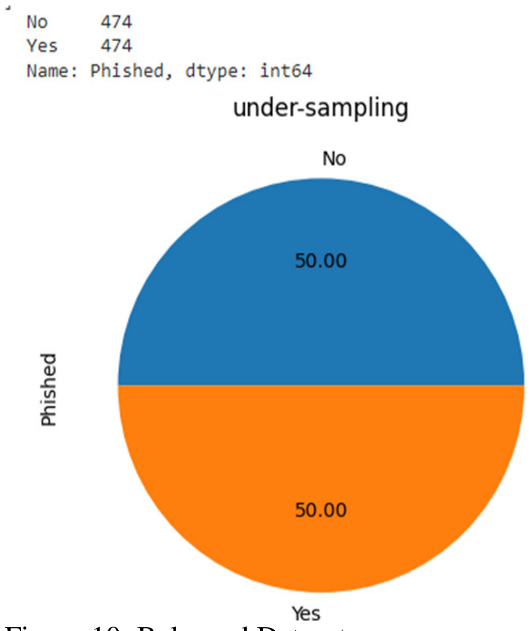


Figure 9: Imbalanced Phishing Dataset

The dataset includes a large portion of respondents who have been phished (77.7%) compared to those who haven't (22.3%). Imbalanced datasets pose challenges in machine learning tasks because models tend to be biased towards the majority class. The classifier may achieve high accuracy by simply predicting the majority class for most instances, while performing poorly on the minority class.

Balancing Overfitting phishing dataset

Addressing the Overfitting challenges posed by imbalanced phishing datasets to ensure fair and accurate model, one of the resampling techniques that reduce overfitting called undersampling was used.



```
No     474
Yes    474
Name: Phished, dtype: int64
```

Figure 10: Balanced Dataset

Imbalanced datasets can cause models to overfit, meaning they memorize the majority class rather than learning the underlying patterns. By balancing the class distribution through undersampling, the model is less likely to overfit and can generalize better to unseen data as shown in the below pie chart.

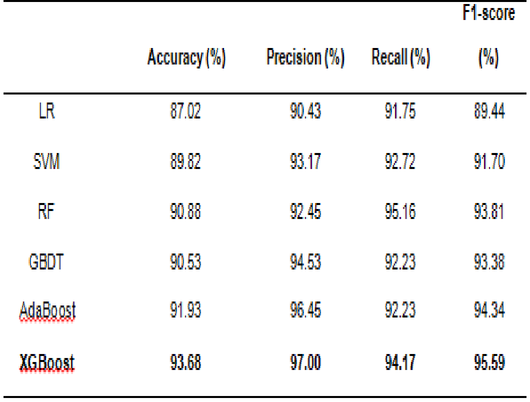
Based on this information in the pie chart, it appears that the phishing dataset is relatively balanced with an equal number of instances for both ―Yes‖ and ―No‖ classes.

This balanced distribution can be advantageous for building our MPSPM models as there is an equal representation of the target variable classes.

Phishing prediction result

Table 2 presents a summary of the binary classification models conducted using various supervised learning algorithms. The models aimed to predict binary outcomes (phished or unphished) based on multiple dimensional features. Performance metrics such as accuracy (ACC), precision, recall, and F1-score were computed for each method.

Table 2 Scores for each learning algorithm

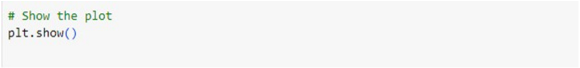


	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
LR	87.02	90.43	91.75	89.44
SVM	89.82	93.17	92.72	91.70
RF	90.88	92.45	95.16	93.81
GBDT	90.53	94.53	92.23	93.38
AdaBoost	91.93	96.45	92.23	94.34
XGBoost	**93.68**	**97.00**	**94.17**	**95.59**

Regarding accuracy (ACC), the XGBoost algorithm demonstrated the highest prediction rate at 93.68%, followed by AdaBoost and RF at 91.93% and 90.88% respectively. The GBDT algorithm achieved an accuracy of 90.53%. On the other hand, the LR and SVM algorithms scored below 90% in ACC and consistently performed poorly across other evaluation metrics.

Consistent with the ACC scores, RF exhibited the highest recall rate at 95.16%, closely followed by XGBoost, SVM, GBDT, AdaBoost, and lastly LR. The remaining metrics exhibited patterns consistent with the mentioned scores, with the most significant differences observed in precision ratings. On average, XGBoost achieved the highest overall score, followed by RF, AdaBoost, and RF.

To visualize the analysis of the algorithm results more effectively, Figure 12 depicts a graph illustrating the performance metrics for the different models. The graph confirms that the GBDT algorithm outperforms others in terms of accuracy (ACC), precision, recall, and F1-score.



```
# Show the plot
plt.show()
```

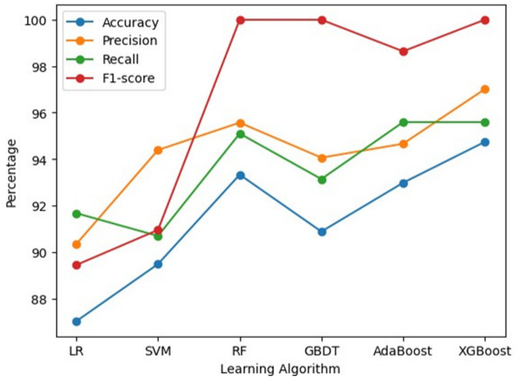


Figure 11: Models Performance Metrics

To evaluate the performance of MPSPM model, stratified cross-validation was used to ensure that each fold of the cross-validation process contains a representative distribution of the target variable's classes with k=10. The dataset was divided into k subsets of equal sizes using stratified sampling. The validation process was repeated k times, with k-1 subsets utilized for training the model and one subset used for testing, with a different subset chosen each time. Using stratified cross-validation, ensures that the MPSPM model is trained and evaluated on representative samples from each class, which helps to mitigate the risk of biased performance evaluation and ensures a more robust assessment of the model's effectiveness across different classes.

RUC Curve

The ROC curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity) at various threshold settings. A perfect classifier would have an ROC curve that passes through the top-left corner of the plot as can be seen in the figure below, indicating the XGboost model high TPR and low FPR for all thresh

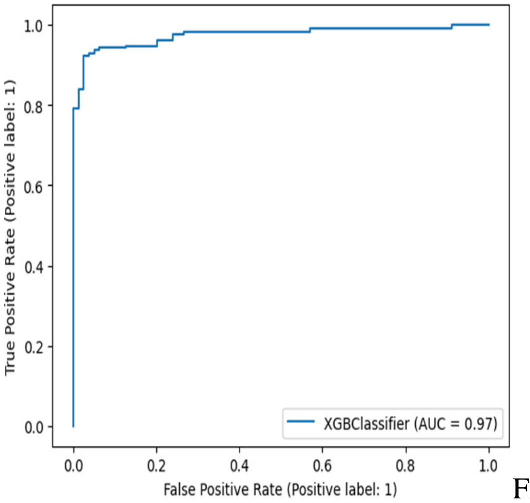


Figure 12: Auc-Roc Curve

Confusion Matrix

The confusion matrix is a valuable tool for evaluating, understanding, and improving the performance of classification models. It provides a comprehensive overview of the model's predictions and enables deeper analysis of its strengths, weaknesses, and biases.

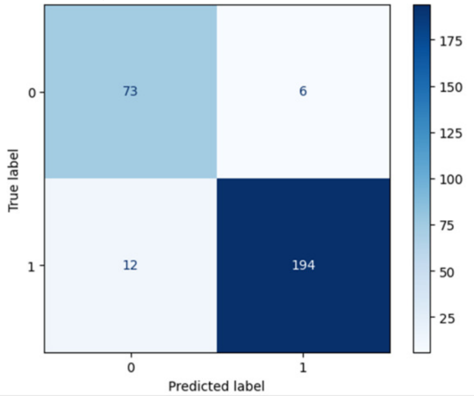
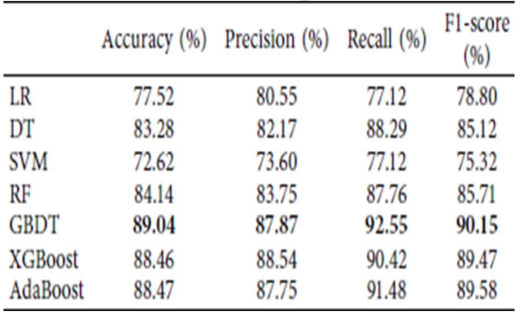


Figure 13: Confusion Matrix

Based on the above confusion matrix, XGBoost model demonstrates high accuracy of (93.68%), precision (97.0%), recall (94.17%), and F1-score (95.59%). These metrics suggest that the XGBoost model performs well in predicting both positive and negative instances, with a relatively low number of false positives and false negatives.

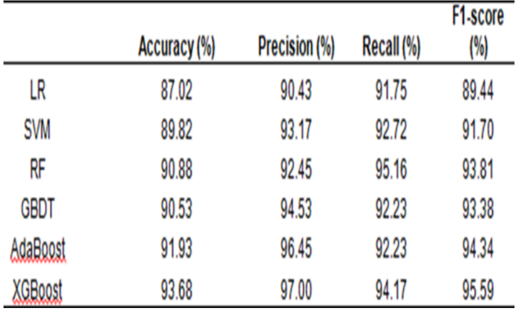
Results comparison

Performance of several machine learning models using the evaluation metrics of accuracy, precision, recall, and F1-score is compared.



	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
LR	77.52	80.55	77.12	78.80
DT	83.28	82.17	88.29	85.12
SVM	72.62	73.60	77.12	75.32
RF	84.14	83.75	87.76	85.71
GBDT	**89.04**	**87.87**	**92.55**	**90.15**
XGBoost	88.46	88.54	90.42	89.47
AdaBoost	88.47	87.75	91.48	89.58

Figure 14: Result From The Existing Work



	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
LR	87.02	90.43	91.75	89.44
SVM	89.82	93.17	92.72	91.70
RF	90.88	92.45	95.16	93.81
GBDT	90.53	94.53	92.23	93.38
AdaBoost	91.93	96.45	92.23	94.34
XGBoost	93.68	97.00	94.17	95.59

Figure 15: New Improved Result

Accuracy: In the result from the existing work, GBDT has the highest accuracy (89.04%), while in the new result, XGBoost has the highest accuracy (93.68%). Overall, the new improved result seems to have higher accuracy values compared to the existing result.

Precision: In both results, XGBoost achieves the highest precision values. However, precision values in the new results (97.00% to 96.45%) are generally higher compared to the existing work (88.54% to 80.55%).

Recall: In both sets, the highest recall is achieved by RF. Recall values in the new result (95.16% to 91.75%) are generally higher compared to the existing work of (92.55% to 77.12%).

F1-score: In both results, XGBoost achieves the highest F1-scores of (95.59%).

In general, the new set of prediction results perform better across all metrics compared to the existing work with XGBoost appears to be the overall best model for phishing susceptibility prediction.

V. CONCLUSIONS

The primary objective of this study was to assess the effectiveness of phishing emails, examine the factors that contribute to phishing vulnerability, and develop a predictive model called MPSPM to identify potential phishing victims. Data were gathered on personal characteristics from 2,121 respondents and conducted phishing tests to evaluate their susceptibility. The collected data was then analyzed using six supervised machine learning techniques.

The results from the study indicate that identifiable models can accurately predict potential phishing victims, with the MPSPM model achieving a correct detection rate of 93.68% on the test set. Furthermore, we investigated the influencing factors behind phishing susceptibility and found that personality, cognitive processes, and network security knowledge play crucial roles. Notably, social engineering knowledge demonstrated a strong correlation with phishing susceptibility.

Additionally, we examined the importance of various factors by analyzing their influence on phishing vulnerability. The analysis revealed that personality, cognitive processes, computer knowledge, password generation behavior, social engineering knowledge, and network security knowledge were among the most influential factors.

The findings further revealed that, people with conscientiousness personality are less likely to fall victims of phishing attack while those with agreeableness personality are the most vulnerable once that can easily fall victims of phishing attack and followed by extraversion people.

Synthetic Minority Over-sampling Technique (SMOTE) or Adaptive Synthetic Sampling (ADASYN) can be used in the future research to balance the phishing dataset.

REFERENCES

[1] [1].Azeez NA, Otudor AE, "Modelling and simulating access control in wireless ad-hoc networks", Fountain J. Natl. Appl. Sci. 5(2), 2016, 18–30.

[2] [2]. Ludl C, McAllister S, Kirda E, Kruegel C, "The effectiveness of techniques to detect phishing sites", In: DIMVA '07 Proceedings of the

4th international conference on Detection of Intrusions and Malware, and Vulnerability; 2007, p. 20–39. doi: 10.1007/978-3-540-73614-1_2.

[3] [3]. Suryavanshi N, Jain A. “A review of various techniques for detection and prevention for phishing attack” Int. J. Adv. Comput. Technol. (IJACT); 4(3), 2015, 41–6.

[4] [4]. Adebowale MA, Lwin KT, Sánchez E, Hossain MA, “Intelligent web- phishing detection and protection scheme using integrated features of Images, frames and text. Expert Syst”, Appl.; 115(2019), 2019, 300–13.

[5] [5]. Khonji M, Iraqi Y, Jones A. “Phishing Detection: A Literature Survey”, IEEE Commun. Surv. Tutor. 15(4): 2013, 2091–121. doi: 10.1109/SURV.2013.032213.00009.

[6] [6]. Biedermann S, Ruppenthal T, Katzenbeisser S., “Data-centric phishing detection based on transparent virtualization technologies”, In: Twelfth Annual Conference on Privacy, Security and Trust (PST); 2014, p. 215–23.

[7] [7]. Rundong Yang, KangfengZheng, Bin Wu, Chunhua Wu, and Xiujuan Wang, “Prediction of Phishing Susceptibility Based on a Combination of Static and Dynamic Features”.

[8] [8]. Edwin Donald Frauenstein, and Stephen Flowerday, “Susceptibility to phishing on social network sites: A personality information processing model”, Computers & Security. Vol. 93. 101862. 2020.

[9] [9]. Kathryn Parsonsa, Marcus Butaviciusb, Paul Delfabbroa, Meredith Lillie, “Predicting susceptibility to social influence in phishing emails”, International journal of human-computer studies. Vol. 128. 2019, p. 17-26

[10] [10].Edwin Donald Frauenstein, and Stephen Flowerday, “Susceptibility to phishing on social network sites: A personality information processing model”, Computers & Security. Vol. 93. 101862. 2020.

[11] [11].A. Apwg, ―Phishing activity trends report: 4rd quarter 2020. Anti-Phishing Working Group, Association for Computing Machinery, San Francisco, Retrieved April, 30, 2020.

[12] [12].Carlo Marcelo Revoredo da Silva, Bruno José Torres Fernandes, Eduardo LuzeiroFeitosa, ViniciusCardosoGarcia, “Piracema.io: A rules-based tree model for phishing prediction”, 2022.

[13] 13]. Rundong Yang, KangfengZheng, Bin Wu, Di Li, Zhe Wang, and Xiujuan Wang, “Predicting User Susceptibility to Phishing Based on Multidimensional” 2022.

[14] [14].J. L. Parrish Jr, J. L. Bailey, and J. F. Courtney, “A Personality Based Model for Determining Susceptibility to Phishing Attacks”, University of Arkansas, Fayetteville, Arkansas, 2009, pp. 285–296.

[15] [15].J. Wang, T. Herath, R. Chen, A. Vishwanath, and H. R. Rao, “Research article phishing susceptibility: an investigation into the processing of a targeted spear phishing email”, from IEEE Transactions on Professional Communications, vol. 55, no. 4, 2012, pp. 345–362.

[16] [16].E. R. Leukfeldt, “Phishing for suitable targets in the Netherlands: routine activity theory and phishing victimization”, Cyberpsychology, Behavior, and Social Networking, vol. 17, no. 8, 2014, pp. 551–555.

[17] [17].G. D. Moody, D. F. Galletta, and B. K. Dunn, “Which phish get caught? an exploratory study of individuals′ susceptibility to phishing”, European Journal of Information Systems, vol. 26, no. 6, 2017, pp. 564–584.

[18] [18].W. R. Flores, H. Holm, N. Marcus, and M. Ekstedt, “Investigating personal determinants of phishing and the effect of national culture, Information & Computer Security, vol. 23, no. 2, 2015.

[19] [19].S. Goel, K. Williams, and E. Dincelli, “Got phished? Internet security and human vulnerability”, Journal of the Association for Information Systems, vol. 18, no. 1, 2017, p. 2.

[20] [20].T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, “Social phishing, Communications of the ACM, vol. 50, no. 10, 2007, pp. 94–100.

[21] [21].R. T. Wright and K. Marett, “The influence of experiential and dispositional factors in phishing: an empirical investigation of the deceived”, Journal of Management Information Systems, vol. 27, no. 1, 2010, pp. 273–303.

[22] [22].X. Luo, W. Zhang, S. Burd, and A. Seazzu, “Investigating phishing victimization with the heuristic-systematic model: a theoretical framework and an exploration”, Computers & Security, vol. 38, 2013, pp. 28–38.

[23] [23].A. Vishwanath, T. Herath, R. Chen, J. Wang, and H. R. Rao, “Why do people get phished? testing individual differences in phishing vulnerability within an integrated, information processing model”, ‖ Decision Support Systems, vol. 51, no. 3, 2011, pp. 576–586.

[24] [24].A. Vishwanath, “Getting phished on social media”, Decision Support Systems, vol. 103, 2017, pp. 70 81.

[25] [25].A. Aleroud, E. Abu-Shanab, A. Al-Aiad, and Y. Alshboul, “An examination of susceptibility to spear phishing cyber-attacks in non-English speaking communities”, Journal of Information Security and Applications, vol. 55, 2020, Article ID 102614.

[26] [26].M. P. Steves, K. K. Greene, and M. F. Theofanos, “A phish scale: rating human phishing message detection difficulty”, in Proceedings of the Workshop on usable security (USEC), New York, NY, USA, 2019.

[27] [27].R. W. Rogers, “A protection motivation theory of fear appeals and attitude change1”, Journal of Psychology, vol. 91, no. 1, 1975, pp. 93–114.

[28] [28].S. Chen and S. Chaiken, “The heuristic-systematic model in its broader context”, Dual-process theories in social psychology, 1999, pp. 73–96.

[29] [29].D. Modic, R. Anderson, and J. Palom¨aki, “We will make you like our research: the development of a susceptibility-topersuasion scale”, PLoS One, vol. 13, no. 3, 2018, Article ID e0194119.

[30] [30].S. Egelman and P. Eyal, “Scaling the security wall: developing a security behavior intentions scale (sebis)”, in Proceedings of the 33rd annual ACM conference on human factors in computing systems, 2015, pp. 2873–2882, Seoul, Republic of Korea.