# IMPROVEMENT OF CLASSIFICATION FEATURES TO INCREASE PHISHING TWEETS DETECTION ACCURACY

[1]SEOW WOOI LIEW, [2]NOR FAZLIDA MOHD SANI, [3]MOHD. TAUFIK ABDULLAH, [4]RAZALI YAAKOB, [5]MOHD YUNUS SHARUM

[1, 2, 3, 4, 5]Department of Computer Science, Faculty of Computer Science and Information Technology,
Universiti Putra Malaysia, Malaysia

E-mail: [1]liewsw28@gmail.com, [2]fazlida@upm.edu.my, [3]taufik@upm.edu.my,
[4]razaliy@upm.edu.my, [5]m_yunus@upm.edu.my

## ABSTRACT

Phishing can be defined as a form of social engineering crime that uses to deceive victims by directing them to the fraudulent websites that appear legitimate which will then collect their personal and sensitive information. Phishing attacks use to target email users traditionally but now, target to Online Social Networks (OSN)s typically Twitter. Therefore, a research study of improving classification features for machine learning technique to classify a dataset collected from Twitter is required. In this study, 3 supervised machine learning techniques - Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Random Forest (RF) and classification features were used to test on a dataset collected from Twitter. The result of our experiment showed that with only 11 selected features, we managed to yield 94.75% classification accuracy higher than 94.56% achieved by other researchers who made use of more than 11 features for the same dataset. From the experiment, we also found that RF remained the best machine learning technique compared to SVM and KNN.

**Keywords:** *Phishing, Online Social Networks (OSN)s, Twitter, Classification Features, Machine Learning Techniques*

## 1. INTRODUCTION

Social engineering is an art of getting users to compromise information systems [1] and a form of information gathering involving human intervention to breach security without victims realize that they have been manipulated. In general, it can be interpreted as a method of launching attacks against information and information systems [2] and regarded as "people hacking" [3]. Social engineering is always overlooked due to low awareness and lack of proper training to people [4]. In 2013, a study conducted by Verizon of security breaches showed that 29% out of the security breaches were contributed by social engineering [5]. Social engineers always take advantage from the existing security breaches or vulnerabilities especially employees' poor training, ineffective segregation of duties and faulty supervision of tasks [6]. Social engineering attacks consist of 4 iterative phases - information gathering, development of relationship, exploitation and execution called "Social Engineering Attack Cycle" [7] as showed in Figure 1.

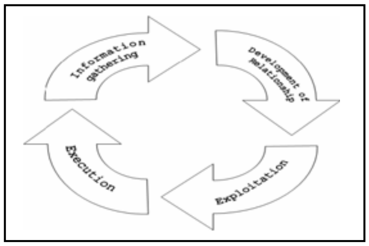

*Figure 1: Social Engineering Attack Cycle*

*Source :* [7]

There are 2 social engineering approaches - Computer (Technology) based and Human (Non-technology) based [8], [9], [10]. Both of these social engineering approaches are exhaustive approaches where the former deceives people via application or system and the latter deceives people via human behaviour weaknesses. Current studies in social engineering reveal that human is always the weakest link to this Human (Non-technology) based social engineering attacks [11], [12]. With regard to whichever social engineering approaches, the main objective of these 2 approaches is to obtain user's personal and sensitive information for further illegal actions.

Phishing is a most significant Computer (Technology) based social engineering attack. Basically, it is a form of social engineering crime so called semantic attack and well known as online identity theft normally uses to deceive victims by directing them to a fraudulent websites appear legitimate [13], [14], [15] which will then collect their personal and sensitive information and an effort for retrieving people critical information [16]. Phishing has become a prevalent problem due to its huge profit margins, ease to be conducted and difficult to be brought those responsible to justice [17] and a serious threat to information security and Internet privacy [18]. According to the Anti-Phishing Working Group [19] - a non-profit organization, 789,068 unique Phishing websites were reported in 2015. Phishing Web pages basically are forged Web pages created by malicious people who mimic the Web pages of real websites [20] and are not isolated which often associated with their targets referred as a "parasitic relationship" [21]. Many thousands of websites each month are compromised by criminals as well as repurposed to host Phishing websites [22]. In summary, Phishing is a most common approaches to obtain personal and sensitive information from victims [23] and a massive problem that getting bigger each month [24].

Phishing attacks traditionally target email which serves as the primary vector [25] but nowadays, they have exposed into popularity of OSNs. In [19] survey, it remarked that OSNs have become significant platforms where Phishers launching the Phishing attacks. Unlike email, Phishers have started using OSNs typically Twitter as a medium to spread Phishing due to its vast information dissemination and difficult to be detected because of its fast spread in the network, short content size and short Uniform Resource Locator (URL) [26], [27]. Twitter basically is an immensely popular micro-blogging network where people post short messages of 140 characters called tweets [26], [27], [28], [29]. [30] pointed out that the relationship of Twitter in term of following and being followed do not require reciprocation unlike other OSNs such as Facebook or Myspace. There are a number of researches adopting machine learning techniques to detect Phishing tweets on Twitter [26], [27], [31] conducting in the past. However, the number of classification features used by them are considered high and can be improved further. As such, a research study of improving classification features for machine learning technique to classify a dataset containing Phishing and safe URLs collected from Twitter is deem required.

The objective of this study is to explore whether there is any possibility to reduce the number of classification features used in order to increase the Phishing tweets detection accuracy. In this context, 94.56% accuracy achieved by [31] using a number of classification features to classify a dataset containing Phishing and safe URLs collected from Twitter was selected as an improvement target to explore whether what they achieved, can be improved further using the reduction of classification features. The second objective is to determine whether RF as claimed by [31] the best machine learning technique remains the best machine learning technique when compared to SVM and KNN in our experiment.

This paper proceeds with the literature study that discussing some related works in the past and methodology that detailing out the entire work activities conducted. Follow by the next section of focusing on findings and discussions with respect to the objectives of the study. The paper is ended with a conclusion.

## 2. LITERATURE STUDY

We carried out the literature study on some related works pertaining to the collection of data, techniques of Phishing detection and machine learning, and analysis of classification features before the experiment is conducted.

### 2.1 Collection of Data

According to [31], they managed to achieve detection accuracy of 94.56% tested on a dataset containing 1573 labelled as Phishing URLs and 1400 labelled as safe URLs collected from Twitter. They determined whether the URL is Phishing or safe URL based on a series of processes started from where the tweet's URL was first checked with PhishTank. If it existed in PhishTank, the particular URL will be labelled as Phishing or safe URL accordingly. Otherwise, they will proceed to check the tweet's URL with Google Safe Browsing and Web of Trust (MyWOT) until it was labelled. Table 1 lists the training data based on URL type used by [31].

*Table 1: Training Data.*

| Type Of URL | No. Of Training |
|---|---|
| Phishing | 1573 |
| Safe | 1400 |
| Total | 2973 |

There is another Phishing tweets detection experiment conducted by [26] earlier, on dataset containing 1473 Phishing URLs and 1500 safe URLs collected from Twitter too. They labelled the URL as Phishing or safe based on 2 blacklists from PhishTank and Google Safe Browsing respectively. Besides, they labelled the URL as Phishing declared by Twitter itself. The detection accuracy achieved by them is 92.52%.

## 2.2 Techniques of Phishing Detection

Phishing is a broad problem where no single effective solution exists to mitigate its vulnerabilities [32]. As such, many technologies and techniques with regard to software-based (blacklists), heuristics, visual similarity and data mining respectively are often studied and implemented to mitigate and detect such Phishing attacks by researchers and commercial companies.

[20] proposed an effective approach to detect Phishing Web pages using Earth Mover's Distance (EMD). Basically, it calculated the visual similarity of two Web pages with corresponding centroid by first converting the Web pages into normalized images followed by representing the image signatures with features composed of dominant color category.

A Phishing filter to determine an incoming email called "Phishwish" proposed by [33] and a heuristic based algorithm relied on a set of Phishing rules or filters to detect and alert users about the Phishing emails called "PhishCatch" proposed by [34] made use of hard-coded rules to detect Phishing.

Another heuristic based Phishing attacks detection technique called "PhishNet" proposed by [35] consists of 2 components - URL prediction that made use of 5 heuristics to discover new Phishing URL and URL matching that made use of the existing blacklist to determine whether a given URL is Phishing.

According to [35], blacklisting is a most common technique to defend against Phishing attacks but it's major problem or challenge is incompleteness. In addition, [36] claimed that blacklisting is inefficient due to Phishing websites are short lifetime. [37] pointed out that blacklists are increasingly ineffective. Thus, they proposed "PhishBlock" - a hybrid of lookup making use of 3 lists - blacklist, whitelist and suspicious list and classifier systems in one simple browser independent user friendly application.

Despite many Phishing attacks detection solutions were proposed as discussed, they have issues respectively such as had a dependency on the tested Web page where it must be similar to the real or legitimate Web page, introduced challenges to their own solution, unable to detect Phishing patterns change as they depended very much on hard-coded rules, increased size of the expanded blacklist from the result of URL prediction and bandwidth demands, cannot justify completely as they focused and examined only on some specific anti-Phishing tools and did not achieve zero-day attacks of Phishing detection.

## 2.3 Techniques of Machine Learning

In view of the issues for the existing proposed solutions as well as unable to detect Phishing attacks on zero-day, researchers are now embarking and focusing on machine learning techniques.

According to [38], Phishing email classification using RF machine learning technique allowed them to yield high classification accuracy. Similar to [26] and [31] who managed to yield high classification accuracy when tested on a dataset containing Phishing and safe URLs collected from Twitter using RF machine learning technique. [27] also managed to achieve high classification accuracy when making use of machine learning technique to detect Phishing on Twitter in real-time. Another experiment conducted by [39] using machine learning technique of KNN tested on a dataset containing Phishing and safe URLs yielded a high classification accuracy. As such, this could be concluded that machine learning technique is an appropriate solution to detect Phishing attacks as well as on zero-day.

Many machine learning techniques are used by researchers for Phishing attacks detection accuracy experiments but the best among them has yet to be discovered, determined and consented as every researcher has their own preferences about the best machine learning technique.

3 supervised machine learning techniques namely SVM, KNN and RF were studied as they have been selected as best machine learning techniques and used by [26], [31], [39], [40] and [41] in their experiments respectively.

### 2.3.1 SVM

SVM is a supervised learning model that associates with learning algorithms use for classification [42]. It is suitable for binary classification [39]. By given a set of training examples for learning, a SVM training algorithm will build a model to predict a new example whether it falls into one category or the other [42].

### 2.3.2 KNN

KNN is a simplest instance-based learning algorithm among other machine learning techniques [43], [44] and a non-parametric method uses for classification [41]. It stores all available cases and classifies new cases based on a similarity measure such as distance and has been used in statistical estimation and pattern recognition in the beginning of 1970's [43], [44]. In KNN, "K" value is used as the number of instances after which the majority class is selected to classify the new instance [39].

### 2.3.3 RF

RF is an ensemble learning classification method uses to handle problems involving data grouping into classes [41], [45]. It is one of the most accurate classifiers which works efficiently for large databases and the most effective method of machine learning algorithms [26], [31]. Basically, RF uses decision trees to predict an output by considering the voted classes from each of the trees where the highest voted class is considered to be the output [45].

## 2.4 Analysis of Classification Features

Selection of classification features is an important activity as it will determine the accuracy level of a machine learning technique. Some related works in this aspect were explored and studied.

In [28], they selected user based and content based features in their machine learning classification. The user based features basically refer to a user's relationships such as follower and followee or user behaviors, and content based features refer to the average length of a tweet, number of URLs, replies or mentions, keywords or wordweight, retweets or tweetlen and hashtags.

According to [31], they made use of 6 sets of features containing URL, tweet, WHOIS, user and network-based data in their machine learning classification experiment. They carried out the classification with 1 feature set at a time and add on another feature set in the next classification activity. From their experiment, they concluded that the performance of Phishing detection significantly improved when more feature sets were added typically the tweet based feature sets for classification.

[46] grouped features into 4 broad categories of lexical based, keyword based, search engine based and reputation based. Each of the categories contains a number of features. In this case, they made use of 138 features in their experiment. Table 2 shows the number of features in each group and some example of features used in the respective groups.

*Table 2: No. Of Features In Each Group And Some Example Of Features Used In The Respective Groups.*

| Feature Group | No. Of Features | Example Of Features |
|---|---|---|
| Lexical based | 24 | Length of URL, Length of host, Length of path, digit in host, etc. |
| Keyword based | 101 | Login, Signin, Confirm, Verify, etc. |
| Search engine based | 6 | Google pagerank, Age of domain, etc. |
| Reputation based | 7 | PhishTank top 10 domain/target in URL, PhishTank top 10 target in URL, IP in PhishTank top 10 IPs, IP in StopBadware top 50 IPs, URL in Phishing blacklist, URL in malware blacklist, etc. |

As for [47], they used 24 lexical based features, 48 WHOIS based features, pagerank feature, Alexa rank feature and PhishTank-based features in their experiment.

i. Lexical based features
- URL properties such as listed in Table 2.

ii. WHOIS based features
- Properties that explained who manages the websites, where the websites are hosted and how the websites are administered.

iii. Pagerank feature
- Technique to determine the number and quality of links to a page so that a learning model can decide how essential the particular website can be. If the page is important, it must have more links from other websites. Pagerank value from Google is robust and updated frequently.

iv. Alexa rank feature
- Ranking set by Alexa to audit the frequency of visits on numerous websites and makes it public for reference. The parameters for the traffic record are based on the number of reaches (the number of Alexa users visiting a particular site in one day) and page views (the number of times a particular URL is viewed by Alexa users).

v. PhishTank-based feature
- PhishTank's statistical reports on Phishing websites recorded every month. The host of the URL, which belongs to the top IP or domain that hosts Phishing websites listed in the reports will be used as reference to determine whether the URL is Phishing.

Another experiment conducted by [39] indicates that they managed to yield high accuracy by using 9 features. Table 3 lists the 9 features used.

*Table 3: 9 Features.*

| No | Feature | Feature Description |
|---|---|---|
| 1 | Long URL | Length of URL |
| 2 | Dots | No. of dots existed in a URL |
| 3 | IP-address | IP address existed in a URL |
| 4 | SSL connection | Https connection existed in a URL |
| 5 | At "@" symbol | Symbol "@" existed in a URL |
| 6 | Hexadecimal | Symbol "%" existed in a URL e.g. "http://%30%31%30%/paypal/cgi=bin/webscrcmd_login.asp" |
| 7 | Frame | Frame existed in a URL |
| 8 | Redirect | Redirect existed in the URL e.g. "www.facebook.com/2/12432;phish.com" |
| 9 | Submit | Submit button existed in the URL and source code |

As compared to other researchers, the 9 features used by [39] for machine learning could be the effective features for Phishing classification that can be explored in our study.

From the literature study, it was noted that machine learning techniques, compared to other discussed Phishing detection techniques are the preferred techniques focused and used to detect Phishing attacks by researchers nowadays. In addition, it was also noted that most of the researchers made use of high number of classification features in their machine learning classification for Phishing tweets detection.

In view of this, classification features improvement in term of reducing the number of classification features used to increase the accuracy of Phishing tweets detection adopting machine learning technique is initiated in this study.

## 3. METHODOLOGY

The entire process of this study was divided into 3 main aspects. They are selection of dataset, selection of machine learning techniques and selection of classification features. Figure 2 shows the entire components used in our experiment.

i. Selection of dataset
In this study, a supervised machine learning approach was adopted to train a model. The dataset containing 1573 labelled as Phishing URLs and 1400 labelled as safe URLs collected from Twitter [31] was selected and used.

ii. Selection of machine learning techniques
3 supervised machine learning techniques namely SVM, KNN and RF were selected and used in the experiment. We selected these 3 machine learning techniques based on the literature study conducted on the articles from [26], [31], [39], [40] and [41] respectively.

iii. Selection of classification features
Classification features used by [28], [31], [39], [46] and [47] were selected as the basis of features to be explored for determining the best features in the experiment.

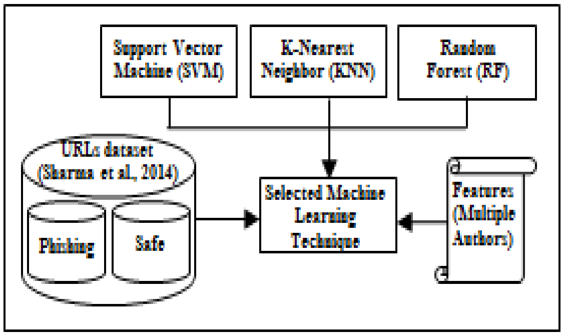
Support Vector Machine (SVM)
K-Nearest Neighbor (KNN)
Random Forest (RF)
URLs dataset (Sharma et al, 2014)
Phishing
Safe
Selected Machine Learning Technique
Features (Multiple Authors)

*Figure 2: Entire Components Used*

In this study, we adopted Standard Information Retrieval Metrics viz. Accuracy, Precision and Recall to evaluate the effectiveness of the classification and a Confusion Matrix as Table 4 to explain further on the experiment.

*Table 4: Confusion Matrix.*

| | | PREDICTED | |
|---|---|---|---|
| | | Phishing | Safe |
| ACTUAL | Phishing | TP | FN |
| | Safe | FP | TN |

Where TP - True Positive,
FP - False Positive,
TN - True Negative,
FN - False Negative

Accuracy = (TP+TN) / (TP+FP+TN+FN)
Precision (Phishing) = TP / (TP+FP)
Recall (Phishing) = TP / (TP+FN)
Precision (Safe) = TN / (TN+FN)
Recall (Safe) = TN / (TN+FP)

We selected Weka tool to assess the 3 machine learning techniques - SVM, KNN and RF, and used cross validation with 10 folds test mode in the experiment as the amount of data available was limited to 2973. The purpose of having this test mode is to avoid the possible bias on any particular division into train and test components. All the available data were used to train as well as to compare on the test set in a particular division respectively in a rotation manner. In summary, all data in the dataset were used for both training and testing.

The result of the classification from the selected machine learning technique was saved in a model and used for subsequent testing. 2 new testing datasets containing Phishing and safe URLs of 1500 and 3000 respectively were used for prediction testing using the earlier saved model. These 2 datasets were extracted from a URLs dataset where its data were collected from the source of PhishTank, Google Search with McAFee WebAdvisor, Google Safe Browsing and MyWOT. Figure 3 shows the process flow of URLs dataset compilation.

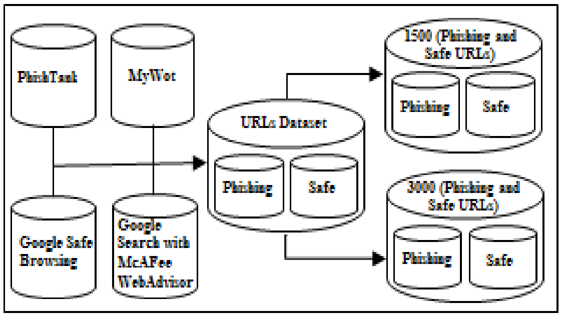
PhishTank
MyWot
Google Safe Browsing
Google Search with McAFee WebAdvisor
URLs Dataset
Phishing
Safe
1500 (Phishing and Safe URLs)
Phishing
Safe
3000 (Phishing and Safe URLs)
Phishing
Safe

*Figure 3: Process Flow Of URLs Dataset Compilation*

## 4. FINDINGS AND DISCUSSIONS

From the experiment, 11 features as listed in Table 5 were identified as the best classification features. 7 out of 11 features were suggested by Weka tool under Attribute Evaluator as “CfsSubsetEval” and Search Method as “BestFirst”. The remaining 4 were selected from the pool of 22 features gathered from the set used by [28], [31], [39], [46] and [47] based on the evaluation testing according to the typical feature selection method [48], [49], [50].

*Table 5: 11 Best Classification Features.*

| No | Feature | Suggested | Description |
|---|---|---|---|
| 1 | URL length (F1) | Week tool | Length of URL |
| 2 | SSL connection (F2) | Weka tool | Https connection existed in a URL |
| 3 | Hexadecimal (F3) | Weka tool | Symbol “%” existed in a URL e.g. “http://%30 %31%30%/paypal/cgi=bin/webscrcmd _login.asp” |
| 4 | Alexa rank (F4) | Weka tool | A metric that ranks websites in order of popularity or how well a |

| | | | |
|---|---|---|---|
| | | | website is doing over the last 3 months |
| 5 | Age of domain - Year (F5) | Weka tool | Length of time a website has been registered and active |
| 6 | Equal (F6) | Evaluation testing | Symbol "=" existed in a URL e.g. "http://%30%31%30%/paypal/cgi=bin/webscrcmd_login.asp" |
| 7 | Digit in host (F7) | Weka tool | Digits existed in Host e.g. www.3sports.com |
| 8 | Host length (F8) | Evaluation testing | Length of host |
| 9 | Path length (F9) | Evaluation testing | Length of path |
| 10 | Registrar (F10) | Weka tool | Registrar existed in WHOIS |
| 11 | No of Dots in Host (F11) | Evaluation testing | No of dots existed in host |

RF was determined as a best machine learning technique with the 11 best classification features compared to SVM (79.95%) and KNN (90.48%) because it allowed us to achieve 94.75% and also higher than the classification accuracy of 94.56% achieved by [31]. In other words, RF is the selected machine learning technique. Table 6 shows the classification result of accuracy achieved based on features for SVM, KNN and RF respectively and Figure 4 shows the accuracy achieved for the 3 machine learning techniques in more detail.

*Table 6: Classification Result Of Accuracy Achieved Based On Features For SVM, KNN And RF.*

| No | Features | No. Of Features | SVM (%) | KNN (%) | RF (%) |
|---|---|---|---|---|---|
| 1 | F1+F2+F3+F4+F5+F7+F10 | 7 | 82.71 | 89.67 | 92.70 |
| 2 | F1+F2+F3+F4+F5+F7+F8+F10 | 8 | 80.63 | 90.01 | 94.15 |
| 3 | F1+F2+F3+F4+F5+F7+F8+F9+F10 | 9 | 79.52 | 90.45 | 94.28 |
| 4 | F1+F2+F3+F4+F5+F7+F8+F9+F10+F11 | 10 | 79.72 | 90.68 | 94.55 |
| 5 | F1+F2+F3+F4+F5+F6+F7+F8+F9+F10+F11 | 11 | 79.95 | 90.48 | 94.75 |

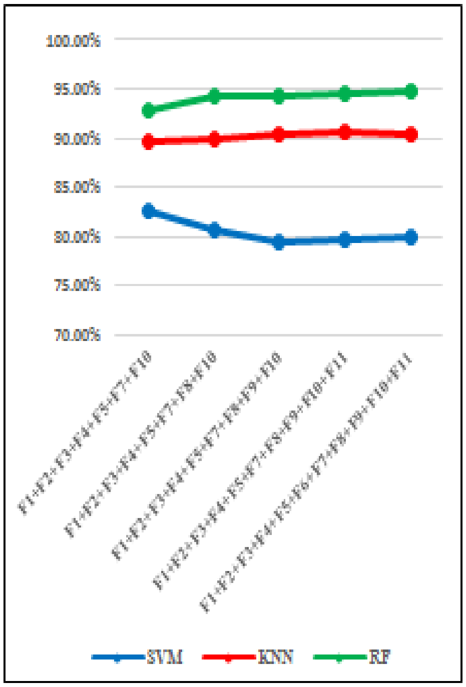

*Figure 4: SVM, KNN And RF Achieved Classification Accuracy Results By Features*

Machine learning technique of RF managed to predict 1502 Phishing URLs as Phishing URLs correctly. Similarly, 1315 safe URLs were predicted as safe URLs correctly. As such, this contributed to an accuracy of 94.75% achieved by RF for the entire classification process. Table 7 shows the Confusion Matrix of RF.

*Table 7: RF Confusion Matrix.*

| | | PREDICTED | |
|---|---|---|---|
| | | Phishing | Safe |
| ACTUAL | Phishing | 1502 | 71 |
| | Safe | 85 | 1315 |

In term of precision for Phishing and safe, RF achieved 94.64% and 94.88% respectively. As of the recall for Phishing and safe, they were 95.49% and 93.93% achieved respectively. Table 8 shows the RF precision and recall for Phishing and safe in more details.

*Table 8: RF Precision And Recall For Phishing And Safe.*

| No | Description | Achieved (%) |
|---|---|---|
| 1 | Precision (Phishing) | 94.64 |
| 2 | Precision (Safe) | 94.88 |
| 3 | Recall (Phishing) | 95.49 |
| 4 | Recall (Safe) | 93.93 |

The accuracy achieved using the model derived from RF and the 11 best classification features, tested on the 2 new testing datasets containing 1500 and 3000 URLs were 94.13% and 94.30%. Table 9 summarises the accuracy achieved for the training dataset and the 2 new testing datasets and Figure 5 shows the accuracy achieved comparison in more detail.

*Table 9: The Accuracy Achieved For Training And New Testing Datasets.*

| Dataset | No. Of URLs | Accuracy Achieved (%) |
|---|---|---|
| Training | 2973 | 94.75 |
| Test dataset 1 | 1500 | 94.13 |
| Test dataset 2 | 3000 | 94.30 |

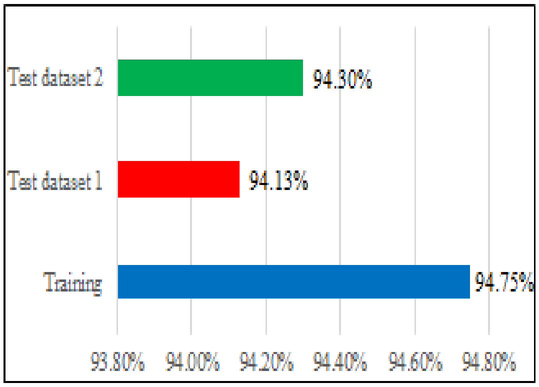

*Figure 5: Achieved Accuracy Comparison*

## 5. DIFFERENT FROM PREVIOUS WORK

Unlike the works conducted by other researchers who focusing on Phishing detection using various different detection techniques, our main study focus is to improving the classification features in order to increase the accuracy of Phishing tweets detection using machine learning technique.

From the detail analysis, the significant pro of our study is we managed to reduce the number of classification features used in Phishing tweets detection to increase its accuracy higher than what achieved by other researchers [26], [27], [31] in the past. As for the con, we are unable to achieve 100% accuracy of Phishing tweets detection using such reduced number of classification features as there are still false positive and negative observed in the detection.

## 6. CONCLUSION

Based on the experiment, it shows significantly that just having the 11 best classification features of URL length, SSL connection, hexadecimal, Alexa rank, age of domain, equal, digit in host, host length, path length, registrar and no. of dots in host for classifying a dataset containing Phishing and safe URLs collected from Twitter, the machine learning technique of RF is able to yield accuracy of 94.75% higher than the accuracy achieved by [31].

Besides, it can be concluded that the model derived from RF and the 11 best classification features is justifiable and practical enough to be an acceptable model because the accuracy achieved for these 2 new testing datasets are almost closer to the accuracy achieved using the training dataset.

The experiment also shows that RF remained the best machine learning technique that yielded higher accuracy compared to SVM and KNN.

There are, however, limitations in this study where we are unable to use Google pagerank feature via toolbar automatically because it has been shuttered and no longer available, and tweet based features because the dataset containing Phishing and safe URLs used is from [31] who already collected them from Twitter directly, which may help to reduce further the number of classification features explored in our experiment.

## ACKNOWLEDGEMENT:

Our special thanks go to Ms. Smriti Maheshwari for her assistance on helping us to using her dataset containing 1573 Phishing and 1400 safe URLs collected from Twitter in this study. In additon, we would like to thank the editor and the anonymous reviewer(s) for the useful comments and suggestions on this paper. This work supported by Faculty of Computer Science and Information Technology, Universiti Putra Malaysia.

## REFERENCES:

[1] K. Krombholz, H. Hobel, M. Huber, and E. Weippl, "Advanced Social Engineering Attacks," *J. Inf. Secur. Appl.*, vol. 22, pp. 113–122, 2015.

[2] L. J. Janczewski and L. (Rene) Fu, "Social Engineering-Based Attacks: Model And New Zealand Perspective," *Proc. Int. Multiconference Comput. Sci. Inf. Technol.*, pp. 847–853, 2010.

[3] H. Hasle, Y. Kristiansen, K. Kintel, and E. Snekkenes, "Measuring Resistance To Social Engineering," *Inf. Secur. Pract. Exp.*, pp. 132–143, 2005.

[4] A. Chitrey, D. Singh, M. Bag, and V. Singh, "A Comprehensive Study Of Social Engineering Based Attacks In India To Develop A Conceptual Model," *Int. J. Inf. Netw. Secur.*, vol. 1, no. 2, pp. 45–53, 2012.

[5] R. Bhakta and I. G. Harris, "Semantic Analysis Of Dialogs To Detect Social Engineering Attacks," *Proc. 2015 IEEE 9th Int. Conf. Semant. Comput. IEEE ICSC 2015*, pp. 424–427, 2015.

[6] J. J. Gonzalez, J. M. Sarriegi, and A. Gurrutxaga, "A Framework For Conceptualizing Social Engineering Attacks," *CRITIS*, pp. 79–90, 2006.

[7] S. Heikkinen, "Social Engineering In The World Of Emerging Communication Technologies," *Proc. Wirel. World Res. Forum*, pp. 1–10, 2006.

[8] R. Gulati, "The Threat Of Social Engineering And Your Defense Against It," *Inf. Secur.*, pp. 1–15, 2003.

[9] P. S. Maan and M. Sharma, "Social Engineering: A Partial Technical Attack," *IJCSI Int. J. Comput. Sci.*, vol. 9, no. 2, pp. 557–559, 2012.

[10] F. L. Greitzer, J. R. Strozer, S. Cohen, A. P. Moore, D. Mundie, and J. Cowley, "Analysis Of Unintentional Insider Threats Deriving From Social Engineering Exploits," *Proc. - IEEE Symp. Secur. Priv.*, pp. 236–250, 2014.

[11] E. Nyamsuren and H.-J. Choi, "Preventing Social Engineering In Ubiquitous Environment," *Futur. Gener. Commun. Netw. (FGCN 2007)*, pp. 573–577, 2007.

[12] T. Mataracioglu and S. Ozkan, "User Awareness Measurement Through Social Engineering," *ArXiv e-prints*, pp. 1–7, 2011.

[13] N. A. G. Arachchilage and S. Love, "A Game Design Framework For Avoiding Phishing Attacks," *Comput. Human Behav.*, vol. 29, no. 3, pp. 706–714, 2013.

[14] N. A. G. Arachchilage and S. Love, "Security Awareness Of Computer Users: A Phishing Threat Avoidance Perspective," *Comput. Human Behav.*, vol. 38, pp. 304–312, 2014.

[15] N. A. G. Arachchilage, S. Love, and K. Beznosov, "Phishing Threat Avoidance Behaviour: An Empirical Investigation," *Comput. Human Behav.*, vol. 60, pp. 185–197, 2016.

[16] M. Dadkhah, T. Sutikno, M. D. Jazi, and D. Stiawan, "An Introduction To Journal Phishings And Their Detection Approach," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 13, no. 2, pp. 373–380, 2015.

[17] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, "Social Phishing," *Commun. ACM*, vol. 50, no. 10, pp. 94–100, 2007.

[18] K.-T. Chen, C.-R. Huang, C.-S. Chen, and J.-Y. Chen, "Fighting Phishing With Discriminative Keypoint Features," *IEEE Internet Comput.*, vol. 13, no. 3, pp. 56–63, 2009.

[19] "Anti-Phishing Working Group (APWG)." [Online]. Available: http://www.antiphishing.org/. [Accessed: 01-May-2016].

[20] A. Y. Fu, L. Wenyin, and X. Deng, "Detecting Phishing Web Pages With Visual Similarity Assessment Based On Earth Mover's Distance (EMD)," *IEEE Trans. Dependable Secur. Comput.*, vol. 3, no. 4, pp. 301–311, 2006.

[21] L. Wenyin, G. Liu, B. Qiu, and X. Quan,

"Antiphishing Through Phishing Target Discovery," *IEEE Internet Comput.*, vol. 16, no. 2, pp. 52–60, 2012.

[22] M. Vasek, J. Wadleigh, and T. Moore, "Hacking Is Not Random: A Case-Control Study Of Webserver Compromise Risk," *IEEE Trans. Dependable Secur. Comput.*, vol. 13, no. 2, pp. 206--219, 2015.

[23] N. M. Shekokar, C. Shah, M. Mahajan, and S. Rachh, "An Ideal Approach For Detection And Prevention Of Phishing Attacks," *Procedia Comput. Sci.*, vol. 49, no. 1, pp. 82–91, 2015.

[24] F. Toolan and J. Carthy, "Phishing Detection Using Classifier Ensembles," *2009 eCrime Res. Summit, eCRIME '09*, 2009.

[25] H. Wilcox and M. Bhattacharya, "Countering Social Engineering Through Social Media: An Enterprise Security Perspective," *ICCCI*, pp. 54–64, 2015.

[26] A. Aggarwal, A. Rajadesingan, and P. Kumaraguru, "PhishAri: Automatic Realtime Phishing Detection On Twitter," *eCrime Res. Summit, eCrime*, pp. 1–12, 2012.

[27] M. C.Nair and S.Prema, "A Distributed System For Detecting Phishing In Twitter Stream," *Int. J. Eng. Sci. Innoavtive Technol.*, vol. 3, no. 2, pp. 151–158, 2014.

[28] M. McCord and M. Chuah, "Spam Detection On Twitter Using Traditional Classifiers," *ATC*, pp. 175–186, 2011.

[29] S. Lee and J. Kim, "Warningbird: A Near Real-Time Detection System For Suspicious URLs In Twitter Stream," *IEEE Trans. Dependable Secur. Comput.*, vol. 10, no. 3, pp. 183–195, 2013.

[30] H. Kwak, C. Lee, H. Park, and S. Moon, "What Is Twitter, A Social Network Or A News Media?," *Int. World Wide Web Conf. Comm.*, pp. 1–10, 2010.

[31] N. Sharma, N. Sharma, V. Tiwari, S. Chahar, and S. Maheshwari, "Real-Time Detection Of Phishing Tweets," *Fourth Int. Conf. Comput. Sci. Eng. Appl.*, pp. 215–227, 2014.

[32] M. Khonji, Y. Iraqi, and A. Jones, "Phishing Detection: A Literature Survey," *IEEE Commun. Surv. Tutorials*, vol. 15, no. 4, pp. 2091–2121, 2013.

[33] D. L. Cook, V. K. Gurbani, and M. Daniluk, "Phishwish: A Stateless Phishing Filter Using Minimal Rules," *FC*, pp. 182–186, 2008.

[34] W. D. Yu, S. Nargundkar, and N. Tiruthani, "PhishCatch - A Phishing Detection Tool," *2009 33rd Annu. IEEE Int. Comput. Softw. Appl. Conf.*, vol. 2, pp. 451–456, 2009.

[35] P. Prakash, M. Kumar, R. R. Kompella, and M. Gupta, "PhishNet: Predictive Blacklisting To Detect Phishing Attacks," *Proc. - IEEE INFOCOM*, 2010.

[36] S. Marchal, J. Francois, R. State, and T. Engel, "PhishStorm: Detecting Phishing With Streaming Analytics," *IEEE Trans. Netw. Serv. Manag.*, vol. 11, no. 4, pp. 458–471, 2014.

[37] H. M.A.Fahmy and S. A.Ghoneim, "PhishBlock: A Hybrid Anti-Phishing Tool," *2011 Int. Conf. Commun. Comput. Control Appl. CCCA 2011*, pp. 2–6, 2011.

[38] A. A. Akinyelu and A. O. Adewumi, "Classification Of Phishing Email Using Random Forest Machine Learning Technique," *J. Appl. Math.*, vol. 2014, 2014.

[39] O. A. Akanbi, I. S. Amiri, and E. Fazeldehkordi, *A Machine-Learning Approach To Phishing Detection And Defense*. 2015.

[40] M. Olalere, M. T. Abdullah, R. Mahmod, and A. Abdullah, "Identification And Evaluation Of Discriminative Lexical Features Of Malware URL For Real-Time Classification," *Proc. - 6th Int. Conf. Comput. Commun. Eng. Innov. Technol. to Serve Humanit. ICCCE 2016*, pp. 90–95, 2016.

[41] M. A. Al-Garadi, K. D. Varathan, and S. D. Ravana, "Cybercrime Detection In Online Communications: The Experimental Case Of Cyberbullying Detection In The Twitter Network," *Comput. Human Behav.*, vol. 63, pp. 433–443, 2016.

[42] "Support Vector Machine (SVM)." [Online]. Available: https://en.wikipedia.org/wiki/Support_vector_machine. [Accessed: 15-May-2016].

[43] "K-Nearest Neighbor (KNN)." [Online]. Available: https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm. [Accessed: 15-May-2016].

[44] "K-Nearest Neighbor (KNN) - 2." [Online]. Available: http://www.fon.hum.uva.nl/praat/manual/k

NN_classifiers_1__What_is_a_kNN_classifier_.html. [Accessed: 15-May-2016].

[45] "Random Forest (RF)." [Online]. Available: https://en.wikipedia.org/wiki/Random_forest. [Accessed: 15-May-2016].

[46] R. B. Basnet, A. H. Sung, and Q. Liu, "Learning To Detect Phishing URLs," *IJRET Int. J. Res. Eng. Technol.*, vol. 3, no. 6, pp. 11–24, 2014.

[47] B. E. Sananse and T. K. Sarode, "Phishing URL Detection: A Machine Learning And Web Mining-Based Approach," *Int. J. Comput. Appl.*, vol. 123, no. 13, pp. 46–50, 2015.

[48] M. Dash and H. Liu, "Feature Selection For Classification," *Intell. Data Anal.*, vol. 1, no. 3, pp. 131–156, 1997.

[49] H. Liu and L. Yu, "Toward Integrating Feature Selection Algorithms For Classification and Clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 4, pp. 491–502, 2005.

[50] J. Tang, S. Alelyani, and H. Liu, "Feature Selection for Classification : A Review," *Data Classif. Algorithms Appl.*, pp. 37–64, 2014.