Phishing Classification Models: Issues and Perspectives

Hiba Zuhair
College of Information Engineering
Al-Nahrain University
Baghdad, Iraq
hiba.zuhair.pcs2013@gmail.com,
hiba.zuhir@coie-nahrain.edu.iq

Ali Selamat
Center of Information and Communication Technologies,
Software Engineering Department, Faculty of Computing
Universiti Teknologi Malaysia (UTM),
Johor, Malaysia
aselamat@utm.my

***Abstract*— Never-ending phishing threats on cyberspace motivate researchers to develop more proficient phishing classification models to survive a supreme cyber-security with safe web services. However, such achievements remain incompetent in their performance against novel phish attacks. This is attributed to the induction factors of the classification model itself such as hybrid feature space, inactive learning on up-to-date data flows, and limited adaptation to the evolving phish attacks. In this light, this paper surveys the current achievements, studies their limitations, restates what induction factors need to boost for a successful real-time application. Consequently, future outlooks are recommended on how to devote well-performed anti-phishing scheme.**

Keywords: novel phish; feature-base classifier; adaptive detection; active learning.

I. INTRODUCTION

Motivating by the more illegitimate gains, phishers targeting users' credentials and industries reputation on the cyberspace. They deploy social engineering technology to impersonate trustworthy websites with spoofed links for users misleading. Victim users catch the bait, submit their own credentials via spoofed links, and then phishers acquire their credentials for theft and illegal profits [1 and 2]. Day after day, the swift increasing and rapid advancement of phishing activities threaten cyber-security and economy [2]. To mitigate them, many efforts have been made by researchers in academia and industry to achieve effective anti-phishing [1-4]. Most anti-phishing schemes, for example machine learning-based phishing classification models; adopt client side filtering that inspect phish websites and warn users online [1, 4]. Machine learning-based anti-phishing schemes rely on feature vectors and feature-base classifiers (FBC). FBCs are designated with typical machine learning algorithms to assert accurate classification with least faults in practice. However, FBCs vary in their classification outcomes and performance specifically versus continually evolved phish websites [5-8]. This becomes an intricate challenge to researchers to detect phishing on big and strongly inter-related web data. Mainly, this is attributed to the deficiency of prominent induction factors to leverage training and testing tasks [5-8]. In this light, this study revisits the current machine learning-based phishing classification models and it highlights the causality between their critical issues and their lack to induction factors. Accordingly, perspectives for future work are introduced towards proficient phishing classification.

To point out the aforesaid, this paper as follows: Section II introduces machine learning algorithms and feature-base classifiers. Whereas, Section III surveys and appraises the related literatures critically. Section IV restates what induction factors to boost and what perspectives to contribute in the future work. Finally, Section V drawn several concluding remarks.

II. MACHINE LEARNING ALGORITHMS

In the anti-phishing domain, many typical machine learning algorithms were applied either in single feature-base classifier (FBC) or in an ensemble features-base classifier (EFBC); an assembly classifier integrates several machine learning algorithms of different induction settings [6-8] Commonly used machine learning algorithms include Naïve Bayes (NB), Logistic Regression (LR), Sequential Minimal Optimization (SMO), Support Vector Machine (SVM), and Transductive Support Vector Machine (TSVM), etc. Table I in Appendix, presents them briefly.

Typically, FBC maps the input feature vector to the output classes by attributing the input feature vector $V = (v_1 \ldots v_n)$ and inducts its relevance to either phish or not phish classes with $Y = f(V, \gamma)$. All input feature vectors that extracted from the m-dimensional training dataset ($V_1, V_2, \ldots, V_m$) are induced in the training phase to classify the incoming instance V_{new} in the testing phase into either phish or legitimate label [9-11]. Whereas, EFBC learn the training dataset by its constituent machine learning algorithms that may vary in their inspecting features and induction settings [12, 13]. In practice, EFBC outperform FBC because the final judgment is obtained from the average of all its constituent

978-1-5386-0731-2/17/$31.00 ©2017 IEEE

algorithms' predictions [12, 13] However, existing anti-phishing schemes that assisted by either FBCs or EFBCs varied in their performance due to their limitations and divergence in induction settings [4, 12, 13].

III. Machine Learning-Based Anti-Phishing

Over the last few years, the most salient anti-phishing schemes are those machine learning-based [2-4]. Among them, was a phishing classification model that developed in [14] that attained 12 identities and textual features along with Support Vector Machine (SVM). Throughout the classification, the developed model predicted phishness on redirecting web page, login forms, e-business, and English hosting web pages with performance of 97.33% and 1.45% as True Positive and False Positive rates respectively. Because it deployed textual features and textual information retrieval method which were wholly dedicated for English language, it leveraged phish language exploits partially.

Also, [15] at Carnegie Mellon upgraded a former anti-phishing scheme (CANTINA) to a hybrid feature-based scheme CANTINA$^+$. The latter version was developed as an ensemble feature-base classifier including Naïve Bayes (NB), Support Vector Machine (SVM), and Logic Regression (LR) etc. Around 15 textual and structural features were derived from web page URL and web page contents as well as some online features were devoted to accurately classify phish exploits (92% True Positive Rate and 1.4% False Positive Rate) on redirecting web page, login form handler, and web pages hosting in English. However, CANTINA$^+$ encountered a trade-off in leveraging up-to date phish webpages due to the use of limited feature space to English textual features as well as re-learning on defaults settings.

Likewise, a phishing classification model proposed by researchers in [16] relied on six ordinary features including visual and DOM features along with Semi-Transductive Support Vector Machine (TSVM) algorithm. It could capture frequently evolved phish websites on a high-dimensional training and testing data set. It achieved (96.4 %) of the accuracy rate and (3.5%) mistake rates respectively. However, it lacks to adapt novel phishes and up-to-date data sets due to limited feature space.

Later, the authors in [17-18] leveraged 17 features to examine login form phish webpages via a developed classification model by using Support Vector Machine (SVM) classifier. Their model achieved a rationale performance with (99.6%) of True Positive Rate and (0.44%) of False Positive Rate. However, it was computationally intensive and time-consuming due to the use of external resources and less adaptive to present training data sets.

Meanwhile, these researchers [19-20] learnt 12 URL features on an EFBC with Support Vector Machine (SVM), Random Forest (RF), C4.5, and JRip algorithms. Their EFBC achieved (94.91%) and (1.44%) as classification accuracy and faults. In spite of using big training and testing data sets, the used data set was imbalanced in classes and it included e-Commerce websites exclusively.

Oppositely, a phishing classification model was devoted in [21] to catch phishing in e-commerce, login form, and English and French webpages by using 17 ordinary various features and Neural Network (NN) classifier. Even though, achievements yielded up to 94.07% accuracy rates, high misclassification rates were reported. The model scarcely detected novel phish websites due to its inactive learning on imbalanced training data set.

On the other hand, the authors in [22] identified phishing on Chinese e-business websites via phishing Chinese website detection model. They selected 15 language independent features exclusively to identify Chinese websites. Four machine learning algorithms including Sequential Minimum Optimization (SMO), Logic Regression (LR), Naïve Bayes (NB), and Random Forests (RF) were applied individually in an FBC. Their model performed (95.83%) accuracy rate on Chinese e-business websites solely. Thus, it was not reliable for generic phish websites classification due to its exclusive features and data sets.

Unlikely, the researcher in [23] optimized a former version of phishing classification model with an EFBC. Optimized EFBC attained with Support Vector Machine (SVM), Decision Tree (C4.5), and Random Forests (RF). It actively learned EFBC with 212 different features to perform effective classification in the testing task. However, attaining typical features on large and imbalanced data sets which varied in webpage exploits revealed notable misclassification rates versus novel phish variants. However, long execution time, complex computations due to data query from external resources like GoogleTrends and YahooClues, and an inactive learning on up-to-date data caused limited adaptation in real-time practice.

Overall, the aforementioned achievements have outperformed their competitors in phishing classification. However, they lacked to attain whole or some induction factors such as data set's size and imbalance, various and numerous features, active learning of the feature-base classifier. Altogether might deteriorate anti-phishing and its adaptation to up-to-date and big data on the Web. So far, they become unaware of rapid phishing evolution which would enable phishers to intrude existing anti-phishing schemes and threaten both users and industries. Appendix Table II characterizes the related literatures in terms of their classifiers and limitations.

IV. Induction Factors and Perspectives

The aforesaid review acknowledged that more effective machine learning-based phishing classification model can be attained by boosting its induction power. In greater detail, powerful induction can be maintained via the following factors:

- Feature and feature category. Mainly, numerous typical features were deployed to detect phishing. Such features either belonged to similar feature

categories or different feature categories [4]. Examples of feature categories are: cross site (XSS) scripting features, embedded objects features, language independent features, and hybrid features, i.e. a set of features belonging to multiple former categories [3, 4, 24]. Today, phishers exploit many and hybrid as well as new features to bypass existing anti-phishing schemes for more damages [2, 4, and 24] Therefore, it is essential to inspect new phishing features that will promote the classifiers induction [25].

- Features Selection. Some machine learning-based anti-phishing schemes integrated typical feature selection methods to select a minimal subset of most significant features for induction purposes [26-28]. However, the applied feature selection methods do not always select the optimal feature sets. Because of their highly constrained power to the features heterogeneity, relevance and redundancy in the feature space of the training datasets [28]. Also, their own search strategy and evaluation criteria limit to the best selection of features on large scale feature space [28, 29]. Therefore, some aforesaid anti-phishing schemes varied in their performance due to variant outcomes of features selection methods that falling into either features weighting, or ranking, or nominating features in terms of their interdependencies [24, 28, and 29].
- Data set. Size and imbalance problems of the training dataset caused misleading induction on the incoming web flow during testing. For example, when FBC classifies a phish webpage mistakenly as a legitimate webpage, a suboptimal phishing classification is revealed in real-time application [15, 25, 30] Furthermore, the big web include numerous webpages of various classes and then various attacks to classify. For example, ham, spam, phish, and legitimate etc. [7, 13, 15, 25 and 26]. Such attacks might exploit different and/or common features with phishing that in turn yielded inaccurate and high computational phishing classification [13, 31, 32].
- Active learning. Active learned FBC can instate the expected future error and select the batch of instances from training dataset which are expected to decrease that future error. That implies minimal rates of classification error and maximal classification accuracy [9, 11 and 13]. Therefore, active learned FBC is required to train data set artificially and frequently.
- Adaptation. Adaptable classification model is that reconfigure its function settings dynamically and quickly for better new attacks detection in real-time constraints. Almost machine learning-based anti-phishing schemes as well as their competitors lacked to this important induction factor. That implied misclassification models against novel phish variants emerging periodically with new and undiscovered deceptions [11, 13, 33].

As time progresses, the utilization of machine learning-based classification models in anti-phishing domain still appears more challenging in realistic situations. Lack of all or some of the aforementioned inductive factors could cause substantial trade-off between computational performance and time as well as misclassification of any smaller variation in phish attacks. That, in turn, enforces the machine learning-based phishing classification models to be convolved with proper induction biases. As such, performance overhead could be monitored in minority with the large scale training and testing datasets in offline and/or online situations.

Herewith, a research question is raised to solve based on the above insights: “How to optimize the biases of induction of the existing machine learning-based phishing classification models?” To answer this question, it is highly recommended to take the advantage of the aforesaid induction factors seriously via the following perspectives:

- A great care must be put on exploring new features and enrich the currently used ones. That will provide present characterization on novel phishes. It is worthy to note that researchers in [34] have proposed 58 new and hybrid features and deployed them throughout a hybrid features-based phish website prediction scheme. Their proposed scheme revealed well-performed feature-base classifier with high efficiency.
- Real-time detection mechanism must be attained to adapt the modest phishing exploits and update recent training data for predicting future changes in phishing deceptions. By buffering a stream of modest data, validating them, excluding the false negatives data, and patching the past training data with them to update the FBC settings accordingly.
- Variations in classification outcomes could be monitored with heavy dependence on selected features over big datasets. Specifically, unsatisfactory outcomes can be caused by the use of irrelevant and redundant features. Thus, phishing classification in reality could not promoted with the best method of feature selection that produce similar outputs on the different and large scale data sets.
- Up-to-date data collection. Based on the assumption that FBC is learned with the training data set at a certain time T. A given website w at T could be predicted as phish website in the future time ($T+\check{T}$) that can rarely being estimated; i.e. archived data can be used to learn FBC, and then update the induction settings by which incoming data could be validated versus phishing in the present,
- Chronological validation. Aggregating data sets periodically to validate FBC and qualify its outcomes chronologically, is a complementary factor in predicting phishing. This, in turn, will form a data resampling and minimize the problems of imbalanced data sets.

Overall, a framework of adaptable phishing classification model could be devoted with three functionally inter-related modules in a synchronized mechanism; prediction, validation and detection modules. That implies ability to conquer novel phish-aware detection in real-time practice. Prediction module classifies the training data set and actively learn its FBC offline. Detection module adapts its default settings whenever a new phish activity occurs online. While, Validation module regulates the new settings and updates the prediction module with feedback data.

V. CONCLUDING REMARKS

By revisiting the current achievements in machine learning-based anti-phishing domain, it is observed that they affirmed to be either computationally ineffective or well-performed but complex to accomplish real-time practice. That is due to their full or partial lack to induction factors such as the features and feature categories in use, integrated feature selection methods, data set size and imbalance, active learning, and adaptable modelling. By restating the causality between their related limitations and their induction deficiency; future outlooks are suggested to promote anti-phishing campaign. Furthermore, scope of solutions could be extended in the future via a high level assembly of phishing classification model. The proposed assembly integrates multiple modules working offline and/or online to learn, actively learn, adaptively induct phishing variants on the vast and imbalanced web data. Regarding the survey, this paper demonstrates that machine learning-based classification models with powerful induction elevates classification performance and devote substantial outcomes on the big web data. Thus, it is hoped to serve as a navigating taxonomy to the reseachers for future work.

REFERENCES

[1] M. Khonji, Y. Iraqi & A.Jones, "Phishing detection: a literature survey," Comm. Surveys & Tutorials, vol. 15, no. 4, pp. 2091–2121, 2013.

[2] H. Z., Zeydan, A. Selamat, M. Salleh, "Survey of anti-phishing tools with detection capabilities," *In the proceedings of 14 Int. Symposium on Biometrics and Security Technologies (ISBAST'2014)*, Kuala Lumpur, Malaysia.

[3] H. Shahriar, "Trustworthiness testing of phishing websites: a behavior model-based approach," Future Generation Comput. Syst., vol. 8, no. 28, pp. 1258–1271, 2012.

[4] H. Z., Zeydan, A. Selamat, M. Salleh, "Current state of anti-phishing approaches and revealing competencies," *Journal of Theoretical and Applied Information Technology*, 70(3), 2014, 507-515.

[5] C. Whittaker, B. Ryner, and M. Nazif, "Large-scale automatic classification of phishing pages," *Proceedings of 17th Annual Internet Society on Networks and Distributed System Security Symposium* (NDSS2010). March 2010, San Diego, California, USA.

[6] B. Wardman, J. Britt, and G. Warner, "New tackle to catch a phisher," *International Journal of Electronic Security and Digital Forensics*. 6(1), 2014, 62-80.

[7] A. Abbasi, and H. Chen, "A comparison of fraud cues and classification methods for fake escrow website detection," *Information Technology and Management*, 10(2-3), 2009, 83-101.

[8] R. Islam, and J. Abawajy, "A multi-tier phishing detection and filtering approach," *Journal of Network and Computer Applications*, 36(1), 2013, 324-335.

[9] G. Kumar, K. Kumar, and M. Sachdeva, "The use of artificial intelligence based techniques for intrusion detection: a review," *Artificial Intelligence Review*, 34(4), 2010, 369-387.

[10] S., Kotsiantis, "Supervised machine learning: a review of classification techniques," Informatica, 31, 2007, 249-268.

[11] T. T. Nguyen, and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," *Communications Surveys & Tutorials*, *IEEE,* 10(4), 2008, 56-76.

[12] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches," *IEEE Transactions On Systems, Man And Cybernetics*, *Part C, Applications and reviews,* 42(4), 2012, 463-484.

[13] A. Shabtai, R. Moskovitch, Y., Elovici, and C. Glezer, "Detection of malicious code by applying machine learning classifiers on static features: a state-of-the-art survey." *Information Security Technical Report*. 14(1), 2009, 16-29.

[14] M. He, S.-J. Horng, P. Fan, M. K. Khan, R.-S. Run, J.-L Lai, and A. Sutanto, "An efficient phishing webpage detector," *Expert Systems with Applications*, 38(10), 2011, 12018-12027.

[15] G. Xiang, "Towards a phish free world: a cascaded learning framework for phishing detection," Doctoral Dissertation, *Carnegie Mellon University*, 2013, Pittsburgh, PA 15213.

[16] Y. Lie, R. Xiao & J.Feng, "A semi-supervised learning approach for detection of phishing webpages," Optik-Int. J. for Light Electron Optics, vol. 14, no. 23, pp. 6027–6033, 2013.

[17] R. Gowtham, and I. Krishnamurthi, "A comprehensive and efficacious architecture for detecting phishing webpages," *Computers & Security*, 40, 2014, 23-37.

[18] R. Gowtham, and I. Krishnamurthi, "PhishTackle-a web services architecture for anti-phishing," *Cluster Computing*, 17(3), 2014, 1051-1068.

[19] S. Marchal, J. François, R. State, and T. Engel, "PhishScore: hacking phishers' minds," *Proceedings of 10th International Conference on Network and Service Management* (CNSM2014). Rio de Janeiro: IEEE, 17-21 Nov. 2014, 46-54.

[20] S. Marchal, S., J. François, R. State, and T. Engel, "PhishStorm: detecting phishing with streaming analytics," *IEEE Transactions on Network and Service Management*, 11(4), 2014, 458-471.

[21] R. M. Mohammad, F. Thabtah, F., and L. McCluskey, "Predicting phishing websites based on self-structuring neural network," *Neural Computing and Applications*, 25(2), 2014, 443-458.

[22] D. Zhang, Z. Yan, H. Jiang, H., and T. Kim, "A domain-feature enhanced classification model for the detection of Chinese phishing e-Business websites," *Information & Management*, 51(7), 2014, 845-853.

[23] S. Marchal, "DNS and semantic analysis for phishing detection," Doctoral Dissertation. *University of Luxembourg*, 22 June 2015.

[24] H. Zuhair, A. Selamat, M. Salleh, "Selection of robust feature subsets for phish webpage prediction using maximum relevance and minimum redundancy criterion," *Journal of Theoretical and Applied Information Technology*, 81(2), 2015, 188-205.

[25] O. Kwon, and J. M. Sim, "Effects of data set features on the performances of classification algorithms," *Expert Systems with Applications*, 40(5), 2013, 1847-1857.

[26] E. Uzun, H. V. Agun, and T. Yerlikaya, "A hybrid approach for extracting informative content from web pages," *Int. Journal of Information Processing and Management: an International Journal.* 49(4), 2013, 928-944.

[27] F. Toolan, and J. Carthy, "Feature selection for Spam and Phishing detection," *Proceedings of eCrime Researchers Summit* (eCrime), 2010.

[28] H. Zuhair, A., Selamat, M. Salleh, "Feature Selection for phishing detection: a review of research," *Int. Journal of Intelligent Systems Technologies and Applications.* 15(2), 2016, 147-162.

[29] H. Zuhair, A. Selamat, M. Salleh, "The effect of feature selection on phish website detection: an empirical study on robust feature subset selection for effective classification," *Int. Journal of Advanced Computer Science and Applications. 6(10), 2016, 221-232.*

[30] S. Amiri, O. A. Akanbi, and E. Fazeldehkordi, "A machine-learning approach to phishing detection and defense", Syngress, 2014.

[31] C. K. Olivo, A. O. Santin, and L. S. Oliveira, "Obtaining the threat model for e-mail phishing," Applied Soft Computing, 13(12), 2013, 4841-4848.

[32] M. Aburrous, M. Hossain, K. Dahal, and F. Thabtah, "Associative classification techniques for predicting e-Banking phishing websites," *Int. Conference Multimedia Computing and Information Technology (MCIT'10)*, 2010. R. B.

[33] C. Tsai, Y.-F. Hsu, C.-Y. Lin, and W.-Y. Lin, "Intrusion detection by machine learning: a review," *Expert Systems with Applications*, 36(10), 2012, 11994-12000.

[34] H., Zuhair, M. Salleh, and A. Selamat,"Hybrid features-based prediction for novel phish website," Jurnal Teknologi, 78(12-3), 2016.

Appendix

TABLE I. EXAMPLES OF MACHINE LEARNING ALGORITHMS USED IN ANTI-PHISHING DOMAIN[29]

Algorithm	Description
C4.5	It depends on Decision Tree hypothesis that traces the node paths, their branches until terminating leafs.
Decision Tree (DT)	It models the unknown instances as nods in a rooted tree, and the feature values as edges. Induction starts from the root node approaching to leafs and passing through edges. Test is applied at each node to re-order feature values which determine the next edge to go. Final decision found at the end-up leaf node.
Naïve Bayes (NB)	A probabilistic judgment done conditionally with independent attributes of all instances belonging to a given class: $P(C\|X) = P(C\|x_1, \dots, x_n) = \frac{P(C)P(x_1,\dots,x_n\|C)}{P(x_1,\dots,x_n)}$ (1) Where X is an instance with a vector of *n* features($x_1, \dots, x_n$), C is the class label that the classifier seeks for.
Support Vector Machine (SVM)	A separating hyper-plane maximizes the margins between closest points of two classes to estimate the induction function: $min\frac{1}{2}w^T w + C\sum_i \xi_i$ (2) That subjects to: $y_i\big((w^T \cdot x_i) + b\big) \geq 1 - \xi_i, \xi \geq 0, i = 1, 2, \dots, m$ (3) $max\sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{m} y_i y_j \alpha_i \alpha_j K\big(x_i, x_j\big)$ (4) Which is Subject to :$0 \leq \alpha_i \leq C, i = 1, 2, \dots, m\ and\ \sum_{i=1}\alpha_i y_i = 0$ (5) Where: x_i is m-dimensional data vector $x_i \in R^m$ with samples belong to either one of two classes labeled as $y \in \{-1, +1\}$that it is separated by a hyper-plane of $(w \cdot x) + b = 0$, $\alpha_i$denotes the lagrange multipliers for each vector in the training dataset
Transductive Support Vector Machine (TSVM)	It separates positive and negative samples of training dataset with a maximal margin of SVM hyper-plane, such that it minimizes over $(y_1^*, \dots, y_k^*, w, b, \xi_1, \dots, \xi_n, \xi_1^*, \dots, \xi_k^*)$ into:$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_1 + C'\sum_{j=1}^{k}\xi_j^*$ (6) Which subjects to: $\forall_{i=1}^{n}: y_i[wv_i + b] \geq 1 - \xi_i$, (7) $\forall_{j=1}^{k}: y_i[wv_i^* + b] \geq 1 - \xi_j^*, \forall_{i=1}^{n}: \xi \geq 0, \forall_{j=1}^{k}: \xi_j^* \geq 0$ (8)
Logistic Regression (LR)	Use probabilistic induction that evaluates relationship between a categorical dependent variable and a continuous independent variable (s): $\pi(x) = \frac{e^{(\beta_0+\beta_1 x)}}{e^{(\beta_0+\beta_1 x)}+1} = \frac{1}{e^{-(\beta_0+\beta_1 x)}+1}$, (9) $g(x) = ln\frac{\pi(x)}{1-\pi(x)} = \beta_0 + \beta_1 x$, (10) $\frac{\pi(x)}{1-\pi(x)} = e^{(\beta_0+\beta_1 x)}$ (11) Where: $g(x)$ is the logistic function of a given predictor *X*, *ln* and, $\pi(x)$ denote natural logarithm and case probability, $\beta_0$and β_1 denote criterion of *X* , and $\beta_1 x$ is the regression coefficient
Random Forests (RF)	Forest constructed for randomly selected set of instances on training dataset. It comprises of many combined tree predictors that are distributed similarly. Each tree predictor is learned on feature vector belongs to independent sample.
K-Nearest Neighbor (K-NN)	Nonparametric classifier estimates class conditional densities by using a discriminant function $g_i(x) = P(x\|C_i)P(C_i)$, (12) $P(x\|C_i) = \frac{k_i}{(N_i V^k(x))}$ (13) Where $P(x\|C_i)$, k_i and $V^k(x)$are the class conditional densities, the number of nearest neighbors that belong to C_i, and the volume of *n*-dimensional hyper-sphere centered at *x* with radius of $r = \|x - x_k\|$ and x_k is the nearest observation to *x*.

TABLE II. MERITS AND DEMERITS OF NOTABLE MACHINE LEARNING-BASED ANTI-PHISHING MODELS

Study / **Issues**	[14]	[15]	[16]	[17-18]	[21]	[22]	[19-20, 23]
Machine Learning	SVM	SVM, LR, BN, DT, Adaboost	SVM/TSVM	SVM	Neural Network	SMO, LR, RF, NB	SVM, C4.5, RF, JRip
FBC design	Single	Ensemble	Single	Single	Ensemble	Single	Single
No. of Features	12	15	6	17	17	15	212
Type of Features	Hybrid features	Hybrid features	Textual/visual features	Hybrid features	URL/Textual features	URL features	URL/Textual features
New Features	Not	8 new features	Not	3 new features	Not	5 new features	Not
Data Set (Phish/Not Phish)	325/200	8118/4883	200/200	1764/700	800/600	1416/1462	48,000/48,000
Data Set Source	PhishTank, CastleCops, Millersmiles	PhishTank, Alexa, 3Sharp	PhishTank, GoogleWhite	NetCraft, MillerSmile, GoogleTop, Alexa	PhishTank, Millersmiles,	Chinese Data Archives	PhishTank, DMOZ
Detected Phish Types	Redirecting, Login Form, Homepage, e-Business, English	Redirecting, Login Form, e-Business, Social Networking, English	e-Commerce Login Forms, English	e-Commerce, Login Forms, Redirecting, English	e-Commerce, Login Forms, Redirecting, English, French	e-Business, Chinese	e-Commerce, English, French, German, Italian, Spanish, Portuguese
Overall Performance*	TP 97.33% FP 1.45%	TP 92%; FP 1.4%	Accuracy 96.4%; Miss Rate 3.5%; Mistake Rate 5.4%	TP 99.6%; FP 0.42%	Accuracy Rates 91.31-94.7%	Accuracy Rate 95.83%	Accuracy Rate 99%, FP 0.004%
Demerits	Typical features; Few feature; Small data set, No feature selection; Inactive learning, No Adaptation, Time consuming	Imbalanced data; Typical features; Few features; No feature selection; No adaptation; Time consuming	Imbalanced data; Typical features; Few features; Small data set, No feature selection No adaptation	Typical features, Few features; Small data set, No feature selection, No adaptation	Typical features, Few features; Small data set, No feature selection, Inactive learning, No adaptation	Typical & few features, One feature category, Imbalanced data; Inactive learning, No adaptation	Typical features, One feature category, No features selection, No adaptation
Notes	FBC=Feature-Base Clasifier, *SVM =Support Vector Machine; LR=Logistic Regression; BN=Bayesian; DT, C4.5, and JRip are types of Decision Tree Classifier; RF=Random Forest; NN=Neural Network; SMO=Sequential Minimal Optimization, NB=Naïve Bayes.* *Overall performance outcomes are adopted as they presented in the related works.						