Applied Soft Computing 48 (2016) 729–734

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

A new fast associative classification algorithm for detecting phishing websites

Wa'el Hadi[a,*], Faisal Aburub[a], Samer Alhawari[b]

[a] University of Petra, MIS Department, Jordan
[b] The World Islamic Science & Education University, MIS Department, Jordan

ARTICLE INFO

Article history:
Received 18 April 2016
Received in revised form 27 July 2016
Accepted 2 August 2016
Available online 6 August 2016

Keywords:
Associative classification
Phishing websites
Classification
Data mining

ABSTRACT

Associative classification (AC) is a new, effective supervised learning approach that aims to predict unseen instances. AC effectively integrates association rule mining and classification, and produces more accurate results than other traditional data mining classification algorithms. In this paper, we propose a new AC algorithm called the Fast Associative Classification Algorithm (FACA). We investigate our proposed algorithm against four well-known AC algorithms (CBA, CMAR, MCAR, and ECAR) on real-world phishing datasets. The bases of the investigation in our experiments are classification accuracy and the F1 evaluation measures. The results indicate that FACA is very successful with regard to the F1 evaluation measure compared with the other four well-known algorithms (CBA, CMAR, MCAR, and ECAR). The FACA also outperformed the other four AC algorithms with regard to the accuracy evaluation measure.

1. Introduction

The rapid growth in the number of internet users has dramatically increased e-shopping as a practical alternative to traditional shopping. According to Ingham et al. [11], worldwide sales from e-shopping increased by 20.1% in 2014 to $1500 trillion. Therefore, the main objective of e-shops is to attract more and more consumers. This has led to increased competition among e-shops to introduce quality services for consumers. However, the increasing number of e-shops and websites has been accompanied by a rapid growth in the number of phishing websites. Although internet users are aware of phishing, many fall victim to such attacks. The aim of phishers is to make internet users believe that they are interacting with trusted online sites. Phishing websites can appear to be any type of website, including online payment or auction websites. An efficient method of identifying phishing websites is required in order to protect users' sensitive data.

Phishing applies both technical tricks and social engineering to access private information illegally. The purpose of the phishing is to take private information by publishing a forged website that appears to be a legitimate one, such as a real website of a company or bank, and requesting a person to input private information, such as account number, credit card number, username, and password. As well as being harmful to customers, phishing attacks also damage the reputation of the financial institutions concerned, since customers become less confident that they can securely access their accounts. Phishing websites are considered to be one of the most common electronic crimes [7,13]. According to a report from APWG [7], the number of distinct phishing reports submitted to the organization during quarter 4 of 2014 was 197,252. This was an increase of 18% on the 163,333 received in quarter 3 of 2014.

Data mining is a field of study that aims to find useful information in large databases in order to help decision makers to make correct decisions. Data mining involves many tasks, such as classification, association rules and clustering. Classification is the task of forecasting, assigning or predicting unseen instances to their predefined classes, for example, forecasting incoming email as either inbox or spam. Association rules is the task of finding relationships between attributes (features) in a large database. For example, if a consumer buys soda and potatoes together, they are also likely to buy meat; this relationship is represented as a rule of the form (*soda, potatoes* → *meat*), where "*soda and potatoes*" is called the rule body and "*meat*" the head of the rule. Associative classification (AC) is a new task in data mining and machine learning, and aims to forecast unseen instances based on association rules. AC is a promising approach because many researchers have indicated that it produces more accurate results than other traditional data mining classification techniques [19,2,1,5].

* Corresponding author.
E-mail addresses: whadi@uop.edu.jo (W. Hadi), faburub@uop.edu.jo (F. Aburub), samer.alhawari@yahoo.com (S. Alhawari).

http://dx.doi.org/10.1016/j.asoc.2016.08.005
1568-4946/

The main goals of this paper are to present a new, fast, and very efficient AC classifier, and to compare this new AC classifier with four well-known AC algorithms with reference to classification accuracy and F1 evaluation measures on a new phishing dataset proposed by Mohammad et al. [16].

The rest of this paper is organized as follows. Related works are explained in Section 2 and the proposed AC classifier is described in Section 3. In Section 4, the experimental results are discussed, and finally, in Section 5, conclusions are presented.

2. Related works

Over the past decade, many researchers have investigated the problem of detecting phishing websites using data mining techniques, but there are a limited number of research articles relating to the AC approach. In this section, we shed light on both traditional data mining techniques and AC approaches.

Abdelhamid et al. [3] investigated the problem of website phishing using a new proposed multi-label classifier-based associative classification, MCAC. The main goal of the MCAC algorithm developed is to recognize attributes or features that distinguish phishing websites from legitimate ones. The results showed that the MCAC algorithm forecasted phishing websites better than traditional data mining algorithms.

Dadkhah et al. [8] developed a new method to forecast and detect phishing websites using classification algorithms based on the weight of web page features. The results showed that the proposed method produced a lower error rate than other data mining methods.

Abdelhamid [1] proposed an enhanced multi-label classifier-based associative classification algorithm, eMCAC. This generates rules associated with a set of classes from single-label datasets using the transaction ID list (Tid-list) vertical mining approach. The algorithm employs a novel classifier building method that reduces the number of generated rules. The experiments indicated that the eMCAC algorithm outperformed other algorithms with regard to the accuracy evaluation measure.

Jabri and Ibrahim [12] proposed an enhanced PRISM algorithm for forecasting phishing websites. The experimental results revealed that the modified PRISM algorithm outperformed the original PRISM algorithm in terms of the number of rules, accuracy (87%), and lower error rate (0.1%).

Alazaidah et al. [5] proposed a new multi-label classification algorithm based on correlations among labels, MLC-ACL. The MLC-ACL utilizes both problem transformation techniques and algorithm adaptation techniques. The proposed algorithm starts by converting a multi-label dataset into a single-label dataset using the least frequent label criteria, and then employs the PART machine learning classifier on the converted dataset. The output of the classifier is multi-label rules. In addition, MLC-ACL attempts to gain advantage from positive correlations among labels using the predictive Apriori algorithm. The MLC-ACL algorithm was investigated using two multi-label datasets named Emotions and Yeast. The experiments revealed that the MLC-ACL algorithm outperformed other machine learning algorithms in terms of three well-known evaluation measures (Hamming Loss, Harmonic Mean, and Accuracy).

Taware et al. [17] proposed a new MCAC that aims to recognize attributes that differentiate phishing websites from legitimate ones. The MCAC algorithm produced better results than other data mining algorithms with regard to accuracy.

Antonelli et al. [6] developed a new efficient AC algorithm using a fuzzy frequent pattern method. The experiment results showed that the new fuzzy AC algorithm outperformed the well-known CMAR algorithm and generated accuracies similar to two recent AC algorithms, namely FARC-HD and D-MOFARC, on 17 real-world datasets.

Table 1
Training data.

Tid	Age	Income	Has a car	Buy/class
1	Senior	middle	yes	yes
2	Youth	low	yes	no
3	Junior	high	yes	yes
4	Youth	middle	yes	yes
5	Senior	high	no	yes
6	Junior	low	no	no
7	Senior	middle	no	no

The problem with the current AC algorithms is that the set of candidate rules produced from the training data is typically large, which consumes time and Input/Output resources. This problem motivated us to propose a new AC algorithm that generates all frequent rules using an efficient association rule mining method that reduces the time and memory required. Moreover, we propose a new prediction method to forecast unseen instances more accurately than other methods. Unlike most of the current AC prediction methods, our proposed method considers multiple rules to assign the class in the prediction step.

3. Proposed algorithm

AC mining is a new branch of data mining and machine learning that classifies unseen instances based on association rules. AC mining algorithms have been investigated during the past two decades by many specialist researchers in real-world fields such as phishing websites, text classification, medical diagnoses, and fraud detection. The AC mining approach is widely used for a number of reasons. First, it produces higher classification accuracy rates for the outputting classifier than other data mining and machine learning approaches. Second, rules produced by AC mining classifiers are simple and illustrated by simple "*if–then*" rules, so the user can easily read, remove, modify, and understand the produced rules. The primary goal of an AC mining approach is to build a classifier (model) from a huge database (training data) to forecast (detect or predict) the type of unseen instances (testing data).

In this paper, a new, fast, and efficient AC mining classifier called the Fast Associative Classification Algorithm (FACA) is developed. FACA is an efficient AC mining classifier, the difference between FACA and all other AC classifiers in the literature. It employs a vertical mining approach called Diffset [20] for discovering all frequent itemsets, and utilizes a new prediction method to classify unseen instances more accurately than other methods. To the best of the author's knowledge, there is no AC mining algorithm that implements the Diffset method.

Before we discuss the FACA algorithm steps, more elaboration of the Diffset technique is required in order to ensure a better understanding. Diffset is a vertical data approach that keeps track of only the transaction IDs in which a ruleitem does not occur. For example, the Diffset for <Age, senior> → yes is {2, 3, 4, 6, 7}, according to the training data shown in Table 1, because the ruleitem <Age, senior> → yes occurs in two transactions {1, 5}. The support of a candidate K-ruleitem (rule with K items in its body) is the cardinality of the Diffset for the (K − 1)-ruleitem subtracted from the cardinality of the Diffset for the K-ruleitem itself. If a candidate rule is a single-ruleitem, the support is computed by subtracting the number of transactions in the training data from the cardinality of the Diffset for the single-ruleitem itself, i.e., the support for the single-ruleitem <Age, senior> → yes = 7 − 5 = 2. The confidence of the ruleitem A → B is the conditional probability that a transaction contains B, given that it contains A, and is computed by dividing the support of $(A \bigcup B)$ by the support of (A), i.e., the con-

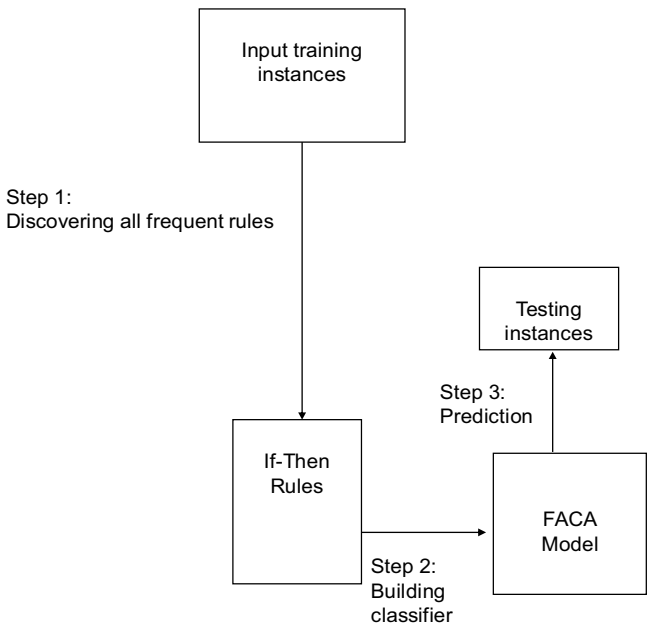

Fig. 1. FACA Main Steps.

fidence for the ruleitem <Age, senior> → yes = 2/3. A ruleitem is frequent if its support is greater than or equal to the minimum support threshold. Once all single-frequent rules have been generated, the Diffset technique iteratively merges frequent (k – 1)-ruleitems to produce a candidate K-ruleitem. For example, to produce a 2-ruleitem like «Income, high>, <has a car, yes> > → yes, two single-ruleitems are merged, i.e., <Income, high> → yes and <has a car, yes> → yes. The Diffset for <Income, high> → yes is {1, 2, 4, 6, 7} and the Diffset for <has a car, yes> → yes is {2, 5, 6, 7}. The Diffset for the 2-ruleitem is a difference of transaction between the second single-ruleitem (<has a car, yes> → yes) and its first single-ruleitem (<Income, high> → yes); thus, the Diffset for «Income, high>, <has a car, yes> > → yes is {5}. To find out if «Income, high>, <has a car, yes> > → yes is frequent, the technique subtracts the support for <Income, high> → yes from the cardinality of the Diffset of «Income, high>, <has a car, yes> > → yes, i.e., 2 – 1 = 1.

3.1. Fast associative classification association algorithm (FACA)

FACA goes through three steps without a pre-processing step: discovering all frequent rules, building the classifier (model), and predicting a new unseen instance, as shown in Fig. 1.

Initially, the end-user specifies minimum support and minimum confidence thresholds using graphical interface. In the first step, the FACA algorithm scans all training instances to discover rules with a single item in their body in the form *(<Attribute, value> → Class)*. Then, it iteratively merges rules with a single item in their body to discover rules with two items in their body, and so on. FACA removes any rule that has support of less than the minimum support threshold. It should be noted that the FACA algorithm utilizes the Diffset vertical mining method for discovering rules.

Once all frequent rules have been discovered, the FACA algorithm investigates their confidence values. If the frequent rule has confidence greater than or equal to the minimum confidence, the rule will be added into class association rules (CARs), otherwise the rule will be removed. The next step is to employ the database coverage pruning method to select a subset of helpful rules to form the FACA model.

The FACA algorithm sorts rules on CARs in order to give the more helpful rules a higher priority for being selected as part of the model.

1) The rule with least number of feature values in its body is given a higher rank; we prefer here a general rule rather than a specific one.
2) If the number of feature values in the body of two or more rules is the same, then the rule with higher confidence is given a higher rank.
3) If the confidence values of two or more rules are the same, then the rules with higher support is given a higher rank.
4) If all the above criteria are similar for two or more rules, then the rule that was produced first is given a higher rank.

Fig. 2. FACA Sorting Rules.

The proposed procedure sorts the discovered rules according to Fig. 2.

FACA builds the model in the following way. It starts with the highest rank rule and investigates it on all training instances; the rule is added to the FACA model if it matches at least one training instance. All training instances that match the rule body and its head are removed and the rule is added to the FACA model. Otherwise, the rule is deleted. The same process is repeated on the remaining rules in order until no more training instances remain, or all rules on CARs have been investigated. This process in the FACA algorithm guarantees that only the most helpful rules remain in the model for the prediction step.

Finally, when the FACA model has been created, our prediction method (All Exact Match Prediction Method) works as follows. Given an unseen instance, the prediction method scans all helpful rules on the model, beginning with the first sorted rule (rule that has the least number of attribute values) and proceeding to the last rule. Our proposed prediction method then divides all rules that match the unseen instance into clusters according to their classes. Finally, the All Exact Match Prediction Method counts the number of rules per cluster, chooses the class with the highest count and assigns it to an unseen instance.

We will describe the discovery of frequent ruleitems and the building of the model in FACA using an example. Consider the training data shown in Table 1, which has four columns: "Age" (senior, youth, junior), "Income" (middle, low, high), "Has a car" (yes, no), and a class "Buy" (yes, no). Assuming minimum support = 20% and minimum confidence = 60%, the frequent 1 and 2 ruleitems produced from Table 1 are displayed in Tables 2 and 3, along with their corresponding supports and confidences. First, FACA scans all the training data in Table 1, and all single-frequent ruleitems are discovered. The ruleitems that are marked with X are removed because the support for these ruleitems is less than the inputted minimum support. The FACA algorithm stores the Diffsets for each ruleitem. The Diffset is a difference of transaction between a candidate K-ruleitem and its prefix (K – 1)-ruleitem. For example, the Diffset for *«Age, senior>,<Income, middle> > → yes* is {5}, because the Diffset for *<Income, middle> → yes* is {2,3,5,6,7} and the Diffset for *<Age, senior> → yes* is {2,3,4,6,7}. So the difference between *<Income, middle> → yes* and *<Age, senior> → yes* is {5}. FACA computes the support for a candidate K-ruleitem by subtracting the cardinality of Diffset between the first ruleitem ((K – 1)-ruleitem) and K-ruleitem itself. The support for *«Age, senior>, <Income, middle> > → yes* is 1, because the support for *<Age, senior> → yes* is 2 and the support for *«Age, senior>, <Income, middle> > → yes* is 1, and 2 – 1 = 1. After all frequent rule itemsets have been discovered, the FACA algorithm builds a model. The FACA sorts all discovered rule itemsets according to Fig. 2, and Table 4 illustrates the results of the sorting step. The highest-ranked rule is *<Income, low> → no* and the rule *«Income, middle>, <Has a car, yes> > → yes* is the lowest ranked.

Table 5 illustrates the ruleitem pruning procedure: when rule (R1) is applied to the training data shown in Table 1, we find that this rule matches instances 2 and 6, so instances 2 and 6 are removed from the training data, and the rule is added to the FACA model. When rule (R2) is applied to the remaining training data

Table 2
Discovering frequent single-ruleitems.

Age						Income						Has a car			
Senior		Youth		Junior		Middle		Low		High		Yes		No	
Yes	No (X)	Yes (X)	No (X)	Yes (X)	No (X)	Yes	No (X)	Yes (X)	No	Yes	No (X)	Yes	No (X)	Yes (X)	No
2	1	1	1	1	1	2	1	1	1	1	1	2	1	1	1
3	2	2	3	2	2	3	2	2	3	2	2	5	3	2	2
4	3	3	4	4	3	5	3	3	4	4	3	6	4	3	3
6	4	5	5	5	4	6	4	4	5	6	4	7	5	4	4
7	6	6	6	6	5	7	5	5	7	7	5		6	6	5
	7	7	7	7	7		6	6			6		7	7	
								7			7				

Table 3
Discovering frequent two-ruleitems.

«Age, senior>,<Income, middle> > → yes (X)	«Age, senior>,<Income, high> > → yes (X)	«Age, senior>,< Has a car, yes > > → yes (X)	«Income, middle>,<Has a car, yes> > → yes	«Income, high>,<Has a car, yes> > → yes (X)	«Income, low>,<Has a car, no> > → no (X)
5	1	5		5	2

Table 4
Sorting all ruleitems.

Rule	Support	Confidence	Rank
<Age, senior> → yes	3	0.67	R4
<Income, middle> → yes	3	0.67	R5
<Income, low> → no	2	1	R1
<Income, high> → yes	2	1	R2
<Has a car, yes> → yes	4	0.75	R3
<Has a car, no> → no	3	0.67	R6
«Income, middle>,<Has a car, yes> > → yes	2	1	R7

Table 5
Pruning rules.

Rule-id	Rule	Instances covered in training data	FACA Model?
R1	<Income, low> → no	2,6	Yes
R2	<Income, high> → yes	3,5	Yes
R3	<Has a car, yes> → yes	1,4	Yes
R4	<Age, senior> → yes		No
R5	<Income, middle> → yes		No
R6	<Has a car, no> → no	7	Yes, and stop
R7	«Income, middle>,<Has a car, yes> > → yes		No

(five instances), we find that the rule matches instances 3 and 5, so these instances are removed and rule (R2) is added to the FACA model. Similarly, rule (R3) is added to the FACA model and instances 1 and 4 are removed because the rule matches them. When rule (R4) is applied to the remaining training data (one instance), we find that it does not match the remaining instance, so the rule is removed. Rule (R5) does not match the remaining instance, so is deleted. Finally, rule (R6) matches instance 7, so this instance is removed and the rule (R6) is added to FACA. Now the training data is empty, so the FACA deletes the remaining rules, namely (R7), and stops building the model. The outcome of this pruning method is a model that contains the best four rules that will help to obtain high-accuracy results.

Lastly, the FACA algorithm tests instances using our prediction method. The All Exact Match Prediction Method counts the class of all rules that match the body of the unseen test instance, then chooses the class with the highest count and forecasts it for unseen instances. For example, consider the test data shown in Table 6. FACA applies the highest-ranked rule (*<Income, low> → no*) from Table 5 and finds that this rule matches the test instance body. The

Table 6
Testing Data.

Tid	Age	Income	Has a car	Buy/class
1	youth	Low	no	?

algorithm also applies the second highest ranked rule (*<Income, high> → yes*), and finds that this rule does not match the body of the test instance. FACA tries the third highest ranked rule (*<Has a car, yes> → yes*), and finds that this rule, too, does not match the body of the test instance. Finally, the last rule (*<Has a car, no> → no*) matches the body of the test instance. Consequently, the largest number of valid rules associated with class is "No," so our FACA prediction method forecasts "No" for the test instance.

4. Experimental results

We investigated a dataset of 11,055 phishing websites in our experiments [16]. These phishing website cases belong to two classes (Phishy (−1), Legitimate (1)), and the dataset contains 30 features (attributes) to differentiate phishing websites. Some of the features hold two values, "Phishy" and "Legitimate," and the rest of the features hold three values, "Phishy," "Suspicious," and "Legitimate." We replaced these values with the numerical values −1, 0, and 1, respectively.

In all the experiments, we used a well-known 10-fold cross-validation testing method for rational evaluation of the models derived by the algorithms considered and to reduce overfitting. In addition, four well-known AC algorithms were evaluated with FACA. These algorithms are CBA [15], CMAR [14], MCAR [18], and ECAR [9].

The experiments were conducted on an I7 machine with 16G main memory, and the experiments of all algorithms were conducted using a WEKA software [10] environment. Our comparisons are based on accuracy and F1 evaluation measures.

We set the minimum support and minimum confidence thresholds for all AC algorithms (FACA, CBA, CMAR, MCAR, and ECAR) to 2% and 50%, respectively, for all experiments. The ultimate aim of having the minimum support of 2% is that previous studies, for example, Abdelhamid [1], Ajlouni et al. [4], and Thabtah et al. [19], have recommended minimum support values ranging between 2% and 5%.

After analyzing Fig. 3, we found that the FACA outperformed all AC algorithms with reference to the accuracy evaluation measure.

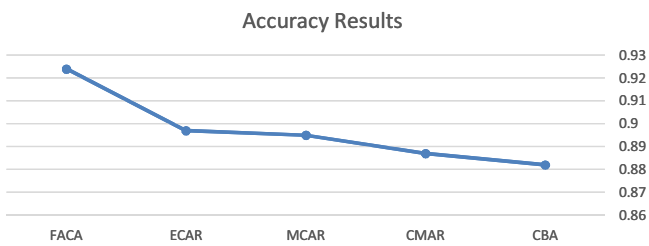

Fig. 3. The outputted classification accuracy.

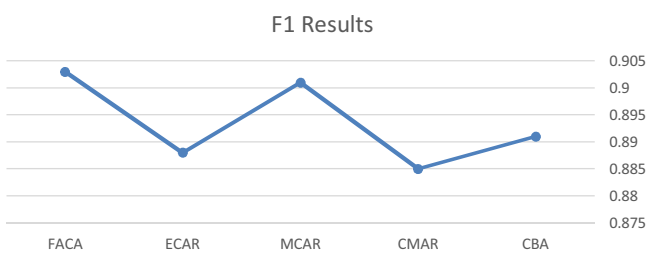

Fig. 4. The outputted F1.

In particular, the FACA outperformed CBA, CMAR, MCAR, and ECAR by 4.2%, 3.7%, 2.9%, and 2.7%, respectively.

As depicted in Fig. 4, the FACA produced F1 results that were better than those of other AC algorithms. In particular, the FACA outperformed CBA, CMAR, MCAR, and ECAR by 1.2%, 1.8%, 0.2%, and 1.5%, respectively.

One major reason that the FACA classifier generated higher classification accuracy and outputted F1 is its ability to forecast unseen test instances using all generated rules. This differs from other AC algorithms such as CBA, MCAR, and ECAR, which use only one rule to forecast the unseen test instance. Furthermore, our proposed ranking method guarantees higher quality and the selection of helpful rules as a part of the FACA model. Finally, the results show that there is potential for computerized data mining techniques to be used in predicting the complex problem of phishing websites.

4.1. Feature selection

To avoid high dimensionality, we used the well-known chi-square (χ^2) feature selection filter method on the phishing websites dataset. Feature selection is the process of selecting the highest N features as a subset of the features found in the training datasets. Feature selection achieves two main goals. First, it makes the training datasets applied to a classifier more efficient by eliminating a number of features. Second, it often generates a classification accuracy rate that is comparable to that generated by a classifier without feature selection if the feature selection method eliminates only the irrelevant features. To accomplish this task, we used the χ^2 filter in WEKA software [10].

We evaluated the phishing websites dataset in order to find the smallest important set of features that allow the type of website to be predicted. Investigation of the 30 features indicated that 12 features correlate with the class feature values and, therefore, may impact positively on the task of detecting phishing websites. These are: SSL Final State, URL of Anchor, Web Traffic, Sub Domain, Prefix Suffix, Links in Tags, Request URL, Domain Registration Length, Server Form Handler, Google Index, Age of Domain, and Page Rank.

Fig. 5 shows the evaluation of the FACA algorithm with ECAR, MCAR, CMAR, and CBA algorithms on the reduced feature set of the data in terms of the error rate. This indicates that the FACA produced the lowest error rate results of all algorithms. Specifically, the FACA error rate is lower than that of CBA, CMAR, MCAR, and ECAR by 4%, 3.7%, 3.2%, and 2.9%, respectively.

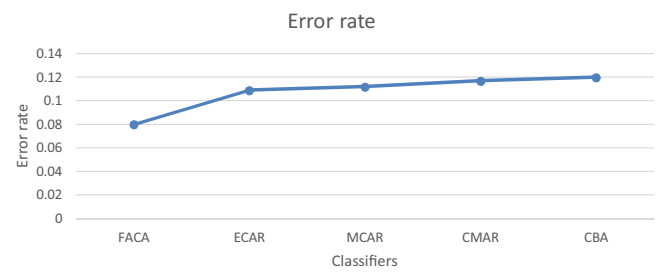

Fig. 5. The outputted error rate.

Finally, reducing some of the features impacted very slightly for all classifiers used: the classification accuracy was reduced for FACA, CBA, CMAR, MCAR, and ECAR by only 0.04%, 0.02%, 0.04%, 0.07%, and 0.06%, respectively. This reflects the fact that the features eliminated are irrelevant and not helpful in distinguishing the type of website.

5. Conclusions

Phishing is an attempt, usually made through fake websites or emails, to steal an individual's private information. Phishing websites are a significant problem, preventing users from carrying out activities via the internet.

This paper aims to develop a new, fast AC algorithm called FACA and investigate FACA against four well-known AC algorithms (CBA, CMAR, MCAR, and ECAR) with regard to classification accuracy and F1 evaluation measures toward the phishing dataset. The results demonstrate that the FACA algorithm outperformed all AC algorithms with regard to accuracy and F1. Our findings also indicate that there is potential for the use of computerized data mining techniques in predicting the complex problem of phishing websites.

References

[1] N. Abdelhamid, Multi-label rules for phishing classification, Appl. Comput. Inf. 11 (1) (2015) 29–46.
[2] N. Abdelhamid, A. Ayesh, W. Hadi, Multi-label rules algorithm based associative classification, Parallel Process. Lett. 24 (1) (2014) 1450001-1–1450001-21.
[3] N. Abdelhamid, A. Ayesh, F. Thabtah, Phishing detection based Associative classification data mining, Expert Syst. Appl. 41 (13) (2014) 5948–5959.
[4] M. Ajlouni, W. Hadi, J. Alwedyan, Detecting phishing websites using associative classification, Eur. J. Bus. Manage. 5 (15) (2013) 36–40.
[5] R. Alazaidah, F. Thabtah, Q. Al-Radaideh, A multi-label classification approach based on correlations among labels, Int. J. Adv. Comput. Sci. Appl. 6 (2) (2015) 52–59.
[6] M. Antonelli, P. Ducange, F. Marcelloni, A. Segatori, A novel associative classification model based on a fuzzy frequent pattern mining algorithm, Expert Syst. Appl. 42 (4) (2015) 2086–2097.
[7] APWG (Anti-Phishing Working Group), Phishing Activity Trends Report Quarter 4, APWG, 2014.
[8] M. Dadkhah, M. Jazi, V. Lyashenko, Prediction of phishing websites using classification algorithms based on weight of web pages characteristics, J. Math. Technol. 5 (2) (2014) 24–35.
[9] W. Hadi, ECAR: a new enhanced class association rule, Adv. Comput. Sci. Technol. 8 (1) (2015) 43–52.
[10] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The WEKA data mining software: an update, ACM SIGKDD Explor. Newsl. 11 (1) (2009) 10–18.
[11] J. Ingham, J. Cadieux, A.M. Berrada, e-Shopping acceptance: a qualitative and meta-analytic review, Inf. Manage. 52 (1) (2015) 44–60.
[12] R. Jabri, B. Ibrahim, Phishing websites detection using data mining classification model, Trans. Mach. Learn. Artif. Intell. 3 (4) (2015) 42–51.
[13] M. Khonji, Y. Iraqi, A. Jones, Lexical URL analysis for discriminating phishing and legitimate e-mail messages, Proceedings of the 6th International Conference for Internet Technology and Secured Transactions (ICITST), 11–14 December (2011) 422–427.

[14] W. Li, J. Han, J. Pei, CMAR: accurate and efficient classification based on multiple-class association rule, in: Proceedings of the ICDM'01, San Jose, CA, 2001, pp. 369–376.
[15] B. Liu, W. Hsu, Y. Ma, Integrating classification and association rule mining, in: Proceedings of the KDD, New York, NY, 1998, pp. 80–86.
[16] R. Mohammad, F.A. Thabtah, T.L. McCluskey, Phishing Websites Dataset, 2015 (Accessed November 2015) https://archive.ics.uci.edu/ml/datasets/Phishing+Websites.
[17] S. Taware, C. Ghorpade, P. Shah, N. Lonkar, S.M.S.C.M. Bk, Phish detect: detection of phishing websites based on associative classification (AC), Int. J. Adv. Res. Comput. Sci. Eng. Inf. Technol. 4 (3) (2015) 384–395.
[18] F. Thabtah, P. Cowling, Y. Peng, MCAR: multi-class classification based on association rule approach, in: Proceeding of the 3rd IEEE International Conference on Computer Systems and Applications, Cairo, Egypt, 2005, pp. 1–7.
[19] F. Thabtah, W. Hadi, N. Abdelhamid, A. Issa, Prediction phase in associative classification mining, Int. J. Softw. Eng. Knowl. Eng. 21 (6) (2011) 855–876.
[20] M.J. Zaki, K. Gouda, Fast vertical mining using diffsets, in: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining-KDD '03, New York, New York, USA: ACM Press, 2003, http://dx.doi.org/10.1145/956750.956788 (p. 326).