Detection of Phishing Websites Using Machine Learning

Neetu Singh Rathore and Mr. Manish Tiwari

Geetanjali Institute of Techanical Studies, Udaipur, India.

Abstract. Phishing is when cybercriminal send malicious email designed to trick people into falling for a scam. The intent is often to get users to reveal financial information, system credentials, or other sensitive data. In phishing process attackers used fake websites page known as phishing page which looks like same as real websites, from that attackers theft all sensitive and private data of the banking websites and any other financial websites data. Phishing web pages used for theft user name, password and account detail of social networking sites. To secure websites many anti-phishing methods are applied. In this paper JRip, J48, Naïve Bayes (NB) and BayesNet algorithms are applied in Weka for phishing website detection using data mining classification model. There are 31 attributes are used in different algorithms to show terms and error rates with accuracy. J48 algorithm gives best results in all algo with minimum error rate.

Keywords: Machine Learning, Data Mining, Data set, Weka Algorithm.

1 Introduction

Phishing is a form of identity theft in which deception is used to trick a user into revealing confidential information with economic value. Similar forms of identity theft, in which worms or viruses install key loggers, are sometimes also referred to as phishing. This report focuses on phishing involving deceptive electronic messages. Attackers trying to track the server connection when the sender and receiver process of messages take place. they trying to hack the system through phishing pages in which user enter their ID and password mistakenly.

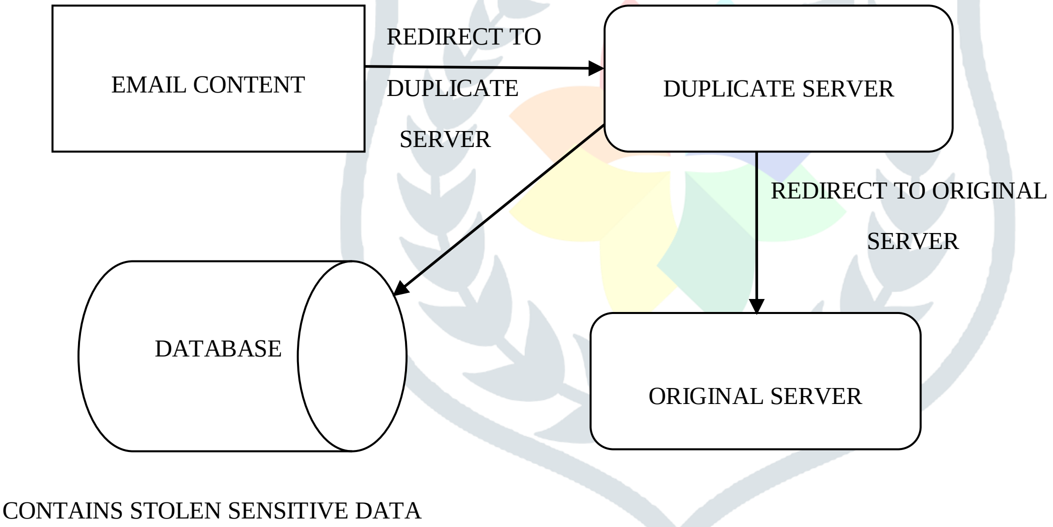

EMAIL CONTENT
REDIRECT TO DUPLICATE SERVER
DUPLICATE SERVER
REDIRECT TO ORIGINAL SERVER
DATABASE
ORIGINAL SERVER
CONTAINS STOLEN SENSITIVE DATA

Fig.1. Architecture of data stealing

Criminals such as phishers can afford to invest in technology commensurately with the illegal benefits gained by their crimes.

Phishers sometimes construct elaborate information flows to cover their tracks and conceal the ultimate destination of compromised information. In some cases, these information flows can contain multiple media, such as compromised “zombie” machines, instant messaging, and anonymous peer-to-peer data transfer mechanisms. In this paper, we will use data mining techniques and exploit the websites features to establish a large labeled training data, and then yield a classifier integrating these features in such a way that new coming websites could be classified correctly. We propose an algorithm for phishing websites detection using data mining classification model.

This paper presented in section 2 related works of Phishing Detection, section 3 presented the classification model and the proposed algorithms, section 4 presented experimental evaluation. A discussion and conclusion are given in section 5 and section 6 respectively.

2 Literature Survey

ZOU FUTAI, PEI BEI and PAN LI[1] Uses Graph Mining technique for web Phishing Detection. It proposed system to detect potential phishing that can't be detect by URL analysis. There is visiting relation utilization between user and website. After anonym zing these data, they have cleansing dataset and each record includes eight fields: User node number (AD), User SRC IP(SRC-IP) access time (TS), Visiting URL (URL), Reference URL(REF), User Agent(UA), access server IP (DSTIP), User cookie (cookie). For a client user, he is assigned a unique AD but a variable IP selected from ISP own IP pool.

In NICK WILLIAMS and SHUJUN LI[2] proposed a system which analysis ACT-R cognitive behaviour architecture model. There is a process which represents validity of web pages following by the HTTPS padlock security indicator. ACT-R map well onto the phishing use case and that work to more fully represent the human security.

In GIOVANNI ARMANO, SAMUEL MARCHAL and N.ASOKAN[3] proposed a use of add on in the browser which is Real-Time Client-Side Phishing Prevention. If any information extracted by any other person, it can use that extracted information and can track if it is phish and warn user. A warning message is displayed in the foreground while the background displays the phishing webpage darkened by a black semi-transparent layer preventing interactions with the website.

In RIAD JABRI and BORAN IBRAHIM[4] proposed a system which analysis Classification model. Proposed system with PRISM algorithm with different attributes and data set. Shows experimental results with data set. Preparatory experiments were conducted on that dataset using the well-known data mining software WEKA to show the classification effectiveness of the machine learning algorithms and to justify the proposed algorithm. Also shows Accuracy and error rate, Number of derived rules.

In HIMANI THAKUR and Dr. SUPREET KAUR[5] proposed a uses technique for web Phishing Detection. It uses Classification of protection against phishing: User education, Software-based defence approaches, Protection at network level, Authentication-based mechanisms, Client-side tools. It also determines phishing categorization: Body-based feature, Subject-based features, URL-based features, Script-based features, Sender-based features.

In HEMALI SAMPAT, MANISHA SAHARKAR, AJAY PANDEY and HEZAL LOPES[6] proposed a use of add on in the browser which is Client-Side Phishing Prevention using the IP Address. If an IP address is used as an alternative of the domain name in the URL, users can be sure that someone is trying to steal their personal sensitive information. Sometime phishers use long URL to hide some doubtful information in address bar. It can detect that things.

3 The Proposed Algorithm

In Bayes algorithm there are BayesNet and NaiveBayes algorithms are applied. Bayesian classifiers are statistical classifiers. There are class membership portabilities, in that portability given sample belongs to a particular class. There are various classifier, they known as Bayesian classifier, based on Bayes' theorem. In Naive Bayesian classifiers effect of an attribute value on given class is independent of the values of other attributes. This type of assumption is known as class conditional independence. In Bayesian classifiers there is high accuracy and speed when we applied large databases.

4 Experimental Evaluation

Experimental Evaluation shows the description of data set and evaluation of the model with accuracy and error rate.

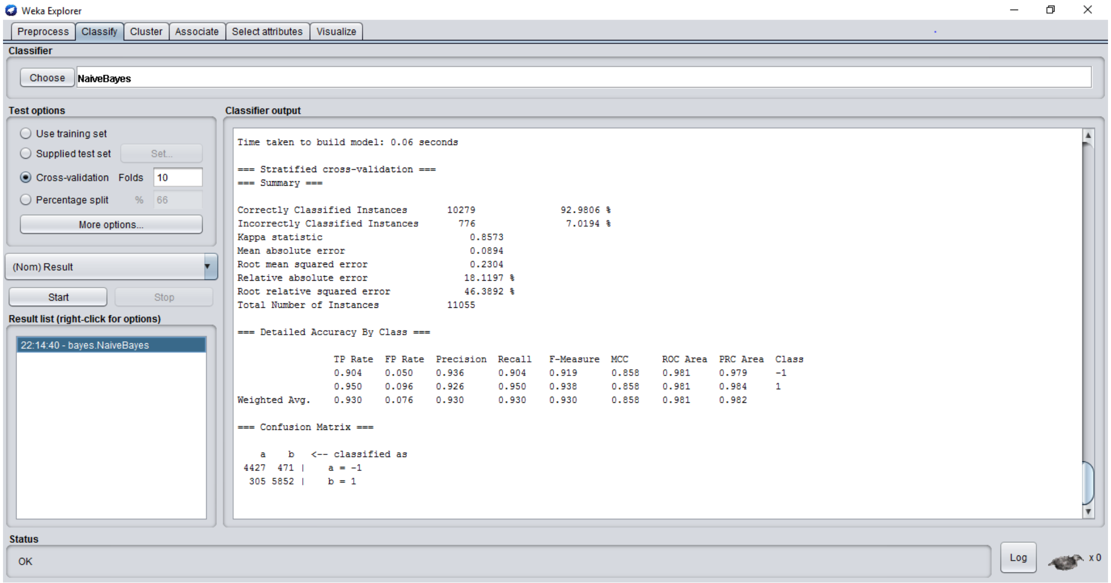
4.1 Data Set

A large number of phishing pages with their respective features were explored for this paper. there are following feature- value pairs were selected based on the frequency analysis. There are 31 attributes we have used and applied it with various algorithms to get results with the accuracy and less error rate.

Experimental results were conducted on this data set using the well-known data mining software WEKA to show the classification effectiveness of the machine learning algorithms and to justify the proposed algorithm. We have tested four algorithms (JRip, J48, Naïve Bayes (NB), BayesNet) to show best results with Accuracy and error rate and Number of derived rules.

5 Experimental Results

We have shown experimental results with NaiveBayes, BayesNet, JRip and J48 algorithms. From all of these J48 gives best result which is 95.88 %. All experiment did with 31 attributes with applying all algorithms.



Classifier: NaiveBayes

Test options: ○ Use training set ○ Supplied test set ● Cross-validation Folds 10 ○ Percentage split % 66

(Nom) Result

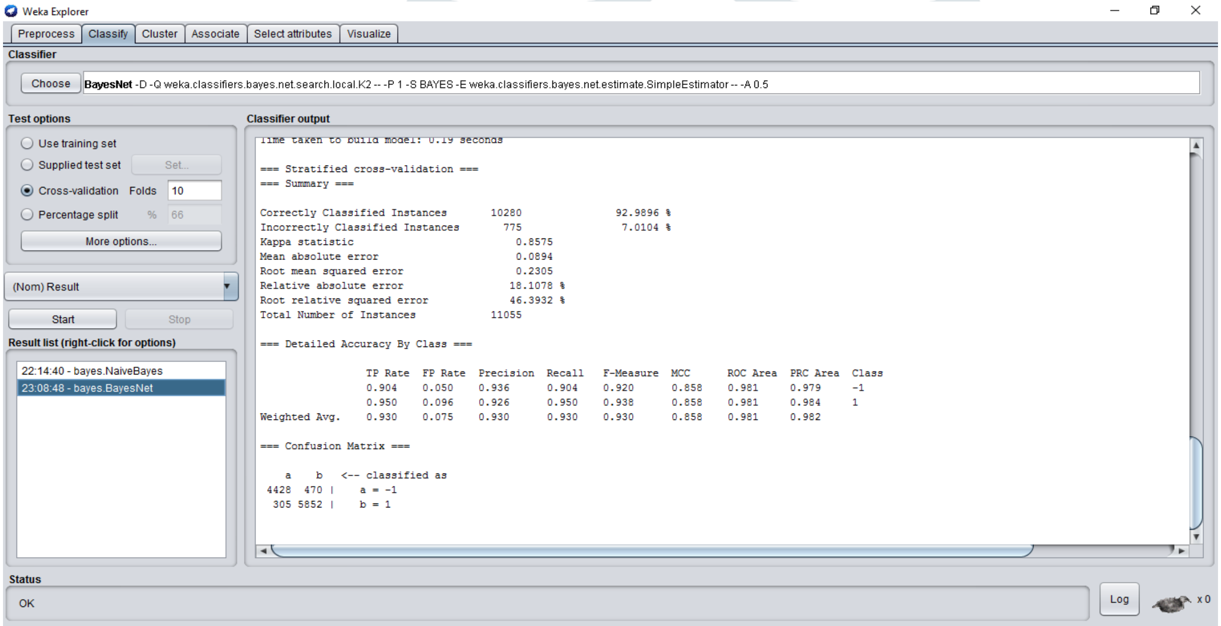
Result list (right-click for options): 22:14:40 - bayes.NaiveBayes

Status: OK

Classifier output

```
Time taken to build model: 0.06 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances       10279               92.9806 %
Incorrectly Classified Instances       776                7.0194 %
Kappa statistic                          0.8573
Mean absolute error                      0.0894
Root mean squared error                  0.2304
Relative absolute error                 18.1197 %
Root relative squared error             46.3892 %
Total Number of Instances            11055

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.904    0.050    0.936      0.904    0.919      0.858    0.981     0.979     -1
                 0.950    0.096    0.926      0.950    0.938      0.858    0.981     0.984     1
Weighted Avg.    0.930    0.076    0.930      0.930    0.930      0.858    0.981     0.982

=== Confusion Matrix ===

    a    b   <-- classified as
 4427  471 |    a = -1
  305 5852 |    b = 1
```

Fig.2. Result with NaiveBayes Algo.



Classifier: BayesNet -D -Q weka.classifiers.bayes.net.search.local.K2 -- -P 1 -S BAYES -E weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.5

Test options: ○ Use training set ○ Supplied test set ● Cross-validation Folds 10 ○ Percentage split % 66

(Nom) Result

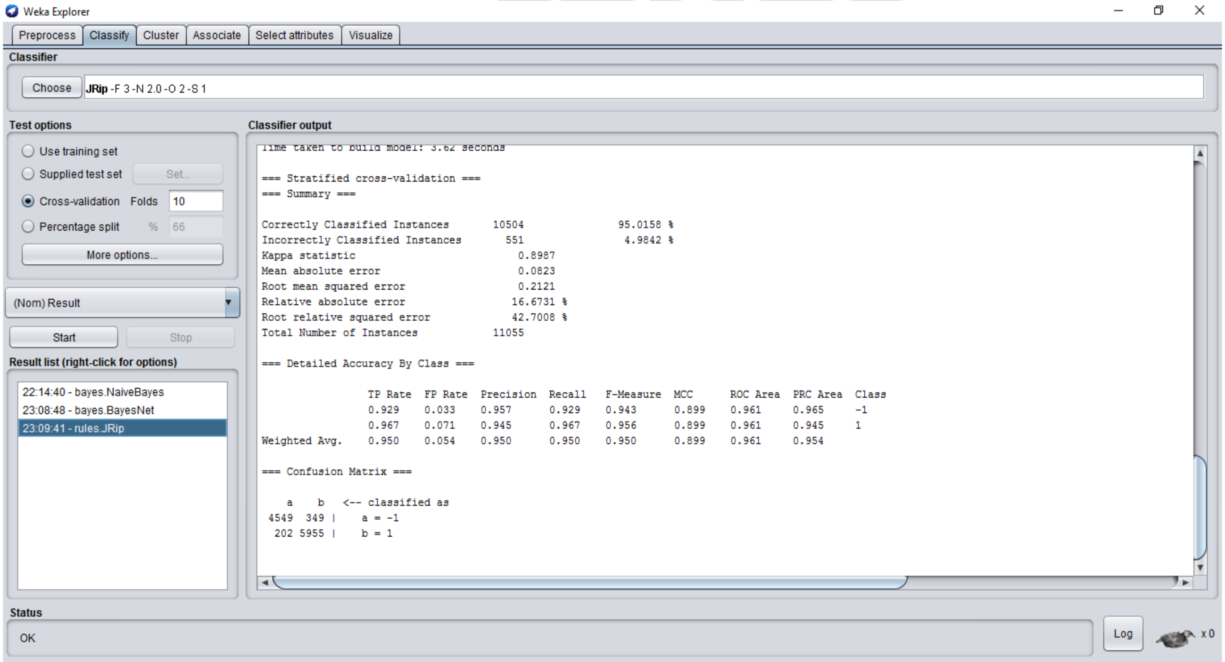
Result list (right-click for options): 22:14:40 - bayes.NaiveBayes; 23:08:48 - bayes.BayesNet

Status: OK

Classifier output

```
Time taken to build model: 0.19 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances       10280               92.9896 %
Incorrectly Classified Instances       775                7.0104 %
Kappa statistic                          0.8575
Mean absolute error                      0.0894
Root mean squared error                  0.2305
Relative absolute error                 18.1078 %
Root relative squared error             46.3932 %
Total Number of Instances            11055

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.904    0.050    0.936      0.904    0.920      0.858    0.981     0.979     -1
                 0.950    0.096    0.926      0.950    0.938      0.858    0.981     0.984     1
Weighted Avg.    0.930    0.075    0.930      0.930    0.930      0.858    0.981     0.982

=== Confusion Matrix ===

    a    b   <-- classified as
 4428  470 |    a = -1
  305 5852 |    b = 1
```

Fig.3. Result with BayesNet Algo.



Classifier: JRip -F 3 -N 2.0 -O 2 -S 1

Test options: ○ Use training set ○ Supplied test set ● Cross-validation Folds 10 ○ Percentage split % 66

(Nom) Result

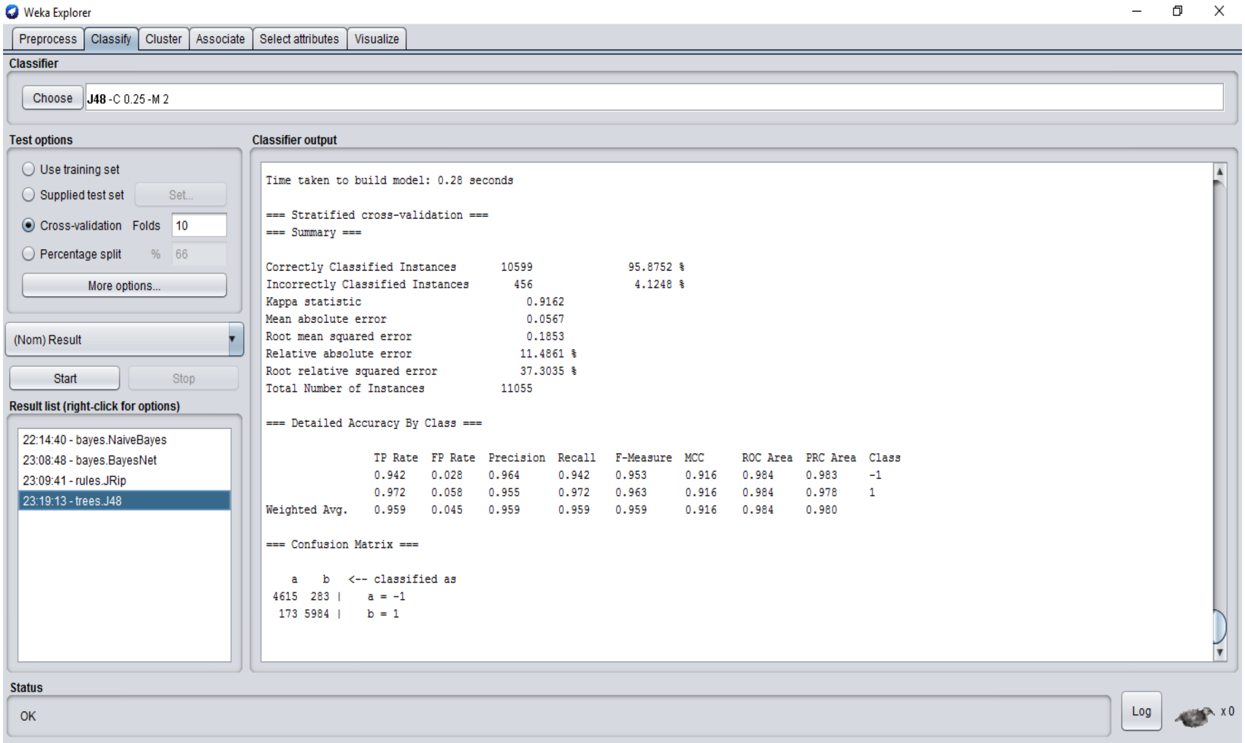
Result list (right-click for options): 22:14:40 - bayes.NaiveBayes; 23:08:48 - bayes.BayesNet; 23:09:41 - rules.JRip

Status: OK

Classifier output

```
Time taken to build model: 3.62 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances       10504               95.0158 %
Incorrectly Classified Instances       551                4.9842 %
Kappa statistic                          0.8987
Mean absolute error                      0.0823
Root mean squared error                  0.2121
Relative absolute error                 16.6731 %
Root relative squared error             42.7008 %
Total Number of Instances            11055

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.929    0.033    0.957      0.929    0.943      0.899    0.961     0.965     -1
                 0.967    0.071    0.945      0.967    0.956      0.899    0.961     0.945     1
Weighted Avg.    0.950    0.054    0.950      0.950    0.950      0.899    0.961     0.954

=== Confusion Matrix ===

    a    b   <-- classified as
 4549  349 |    a = -1
  202 5955 |    b = 1
```

Fig.4. Result with JRip Algo.



Classifier: J48 -C 0.25 -M 2

○ Use training set
○ Supplied test set
● Cross-validation Folds 10
○ Percentage split % 66

(Nom) Result

22:14:40 - bayes.NaiveBayes
23:08:48 - bayes.BayesNet
23:09:41 - rules.JRip
23:19:13 - trees.J48

```
Time taken to build model: 0.28 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances       10599               95.8752 %
Incorrectly Classified Instances       456                4.1248 %
Kappa statistic                          0.9162
Mean absolute error                      0.0567
Root mean squared error                  0.1853
Relative absolute error                 11.4861 %
Root relative squared error             37.3035 %
Total Number of Instances            11055

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.942    0.028    0.964      0.942    0.953      0.916    0.984     0.983     -1
                 0.972    0.058    0.955      0.972    0.963      0.916    0.984     0.978     1
Weighted Avg.    0.959    0.045    0.959      0.959    0.959      0.916    0.984     0.980

=== Confusion Matrix ===

    a    b   <-- classified as
 4615  283 |    a = -1
  173 5984 |    b = 1
```

Status: OK

Fig.5. Result with J48 Algo.

Table 1 Results of WEKA's algorithm

	WEKA Algorithms	Accuracy Rate %	Error Rate %
Websites Data	NB	92.98	18.12
	BayesNet	93.03	18.11
	JRip	95.02	16.67
	J48	95.88	11.49

We have applied four theorems in Weka algorithms that represents the accuracy rate and error rate. J48 Algorithm represents the best results in comparison of other algorithms. There are two types of error rate Relative absolute error and Root relative squared error. Absolute error rate shown in the tabular format. Graph representation shown with accuracy rate in NB, BayesNet, JRip and J48 algorithm.

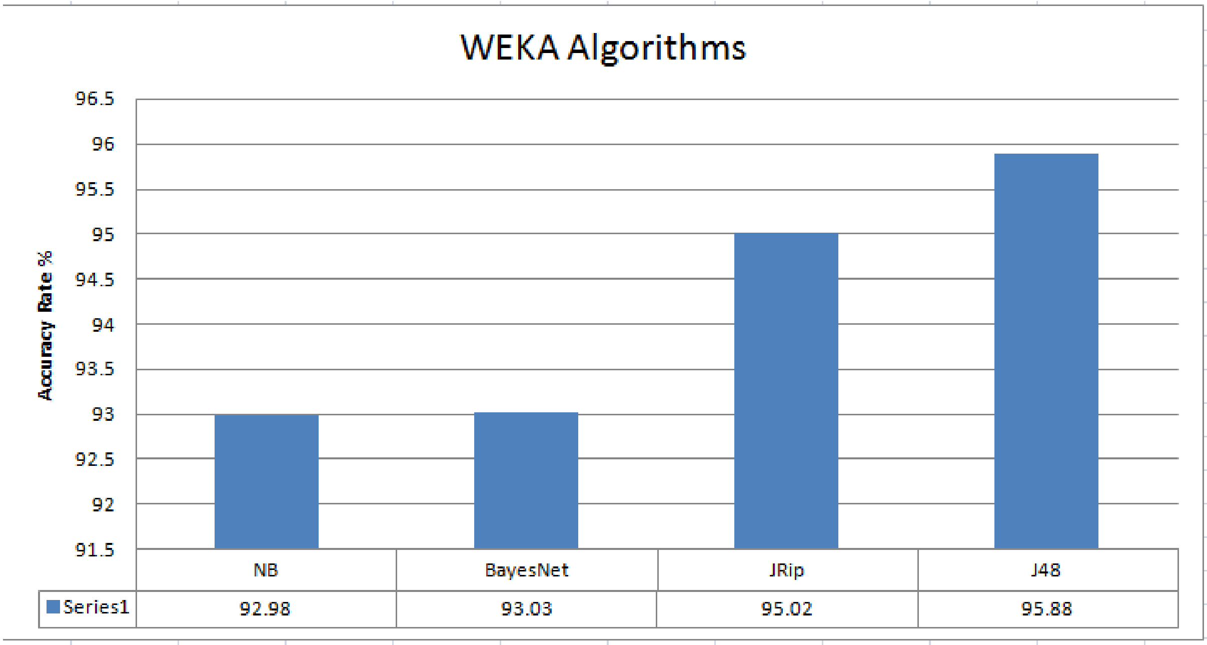


Fig.6. Graphical representation of Accuracy rate

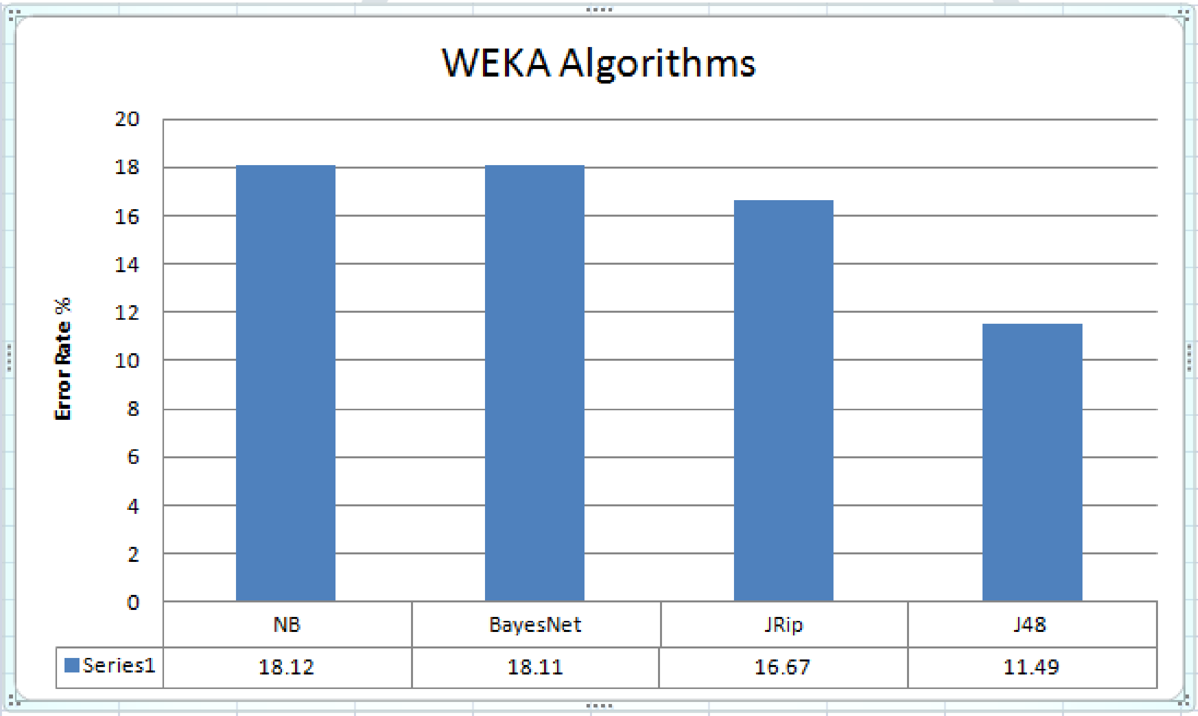


Fig.7. Graphical representation of Error rate

6 Conclusion

We have proposed new phishing websites detection model based on Bayes algorithm. an automatic phishing detection system is presented. The system utilizes a rich set of features, including hard-to-forge features, captured from many different aspects of the corresponding URLs, as well as from distributed vantage points. With using the Majestic top one million list as legitimate URLs and blacklists downloaded hourly from PhishTank as phishing URLs, data collection and machine learning experiments are conducted to evaluate the proposed methodology. In addition, a previous study is analyzed based on our contemporary dataset.

Experiment results show that our approach achieves good accuracy for phishing detection, indicating the effectiveness of the proposed mechanism. The previous study that is analyzed, on the other hand, exposes a performance decline due to the evolution of the phishing ecosystem, while our proposed methodology and feature set demon strates significant superiority. Meanwhile, differences between legitimate and phishing websites are revealed based on the case study of some features.

In this paper there are three to four algorithms are applied to show the highest accuracy and less error rate. Algorithm J48 gives the 95.88% accuracy which is highest then others algorithms. It gives minimum error rate which is 11.49%. We have shown accuracy and error rate comparisons in graphical representation. In which BayesNet, NaiveBayes and JRip represents 93.03, 92.98 and 95.02 % accuracy respectively. And J48 shows 95.88 highest percentage of accuracy.

References

1. Zou Futai, Gang Yuxiang, Pei Bei, Pan Li, Li Linsen. Web Phishing Detection Based on Graph Mining
2. Nick Williams, Shujun Li detection of human in phishing websites: proposed system of ACT-R cognitive behaviour architecture model.
3. Giovanni Armano, Samuel Marchal and N. Asokan RealTime Client-Side Phishing Prevention Add-on.
4. RIAD JABRI and BORAN IBRAHIM analysis Classification model.
5. HIMANI THAKUR and Dr. SUPREET KAUR Web Phishing Detection, Software-based defence approaches.
6. HEMALI SAMPAT, MANISHA SAHARKAR, AJAY PANDEY and HEZAL Client-Side Phishing Prevention using the IP Address.
7. http://docs.apwg.org/reports/apwg_trends_report_q4_2017.pdf.
8. Sidharth Chhabra, Anupama Aggarwal, Fabrício Benevenuto, and Ponnurangam Kumaraguru. Phi.sh/$ocial: the phishing landscape through short urls. In CEAS, 2011.
9. https://majestic.com/reports/majestic-million.
10. Guang Xiang, Jason Hong, Carolyn P. Rose, and Lorrie Cranor. Cantina+: A featurerich machine learning framework for detecting phishing web sites. ACM Trans. Inf. Syst.Secur.,14(2):21:1–21:28,September2011.
11. Angelo PE Rosiello, Engin Kirda, Fabrizio Ferrandi, et al. A layout-similarity-based approachfordetectingphishingpages.InSecurityandPrivacyinCommunications Networks and the Workshops, 2007. SecureComm 2007. Third International Conferenceon,pages454–463.IEEE,2007.
12. Justin Ma, Lawrence K. Saul, Stefan Savage, and Geoffrey M. Voelker. Beyond blacklists: learning to detect malicious websites from suspicious urls. InKDD, 2009.
13. Tu Ouyang, Soumya Ray, Mark Allman, and Michael Rabinovich. A large-scale empiricalanalysisofemailspamdetectionthroughnetworkcharacteristicsinastandaloneenterprise.ComputerNetworks,59:101–121,2014.
14. Ram Basnet, Srinivas Mukkamala, and Andrew H Sung. Detection of phishing attacks: A machine learning approach. In Soft Computing Applications in Industry, pages373–383.Springer,2008.
15. Alexander Moshchuk, Tanya Bragin, Steven D Gribble, and Henry M Levy. A crawler-basedstudyofspywareintheweb.
16. Mahmoud T Qassrawi and Hongli Zhang. Detecting malicious web servers with honeyclients.JournalofNetworks,6(1):145,2011.
17. Andreas Dewald, Thorsten Holz, and Felix C Freiling. Adsandbox: Sandboxing javascript to fight malicious websites. In Proceedings of the 2010 ACM Symposium onAppliedComputing,pages1859–1864.ACM,2010.