6. CONCLUSION

In this paper we did performance analysis of three algorithms namely Support Vector Machine, Naive Bayes and Extreme Learning Machine. The algorithms are analyzed on four parameters they are Accuracy, Precision, Recall and F1 Score.

SVM achieved highest accuracy but ranked 2nd in terms of predicting Phishing URLs after ELM algorithm. ELM algorithm has the highest Precision for predicting URLs as Phishing URLs but ranks 2nd in the race of achieving Accuracy. Naïve Bayes algorithm ranks third in the race of achieving Accuracy after SVM and ELM, but has the highest recall value for predicting URLs as Phishing. Even though ELM ranks 2nd in terms of Accuracy, it ranks 3rd in terms for Recalling URLs as Phishing.

Every algorithm ranks different when analyzed on different parameters. Not necessary that algorithm with highest accuracy will also rank high in terms of other parameters like Precision and Recall.

The Accuracy parameter is not enough when we are analyzing performance of models identifying URL phishing and thus need to focus on Recall and Precision parameter as well. The Recall parameter need to be focused more as it detects False Negatives i.e., predicting Phishing URL as Non-Phishing. Thus, achieving high recall for URL phishing detection model is very important.

7. FUTURE WORK

In this paper, performance analysis is done on three algorithms using 30 features dataset. In future, out of these 30 features, best minimum features will be identified using Forward selection technique and then performance analysis will be done on it. Also, a hybrid approach for URL phishing detection will be implemented with 30 features and also with best minimum features identified earlier and its performance analysis will be done based on parameters like Accuracy, Precision, Recall and F1 Score.

8. REFERENCES

[1] Yasin Sönmez, Türker Tuncer, Hüseyin Gökal, Engin Avci, "Phishing Web Sites Features Classification Based on Extreme Learning Machine", 2018 6th International Symposium on Digital Forensic and Security (ISDFS).

[2] T. Peng, I. Harris, and Y. Sawa, "Detecting Phishing Attacks Using Natural Language processing and Machine Learning", Proc. – 12th OEEE Int, Conf. Semant.Comput. ICSC 2018, vol. 2018-Janua, pp. 300-301, 2018.

[3] Aniruddha Joshi, Tanuja Pattanshetti "Phishing Attack Detection using Feature Selection Techniques". Proeedings on International Conference on Communication and Information processing (ICCIP) 12 Jul 2019 Last revised: 30 Sept 2019.

[4] Bhagyashree E. Sananse, Tanuja K. Sarode "Phishing URL Detection: A Machine Learning and Web Mining-based Approach". International Journal of Computer Applications (0975 - 8887) Volume 123 - No.13, August 2015.

[5] L. MacHad0 and J. Gadge, "Phishing Sites Detection Based on C4.5 Decision Tree Algorithm," in 2017 International Conference on Computing, Communication, Control and Automation, ICCUBEA 2017, 2018, pp. 1-5.

[6] S. Parekh, D. Parikh, S. Kotak, and P. S. Sankhe, "A New Method for Detection of Phishing Websites: URL Detection," in 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 2018, vol. O, no. Icicct, pp. 949-952.

[7] Anjum N Shaikh, Antesar M Shabut, M. A. Hossain, "A Literature Review On Phishing Crime, Prevention Review and Investigation of Gaps", 2016 10th Internation conferences on Software, knowledge, Information Management and Applications (SKIMA).

[8] Joby James, Sandhya L, Ciza Thomas, "Detection of Phishing URLs Using Machine Learning Techniques", 2013 International conference on Control Communication and Computing.

[9] Sneha Mande, Prof. D.S.Thosar, "Detection Of Phishing Web Sites Based On Extreme Machine Learning", Vol-4 Issue-6 2018, IJARIIE-ISSN(O)-2395-4396.