**LITERATURE SURVEY**

1. **J. Shad and S. Sharma, “A Novel Machine Learning Approach to Detect Phishing Websites Jaypee Institute of Information Technology’’.**

In the last few years, many fake websites have developed on the World Wide Web to harm users by stealing their confidential information such as account ID, user name, password, etc. Phishing is the social engineering attacks and currently attacks on mobile devices. That might result in the form of financial loses. In this paper, we described many detection techniques using URL, Hyperlinks features that can be used to differentiate between the defective and non-defective website. There are six main approaches such as heuristic, blacklist, Fuzzy Rule, machine learning, image processing, and CANTINA based approach. It delivers a good consideration of the phishing issue, a present machine learning solution, and future study about Phishing threats by using machine learning Approach.

1. **Y. Sönmez, T. Tuncer, H. Gökal, and E. Avci, “Phishing web sites features classification based on extreme learning machine,” 6th Int. Symp. Digit. Forensic Secur. ISDFS - Proceeding.**

Phishing is a common attack on credulous people by making them to disclose their unique information using counterfeit websites. The objective of phishing website URLs is to purloin the personal information like user name, passwords and online banking transactions. Phishers use the websites which are visually and semantically similar to those real websites. As technology continues to grow, phishing techniques started to progress rapidly and this needs to be prevented by using anti-phishing mechanisms to detect phishing. Machine learning is a powerful tool used to strive against phishing attacks. This paper surveys the features used for detection and detection techniques using machine learning.

1. **T. Peng, I. Harris, and Y. Sawa, “Detecting Phishing Attacks Using Natural Language Processing and Machine Learning,” Proc. - 12th IEEE Int. Conf. Semant. Comput.**

Phishing attacks are one of the most common and least defended security threats today. We present an approach which uses natural language processing techniques to analyse text and detect inappropriate statements which are indicative of phishing attacks. Our approach is novel compared to previous work because it focuses on the natural language text contained in the attack, performing semantic analysis of the text to detect malicious intent. To demonstrate the effectiveness of our approach, we have evaluated it using a large benchmark set of phishing emails.

1. **M. Karabatak and T. Mustafa, “Performance comparison of classifiers on reduced phishing website dataset,” 6th Int. Symp. Digit. Forensic Secur. ISDFS - Proceeding.**

These days, numerous enemy of phishing frameworks are being created to recognize phishing substance in online correspondence frameworks. In spite of the accessibility of hordes hostile to phishing frameworks, phishing proceeds with unabated because of lacking recognition of a zero-day assault, pointless computational overhead and high bogus rates. In spite of the fact that Machine Learning approaches have accomplished promising exactness rate, the decision and the exhibition of the component vector limit their successful location. Phishing is a typical assault on guileless individuals by making them to unveil their one of a kind data utilizing fake sites. In this work, an upgraded AI based prescient model is proposed to improve the effectiveness of against phishing plans. The prescient model comprises of Feature Selection Module which is utilized for the development of a successful element vector. These highlights are removed from the URL, website page properties and site page conduct utilizing the gradual segment-based framework to introduce the resultant component vector to the prescient model. The proposed framework utilizes CNN, KNN AND SVM which have been prepared on a 30-dimensional list of capabilities. AI is an incredible asset used to endeavor against phishing assaults

1. **K. Shima et al., “Classification of URL bitstreams using bag of bytes,” in 21st Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN).**

In present days ,websites are main responsible for the rapid growth of criminal activities in the internet and corresponding activities which results in the many illegal things. So there are many preventive steps to be taken to stop these kind of activities. Here we propose a model which will classify the given URL into any of the three possible classes ,.i.e. Benign, spam and malware. Our model will the detect the classification of the URL without using any websites content.