

Literature Survey

Team No : C01
Team ID : PNT2022TMID08066
College Name : Adhiyamaan College of Engineering(Autonomous)
Department : Department of Computer Science And Engineering
Team Leader : Shree Bhargav RK
Team Member : Santhosh Priyan S
Team Member : Shashank S
Team Member : Yogapriyan M

S.No	TITLE	PROPOSED WORK	TOOLS USED/ ALGORITHM	TECHNOLOGY	ADVANTAGES/ DISADVANTAGES
1	Phishing Website Detection Using Machine Learning Algorithms.	URLs extracting and analyze Various link by check with Backlisting with help of Machine Learning to increase accuracy.	<ul style="list-style-type: none"> Decision Tree Algorithm Random Forest Algorithm Support Vector Machine Algorithm 	<ul style="list-style-type: none"> Machine Learning 	<p>Disadvantage is that the Characteristics are not guaranteed to always exist in such attacks and false positive rate in detection is very high .</p> <p>Advantage is 97.14% detection accuracy using random forest algorithm with lowest false positive rate.</p>
2	Detecting phishing websites using machine learning technique	URL-based anti-phishing machine learning and method URL Net, a CNN-based deep-neural URL detection network.	<ul style="list-style-type: none"> Support Vector Machine K-NN Random forest classification Artificial Neural Network 	<ul style="list-style-type: none"> Machine Learning 	<p>Advantages- Reduces over fitting in decision trees and helps to improve the accuracy</p> <p>Disadvantages- Requires much computational power as well as resources as it builds numerous trees to combine their outputs.</p>

S.NO	TITLE	PROPOSED WORK	TOOLS USED/ ALGORITHM	TECHNOLOGY	ADVANTAGES/ DISADVANTAGES
3	Phishing Website detection using machine learning and deep learning techniques	It discuss the machine learning and deep learning algorithms and apply all these algorithms on our dataset and the best algorithm having the best precision and accuracy is selected for the phishing website detection.	<ul style="list-style-type: none"> Regression Techniques K nearest neighbor Decision Tree Random Forest XG Boost Ada Boost. 	<ul style="list-style-type: none"> Machine Learning , Deep Learning 	<p>Advantage is to eliminate the cyber threat risk level. Increase user alertness to phishing risks.</p> <p>Disadvantage is Negative effects on a business, including loss of money, loss of intellectual property</p>
4	Phishing Website Detection Based on URL	To preserve the confidentiality. develop a user-friendly environment and to prevent or mitigate harm or destruction of computer networks, applications, devices, and data.	<ul style="list-style-type: none"> Learning Model Algorithm Naive Bayes Algorithm Decision tree, Support Vector Machine Artificial Neural Network Sequential Minimal Optimization 	<ul style="list-style-type: none"> Machine Learning 	<p>Advantages – Provide clear idea about the effective level of each classifier on phishing email detection.</p> <p>Disadvantages- Non standard classifier</p>

S.No	TITLE	PROPOSED WORK	TOOLS USED/ ALGORITHM	TECHNOLOGY	ADVANTAGES/ DISADVANTAGES
5	Detection of Phishing Websites by Using Machine Learning-Based URL Analysis	It is seen that current works tend on the use of machine learning-based anomaly detection due to its dynamic structure, especially for catching the zero day attacks	<ul style="list-style-type: none">• Logistic Regression• K-Nearest Neighbor• Support Vector Machine• Decision Tree• Naive Bayes• XG boost• Artificial Neural Network	<ul style="list-style-type: none">• Machine Learning	The biggest disadvantage of these systems is that the small change in the URL prevents matching in the list . Advantage is the strong generalization ability.
6	Phishing Website Detection Based on Deep Convolutional Neural Network and Random Forest Ensemble Learning	It proposes an integrated phishing website detection method based on convolutional neural networks (CNN) and random forest (RF)	<ul style="list-style-type: none">• Linear Regression• K nearest neighbor• Support Vector Machine• Random Forest• XG Boost• Naïve Bayes• RNN Model• CNN Model	<ul style="list-style-type: none">• Machine Learning , Deep Learning	Disadvantage is that the model cannot determine whether the URL is active or not, so it is necessary to test whether the URL is active or not before detection. Advantage is that the Third-party service is independent

THANK YOU