International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 05, May-2022

Cyber Fraud Social Media Application

[1] Mr. Lonari Prathmesh Ganesh
Department Of Informational TechnologyJspm's
Jscoe Hadapsar, Pune, India

[3] Miss. Pandhare Jyotsana Satish
Department Of Informational Technology
Jspm's Jscoe Hadapsar, Pune, India

[2] Miss. Gote Anisha Anand
Department Of Informational TechnologyJspm's
Jscoe Hadapsar, Pune, India

[4] Miss. Divekar Apeksha Vijay
Department Of Informational Technology
Jspm's Jscoe Hadapsar, Pune, India

Under The Guidance of
Prof. D. A. Gaikwad

***Abstract* — In the 20th century, most people are depends on the internet. So the chances of cyber fraud are increased rapidly. The purpose of these frauds is to get the information used by individuals and organizations for transactions and misuse of information. We proposed social media application for detecting phishing URLs or IP addresses based on Extreme Learning Machine Algorithm. Where people can communicate with each other to get more information about these frauds. We proposed an application based on machine learning techniques to detect phishing websites. It compares with 23 parameters and gives us one output based on these parameters. The output is in the form of normal, suspicious, and phishing.**

Keywords: Phishing, Extreme Learning Machine Algorithm, Machine Learning

I. INTRODUCTION

Technology is growing rapidly day by day and with this rapidly growing technology internet has become an essential part of humans daily activities the use of the internet has grown due to the rapid growth of technology and intensive use of digital systems and thus data security has gained great importance the primary objective of maintaining security in information technologies is to ensure that necessary precautions are taken against threats and dangers likely to be faced by users during the use of these technologies phishing is a fraudulent attempt to obtain sensitive information such as usernames passwords and credit card details by disguising as a trustworthy entity in an electronic communication typically carried out by email spoofing or instant messaging it often directs users to enter personal information at a fake website the look and feel of which is identical to the legitimate site information security threats have been seen and developed through time along with development in the internet and information systems the impact is the intrusion of information security through the compromise of private data and the victim may lose money or other kinds of assets in the end internet users can be affected by different types of cyber threats such as private information loss identity theft and financial damages hence using the internet may be suspect for home and official environments identify and defend against privacy leakage efficient analytical tools are required for users to reduce security threats effective systems that can improve self-intervention must be formed using an artificial intelligence-based information securitymanagement system at the time of an attack.

II. RELATED WORK

Santhana Lakshmi and Vijaya used techniques of Machine learning to verify supervised learning algorithms and model the prediction task that Multi-Layer Perceptron. Decision tree and Naive Bayes classifications were used for observing techniques for web Phishing Detection. It can detect. As compared to other learning algorithms the choice tree classifier is more accurate. [1]

Zou Futai, Pei Bei, and Panli projected uses. Graph Mt for some potential phishing that can't be detected by uniform resource locator analysis. It uses contact of the user and website. To induce the dataset from the real traffic of an oversized ISP. After anonyms zing these data, they need a cleansing dataset. Every record includes eight fields: User node number(AD), Visiting URL (URL), User-Agent(UA), User SRC IP(SRC-IP)
access time(TS), Reference URL(REF), access server IP(DSTIP), User cookie(cookie). [2]

Kaytan and Hanbay proposed determining whether phishing websites supported neural networks. Around 30 inputs attribute, and output attribute 1 are used for that experiment. The values 1, 0, and - 1 were used for input attributes, and hence for output attributes Values 1, and - 1 are used. To evaluate the system performance 5-fold cross- validation method was used. The classification accuracy has been measured as 92.45%. And hence the average accuracy has been measured as90.61%. [4]

Yasin Sonmez, Turker Tuncer performs Extreme Learning Machine (ELM) for 30 features. That has phishing websites in a database of machine learning repositories. They compare ELM with SVM and Naive Bayes. These are other methods ofmachine learning. [3]

Li et al proposed a novel approach based on a minimum enclosing ball support vector machine (BVM) to detect phishing websites. It has been aimed at achieving high

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 05, May-2022

speed and high accuracy to detect phishing websites. Studies were done to enhance the integrity of the feature vectors. Firstly, an analysis of the topology structure of the website was performed according to the Document Object Model (DOM) tree. Then, the web crawler was used to extract 12 topological features of the website. Later, the feature vectors were detected by the BVM classifier. The proposed method was compared to the DVM. It was observed that the proposed method has relatively high precision of detecting. In addition, it was observed that the proposed method complements the disadvantage of the slow speed of convergence on large-scale data. It has been shown that the proposed method has better performance than SVM in the experimental results. The accuracy and validity of the proposed system have been evaluated. [5]

III. SYSTEM ARCHITECTURE

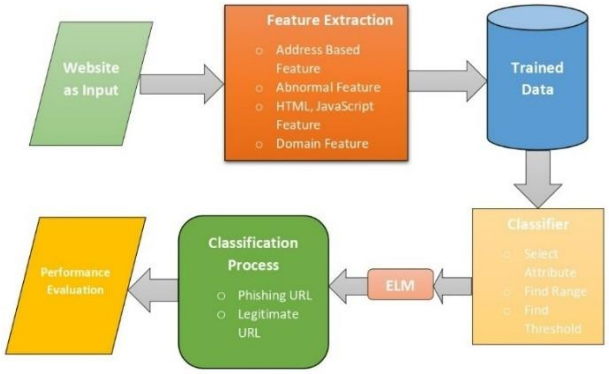

Fig.1. System Architecture

IV. ALGORITHMS

1. Extreme Learning Machine (ELM)

- Extreme Learning Machines (ELMs) are single-hidden layer feedforward neural networks (SLFNs) capable to learn faster compared to gradient-based learning techniques.
- It's like a classical one hidden layer neural network without a learning process.
- This kind of neural network does not perform iterative tuning, making it faster with better generalization performance than networks trained using the backpropagationmethod.

We can understand the working of the Extreme Learning Machine with the help of the following steps:

1. Assign random input of weights and biases.
2. Calculate the hidden layer output.
3. Calculate the output weight.
4. The weights are used to predict the testingdata.

2. Supervised Vector Machine (SVM)

- Support Vector Machine or SVM is one of the foremost common supervised Learning algorithms that is employed for Classification in addition to Regression issues. However primarily, it's used for Classification issues in Machine Learning.
- The goal of the SVM rule is to make the simplest line or call boundary that may segregate an n-dimensional house into categories. This best call boundary is termed a hyperplane.
- SVM chooses the acute points/vectors that facilitate making the hyperplane. In these extreme cases, area units are referred to as support vectors, and therefore the rule is termed a Support Vector Machine.

We can understand the working of the Supervised Machine Learning with the help of the following steps:

1. Import the dataset
2. Explore the data to figure out what theylook like
3. Pre-process the data
4. Split the data into attributes and labels
5. Divide the data into training and testing sets
6. Train the SVM algorithm
7. Make some predictions
8. Evaluate the results of the algorithm

3. Random Forest (RF)

- Random forest is a supervised learning algorithm that is used for both classifications as well as regression. However, it is mainly used forclassification problems.
- As we know that a forest is made up of trees and more trees mean a more robust forest.
- Similarly, the random forest algorithm creates decision trees on data samples and then gets the prediction from each of them, and finally selects the best solution using voting.
- It is an ensemble method that is better than a single decision tree because it reduces over-fitting by averaging the result.

We can understand the working of the Random Forest algorithm with the help of the following steps:

1. First, start with the selection of random samples from a given dataset.
2. Next, this algorithm will construct a decision tree for every sample. Then it will get the prediction result from every decision tree.
3. In this step, voting will be performed for every predicted result.
4. At last, select the most voted prediction result as the final prediction result.

4. Naive Bayes (NB)

- The naïve Bayes algorithm is a supervised learning algorithm, which is based on the Bayes theorem and used for solving classification problems.
- It is mainly used in text classification that includes a high-dimensional training dataset.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 05, May-2022

- Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building fast machine learning models that can makequick predictions.
- It is a probabilistic classifier, which means it predicts based on the probability of an object.
- Some popular examples of the Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles**.**

We can understand the working of the Random Forest algorithm with the help of the following steps:

1. Data Pre-processing step
2. Fitting Naive Bayes to the Training set
3. Predicting the test result
4. Test accuracy of the result (Creation of Confusion matrix).
5. Visualizing the test set result.

V. RESULTS AND DISCUSSION

A. Analysis

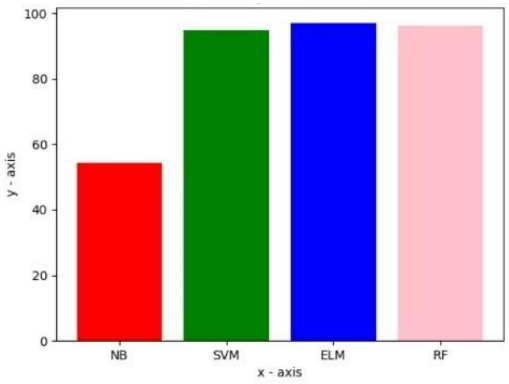

Fig.2. Accuracy Visualization

Four machine learning Algorithms are experimented with to give the accuracy of ELM to be the best and this is used for further performancefor detecting phishing websites.

Fig.3. Accuracy Calculation

B. Results

In this part, we introduce the software design and implementation of each component. We used database and communications among different users of our proposed "Cyber Fraud Social Media Application", system. To develop this application, we have used various technologies. The technologies utilized in this application are:

1. **Android Mobile Operating System:**
   - Android is one of the common important mobile phones
   - OS. Android is available, a free mobile platform, and open source. It has its application security features. The
   - Android products are developed using APIs the same as Java.

The mobile application is an essential part of our system that functions on a smartphone. This system was designed and creates user interfaces based on the Android Operating System.

The first screen is the login screen, as shown in Fig.4. The login screen of this application has three fields, which must be filled in before accessing the application: user email address, password, and IP address. The user will be authenticated by their email address, password, and IP address. It also has a button to direct the user tothe new registration screen.

Fig.4. Login Screen

After correctly adding the information to the login page it gives the message as login successful as shown in fig.5.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 05, May-2022

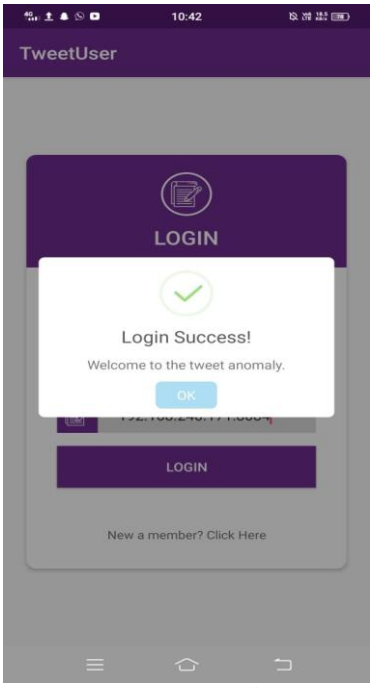

Fig.5. Login Successful

After successful authentication and authorization, the user gets access to the system depending on his/her user type. The home screen has easy options to navigate inside the application by using a menu as shown in Fig.6. And the first option in the navigation drawer is “Home”. In that,we can add a new post.

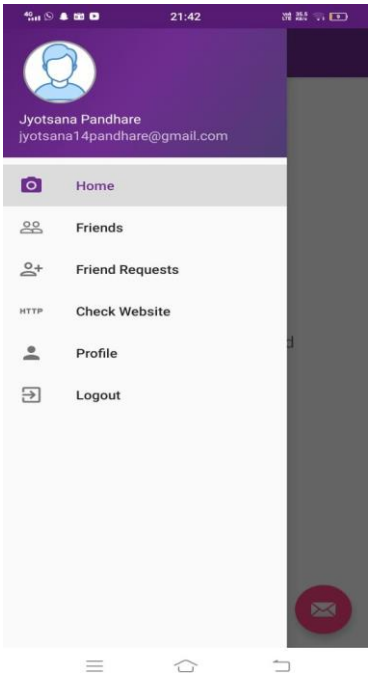

Fig.6. Home Screen

Selecting the image or adding the description to the description box you will be able to add a new post. Then after clicking on add post, it will display the message as “Post Added Successfully” as shown in fig.7. And then you will see the post in the post feed as shown in fig.8.

Fig.7. Post Added Successfully

Fig.8. New Post Feed

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 05, May-2022

As shown in fig.9 and fig.11 you can check whether the URL is phishing or normal. After entering the URL or IP address it gives a message as “This website is safe to browse. Go ahead” as shown in fig.10. Or “This website has phishing features. Do not visit!.” as shown in fig.12.

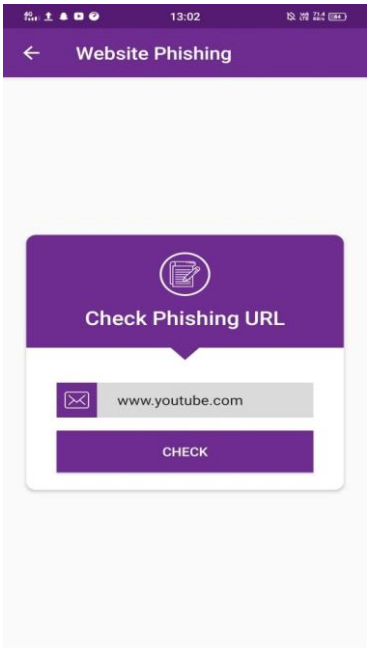

Fig.9. Website Classification

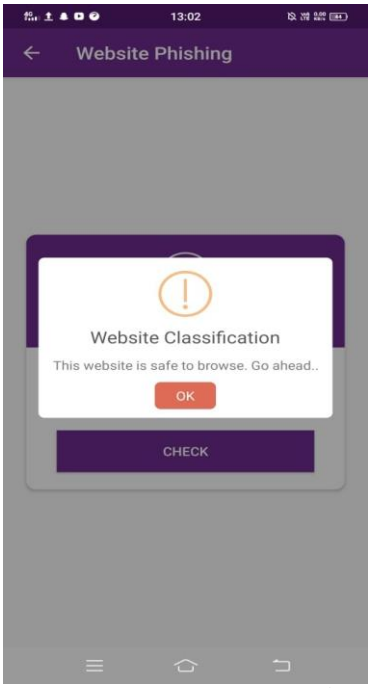

Fig.10. Website Classification

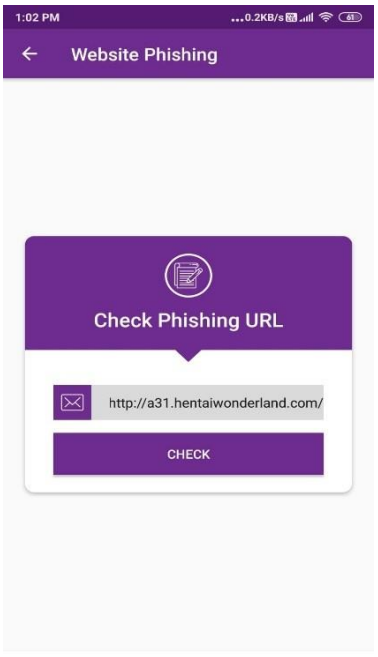

Fig.11. Website Classification

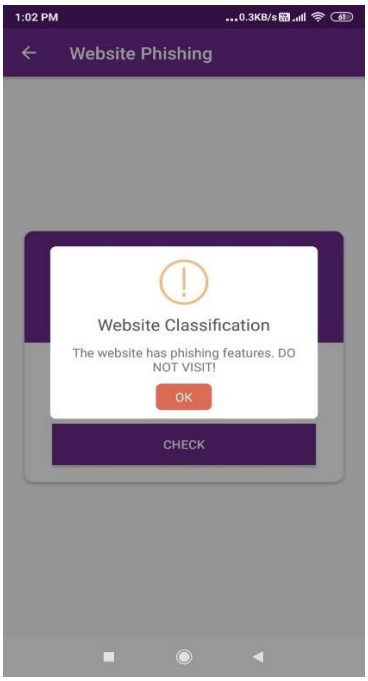

Fig.12. Website Classification

VI. CONCLUSION

The purpose of the application is to make a classification for the determination of one of the types of attacks that cyber threats called phishing. Extreme Learning Machine is used for this purpose. In this study, we can use a data set from the UCI website. In this dataset, input attributes are determined in 30, and the output attribute is determined in 1. Input attributes can be taken in the form of three different values such as 1,0, and -1. Output attribute can take 2 different values which are 1, and -1. As a result of the study, the average

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 05, May-2022

classification accuracy was measured to be 96.93.

VII. FUTURE SCOPE

There are trillions of people using social media in the world however little is known about internet community usage and its determinants that conclusively affect users' vulnerability to phishing attacks on social media networks nevertheless social networking sites are a popular way used by cybercriminals to swindle their targets billions of people log on to their favorite social media accounts that is to say it is a rich source for cybercriminals to gain profit so in the future we will try to extend our work further for social media platforms like Facebook, Instagram, etc. after the successful completion of this proposed work for a bachelor of engineering.

ACKNOWLEDGMENT

I would like to express my special thanks of gratitude to my teachers such as our project guide Prof. D.A.Gaikwad our honorable HOD Prof. Vishwas Kalunge as well as our principal Dr. R.D.Kanphade who gave us the golden opportunity to work on the topic “Cyber Fraud Social Media Application”, which helped us in doing a lot of research and we came to know about so many new things, we are thankful to them. Also, we would like to thank our parents and friends who helped us a lotwith this.

REFERENCES

[1] V. Santhana Lakshmi and M. Vijaya, ”Efficient prediction of phishing websites using supervised learning algorithms”, Procedia Engineering, 30,pp.798- 805, 2012.
[2] Zou Futai, Gang Yuxiang, Pei Bei, Pan Li, Li Linsen “Web Phishing Detection Based on Graph Mining “ 2nd IEEE International Conference on Computer and Communications 978-1-4673-9026-2116 ©2016 IEEE
[3] Yasin Sonmez, TurkerTuncer, based HuseyinGokal, EnginAvci, “Phishing Web Sites Features Classification Based On Extreme Learning Machine ” IEEE 2018 6th International Symposium on Digital Forensic and Security s(ISDFS), DOI: 10.1109/ISDFS.2018.8355342
[4] Mustafa KAYTAN and Davut HANBAY, “Effective Classification of Phishing Web Pages Based On New Rules By Using Extreme Learning Machines”, Anatolian Journal of Computer Sciences, Vol:2 No: 1, pp: 15-36, 2017
[5] Y. Li, L. Yang and J. Ding, "A minimum enclosing ball-based support vector machine approach for detection of phishing websites", Optik, 127(1), pp.345-351, 2016.
[6] Smt.V.Priya Darshini, P.Srilatha, P.Neelima, “Detecting Phishing Website with Machine Learning” International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-3, September 2019
[7] Deepali B.Vaidya1, Poonam R. Dholi2, “Detection of Website Phishing Attack Based on Deep Learning” International Advanced Research Journal in Science, Engineering and Technology Vol. 8, Issue 6, June 2021
[8] Medical Social Media Systems - Implementation of the Android Application|| 2019 11th International Conference on Electronics,Computers and Artificial Intelligence (ECAI) ||publisher: IEEE ||
[9] Cybercrimes in India: trends and prevention, || article published in May 2021 ||
[10] Social media-related cybercrimes and techniques for their prevention || published online: 20 Jun 2019 || © 2019 Tariq Rahim Soomro et al., published by Sciendo this work is licensed under the creative commons attribution 4.0 public license.|| format:-journal | |