Cyber Threat Prediction using ML

Submitted by – Shubham Paikrao, Simranjot Singh Manan, Hrushikesh Jagtap, Shyam Anumalla, Dr. Mahavir Arjun Devmane

(Students of VPPCOE&VA, Mumbai)
Professor Department of Computer EngineeringVasantdada patil college of engineering and visual arts

---

Abstract - *Businesses have long prioritized cybersecurity, but despite expenditures in security procedures and technology, cyberattacks continue to plague all sectors of industry. Cybercriminals have been busy improving their craft over the years, according to evaluations of occurrences, which has led to an increase in cybersecurity incidents overall. The main cybercrime, phishing, needs to be stopped by taking specific safeguards. Phishing is a type of social engineering in which an attacker delivers a false communication intended to deceive a person into giving the attacker critical information or installing harmful software, such as ransomware, on the victim's infrastructure. Phishing attempts are becoming increasingly complex, and many of them replicate the targeted website, enabling the attacker to observe everything the victim does there and to cross any further security barriers with them.*

***Key Words*: Cyber security, cyber attacks, phishing, cybercrime, social engineering, security**

1. INTRODUCTION

Web users today are susceptible to a variety of cyberattacks, including phishing. Phishing is a type of cyberattack that uses social engineering to trick unsuspecting internet users into disclosing personal information such as usernames, passwords, credit card numbers, and other sensitive data by pretending to be a trustworthy website or by sending a malicious URL in an email.

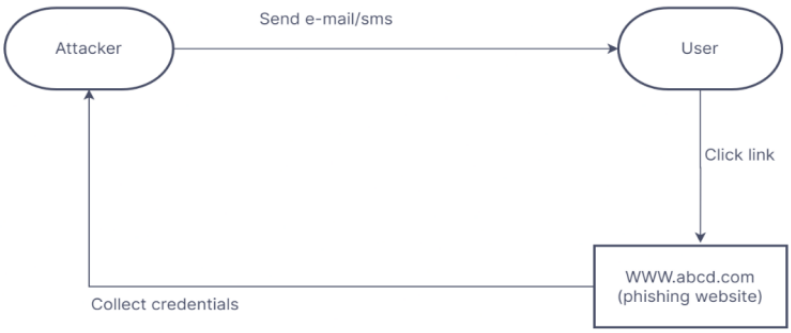
Phishing is an online fraud through which a phisher gains unauthorized access to the user system to lure personal credentials (such as username, password, credit/debit card number, validity, CVV number, and pin) for financial gain. Assaults known as "phishing" incorporate conveying misleading correspondences that appear to be from a solid source. It is usually done through email. The intention is to either steal personal information like credit card numbers and login credentials or infect the victim's computer with malware. Everyone should educate themselves about phishing, a widespread form of cyberattack, to defend themselves. Attackers could be happy to get the credit card number or other personal data from a victim in exchange for cash. Other times, phishing emails are sent to gather employee login credentials or other information for use in a sophisticated attack on a particular firm. Phishing is frequently used as the first step in cybercrime assaults like ransomware and advanced persistent threats (APTs).

Phishing has grown significantly in volume over time, becoming the most usual web threat today. The present economic crisis is an added argument for the great increase in the number of attempts to cheat internet users, both businesses and private ones.

1.1 Aims & Objectives

- To detect phishing and take required precautions to prevent phishing
- To prepare and launch public awareness campaigns to prevent cybercrimes.
- To investigate phishing cases

1.2 Working of Phishing Attack



1.3 Motivation:

The motivation behind creating the cyber threat prediction is the rising phishing attacks although much anti-phishing software is not accessible to everyone

2. System Architecture

Most of the time, the purpose of a phishing attack is to steal data, money, or both. So this cyber threat prediction using ML will provide alerts to the users and make them aware of the cyber attack so they could take the required precaution and avoid any loss of data.

2.1 Current Scenario:

The existing anti-phishing solutions do not work efficiently against phishes because of its continuing growth and day-by-day new tricks. There is a need for rich literature via wider objective, theoretical and practical contributions reeded to meeting cyber security requirements and financial indexes. There is a need to consider new scenarios t-test a deal with novel phishes.

This will help the researcher to stimulate and enhance their interests and attention to the challenges of detection against novel phishes.

The existing anti-phishing solutions do not work efficiently against phishes because of its continuing growth and day-by-day new tricks. There is a need of rich literature via wider objective, theoretical and practical contributions are needed to meet cyber security requirement and financial indexes. There is a need to consider new scenarios to test and deal with novel phishes. This will help the researcher to stimulating and enhancing their interests and attention into the challenges of detection against novel phishes.

The existing anti-phishing solution solutions do not work efficiently against phishes because of their continuing growth and day-by-day new tricks. There is a need for rich literature via wider objective, theoretical and practical contributions needed to meet cyber security requirements and financial indexes. There is a need to consider new scenarios to test and deal with novel phishes. This will help the researcher to stimulate and enhance their interests and attention to the challenges of detection against novel phishes. We reveal that the issues fall into many facts such as features and mechanisms and developed for wider and more effective detection of novel phishes. There is still a big gap in finding an optimum anti-phishing solution against phishes. The existing System works on spam rather than phishing.

2.2 Proposed System:

The recent report gives us a clear picture of how much phishing has increased. The manual approach to investigating cyber-attacks is time-consuming and more error-prone as cyber data attacks proliferate. With the increase in advanced cyber threat attacks with the same patterns, a timely investigation is not possible. There are many systems proposed which analyze and predict threats using various machine learning methods. In this model, we applied machine learning algorithms to analyze and predict Phishing.

In this proposed system:

- Classification algorithms like Linear Regression, Logistic Regression will be used for Spam Filtering,
- Random Forest Algorithm will be used for gaining maximum accuracy.
- Predictive Analytics, is used to study from the data provided, different experiments, and experiences.

2.3 System Backend:

The proposed system consists of different features in the system to predict phishing SMSs, emails, and URLs. This is a system using Machine Learning for the classification of spam SMS, Emails.

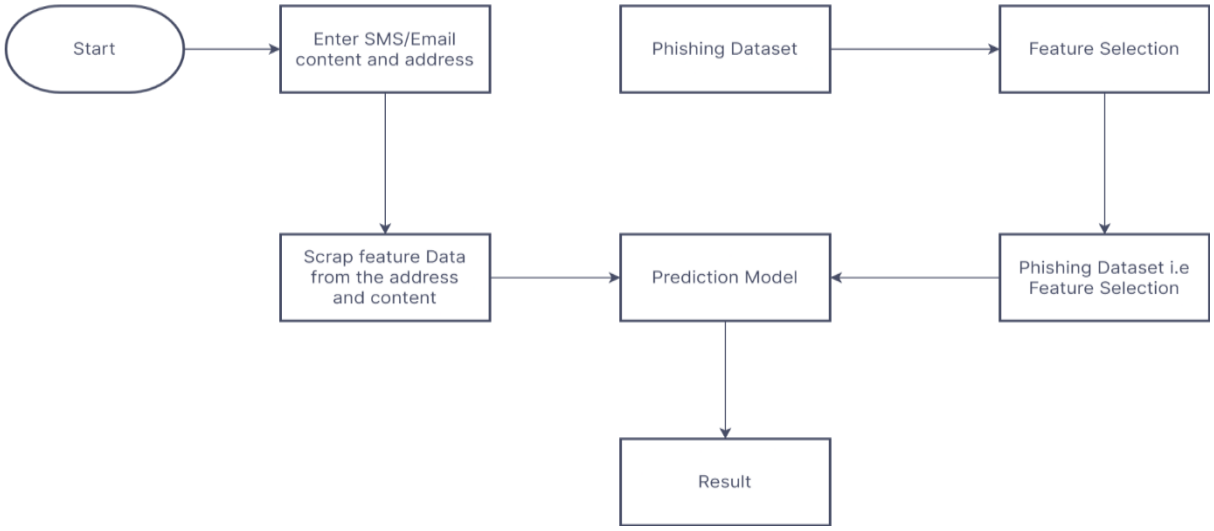
In this cyber threat prediction, when the phishing SMS or mail gets to the user. The Machine learning algorithm checks all the features and datasets for the classification of the mail, and SMS and provides alerts based on the content, URL, address of the SMS, and emails.

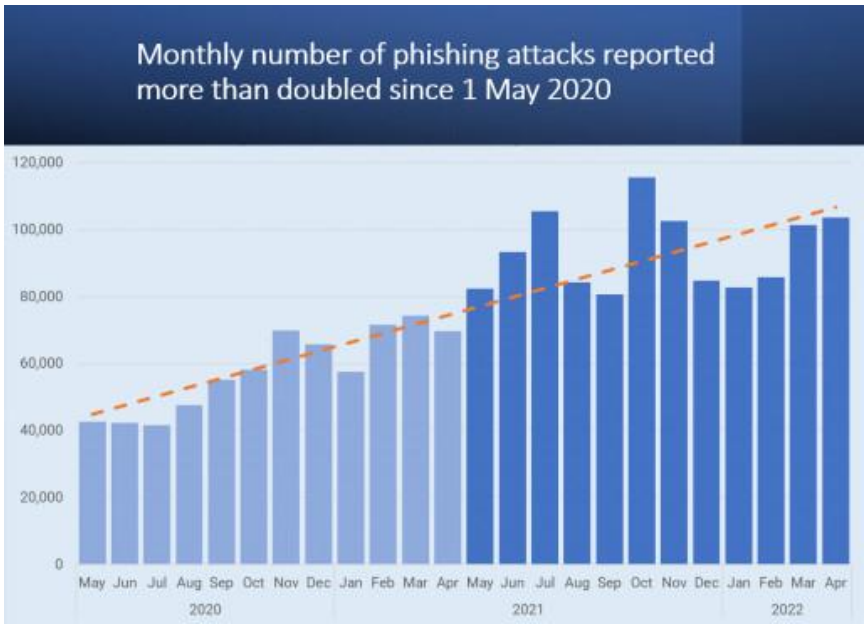
By removing redundant simulation steps, this strategy not only successfully predicts accuracy for unknown datasets, but also reduces the computational cost of the framework by reducing the dimensions of the feature while creating the appropriate structure.

This prediction system predicts whether the received mail or SMS is legitimate or not and provides alerts to the user so that they could take required precautions and avoid the loss of data and control over their devices

- It works on mobile devices
- It blocks malicious URLs
- It scans incoming emails

2.4 Flowchart:





Data Insights of Phishing attacks(2020-22)

3. CONCLUSIONS

3.1 Summary:

We are going to use this model(bot) to predict the various cyber attacks majorly phishing and solve them by predicting them. We will be using ML (Machine Learning) algorithms to overcome these attacks. To overcome such cyber attacks which have occurred previously we have come up with the idea of our project i.e. Cyber Threat Prediction System so the users won't have to face such problems as loss of data and control over the system.

3.2 Future Scope:

This project can be further enhanced to provide greater flexibility and performance with certain modifications whenever necessary. This model can be easily implemented in various situations. We can add new features as and when we require such as an Automated bot using AI and also other anti-phishing features where users connect.

ACKNOWLEDGEMENT

1] Dr. Mahavir Devmane, Department of Computer Engineering, VPPCOE.

2] Prof Asharani Shinde, Department of Computer Engineering, VPPCOE.

REFERENCES

1] https://ieeexplore.ieee.org/document/9448097

2] https://www.researchgate.net/publication/360371905_SMS_Fraud_Detection_Using_Machine_Learning

3]https://www.researchgate.net/publication/349674338_Machine_learning_algorithm_to_identifies_fraud_emails_with_feature_selection