Micro-Project Presentation

Subject Name : Emerging Trends in Computer and Information Technology [22618]

Team Members

SR.No	Roll No	Enrollnment No	Name
1	2251	2100100053	Pratik Pramod Shejwal
2	2254	2100100056	Shravani Bharat Mahajan
3	2260	2100100063	Riya Sunil Kharade

“ MITIGATING PHISHING ATTACKS USING MACHINE LEARNING TECHNIQUES ”

Welcome to this Emerging Trends in Computer and Information Technology micro project. where we'll explore the all details of Mitigating phishing attacks using machine learning.

Index 🧐

1 Introduction

2 Aim and Benefits

3 Output

1)Introduction Of Phishing Attacks
2)Types Of phishing Attacks
3)Impact Of Phishing Attack
4)Introduction Of Machine Learning
5)Deep Learning Techniques For Phishing Detection
6)Feature Selection Methods
7)Real World Datasets for phishing detection
8)Transfer Learning in Phishing Detection
9)Case Studies of Successful Phishing Detection System

4 Conclusion

Introduction :

Phishing tricks online aim to deceive people into sharing private information, and current detection methods often lag behind hackers' innovations. Machine learning offers a solution by training computers to swiftly identify these traps, outsmarting cyber crooks. With machine learning, we stay proactive against threats, safeguarding personal data from falling into the wrong hands. It's like having a digital watchdog, keeping us one step ahead in the cybersecurity game. This approach strengthens our defenses, ensuring our online safety and peace of mind.

1

AIM AND BENEFITS 🏆

◉ AIM:

1. **Spotting Phishing Tricks:**
   Teach computers to quickly identify phishing emails by analyzing sender, content, and language cues.
   1.
2. **Quick Protection:**
   Implement real-time phishing attack detection systems using machine learning for immediate response.

◉ BENEFITS:

1. **Enhanced Security:**
   Machine learning improves email security by swiftly identifying and blocking phishing attempts, protecting personal information.
2. **Cost Savings:**
   Early detection of phishing attacks reduces financial losses and expenses associated with cybersecurity breaches, ensuring financial stability.

.

OUTPUT 🕵️

1) Introduction Of Phishing Attacks

Phishing is a sneaky online threat where scammers trick people into sharing personal info like passwords or credit card numbers by pretending to be trustworthy. But we have a superhero called machine learning! It's like teaching computers to learn from patterns, just like humans. Machine learning analyzes lots of data to catch phishing attempts that humans might miss. It looks at different parts of emails, texts, or websites to figure out if they're real or fake. The cool thing is, machine learning keeps getting better! It learns from each new trick and improves at stopping them. So, by using machine learning, we can make it tough for phishers to trick us and keep our info safe online.

2)Types Of phishing Attacks

1. **Email Phishing:**
   Fake emails from trusted sources.

2. **Spear Phishing:**
   Targeted, personalized email scams.

3. **Clone Phishing:**
   Fake copies with malicious content.

4. **Whaling:**
   Phishing aimed at high-profile individuals

5. **SMiShing (SMS Phishing):**
   Deceptive texts for personal information.

OUTPUT 🕵️

3) Impact Of Phishing Attack

1. **Financial Loss:**
   - Tricks lead to unauthorized transactions.

2. **Identity Theft:**
   - Personal info stolen, fake activities follow.

3. **Data Breaches:**
   - Exposes info, risks, fines, reputation damage.

4. **Reputational Damage:**
   - Phishing tarnishes company image, trust diminishes.

5. **Emotional Distress**:
   - Privacy invasion causes anxiety, distress.

6. **Trust Issues:**
   - Erodes trust, hesitancy in transactions.

4)Introduction Of Machine Learning

Machine learning is like teaching computers to think and learn like humans. It's a way for computers to learn from data and make decisions without being explicitly programmed. Imagine if you could teach your computer to recognize faces in photos or predict which movies you might like to watch. That's what machine learning does! It's used in lots of things we use every day, like voice assistants, recommendation systems, and even self-driving cars. Instead of following strict rules, machine learning algorithms can learn from examples and improve over time. So, it's like giving computers the ability to learn from experience, just like we do.

OUTPUT 🕵️

5)Deep Learning Techniques For Phishing Detection

1. **Microsoft Defender for Office 365:**
   Advanced ML scans emails, isolates threats.
2. **PhishMe (Cofense):**
   Trains users, detects phishing with ML.
3. **Proofpoint Email Protection:**
   ML scans, blocks phishing, provides training
4. **Symantec Email Security:**
   ML detects, blocks evolving phishing tactics.
5. **Google's Safe Browsing:**
   ML spots phishing sites, warns users.

6)Feature Selection Methods

"It is a process of automatically or manually selecting the subset of most appropriate and relevant features to be used in model building."

Below are some benefits of using feature selection in machine learning:

1. It helps in avoiding the curse of dimensionality.
2. It helps in the simplification of the model so that it can be easily interpreted by the researchers.
3. It reduces the training time.
4. It reduces overfitting hence enhance the generalization.

7)Real World Datasets for phishing detection

Machine learning datasets are of data used to train and test machine learning models.

- **Creating a dataset for machine learning (ML) step:**

1. **Training Data Collection:**
   Gather relevant data for training your model.

2. **Data Storage:**
   Keep your collected data in a safe and organized place.

3. **Data Preparation:**
   Clean and format your data to make it usable for training.

4. **Quality Control:**
   Check your data for accuracy, outliers, and biases.

5. **Rescaling:**
   Adjust the scale of your data to ensure fair comparison between features.

- **Various datasets available for training and testing machine learning models for phishing detection**

1. **Phishing Websites Dataset (UCI):**
   Features from URLs aid model training.

2. **Phishing Email Corpus (Enron):**
   Enron emails test phishing algorithms.

3. **Websites Phishing URL Dataset (Kaggle):**
   Labeled URLs for website model training.

4. **Microsoft Malware Prediction Dataset:**
   -Telemetry data for detecting phishing.

OUTPUT 🕵️

8)Transfer Learning in Phishing Detection

Transfer learning involves leveraging knowledge gained from solving one problem and applying it to a different but related problem.

- **Steps to Use Transfer Learning:**

1. **Training a Model to Reuse it:**If you lack data for Task A, train a deep neural network on a related Task B with plenty of data. Then, adapt this model for Task A by either using the entire model or just specific layers, depending on how similar the inputs are.

2. Using a Pre-Trained Model: Utilize existing pre-trained models and decide how many layers to reuse and retrain based on your specific task. Many pre-trained models are available, including those in Keras, which can be fine-tuned for your specific needs.

3. Extraction of Features: Use deep learning to automatically identify important features through representation learning. This method often produces better results than manually designing features, but domain knowledge is still essential for selecting relevant features.

4. Extraction of Features in Neural Networks: Neural networks inherently learn which features are important, reducing the need for manual feature engineering. Representation learning algorithms can efficiently identify significant feature combinations, even for complex tasks, minimizing the need for manual feature selection.

OUTPUT 🕵️

9)Deep Learning Techniques For Phishing Detection

1. **Long Short-Term Memory (LSTM) Networks:** LSTM networks are a type of RNN that can capture long-range dependencies in sequential data, enabling them to identify subtle patterns or anomalies in email content or user interactions indicative of phishing attacks.

2. Deep Autoencoders: These can compare normal and phishing content, like email text or website features, to spot differences and catch phishing attempts.

3. Generative Adversarial Networks (GANs): Can create fake phishing examples to help train anti-phishing models better.

4. Capsule Networks: These can understand the relationships between different parts of emails or websites to find phishing signs based on how things are arranged.

5. Attention Mechanisms: They focus on the important parts of data, like specific words or images, to help catch phishing attempts more accurately.

References

1. **We Referred Book “A MACHINE LEARNING APPROACH TO PHISHING DETECTION AND DEFENSE” by Oluwatobi Ayodejo Akanbi and his team**
    -

1. **We Referred Following Link:**
    -

    a. **https://en.wikipedia.org/wiki/Machine_learning**

    b. **https://www.geeksforgeeks.org/machine-learning/**

    c. **https://www.phishprotection.com/phishing-awareness/machine-learning-helps-fighting-phishing-attacks**

Conclusion 🌎

Machine learning helps spot phishing emails quickly and accurately, protecting us from scams. It saves time and money by automating the process and catching tricky emails that could fool us. But we still need to learn about phishing and use strong passwords to stay safe online. Remember, combining machine learning with other security measures gives us the best protection against scams.

THANKYOU!!