Develop deep learning software with automated filtering of email phishing to counter phishing attacks

Cheah Khai Xuan[1], Dr. Intan Farahana Binti Kamsin[2], Salmiah Binti Amin[3], Nur Khairunnisha Binti Zainal[4]
[1, 2, 3, 4]Asia Pacific University of Technology and Innovation, Technology Park Malaysia, Bukit Jalil, Kuala Lumpur, Malaysia.
[1]cheahkhaixuan@gmail.com, [2]intan.farahana@staffemail.apu.edu.my, [3]salmiah@staffemail.apu.edu.my,
[4]khairunnisha.zainal@staffemail.apu.edu.my

Abstract **- Tech-savvy con artists and identity theft criminals carry out email phishing attacks online. To fool you into giving critical information like bank account passwords and credit card details, they employ spam, bogus websites that seem just like legitimate websites, email, and instant chats. Therefore, this research paper is to develop deep learning software with auto-filtering the phishing email from the sender for prevents unauthorize spoofing and using the survey method to collect the respondent perspective about the phishing email with open and closed questions. Therefore, by using this system will help the users to protect their confidential data.**

Index Terms **- Deep Learning, Software, Phishing attacks**.

1. Introduction

Phishing is a type of cybercrime in which a person acting as a genuine organization contacts a target or targets by email, phone, or text message to persuade them to provide sensitive data such as personally identifying information, account banking and credit card information, and passwords [1]. Besides, phishing has been one of the most common forms of cyberattacks since the 1990s. Furthermore, phishing messages and strategies are becoming more complicated, yet it remains one of the most popular and hazardous scams. Apart from that, the term "phishing" is pronounced exactly how it is written, which is to say like the word "fish" – imagine an angler casting a baited hook out there (the phishing email) and hoping you bite. The "ph" was possibly prompt by the word "phreaking," short for "phone phreaking," an early technique of hacking that included playing sound tones into telephone handsets to gain free phone calls [9]. Furthermore, phishing attacks have twelve different types of attacks which are spear phishing, whaling, vishing, etc [2]. Therefore, this research is focusing on email phishing where the attacker dupes a victim into opening a link or downloading attachments that include malware in the content by utilizing social engineering to mimic a well-known brand.

2. Research Background / Literature Review

2.1. Deep Learning

Deep Learning is a machine learning technique that uses multiple layers of information processing units to exploit classification patterns and attributes or to learn by representation. Besides, Deep learning algorithms can be divided into three classes base on whether they have been trained to produce the desired outcomes. Unsupervised, supervised, and hybrid are the three types of subgroups [6]. In supervised learning, the user instructs the algorithm to provide an answer based on a pre-defined data set that has been labeled. For supervised learning tasks, classification, and regression techniques such as random forests, decision trees, and support vector machines are frequently utilized. However, the methods in unsupervised machine learning create responses from unidentified and unlabelled data. Unsupervised approaches are often used by data scientists to uncover patterns in fresh data sets. Unsupervised machine learning frequently uses clustering methods such as K-means [11]. Furthermore, estimates of memory capacity are critical for accurately sizing models, inputs, and GPUs. In order to evaluate deep learning training performance, it's also crucial to look at GPU compute and memory bandwidth needs [12]. Therefore, based on the research above, this software will be suitable for using hybrid algorithms to develop the detection of phishing emails.

2.2. Software

The instructions that teach a computer what to perform are known as software. The total set of programmes, operations, and routines related with the operation of a computer system is referred to as software. Besides, system software and application software are the two main types of software [15]. System software, often known as an operating system for larger computers devices and firmware for smaller embedded systems, is a form of programme that runs directly on the hardware. Examples of system software include mbed OS, Contiki OS, Zephyr, and others [3]. In addition. application software is the sort of software we use to connect with computers and complete tasks. It is how we interface with computers and conduct actions on their hardware and we may install and remove application software such as games, word processors, browser email, etc [8]. Apart from that, the software is often saving on a hard drive or magnetic diskette, which is an external long-term memory device. The computer reads the programme from the storage device and temporarily stores the instructions in random access memory when it is in use (RAM). The act of storing and then following out instructions is referred to as "running" or "executing" a programme. Firmware, or "hard software," refers to software programmes and operations that are permanently stored in a computer's memory utilizing read-only (ROM) technology. Thus, this software will be developing an application software that the users need to install by using its feature to filter phishing emails.

2.3. Phishing Attack

The phishing attack has become one of the most common threats to internet users, governments, and service providers. In a phishing attack, the attacker uses fake emails or false websites

to obtain the client's data such as user account login details, credit/debit card numbers, etc. Besides, phishing email victims mistakenly believe these sites are affiliated with reputable organizations like Amazon or Google and are thus duped into logging in and giving personal information to the attacker [4]. Furthermore, according to this research paper [5] software as a service (SaaS) and webmail websites were the most popular targets for phishing in the third quarter of 2019. Phishers continue to gather credentials for these types of websites, which they then use to carry out business email compromises (BEC) and get access to corporate SaaS accounts. Moreover, thereare 6 common types of phishing attacks which are deceptive phishing, spear phishing, clone phishing, whaling, link manipulation, and voice phishing [7]. Apart from these types of phishing attacks from the research paper [7], this research paper is mainly focused on email phishing for preventing the attacker obtain the client's data.

2.4. **Similar System**

2.4.1. D-Fence

D-Fence is an email detection system that is effective, flexible, and efficient. Besides, D-Fence used several machine learning and deep learning categorization models in a multi-modular strategy that are structure module, text module, and URL module in Figure 1. Structure module that detects phishing assaults on an email's header and HTML structure using a tree-based classifier, text module trains a classifier using embeddings supplied by a cutting-edge language modelling approach BERT and URL module using a deep-learning model to categorize URL strings detection capability is enhanced by the use of various modules that complement each other [14].

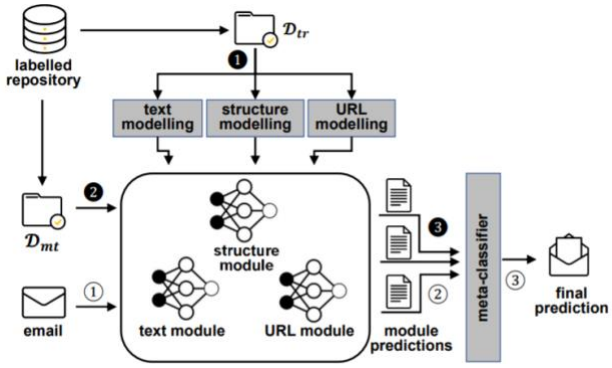

Figure 1.0: D-Fence workflow
Source: ***https://www.researchgate.net/publication/354424369_D-Fence_A_Flexible_Efficient_and_Comprehensive_Phishing_Email_Detection_System***

2.4.2. SAFE-PC

SAFE-PC system is designed for thwart common phishing mail detection techniques. It's been tweaked to extract elements from emails that have been discovered to reoccur in phishing efforts, such as the presence of terms like "account," "suspend," and "expire," a common phishing detecting feature. Besides, it includes engineering to combat misleading phishing methods such intentional misspellings, white space characters, and substitute characters that seem identical to the original characters. Moreover, it employs natural language processing techniques, such as Named Entity Recognition and Freebase, to build "higher level" features [10].

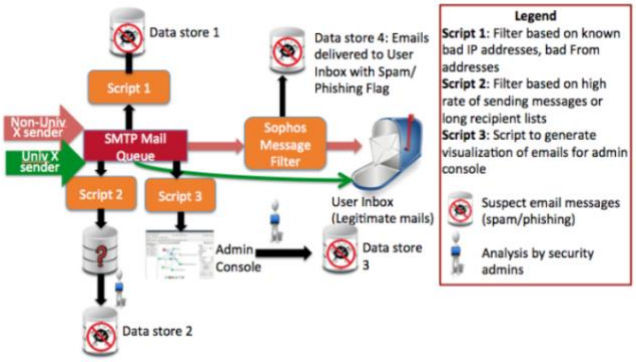

Figure 2.0: SAFE-PC workflow
Source: ***https://ieeexplore.ieee.org/document/8440723***

The table below shows comparison between systems:

	D-FENCE	**SAFE-PC**	**Proposed System**
Machine Learning	✓	✓	✓
Deep Learning	✓	X	✓
Online Learning	X	✓	✓
Accuracy	98.5%	96.5%	99.6%

Table 1.0: Comparison Table

Based on table 1 shows that both systems have implemented the machine learning algorithms in their phishing email detection system. Moreover, the D-Fence system also has applied the deep learning algorithms, but SAFE-PC does not have applied. Furthermore, the SAFE-PC system has online learning that will automatically feed the collected emails into the SAFE-PC periodically. Lastly, the accuracy of detecting the phishing email of D-Fence system is higher than SAFE-PC system. Therefore, this research will fulfil the gap by developing a detection phishing email system.

3. **Problem Statement**

Phishing instances are becoming more common each year, and it is becoming a more popular attack vector for crooks. Email phishing is one of the commonest attacks for victims to easily fall for, the cyber-attack will stratagem the message contains something they want or need to the email recipient such as a letter from their bank or a message from a colleague and open the link or download an attachment. Besides, the attackers pose as a trusted entity of some kind, usually a real or plausibly real person or a corporation with which the victim might do business in order to steal sensitive information from the victim, such as login credentials and credit card numbers [9]. For example, regarding the famous phishing email attack case named Operation Phish Phry. Hundreds of banks and credit card clients were targeted in this well-known cyber-attack, as they got official-looking emails leading them to fraudulent financial websites [13]. Therefore, many of the victims were duped by

the attack because they typed their account numbers and passwords into fake forms, giving cybercriminals easy access to their personal information.

4. Aim and Objectives

4.1. Aim

To develop deep learning software with auto-filtering the phishing email from the sender for prevents unauthorized spoofing.

4.2. Objectives

- To classify each email content by implement the deep learning features and move the email to junk that has the sign of phishing email.
- To prevent users' data from leaking into attackers by labeling the level of the phishing email.
- To improve the user's awareness and email security against phishing emails.
- To reduce the number of victims getting scams by phishing emails.

5. Research Questions

- How the deep learning able to classify the phishing email to move the email to junk?
- How to label the level of phishing email to prevent user's data leaking?
- What awareness do users need to improve?
- What scams will be reducing the number of victims getting scams?

6. Significant of The Research

This research paper aims to investigate the importance of filtering phishing emails that could help any of the organization and users to keep the junk out of the email inboxes and it improves the quality of business communications by ensuring that they function efficiently and are only utilized for their intended purpose. Besides, email spam costs businesses up to $20.5 billion every year, according to Radicati Research Group Inc., and the sum is expected to continue to climb [16] . Therefore, the reason for running this project is to prevent the user and organization data from being expose to an attacker from a phishing link, download an attachment, etc. Thus, the features of this develop application will have labeled the level of the phishing email with the highest percentage the highest the email will be considered as a phishing email. Moreover, the user can decide the application with an auto filtered email that is considering as a phishing email will move it into junk. Thence, this application is developing for the users to prevent opening the of malicious content that contains malware from the email.

7. Methodology

This study paper's main goal is to propose an idea with creating automate filtering of email phishing with deep learning algorithms to counter the email phishing attack with investigate the users using the survey method to captures their perspective about email phishing. Thus, this research paper will be using quantitative research for collecting the data from the user. The reason for using this method is because it is based on tangible data and has a smaller number of variables. This can aid in the removal of biases from the research and improve the accuracy of the results and another benefit is it can easily obtain a large sample of data in a limited time. Moreover, this research will be using probability sampling to select the respondent randomly and each of the respondents has an equal probability that will be chosen for this research.

Apart from that, this research used an online software that allows creating the survey by using the Google Forms software. As the transmission of questions is simple and easy to carry out. Besides, there will be no time limit for the users to fill up the survey. The survey will need to collect with a minimum of 200 respondents to represent it as a large survey and the expectation will have 300 respondents will be conducting this survey. Most of the participants are from APU students and lectures. Furthermore, the questions will involve 2 types of questions which are open questions and closed questions. For open questions the objective questions and subjective questions were used in the survey while for closed questions only the list questions are used.

8. Overview of The Proposed System

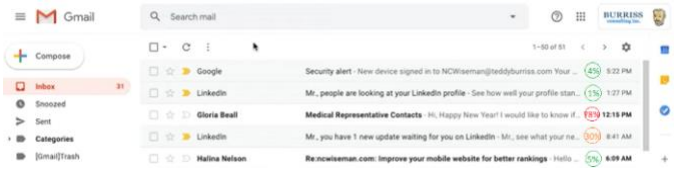

Figure 3.0: Gmail inbox

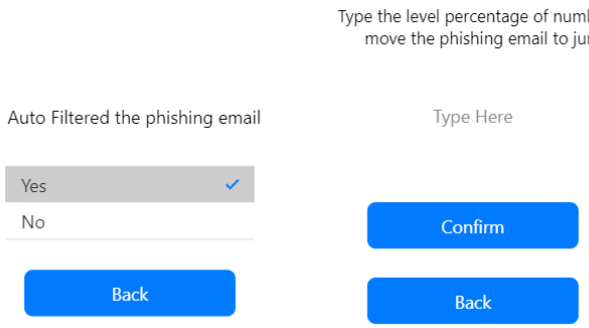

Figure 4.0: Filtering email phishing options

Figures 3 and 4 are examples prototype of users using the filtering phishing email application. Besides in figure 3, the system will display each of the emails with the level of the percentage that has the potential that the email content has contained malware or the email is not from the legitimate companies, etc. Moreover, the users can be easily and quickly identifying the phishing email with the color design shown in figure 3. Apart from that, the other feature of this application is the system will need to approve by the user to automatically filter the phishing email. If the user approves the system to automatically filter the phishing email, then the user is allowing the system to move the phishing email to the junk by entering the level of the percentage shown in figure 4.

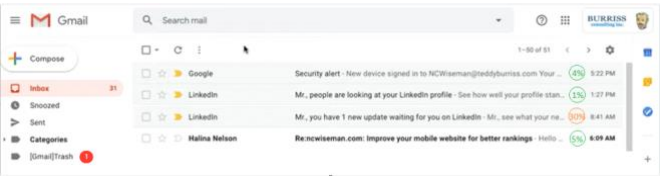

Figure 5.0: Gmail inbox (2)

In figure 5 prototype shown that after the user has been approved the system to auto filtered the phishing email, the system will notify the user with the number beside the trash. In addition, the system can be used in all different type of email platform such as Yahoo Mail, Microsoft Outlook, GMX Mail, etc. Thus, with these features that will decrease the cases of users getting scam from the attackers.

9. Conclusion

As cyberattacks get increasingly sophisticated, so are the strategies for detecting and preventing them. In addition, fileless intrusions continue to elude malware detection. As a result, this article has examined email phishing and produced an original approach with automated email phishing filtering that uses deep learning algorithms to provide users with email security. Although, this system has the limitation of features for countering phishing emails but it will help the users to protect their confidential data without leaking it to the attacker.

References

[1] Phishing.org. (n.d., n.d. n.d.). *Phishing.org*. Retrieved from https://www.phishing.org/what-is-phishing

[2] SecuritySorecard. (n.d., n.d. n.d.). *securitysorecard*. Retrieved from https://securityscorecard.com/blog/types-of-phishing-attacks-and-how-to-identify-them.

[3] Amiri-Kordestani, M. &. (2017). In Free and Open Source Software Conference. *A survey on embedded open source system software for the internet of things*, 27-32.

[4] Bagui, S. N. (2019). 2019 International Conference on Cyber Security and Protection of Digital Services (Cyber Security). *Classifying phishing email using machine learning and deep learning*, 1-2.

[5] Basit, A. Z. (2020). Telecommunication Systems. *A comprehensive survey of AI-enabled phishing attacks detection techniques*, 1-16.

[6] Benavides, E. F. (2020). Developments and advances in defense and security. *Classification of phishing attack solutions by employing deep learning techniques: A systematic literature review*, 51-64.

[7] Bhavsar, V. K. (2018). Int. J. Comput. Appl. *Study on phishing attacks*, 27-29.

[8] Dahar, N. N. (2020). *Interaction Design of Softwares: An HCI Perspective*.

[9] Fruhlinger, J. (2020, September 4). *CSO*. Retrieved from https://www.csoonline.com/article/2117843/what-is-phishing-examples-types-and-techniques.html.

[10] Gutierrez, C. N. (2018). IEEE Transactions on Dependable and Secure Computing. *Learning from the ones that got away: Detecting new forms of phishing attacks*, 988-1001.

[11] Hans-Dieter, W. &. (2017). *Machine Learning, Deep Learning, and AI: What's the Difference?*, 5.

[12] brahim, K. Z. (2021). Benchmarking and Simulation of High Performance Computer Systems. *Architectural Requirements for Deep Learning Workloads in HPC Environments*, 7-17.

[13] Lang, M. (2020, September 7). *BiP SOLUTIONS 37*. Retrieved from https://www.cyberessentialsonline.co.uk/examples-of-phishing-attacks-and-how-they-work/.

[14] Lee, J. T. (2021). IEEE European Symposium on Security and Privacy. *D-Fence: A Flexible, Efficient, and Comprehensive Phishing Email Detection System*, 578-597.

[15] Massarsch, K. (2017). In Geotechnical Engineering for Transportation Infrastructure. *The emergence of information technology: A state of practice report*, 65-82.

[16] Nines, F. (2018, Jun 6). *Five Nines*. Retrieved from https://blog.fivenines.com/spam-filtering-why-its-important-and-how-it-works.