International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Artificial Intelligence (AI) Deep Learning for Phishing Attack

Taif S. Hasan

Computer Science Department, Al-Ma'moon University College, Baghdad, Iraq
Email: *taif.s.hasan[at]almamonuc.edu.iq*

Abstract: ***The highly need usage for the internet has leads to undesired attack that could be considered as dangerous for each of us. That work involves trying to get the sensitive data such as credit card info, Bank account info, personal files and others. There are many aspects for that work. One of the most affected is the phishing attack. Fishing attack involved using much strategy in order to get the sensitive data. Our attention will be devoted to study and suggest efficient AI methods for preventing the attack.***

Keywords: Cipher, Security, Phishing, Attack, Vulnerability, AI

1. Introduction

Recently Internet entered the world of all (Scientists, art, construction, kids, engineering, medicine, sports). The great need for the internet and its services lead all hackers to search for rabid path for earning huge amount of money and get rich. There are a lot of risks corresponding to internet and its services. One of them is the Phishing attack that tries to get sensitive data. [1]

Phishing Attack

A phishing attack is a method of tricking users into unknowingly providing personal and financial information or sending funds to attackers. The most common phishing attacks use some form of electronic messaging such as email to provide a link to what appears to be a legitimate site but is actually a malicious site controlled by the attacker. Phishing is a hybrid attack combining both social engineering and technological aspects and combating phishing attacks requires dealing with both aspects. [2]

Artificial Intelligence (AI)

Artificial Intelligence or sometimes called machine intelligence is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals. Some of the activities that it is designed to do is speech recognition, learning, planning and problem solving. Since Robotics is the field concerned with the connection of perception to action, Artificial Intelligence must have a central role in Robotics if the connection is to be intelligent. Artificial Intelligence addresses the crucial questions of: what knowledge is required in any aspect of thinking; how should that knowledge be represented; and how should that knowledge be used. Robotics challenges Artificial Intelligence by forcing it to deal with real objects in the real world. [3]

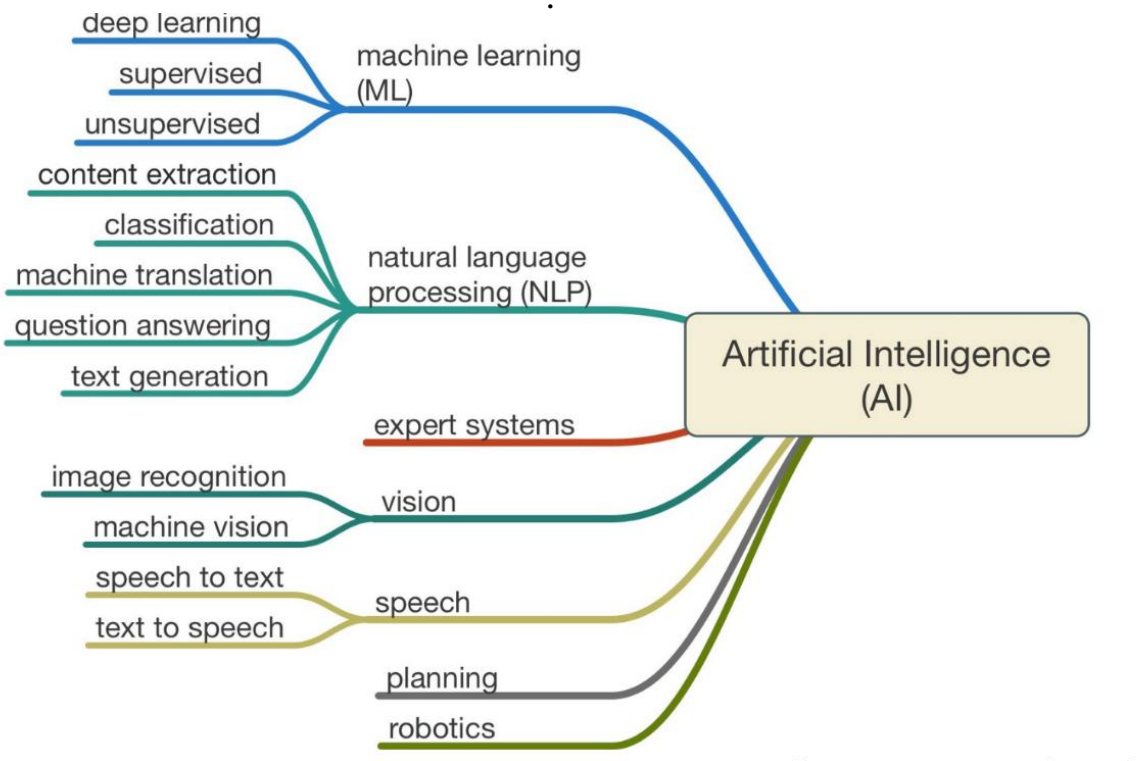

Figure 1: AI Branch

2. Literature Review

- Rutul Patel, Sanjay Kshetry, Sanket Berad, Justin Zirthantlunga; work on paper entitled "Phishing URL Detection using Machine Learning"; DOI: 10.22214/ijraset.2022.39979; May 2022. They presented to more serious dangers as cybercrimes. URL-based phishing assaults are quite possibly the most widely recognized dangers to web client. In this kind of assault, the aggressor takes advantage of the human weakness rather than programming defects. It targets the two people and associations, instigates them to tap on URLs that look secure, and take private data or infuse malware on our framework. [4]
- Jimmy Moedjahedy, Arief Setyanto, Fawaz Khaled Alarfaj, Mohammed Alreshoodi suggest "FuCCrFS: Combine Correlation Features Selection for Detecting Phishing Websites Using Machine Learning"; Future Internet 14 (8): 229; DOI: 10.3390/fi14080229; July 2022 work on Examining the website's URL address is one method for avoiding this type of deception. Identifying the features of a phishing website URL takes specialized knowledge and investigation. Machine learning is one method that uses existing data to teach machines to distinguish between legal and phishing website URLs. In this work, we proposed a method that combines correlation and recursive feature elimination to determine which URL characteristics are useful for identifying phishing websites by gradually decreasing the number of features while maintaining accuracy value. [5]
- Shafana A. R. F., Fanoon Raheem; Predictive Data Mining for Phishing Websites: A Rule Based Approach; December 2020. The rapid advancement in internet has paved way for several serious crimes, of which phishing occupies a very important place. Phishing is a form of cybercrime where an attacker mimicking a legitimate. This project uses the mining strategy to detect the occurrence of phishing [6].
- Shweta Sankhwar, Dhirendra Pandey, Prof. Raees Ahmad Khan, introduce paper "Email Phishing: An Enhanced Classification Model to Detect Malicious URLs"; ICST Transactions on Scalable Information Systems 6 (21): 158529; July 2018. This research focuses on the relevant URLs features that discriminate between legitimate and malicious/phishing URLs. The impact of email phishing can be largely reduced by adopting an appropriate combination of all these features with classification techniques. [7]
- Daisuke Miyamoto, Hiroaki Hazeyama, and Youki Kadobayashi; Human Boost: Utilization of Users' Past Trust Decision for Identifying Fraudulent; 1 National Institute of Information and Communications Technology Traceable Network Group4-2-2 Nukui-Kitamachi, Koganei, Tokyo 184-8795, JAPANdaisu-mi[at]nict. go. jp2. In this paper, we present an approach that aims to study users' past trust decisions (PTDs) for improving the accuracy of detecting phishing sites. [8]
- Chun-Ying Huang, Shang-Pin Ma, Wei-Lin Yeh, introduce Mitigate Web Phishing Using Site Signatures,, TENCON 2010-2010 IEEE Region 10 Conference; December 2010. The researcher explains how to use site signature in phishing detection. [9]
- Brad Wardman, Gary Warner; Chengshan Zhang introduce An Empirical Analysis of Phishing Blacklists; January 2009. A study the effectiveness of phishing black-lists. We used 191 fresh phish that were less than 30 minutes old to conduct two tests on eight anti-phishing toolbars. We found that 63% of the phishing campaigns in our dataset lasted less than [10]
- Oluwatobi Akanbi, Iraj Sadegh Amiri, Elahe Fazeldehkordi introduce "A Machine Learning Approach to Phishing Detection And Defense", which looks at the phishing problem holistically by examining various research works and their countermeasures, and how to increase detection. It composes of three studies. In the first study, focus was on dataset gathering, pre-processing, features extraction and dataset division in order to make the dataset suitable for the classification process. In the second study, focus was on metric evaluation of a set of classifiers (C4.5, SVM, KNN and LR) using the accuracy, precision, recall and f-measure metrics [11].

Phishing Attack Detection Model

MLPs are an excellent place to start learning about deep learning technology. MLPs belong to the class of feed forward neural networks with multiple layers of perceptron that have activation functions. MLPs consist of an input layer and an output layer that are fully connected. They have the same number of input and output layers but may have multiple hidden layers and can be used to build speech-recognition, image-recognition, and machine-translation software.

How Do MLPs Work?

- MLPs feed the data to the input layer of the network. The layers of neurons connect in a graph so that the signal passes in one direction.
- MLPs compute the input with the weights that exist between the input layer and the hidden layers.
- MLPs use activation functions to determine which nodes to fire. Activation functions include ReLUs, sigmoid functions, and tanh.
- MLPs train the model to understand the correlation and learn the dependencies between the independent and the target variables from a training data set.

Below Figure (2) is an example of an MLP. The diagram computes weights and bias and applies suitable activation functions to classify images of cats and dogs.

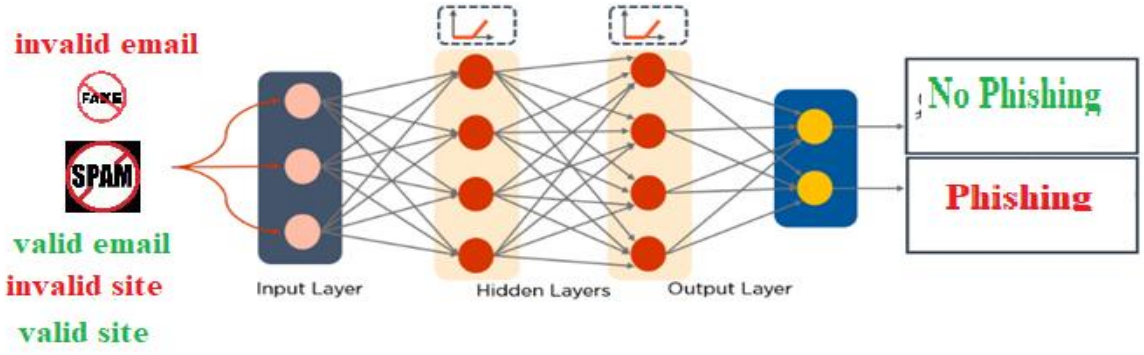

Figure 2: Perceptron Structure

In order to model the problem it must:

- The phishing parameters consists of
- E→ valid email-E→ not fake email
- U→ valid web site-U→ fake web site
- Q→ valid question-Q→ non valid question
- L→ valid location-L →non valid location
- P→ valid port-P→ non valid port
- Input layer contain the parameters
- X1 (email)
- X2 (web site)
- X3 (question)
- X4 (location)
- X5 (port)
- Output layer consists f
- Y1 (True or False)

The neural network must be learned to determine the true case (no phishing) and false case (phishing).

Also the parameters will be updated periodically during usage where some new email, site, and other will be added.

3. Discussion and Suggestion

The phishing cause big problem for all system. The AI include many methods and strategy. There are many researches and papers concerning this field. After survey made in the fields an efficient model consists of AI perceptron that is the engine for learning to distinguish between the existence or not of fishing could be used and it characterize by its adaptation and scalability, fast executing, ability to work in noisy data. The drawback for the method is the learning stage, and could be increase speed by simple adaptation.

References

[1] A Review on Phishing Attacks, International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 9 (2019) pp.2171-2175, © Research India Publications.
[2] Phishing Attacks and Defenses; International Journal of Security and its Applications 10 (1): 247-256;; January 2016.
[3] Ziyad Mohammed; Artificial Intelligence Definition, Ethics and Standards, Electronics and Communications: Law, Standards and Practice
[4] Rutul Patel, Sanjay Kshetry, Sanket Berad, Justin Zirthantlunga; work on paper entitled "Phishing URL Detection using Machine Learning"; DOI: 10.22214/ijraset.2022.39979; May 2022.
[5] Jimmy Moedjahedy, Arief Setyanto, Fawaz Khaled Alarfaj, Mohammed Alreshoodi suggest "FuCCrFS: Combine Correlation Features Selection for Detecting Phishing Websites Using Machine Learning"; Future Internet 14 (8): 229; DOI: 10.3390/fi14080229; July 2022.
[6] Shafana A. R. F., Fanoon Raheem; Predictive Data Mining for Phishing Websites: A Rule Based Approach; December 2020.
[7] Shweta Sankhwar, Dhirendra Pandey, Prof. Raees Ahmad Khan, introduce paper "Email Phishing: An Enhanced Classification Model to Detect Malicious URLs"; ICST Transactions on Scalable Information Systems 6 (21): 158529; July 2018.
[8] Daisuke Miyamoto, Hiroaki Hazeyama, and Youki Kadobayashi; Human Boost: Utilization of Users' Past Trust Decision for Identifying Fraudulent; 1 National Institute of Information and Communications Technology Traceable Network Group4-2-2 Nukui-Kitamachi, Koganei, Tokyo 184-8795, JAPANdaisu-mi[at]nict.go.jp2.
[9] Chun-Ying Huang, Shang-Pin Ma, Wei-Lin Yeh, introduce Mitigate Web Phishing Using Site Signatures,, TENCON 2010-2010 IEEE Region 10 Conference; December 2010.
[10] Brad Wardman, Gary Warner; Chengshan Zhang introduce An Empirical Analysis of Phishing Blacklists; January 2009.
[11] Oluwatobi Akanbi, Iraj Sadegh Amiri, Elahe Fazeldehkordi introduce "A Machine Learning Approach to Phishing Detection And Defense".