National College of Ireland

Phishing Detection in emails using Multi-Convolutional Neural Network Fusion

MSc Research Project
MSc in Cybersecurity

Dharani Kumar Babu
Student ID: X20179197

School of Computing
National College of Ireland

Supervisor: Michael Prior

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Dharani Kumar Babu
Student ID:	X20179197
Programme:	Msc in Cybersecurity **Year:** 2021 - 2022
Module:	Research Project
Supervisor:	Michael Prior
Submission Due Date:	15/08/2022
Project Title:	Phishing Detection in Emails using multi-Convolutional neural network Fusion
Word Count:	**8010 Page Count 24**

I hereby certify the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Dharani Kumar Babu

Date: 14/08/2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Abstract

In today's world, people prefer doing business and other transactions in digital format as the technology and usage of the internet have grown. The high usage of online activities made the attackers discharge criminal activities online. Cybercriminals use the online platform as a tool to steal the user's personal and sensitive information. The large-scale damage had been recorded that the attackers had launched against organizations. This results in a loss of customer trust that they build and millions of dollars in lost data. The victim can be an individual, or it can be an organization. Phishing attacks have become the most common way for an attacker to trick a user into falling into a trap in order to gain access. Hence, this paper proposes phishing detection in emails using multi-convolutional neural network fusion to detect legitimate and phishing URLs. The proposed method validates the URLs without accessing the contents. This study provides guidelines for constructing a robust security defense system so that attackers cannot bypass technical defenses and steal confidential data.

Keywords: ***Emails, phishing detection, Deep learning, Convolutional neural Network.***

Table of Contents

1. Introduction

The high rise in use of the internet by people all around the world allows them to share their personal information online. The data is widespread and easily accessible in public, and this allows attackers to steal sensitive information from the users. Phishing emails are the common tool used by attackers to trick the victim and gain access to the network or organization to steal financial and personal data. Phishing has evolved into a sophisticated attack vector. Phishing attacks may conclude in remarkable losses for their victims, which includes sensitive information, personal data theft, organization, and government secrets. The cybercriminals always target the user and send them a malicious email that contains images, links, attaches a file and is marked as important. That can trick the user into clicking, which redirects the user to the phishing website.

It is a social engineering attack that exploits other mediums such as instant messaging (IM), SMS, and sending mass-email messages. The rise of phishing attacks and spoofing methods in response to the development of advanced defenses is accelerating. In addition to using new tools and technologies to attack system vulnerabilities, attackers often use social engineering method to manipulate unsuspecting users.

With the rise of new technologies such as smartphones and social media, phishing attacks have reached an all-time high in successful attacks. The APWG Phishing Activity Trends Report examines the development, proliferation, and spread of phishing attacks reported to the APWG. According to APWG, in the third quarter of 2019, phishing attacks increased to 266,387, which is marked as the highest level since 2016. It's important to raise awareness among the employees regarding cyber-attacks. The cyber team needs to monitor every incoming and outgoing action to stop unwanted requests from cybercriminals. A firm which contains large amounts of data and sensitive information about both employees and customers doesn't want to lose its integrity in losing and becoming a victim. Every organization must spend nearly 10% of their revenue to build a strong wall against cybercriminals. One phishing attempt can bring the entire firm down, and recovering the data is difficult for the entire team. Numerous methods, which includes social trust-based solutions, greylisting, DNS blacklisting and whitelisting, metaheuristic-based strategies, and machine learning methods, have been developed to detect spam emails.(*Phishing activity trends report 2018| APWG*, no date)

For the purpose of mitigating phishing attacks, new phishing detection algorithms have been developed. Traditional phishing techniques include the use of blacklist and whitelist filters, pattern matching filters, email verification filters, and password filters. The alternate type is the learning machines which is an automated technique. These techniques always follow the pattern and match the patterns that are trained , thus frequent training of Ips and data are required to detect and prevent from phishing attacks. Sometimes traditional techniques fail to detect spam emails. To resolve these issues and to provide better efficiency and performance deep learning technique produces greater results in training and testing large number of datasets. This paper proposes Phishing detection in emails using Multi-Convolutional Neural

network fusion with phishing datasets from Mendeley Data. The Convolutional Neural Network (CNN) is a type of deep learning neural networks used mostly for image and voice recognition applications. This unique feature in using CNN algorithm results in high degree of detecting phishing emails.

This paper consists of literature review which helps in reviewing the previous related work on detecting phishing emails, URL's and the brief details of used models and algorithms are explained. This related work provides a strong foundation about this topic, which will be explained following section. Following literature review the other sections contains Research Methodology , Design & specifications. The last section is conclusion part which provides the information about knowledge gained from previous papers and the strength, improvement, limitations of the proposed paper, and the work that to be carried out in future to overcome the limitations.('Frontiers Phishing Attacks: A Recent Comprehensive Study and a New Anatomy Computer Science', no date)

2. Literature Review

The efforts that have been made by the researchers from the previous papers explain the importance on cyberthreats that occurs due to phishing emails. It is necessary to be aware the current threats that has been faced by the firms and phishing emails are one of the most dangerous parts in cyber filed and it is very difficult to protect the victims from cybercriminals. The previous research papers carried out in detecting spam emails, URLs and keyword information that has been mentioned in the email. Every firm spend lot of investments in cyber to protect their products and trust they build from the customers. Despite all efforts made by cyber team still users eventually fall in trap.

A. Phishing detection using machine learning

The authors (Xing Fang, 2004) have proposed an interesting and challenging paper that connects with biological immune system. They implemented an Artificial Immune system for detecting phishing. The research used memory detectors and mature detectors to capture the phishing emails. The memory detectors are well trained that contains phishing emails that has been previously recorded by the system. The authors have carried out their research that the previous work papers used Naïve Bayesian classifier and Stochastic learning weak estimator to detect the ham and spam email and the proposed paper implemented (AIS) Artificial immune system that replaces the previous model. Although the AIS produces positive results, authors mentioned that the detectors share static fire-threshold value which may not function under certain situations. The future work is based on producing best design by providing dynamic fire-threshold value to individual detectors.

The proposed paper is blended with multiple ML methods to detect phishing emails by ((Ma *et al.*, 2009)). The authors have proposed this paper with implementing five machine learning techniques to identify which ML method performs best in detecting phishing emails. They have used Decision tree, Random Forest, Multi-layer perceptron, Naïve bayers and Support

vector machine(SVM). It has been identified that Random Forest produces best results compared to other machine learning techniques. The researchers would have tried with deep learning approach and to check for the better performance on detecting phishing emails. This research work is restricted to its results where the authors utilize only part of potential features. To enhance the categorization and detection, it can be expanded to add more features. The future idea is to reduce the unnecessarily large value from feature normalization and author intend to improve normalization process by finding a threshold to remove noisy data. The future work continues to build a stable and automatic filter in detecting phishing emails, where less supervision is achieved. Authors planned to develop automatic feature update mechanism where it updates automatically in classifier when necessary.

The author (Zaimi, Hafidi and Lamia, 2020) presented a survey on Taxonomy of website anti-phishing solutions. The survey shows how cybercriminals target the users to gain sensitive information. They classified social engineering attack and technical subterfuge as an idea to trick the victims fall in trap. The survey shows various website classification and the approaches to detect phishing. For URL analysis, machine learning and rule-based approaches are used. The attacks classified into two types, content-based and non-content-based approach. The content-based technique focuses on content of the websites namely URL, image, and any form of text content. The non-content-based technique focuses on blacklisting, whitelisting, DNS based technique. J48 tree, RF and logistic regression are the classifiers used and with the size of the training dataset, performance is determined. The proposed model used CNN and LSTM to deal with large data and to produce higher prediction performance. The combination two algorithms produce higher accuracy with less training time. The future work is to improve better prediction model in detecting phishing websites.

A new and interesting approach by (Park and Stuart M, 2014) differentiate machine and human ability to detect phishing emails. The proposed research has been conducted both on human performance and machine performance that can identify the legitimate and phishing emails. The study tested by classifying human-friendly and machine friendly in testing phishing emails . In the test for machines, a small number of keywords have been altered, resulting in replacements. It is important to classify the keyworks for better identification for the machines to produce the result. With the human-friendly approach, humans are equally cable of identifying phishing emails as machines , except for those that have been altered. Humans are worried about long emails that can make the user to fall in trap and it is difficult to identify whether it is legitimate or phishy and the future work must consider these drawbacks and to work on better experiment in identifying phishing emails. The machine had an advantage at this low level of large non-semantic processing and non-expert level of odd human subjects. The combination of both human and machine has more advantage in identifying phishing emails.

(Feroz and Mengel, 2015) proposed phishing URL detection using ranking of URLs and this method categorize URLs automatically based on their lexical and host-based features. The entire dataset has made clustering and they derived label or cluster ID for individual URL and

these labels are used as predictive feature in classifying the URLs. To produce URL rank, they used URL clustering, URL classification and URL categorization are used. Three types of features are used namely, lexical feature, host-based feature, and cluster label feature in which they used J48 tree for the effectiveness. The URL are guided with rule with URL categorization and URL ranking in displaying with colour type identification. The future works depends on adding filter to detect URL's automatically and to use effective algorithms. The large datasets need to be trained frequently to produce higher accuracy in detecting phishing attacks.

The proposed paper by (Dey *et al.*, 2022), analysis the machine learning algorithms by developing a phishing email and website detection model. This research paper focus on discussion and comparison of various ML algorithms that are more effective in identifying phishing emails and weblinks. The researchers used four machine learning classification algorithms namely, RF, decision tree, logistic regression and Naïve bayes classifier. The idea of this paper is to develop a model that can detect phishing emails and websites with producing higher accuracy. The time needs to be consumed in training the datasets to achieve better performance. Naïve bayes is used to identify spam and ham messages and Multinominal NB is used for classification of phishing emails. Decision tree becomes inaccurate in categorizing new data, as it works on one dimension. Image recognition algorithms produces best performance as they work in one , two and three dimensions. Compared to other algorithms mentioned above random forest classifier has least efficiency in identifying phishing emails and websites. The future work of this research focuses on implementing deep learning approach in identifying phishing attacks.

The proposed paper by (Alkawaz, Steven and Hajamydeen, 2020) Detecting phishing website using ML. The idea of this paper is to detect blacklisted URL from the phishing websites using phishing detection system. The user gets a pop up while accessing malicious websites. This approach can be utilized for identification and authentication and become a better tool to prevent the user from getting tricked by cybercriminals. This alert is pushed by email notification to the user. This paper produces a development environment by using Agile Unified process (AUP) for development and flexible process. This type methos can be useful in monitoring requirement for current situation. The future works produces a text message integration along with above mentioned notifications to the user so that the program can give better view to the user.

(Patil, Rane and Bhalekar, 2017) proposed research on Detecting spam and phishing mails using SVM and obfuscation URL detection algorithm. The researchers have proposed a novel approach in detecting webpages. They have created a similarity detection technique for CSS-related components. The proposed paper uses SVM technique combined with map-reduce paradigm to produce high efficiency in detecting spam email. They use the obURL detection method to identify phishing emails using URL obfuscation. IP address test, DNS test, , shorten URL, whitelist test, URL encode test, blacklist test and pattern matching test are the following test cases that are performed to verify phishing site emails. The advantage of using SVM is it can handle large number data compared to other machine learning techniques. The

training the dataset is much easier but still SVM can't feed with large input. So, they used map-reduce technique to overcome this issue and to deal with large input data.

The interesting and challenging paper by (Li, Zhang and Wu, 2020) proposed paper in detection method of phishing email based on persuasion principle. The idea of this persuasion principle is to add up the respective word of the feature which come out in the email. It verifies the word with features listed. The researchers used 25 features for detecting phishing attack. The methodology consists of three parts namely, feature extraction, persuasion principle related feature extraction, and machine learning based detection. The proposed methos uses KNN algorithm, Decision Tree, and Bayes algorithm. Using a text-based phishing detection mechanism, the persuasion function has been improved. Since this approach produces high values on both TPR and precision, it is possible to control FPR within acceptable range to have effective results in detecting phishing emails. Future work will continue to increase the accuracy of the extraction of phishing email detecting features.

The emergency threat of phishing attack and the detection techniques using machine learning models proposed by (Ripa, Islam and Arifuzzaman, 2021). In this paper the researchers have built spear phishing bot using ML. The analysis is based detecting URL, email, and phishing website. Various classifiers are used to detect phishing URL and they are XGBoost classifier, KNN, Naïve Bayes, Decision tree, Logistic Regression, Gradient Boosting, and Random Forest model. With less time consumed random Forest produces best performance compared to other machine learning algorithms. The Trained dataset are used to test to classify Ham and spam emails. The future work is to build a framework that can improve detection model with higher accuracy in limited time.

The hybrid model proposed by (Sindhu *et al.*, 2020) in detecting phishing using Random Forest, SVM, and neural network with backpropagation. Using the lexical feature extraction, the features of URL are extracted and SVM and neural network with backpropagation is used to classify legitimate or phishing website. SVM proves best compared to other algorithms as it is implemented in chrome extension and gives alerts when loading the site. Web mining technique is used for extraction process. The proposed paper focused on identifying IP address, URL length, sub-domain registration length, @ symbol and request in URL.

Phishing Emails detection using CS-SVM proposed by (Niu *et al.*, 2018). This research paper proposes a different approach in detecting phishing emails. The proposed model named Cuckoo Search SVM (CS-SVM). The hybrid classifier is constructed by the extraction of 23 features that includes body-based features, URL based and header-based features. The Cuckoo search (CS) is combined with SVM to optimize (RBF) Radical Basis Function. On a testing dataset containing both old and new phishing emails, the hybrid classifier combining CS and SVM performs better than SVM classifier with default parameter of Radial Basis Function. The future work is to make optimization of Cuckoo Search -SVM as the research paper presents the model in a single machine and the future idea is to make algorithm run on distributed platform.

The proposed research paper (Bhagwat, 2021) on methodical overview on detection, identification, and proactive prevention of phishing websites. In proposed method instead of using extraction, authors have introduced Fuzzy logic which is best way in dealing with variables. The model is built with fuzziness resolution and an open and intelligent phishing website model is proposed in this assessment. This involves both smooth logic and machine learning algorithms are used in detecting malicious websites. The input was given by feature selection and phishing factors contains in that site are considered. The dataset is more efficient in predicting phishing websites. The future work deals with solving the fuzziness and to build a smart model that distinguish factors on the phishing website.

B. Phishing detection using Deep Learning

The author (Fang *et al.*, 2019) proposed phishing email detection using improves RCNN model with multilevel vectors and attention mechanism. The researchers have analyzed the email structure initially and with RCNN model and multilevel vectors and attention mechanism they proposed a new model named 'THEMIS' that used to detect phishing email. This model is built to analyze the email at the header, body, character level, and the word level. With the help of THEMIS model, the accuracy is achieved, and noise is minimized. This paper focuses on email that contains header and the future work will be to improve the model that can identify phishing emails without email header and only contains body.

The phishing detection from URLs using deep learning approach was proposed by (Singh, Singh and Pandey, 2020). This research paper proposes Convolutional neural network to detect phishing webpage. The CNN extract features automatically from hidden layers; thus, it doesn't require any feature extraction. The research classifies legitimate and phishing URLs. The detection of phishing is carried out in both email and website levels. Phishing is classified into two at website levels namely, list-based, and heuristic approach. The IP address and login user interface are recorded and added in whitelist, so each time it matches the traces listed in whitelist and protects the user when visits the malicious website. If the heuristic method is used, phishing websites may be incorrectly categorized as legitimate websites, which is a significant drawback.

(McGinley and Monroy, 2022) This research paper proposes Convolutional neural network optimization for phishing email classification to classify the text in email messages. The text of the body is given as input and the output produces the malicious or legitimate. In related work, the researchers identified that they had challenges in training datasets and efficiency was not outperformed. Deep learning neural networks are among the most successful models in terms of providing very effective outcomes. The researchers have taken an optimistic approach to identifying fraudulent email text messages in their study report. The proposed model aims to detect beyond text. Depending on the objectives of the email service provider that monitors incoming email messages, the future goal of this research will be to build an architecture that improves the recall or accuracy of deep learning architecture.

A spam email detection mechanism for English language text emails using deep learning approach was proposed by (Kaddoura, Alfandi and Dahmani, 2020). This research paper focuses on both machine learning and deep learning approach in classifying link-less emails. This paper approach spam detection mechanism based on neural network using Enron dataset. The methodology aims to identify ham and spam classes. The model consists of three steps namely, Text pre-processing, feature extraction and training and testing the model. In this paper Feed forward neural network (FFNN) is used and applied. With the trained dataset the FFNN model identifies Ham and spam emails. The future works plays an important role has the paper goes forward working with machine learning algorithms and compare with the proposed model to check which algorithm produces better performance.

2.1 Research Niche

RELATED WORK	STRENGTH	LIMITATIONS
(Xing Fang, 2004)	The proposed paper used detectors namely memory and mature detectors in detecting phishing emails which is an advantage, and this paper is replaced by Artificial immune system (AIS) which is much more useful in security related systems.	The memory detectors need to be trained frequently and needs to be updated. Under certain conditions the detectors will not work has the design is set to be in static fire-threshold value.
(Singh, Singh and Pandey, 2020)	The phishing detection in both website and email levels . The proposed method used DL method that used CNN. The extraction of features is made easier in neural networks.	The proposed method used Heuristic intrusion this might generate false positive, so the detection rate becomes less in classifying legitimate phishing websites.
(Park and Stuart M, 2014)	The proposed method uses both machine and human ability in detecting phishing emails. The efficiency is higher in combing machine with human ability. low-level of large non-semantic processing by machines and identification of non-experts on malicious human subjects	The machine needs to be trained frequently if any keywords get altered it is less chance in detecting the malicious emails and humans can be trapped easily when email body is in large.
(McGinley and Monroy, 2022)	The related work had difficulties in training and efficiency, and this proposed model deep learning algorithm to resolve the issue. Using different layers, the performance and accuracy can be improved.	The research project is limited to build a system that enhance the recall or precision of deep learning architectures, requesting the email service provider the importance in analyzing incoming email messages.

(Ma et al., 2009)	The phishing emails are detected using multiple machine learning methods brings a positive result in terms of producing better accuracy and performance.	The issue of having noisy data is to neglect in future model and to improve the identification process a threshold. Using automatic filter can improve the model in less supervision.
(Zaimi, Hafidi and Lamia, 2020)	The proposed model provides classification of website using classifiers J48 tree, random Forest, and logistic regression. DNS based technique is used in detecting phishing websites which is additional advantage of this research paper. Both CNN and LSTM algorithm are used to solve large training datasets.	Work was restricted in building an effective prediction model to detect phishing websites.
(Feroz and Mengel, 2015)	The proposed paper automatically classifies URLs based on lexical and host-based features. This paper uses J48 tree algorithm for the effectiveness of the model.	This paper is limited to having filters in the model, so building a model with automatic filter will produce high efficiency in detecting URLs.
Kaddoura, Alfandi and Dahmani, 2020)	The proposed research paper detects phishing email contains no-link using deep learning approach. Research paper uses both machine learning and deep learning which produces high efficiency.	The proposal of moth ML and deep learning could take more time in processing and maintenance is very difficult in training and testing the datasets.
(Dey et al., 2022)	The proposed paper implements RF, decision tree, logistic regression, and Naïve bayes algorithm that the model produces efficient outcome. Since their model performs by the way they train their datasets.	Having multiple algorithms, the training and test would be difficult, and it takes more time. Space utilization will more consumed.

(Alkawaz, Steven and Hajamydeen, 2020)	The approach of detecting phishing site and provides an alert to the user in efficient method in notifying before falling into trap.	The datasets need to be trained frequently to detect the malicious sites and filter need to be efficient in notifying to the user.
(Patil, Rane and Bhalekar, 2017)	This paper proposed a method in combination of SVM with map-reduce to obtain more accuracy in identifying spam email. For dealing with large data map-reduce technique is one the best method to resolve the issue.	The research paper is limited to the way that it detects in the email and the future work needs to take in website as the user gets redirected to the phishing sites.
Xue Li (Li, Zhang and Wu, 2020)	This method proposed with persuasion principle to classify the malicious emails. The advantage of KNN, Decision tree, and bayes algorithm gives more efficiency and detects faster rate in text-based phishing emails.	The method uses machine learning and persuasion principle where it takes more time in determining the features, feature selection will be difficult. To resolve this issue future work must develop extraction of features to improve accuracy
Yong Fang (Fang et al., 2019)	This method of detecting phishing email in improved RCNN with multilevel vectors provides more efficiency and the authors created a model named 'THEMIS' which can identify malicious in header, body and word level	Using RCNN model the research would have more filters for extraction. They limited in testing and could tried malicious email with no header and contains only body.
Ripa, Islam and Arifuzzaman, 2021)	The proposed model has a advantage that this model built with spear phishing bot that detects the phishing URLs, emails, and websites. For fast and higher efficiency this proposed paper used XGBoost, Gradient boosting algorithms are used along with Random Forest and logistic regression.	This paper is limited framework that detects the malicious sites, and the future would be the building a framework as an extension with browser that could block the user in entering to phishing websites.

(Sindhu et al., 2020)	The strength of this proposed paper is they have utilised the efficiency of using Artificial neural network along with Deep neural network. One of the advantages is they have added chrome extension to the method it helps in protecting users not to fall in trap.	The training and testing more complicated in working with SVM, Random Forest and in Neural networks all together.
(Niu *et al.*, 2018)	The combination of using Cuckoo search with SVM produces efficiency in rate of this model where it detects in network-based also blacklist, whitelist and content-based technique	The algorithm which they have proposed does not work in all distributed systems, its considered to be an drawback of this model.
(Bhagwat, 2021)	The fuzzy logic and machine learning are used in the proposed model to detect the phishing email and they proposed solution in preventing of phishing websites.	The complexity with combination of fuzzy logic with machine learning will difficult in solving the fuzziness and the future work will propose the model that solves the issue.

2.2 Effectiveness of Deep Learning algorithm

Deep learning is an effective technique that resolves the issues of manual feature extraction. In this proposed research paper, a Convolutional Neural Network (CNN) algorithm is applied to detect phishing emails, and, by this idea, the proposed model is built with a multi-CNN model. It is a combination of two models combined to get the best efficiency in terms of time consuming, performance, and accuracy. The input data to the CNN can be in the form of images in one-, two-, or three-dimensional data. Compared to other neural networks, CNN provides better efficiency in producing results as the input could be in the form of video, audio, text, or image. The convolutional neural network has more than one convolutional layer. All layers are fully interconnected, and regularization is achieved using dropout layers. The unique characteristic of convolutional neural networks is that sections of images are passed at the same level, which allows for recognition of the same patterns in different locations. Convolutional neural networks differ from other neural networks as they learn the filter weights by training on data instead of working the other way around.(*Using Convolutional Neural Network for Image Classification | by Niklas Lang | Towards Data Science*, no date)

3. Research Methodology

The proposed methodology consists of many steps. Firstly, the raw data is converted into csv format and processing it. In pre-processing stage, the data is converted into images and once

it is completed the dataset is spilt into two for testing and training purpose, which is a part of implementation. Finally, once the data is trained and tested it is then evaluated in the basis of model performance.

3.1 Dataset Description:

The phishing dataset for machine learning feature evaluation is downloaded from Mendeley Data website. This dataset consists of 48 features extracted from 5000 phishing websites and 5000 legitimate websites. These datasets were downloaded from January to May 2015 and from May to June 2017. Utilizing the browser automation framework results in a more exact and robust feature extraction strategy than the parsing approach based on regular expressions. This dataset is mentioned as WEKA-ready. This dataset is valuable for analyzing phishing characteristics, performing fast proof-of-concept tests, and evaluating phishing classification algorithms. The datasets consist of important features that includes URL length, Sub domain level, IP address, Path length, Abnormal form action, Fake link in status bar and popup window. The class labels are classified into Legitimate and phishing. For data splitting , the dataset is separated into 50:50 ratio. Having the same phishing dataset, the two model is created and developed a new fusion model. The created new model is then trained and tested which then classifies the URLs legitimate and phishing. (Tan, 2018)

3.2 Data pre-processing

The input for the convolutional neural network requires input in the form of images. The images as input are fed into CNN. The Mendeley datasets are in the form of numerical values , not in the form of strings, so neuralization is not required. The data is in the form of arrays. The two-dimensional matrix represents the arrays, and these arrays are converted into grayscale images, which produces higher efficiency. It is important to use standardized data in the pre-processing technique. We have different sources of data, so we need to improve our processing time by using data normalisation. These scales the range from 0 to 1. Once completing the data clustering process, the data is split into train and test. The batch normalization is performed only for numerical features by a simple linear process.

$$A' = (A - A_{min}) / (A_{max} - A_{min})$$

The different features in the datasets consists of values that range from maximum to minimum value. A_{max} represents the maximum value in the feature and A_{min} represents the minimum value in the feature. Finally, once the normalization is done the data are ready to train and test purpose.

3.3 Converting into matrix

When successfully batch normalization is completed and the dataset is divided into half for training and testing the next step is to convert each row into a square matrix. The 7x7 size matrix is created and matrix is converted into grey scale image, and it is stored. The matrix operation conducts a mathematical operation that iteratively performs element-by-element multiplication on input data before summing the output. This convolution matrix permits the sharing of weights, hence reducing the number of effective parameters and image translation.

This extraction process improves the efficiency the model and produce higher accuracy in feature mapping. The datasets consist of features are converted into images and these converted images are used throughout the process in completing the model and to generate higher accuracy. The data that are stored in the training datasets are divided by 255, which is the highest possible pixel value of the image. This would the final stage and then the CNN model is created.

3.4 Training the Model

The model consists of three layers namely, The convolutional layer, Pooling Layer, ReLU correction layer and fully connected layer. The convolutional layer is the first layer where ethe input is given in the form of images. The features in the image are extracted and filtered. The filters correspond exactly to the features we want to find in the images. The filters correlate precisely to the images we need. In contrast to conventional approaches, features are not pre-defined according to a certain formalism but are instead learnt by the network during the training process. Filter kernels relate to the weights of the convolution layer. The pooling layer performs several feature maps operation to each image they receive. The procedure of pooling consists of lowering the size of the images while keeping their essential qualities. This increases the efficiency of the network. The ReLU function is a real non-linear function since it replaces all negative inputs with 0. It simply acts as activation function. The second layer input is extracted using 32 filters size 3x3 and third layer verifies the data by 62 filters with size 3x3. With the same process the model 1 CNN and model 2 CNN are created and stopped at flatten layer and the output from each model is combined by the function'concatenated' and feed into hidden layer.

3.4.1 Fusion Model

In this fusion model, the output of both model1 and model2 are combined. Each model output called by 'model.output' function. 'concatenate' function is used to combine two models output together. The hidden layer is feed with combined output as input, where hidden layer contains three different dense sizes 512, 256, 128. The 'ReLU' function is used to activate this dense layer in the model. 'SoftMax' function is used to get the output which is the final layer of Convolutional neural network. The final model is created using 'tf.keras.Model' where the input is feed as values and given to the model and the output from the model is received as an output. 'RMSprop' function is used to increase the learning rate. 'final_model.compile' is used to compile the fusion model and loss are maintained by ' binary_crossentropy' function. 'final_model.fit' function is used to measure the model performance on how well it produces output on trained data. The number of iterations is produced by epochs value and batch size provides the size of each batch. By using 'final_model.save' function the final model is saved and where the model represented by 'model.h5'. Using 'matplotlib' function we get the graph that displays the loss and accuracy.

3.5 Testing the Model

In this multi-convolutional neural network model , the 50% of dataset are trained & remaining 50% are tested and that data is given into the model that detects the legitimate and phishing with accuracy. In this we load the trained model and the input from test data is stored as ‘X_test’ and ‘X_test’. The ‘y_test’ variable consists of labels that are stored.

4. Design Specification

The design specifications for detecting phishing in emails is divided into three stages and they are

Preparation of Datasets: The phishing dataset which is used in this paper is Mendeley data and the ARFF file is converted to CSV file to proceed with pre-processing steps. The resources used for processing the data is Jupyter notebook for features selection and python framework. Once the datasets are imported , it is necessary to drop unwanted columns and segregate the results columns. Once the null values are removed it then converted into matrix , and the feed the input to the CNN.

Building CNN Model: It is the next stage , where the implementation takes place. In this proposed paper Model 1 and Model 2 CNN are created and combining these two models , the fusion model is built and for this we use framework ‘Tensorflow’ and ‘Keras’. With the use of same data to built both models the processing time is reduced, and the space utilization is improved with memory and CPU usage.

Visual representation: The process that has been carried out the code is converted into visual view for the better understating for the users , the plot has been displayed by means of graph and finally with confusion matrix.

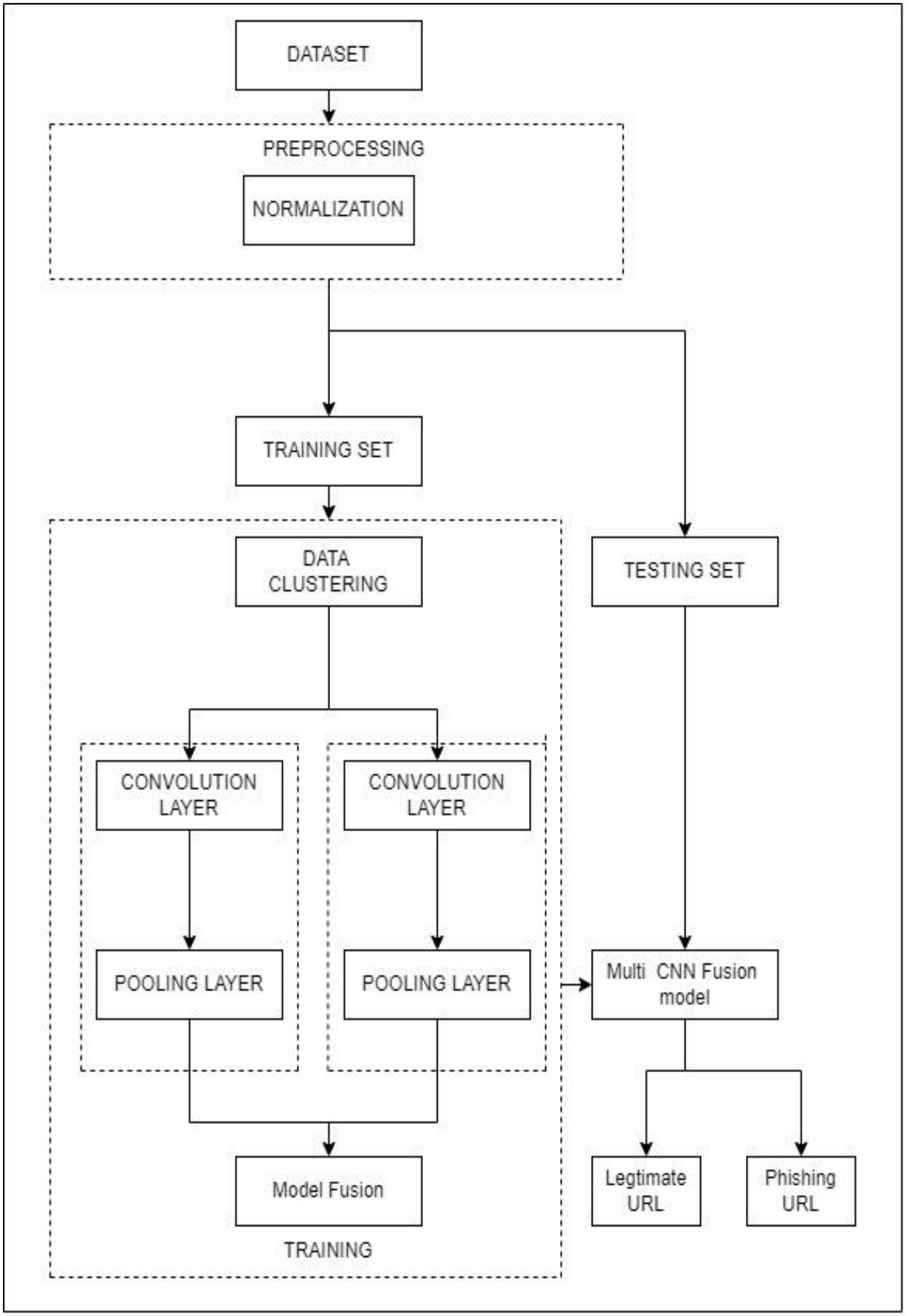


Fig 1: Design Model

5. Implementation

The advantage of using deep learning technique is this method solves large amount of data with better efficiency and improved accuracy. The input fed into the convolutional neural network by means of image and implementation takes place this stage.

5.1 Preparation of Phishing Datasets

The Phishing dataset is converted into csv format are converted into grey scale image in the pre-processing stage. The feature labels are separated, and the results label are separated and displayed. The normalization is carried out with the labels to bring the values 0 to 1. With this the data is converted into matrix. The manual separation of data is not necessary as the normalization provides automatic splitting of data using ‘ sklearn,model_selection’ function and ‘train_test_split’ function. The dataset has 5000 phishing websites and 5000 phishing websites, so it is divides into 50-50% i.e., 50% of data for training and the remaining 50% for testing purpose. The features in the datasets are converted into 7x7 matrix. Once processing is complete, the images are generated and stored in the respective path.

5.2 Training the CNN Model

In the data pre-processing stage, the images are transformed into grey scale image having size 50x50 and the class labels are stored in corresponding ‘Y_Train’ and the grey scale images loaded in X_train’ are stored after divided by 255, which is considered as highest pixel value of any image. With the Sequential() function the model has been initiated along to feed into input layer. Once the model has been created it consists of three layers as shown in Fig12. They are Activation, Maxpooling, and Batch Normalisation. Each filter has its size of 32,64,128. In each layer ‘ReLU’ is the activation function. CNN model stops at flatten layer before hidden layer. With the same procedure the second model2 is created. The final output of CNN model consists of two labels, and they are ‘legitimate’ and ‘Phishy’. ‘softmax’ method is used to get the result which is the final layer of Convolutional neural network.

The program is provided with batch size of 200 for each epoch to run and 40 epoch are fixed to achieve desired output with minimum loss and loss are maintained by ‘binary_crossentropy’ function. The accuracy is obtained by using ‘accuracy_score’ from library ‘sklearn.metrics’. The final model is saved as ‘model.h5’

```
model1 = Sequential()

model1.add(Conv2D(32,(5,5), input_shape=(50,50,1), activation='relu'))
model1.add(MaxPooling2D(pool_size=(2, 2)))
model1.add(Dropout(0.5))

model1.add(Conv2D(filters=32, kernel_size=(3, 3)))
model1.add(Activation('relu'))
model1.add(MaxPooling2D(pool_size=(2, 2)))
model1.add(BatchNormalization())

model1.add(Conv2D(filters=64, kernel_size=(3, 3)))
model1.add(Activation('relu'))
model1.add(MaxPooling2D(pool_size=(2, 2)))
model1.add(BatchNormalization())
model1.add(Dropout(0.1))

model1.add(Conv2D(filters=128, kernel_size=(3, 3)))
model1.add(Activation('relu'))
model1.add(MaxPooling2D(pool_size=(2, 2)))
model1.add(BatchNormalization())


model1.add(Flatten())


model1.compile(loss='binary_crossentropy', optimizer='adam', metrics = ['accuracy'])

#model1.summary()
```

Fig 2: Code snippet of single CNN model

```
model0 = concatenate([model1.output, model2.output])

x = Dense(512, activation='relu')(model0)
x = Dropout(0.4) (x)
x = Dense(256, activation='relu')(x)
x = Dropout(0.4) (x)
x = Dense(128, activation='relu')(x)
x = Dropout(0.4) (x)
output = Dense(num_classes,activation='softmax')(x)
```

Fig 3: Fusion model

In this fusion model, the output of both model1 and model2 are combined. Each model output called by ‘model.output’ function. ‘concatenate’ function is used to combine two models output together.

5.3 Testing the CNN Model

In this multi-convolutional neural network model , the 50% of dataset are tested and that data is given into the model that detects the legitimate and phishy with accuracy. In this we load the trained model and the input from test data is stored as ‘X_test’ and ‘X_test’. The ‘y_test’

variable consists of labels that are stored. ‘load_model’ function is used to call the model from trained model. ‘model.predict' is used to generate the prediction function, which is then given to the variable 'y pred', which contains the value of input given to the loaded model. Form the testing we get the confusion matrix from that we can determine accuracy; F-1 score and precision.

6. Evaluation

Accuracy is one of the important features in classifying the legitimate and phishing emails in-order to evaluate multi-CNN fusion model performance. For this evaluation process we determine three indicators namely Accuracy, precision, and recall. To achieve these indicators, we have four parameters check they are True positive (TP), false positive (FP), false negative (FN), True negative(TN). These metrics are calculated by function’sklearn,metrics’. The values are classified using ‘classification_report’ function from ‘sklearn.metrics’ library, where we get precision, recall and F1- score.

1. Accuracy:

It represents the proportion of accurate classifications throughout the entire phishing dataset, and it is calculated by using below formula,

$$\text{Accuracy} = \frac{\text{TRUE}_{\text{POSITIVE}} + \text{TRUE}_{\text{Negative}}}{\text{TRUE}_{\text{Positive}} + \text{TRUE}_{\text{Negative}} + \text{FALSE}_{\text{Positive}} + \text{FALSE}_{\text{Negative}}}$$

2. Precision:

Precision is defined as the proportion of all retrieved instances that are relevant. i.e., It indicates the ratio of successfully predicted instances to the total number of cases expected for each class.

$$\text{Precision} = \frac{\text{TRUE}_{\text{Positive}}}{\text{TRUE}_{\text{Positive}} + \text{FALSE}_{\text{Positive}}}$$

3. Recall:

Recall measures the phishing emails that are classified from the total predicted instance. It is denoted by,

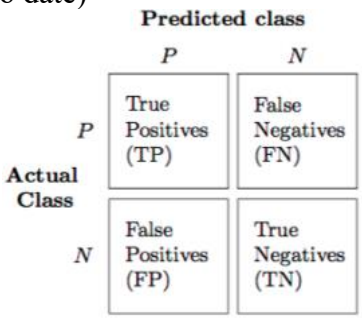
$$\textbf{Recall} = \frac{TP}{TP + FN}$$

4. F1- score

The F1 score represents the average weighted of precision and recall. It needs to be considered both precision and recall as there is FP and FN.

$$\text{F1 Score} = 2*(\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision})$$

('Accuracy, Precision, Recall & F1 Score: Interpretation of Performance Measures @ blog.exsilio.com', no date)



Predicted Class	Actual Class	
	Positive	Negative
Legitimate	425	46
Phishy	0	529

Table 1: Confusion matrix of multi-CNN

```
print(classification_report(y_test,y_pred))
```

	precision	recall	f1-score	support
0	1.00	0.25	0.40	3484
1	0.57	1.00	0.73	3516
accuracy			0.62	7000
macro avg	0.79	0.62	0.56	7000
weighted avg	0.78	0.62	0.56	7000

Fig: Single-CNN

	precision	recall	f1-score	support
0	1.00	0.98	0.99	471
1	0.98	1.00	0.99	529
micro avg	0.99	0.99	0.99	1000
macro avg	0.99	0.99	0.99	1000
weighted avg	0.99	0.99	0.99	1000
samples avg	0.99	0.99	0.99	1000

Fig: Multi-CNN

The phishing dataset is provided for both model1 and model2 and the outcome has proven that multi-CNN model has higher efficiency compared to single-CNN model. The single-

CNN model produces 62% accuracy and multi-CNN produces 98%. Multi-Convolutional neural network is one the best technique to detect phishing emails that produces best results.

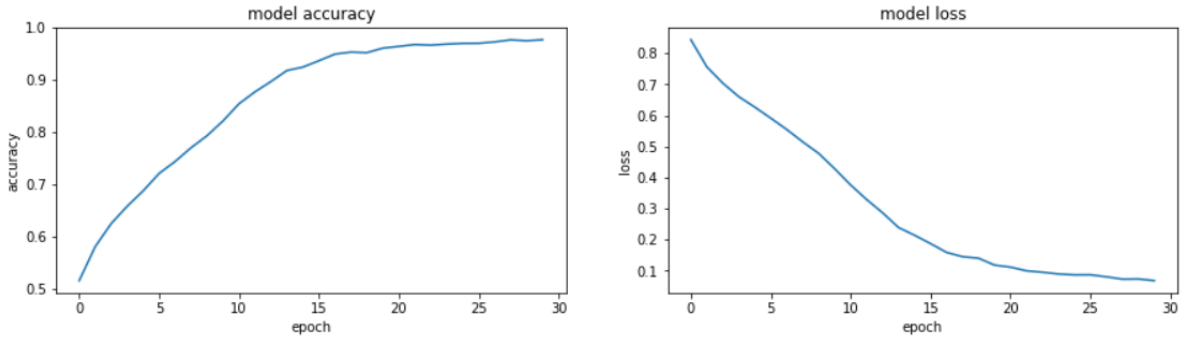


Fig: Multi-CNN Plot

6.1 Discussion

The main idea behind this research paper is create a methodology to detect phishing emails using multi-Convolutional neural network. This approach uses image recognition technique which produces more effective results in classifying legitimate and phishing emails. The cybercriminals use phishing as a tool to trick the users and steal information and get access through the network. It is important detect and prevent the malicious URLs at the server end before entering the system. This implementation in real world helps to secure the corporate security and the individuals can protect their confidential information safe. The awareness should impose to every user and employee so that they can protect themselves in being failing in trap.

7. Conclusion and Future work

This research proposed a deep learning algorithm based multi-convolutional neural network in detecting phishing emails. With the phishing datasets, two models, Model1 and Model2 are created and trained and tested. The multi-CNN produces better accuracy than single CNN. The single-CNN provides 62% and multi-CNN provides 98% accuracy. The advantage of this research is that we have used same dataset that has been given to both the model, thus processing time gets reduced and space utilization is achieved. The more powerful data we collect for this research will help in classifying the phishing emails much more accurate in less time with limited iterations. Due to the limited time , this model cannot be able to differentiate with any other machine learning in this paper. Thus, the future work of this research will be comparing the existing model other machine learning and deep learning techniques. To create a framework that able to detect the malicious request and block them in the network.

References

‘Accuracy, Precision, Recall & F1 Score: Interpretation of Performance Measures @ blog.exsilio.com’ (no date). Available at: https://www.linkedin.com/pulse/confusion-matrix-accuracy-precision-recall-f1-score-measures-silwal/?trk=pulse-article_more-articles_related-content-card (Accessed: 14 August 2022).

Alkawaz, M. H., Steven, S. J. and Hajamydeen, A. I. (2020) 'Detecting Phishing Website Using Machine Learning', *Proceedings - 2020 16th IEEE International Colloquium on Signal Processing and its Applications, CSPA 2020*, (Cspa), pp. 111–114. doi: 10.1109/CSPA48992.2020.9068728.
Bhagwat, M. D. (2021) 'Identification and Proactive Prevention of', (Icicv), pp. 1505–1508.
Dey, N. *et al.* (2022) 'Analysis of Machine Learning Algorithms by Developing a Phishing Email and Website Detection Model', pp. 1–7. doi: 10.1109/csitss54238.2021.9683131.
Fang, Y. *et al.* (2019) 'Phishing Email Detection Using Improved RCNN Model With Multilevel Vectors and Attention Mechanism', *IEEE Access*, 7, pp. 56329–56340. doi: 10.1109/ACCESS.2019.2913705.
Feroz, M. N. and Mengel, S. (2015) 'Phishing URL Detection Using URL Ranking', *Proceedings - 2015 IEEE International Congress on Big Data, BigData Congress 2015*, pp. 635–638. doi: 10.1109/BigDataCongress.2015.97.
'Frontiers Phishing Attacks: A Recent Comprehensive Study and a New Anatomy Computer Science' (no date). Available at: https://www.frontiersin.org/articles/10.3389/fcomp.2021.563060/full (Accessed: 14 August 2022).
Kaddoura, S., Alfandi, O. and Dahmani, N. (2020) 'A Spam Email Detection Mechanism for English Language Text Emails Using Deep Learning Approach', *Proceedings of the Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE*, 2020-Septe, pp. 193–198. doi: 10.1109/WETICE49692.2020.00045.
Li, X., Zhang, D. and Wu, B. (2020) 'Detection method of phishing email based on persuasion principle', *Proceedings of 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference, ITNEC 2020*, (Itnec), pp. 571–574. doi: 10.1109/ITNEC48623.2020.9084766.
Ma, L. *et al.* (2009) 'Detecting phishing emails using hybrid features', *UIC-ATC 2009 - Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing in Conjunction with the UIC'09 and ATC'09 Conferences*, pp. 493–497. doi: 10.1109/UIC-ATC.2009.103.
McGinley, C. and Monroy, S. A. S. (2022) 'Convolutional Neural Network Optimization for Phishing Email Classification', pp. 5609–5613. doi: 10.1109/bigdata52589.2021.9671531.
Niu, W. *et al.* (2018) 'Phishing emails detection using CS-SVM', *Proceedings - 15th IEEE International Symposium on Parallel and Distributed Processing with Applications and 16th IEEE International Conference on Ubiquitous Computing and Communications, ISPA/IUCC 2017*, pp. 1054–1059. doi: 10.1109/ISPA/IUCC.2017.00160.
Park, G. and Stuart M, L. (2014) 'comparing machine and human ability to detect phishing emails', pp. 5–10.
Patil, P., Rane, R. and Bhalekar, M. (2017) 'and Obfuscation URL Detection Algorithm', pp. 1–4.
Phishing activity trends report 2018| APWG (no date). Available at: https://apwg.org/trendsreports/ (Accessed: 14 August 2022).
Ripa, S. P., Islam, F. and Arifuzzaman, M. (2021) 'The emergence threat of phishing attack and the detection techniques using machine learning models', *2021 International Conference on Automation, Control and Mechatronics for Industry 4.0, ACMI 2021*, 0(July), pp. 8–9. doi: 10.1109/ACMI53878.2021.9528204.
Sindhu, S. *et al.* (2020) 'Phishing detection using random forest, SVM and neural network with backpropagation', *Proceedings of the International Conference on Smart Technologies in Computing, Electrical and Electronics, ICSTCEE 2020*, pp. 391–394. doi: 10.1109/ICSTCEE49637.2020.9277256.
Singh, S., Singh, M. P. and Pandey, R. (2020) 'Phishing detection from URLs using deep learning approach', *Proceedings of the 2020 International Conference on Computing,*

Communication and Security, ICCCS 2020, pp. 2020–2023. doi: 10.1109/ICCCS49678.2020.9277459.
Tan, C. L. (2018) 'Phishing dataset for machine learning: Feature evaluation', *Mendeley Data*, 1, p. 2018. Available at: https://data.mendeley.com/datasets/h3cgnj8hft/1 (Accessed: 14 August 2022).
Using Convolutional Neural Network for Image Classification | by Niklas Lang | Towards Data Science (no date). Available at: https://towardsdatascience.com/using-convolutional-neural-network-for-image-classification-5997bfd0ede4 (Accessed: 14 August 2022).
Xing Fang, N. K. (2004) 'An artificial immune system for phishing', *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, 3029, pp. 219–228. doi: 10.1007/978-3-540-24677-0_24.
Zaimi, R., Hafidi, M. and Lamia, M. (2020) 'Survey paper: Taxonomy of website anti-phishing solutions', *2020 7th International Conference on Social Network Analysis, Management and Security, SNAMS 2020*. doi: 10.1109/SNAMS52053.2020.9336559.