FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

EMAIL AND SMS SPAM DETECTION BASED ON DEEP LEARNING

Abdullahi Abba Abdullahi

Master's Thesis

DEPARTMENT OF COMPUTER ENGİNEERİNG

FEBRUARY 2021

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Computer Engineering

Master's Thesis

EMAIL AND SMS SPAM DETECTION BASED ON DEEP LEARNING

Author
Abdullahi Abba Abdullahi

Supervisor
Prof. Mehmet KAYA

FEBRUARY 2021
ELAZIG

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Computer Engineering

Master's Thesis

Title:	Emaıl and SMS Spam Detectıon Based on Deep Learnıng
Author:	Abdullahi Abba Abdullahi
Submission Date:	26 January 2021
Defense Date:	12 February 2021

THESIS APPROVAL

This thesis, which was prepared according to the thesis writing rules of the Graduate School of Natural and Applied Sciences, Fırat University, was evaluated by the committee members who have signed the following signatures and was unanimously approved after the defense exam made open to the academic audience.

		Signature
Supervisor:	Prof. Mehmet KAYA Firat University, Faculty of Engineering	Approved
Chair:	Assoc.Prof. Taner TUNCER Firat University, Faculty of Engineering	Approved
Member:	Assist.Prof. Serpil ASLAN Malatya Turgut Özal University, Faculty of Engineering	Approved

This thesis was approved by the Administrative Board of the Graduate School on

..... / ..... / 20 .....

Signature

Doç. Dr. Kürşat Esat ALYAMAÇ
Director of the Graduate School

DECLARATION

I hereby declare that I wrote this Master's Thesis titled “Emaıl and SMS Spam Detectıon Based on Deep Learnıng” in consistent with the thesis writing guide of the Graduate School of Natural and Applied Sciences, Firat University. I also declare that all information in it is correct, that I acted according to scientific ethics in producing and presenting the findings, cited all the references I used, express all institutions or organizations or persons who supported the thesis financially. I have never used the data and information I provide here in order to get a degree in any way.

12 February 2021

Abdullahi Abba Abdullahi

PREFACE

Communication has evolved through the use of technological medium to transmit information since the emergence of smoke and then symbols by humans to communicate during ancient society. Communication medium such as electronic-mail and short message service (SMS) plays significant role in simplifying the exchange and access to information in the twenty-first century society. But over the last decade, there have been increase in spam to these communication medium which has become a major issue to security of user of these medium. Deep Learning, a subset of Artificial Intelligence is a modern technological approach to curb the threat pose by spam in any communication medium.

Alhamdulillah, Alhamdulillah, Alhamdulillah, All Praise is due to ALLAH (S.W.T) for the gift of life and the ability to finish this thesis.

I would like to thank my supervisor, Prof. Dr Mehmet Kaya, for all the guidance he provided and all he has done for me. I especially enjoyed the patience and help he has render to me. May ALLAH (S.W.T) Bless Him and His Family Abundantly Here in the Duniya and Aljannatul Firdausi in Al-Akhra.

To all my friend here in Elazig and back home, My Dear Wife, My Dear Son, My Dear Parents, My Family back home, My New Families in Turkey and the entire Muslim Umma, May ALLAH (S.W.T) Bless Us All with Aljannatul Firdausi. Ameen Ya Rabbi.

Abdullahi Abba Abdullahi
ELAZIG, 2021

TABLE OF CONTENTS

Page

ABSTRACT

Emaıl and SMS Spam Detectıon Based on Deep Learnıng

Abdullahi Abba Abdullahi

Master's Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences

Department of Computer Engineering

February 2021, Page: xi + 31

Over the last decade, the overwhelming use of smartphone have made users depend on it for digital communication such as Email and SMS. Consequently, the use of email and SMS services as a medium for "two-way communication" and "one-way communication" such as notification, alerts, reminders etc., by individuals and organizations has increased tremendously. Technological advancement, cost-effectiveness, high speed acknowledgment time and viability are factors that have led to their exponential growth in the 21st century. However, these communication mediums have been vulnerable to malicious attacks called spams. The constant rise of spam across these communication medium has questioned the demand for solutions. Thus, the need for ML/DL solution which without specific instructions (rule-based codes) use algorithms and statistical models to filter spam by relying on patterns and inferences in the content of message. In this thesis, the proposed method parses the message, learns from it and decides if it is ham or spam. This approach is economical, faster, and efficient than the other two traditional methods since it can filter huge messages in short time and the ML/DL algorithms learn from the previous messages and apply it to incoming messages.

Keywords: Deep learning, Machine learning, Dense Neural Network, Spam detection.

ÖZET

Derin Öğrenmeye Dayalı İstenmeyen Email ve SMS Belirleme

Abdullahi Abba Abdullahi

Yüksek Lisans Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Şubat 2021, Sayfa: xi + 31

Son on yılda, akıllı telefonların yaygın kullanımı, kullanıcıları e-posta ve SMS gibi dijital iletişim için telefonlara bağımlı hale getirmiştir. Ayrıca, bireyler ve kuruluşlar tarafından bildirim, uyarı, hatırlatma vb. gibi "iki yönlü iletişim" ve "tek yönlü iletişim" için e-posta ve SMS hizmetlerinin bir araç olarak kullanımı muazzam bir artış göstermiştir. Teknolojik ilerleme, maliyet etkinliği, yüksek hızlı bilgilendirme süresi ve uygulanabilirlik, 21. yüzyılda akıllı telefonların kullanımının üstel büyümesine yol açan faktörlerdir. Ancak, bu iletişim araçları spam adı verilen kötü niyetli saldırılara karşı savunmasızdır. Bu iletişim araçlarında sürekli spam artışı, çözüm talebini de artırmıştır. Böylece, belirli komutlar (kural tabanlı kodlar) olmaksızın bir ML / DL çözümü mesajın içeriğindeki kalıplara ve çıkarımlara dayanarak spam filtrelemek için algoritmalar ve istatistiksel modeller kullanır. Bu tezde önerilen yöntem mesajı ayrıştırır, ondan öğrenmeyi gerçekleştirir ve spam olup olmadığına karar verir. Bu yaklaşım, diğer iki geleneksel yöntemden daha ekonomik, daha hızlı ve daha verimlidir. Çünkü çok büyük mesajları kısa sürede filtreleyebilir ve ML / DL algoritmaları önceki mesajlardan öğrenme işlemini gerçekleştirir ve bunu gelen mesajlara uygular.

Anahtar Kelimeler: Derin öğrenme, Makine öğrenmesi, Derin sinir ağı, Spam algılama.

LIST OF FIGURES

Page

LIST OF TABLES

Page

SYMBOLS AND ABBREVIATIONS

Abbreviations

DL	: Deep Learning
ML	: Machine Learning
Email	: Electronic Mail
FN	: False Negative
FP	: False Positive
NLP	: Natural Language Processing
SMS	: Short Message Service
TN	: True Negative
TP	: True Positive

1. INTRODUCTION

Communication has evolved through the use of technological medium to transmit information since the emergence of smoke and then symbols by humans to communicate during ancient society. Communication medium such as electronic-mail and short message service (SMS) plays significant role in simplifying the exchange and access to information in the twenty-first century society [1]. In recent years, there has been an explosion in growth of number of users of these communication medium due to their availability, reliability and low cost of connection. The internet and mobile communication provides the avenue for these mediums to be used to communicate.

Nowadays, the internet offers a diverse range of knowledge, tools, resources, product and services. One of the most common form of the application of the internet is the Email service, which is the generally known computer network web application frequently used by internet end users around the world for private, corporate, organizational, commercial, governmental or official purposes to transmit messages with an option to append all types of files such as images, sounds or recorded sound and image in a form called attachment.

Subsequently, another great technological invention for communication and integral part of human lives is the mobile phone. it contains several components for the purpose of communication. One of its most commonly used component today is Short Message Service (SMS). It allows users of GSM, TDMA, CDMA mobile network to exchange/send messages usually not more than 160 character through a "store and forward" standard [1].

In fact, different communication technologies, such as mobile banking app, voting app, special purpose app, social networking app, and public health initiatives app have all taken advantage of E-Mail and SMS services to communicate with users of this apps. Although, these two medium of communication (E-mail and SMS) have brought a lot of satisfaction in the way people communicate, they are becoming more vulnerable to malicious attacks. One of such threat and popular attack is spam. Thus, the need for solutions to tackle the up rise of spam attack to users of these communication mediums.

1.1. Problem Statement

Over the last decade, the overwhelming use of smartphone have made users depend on it for digital communication such as Email and SMS. Consequently, the use of email and SMS services as a medium for "two-way communication" and "one-way communication" such as notification, alerts, reminders etc., by individuals and organizations has increased tremendously. Technological

advancement, cost-effectiveness, high speed acknowledgment time and viability are factors that have led to their exponential growth in the 21st century.

According to an intelligence forecasts report by GSMA [2], 3.8 billon of the world's populace use mobile internet. Another survey by [3], shows that the major operations on smartphones owners are 91% and 90% accessing Email and SMS messages respectively. However, these communication mediums have been vulnerable to malicious attacks called spams. The constant rise of spam across these communication medium has questioned the demand for solutions.

There are two traditional approach to filter spam in these communication mediums. The first, is through human intervention to scan through potential harmful messages or links that are suspicious. This method is inefficient as it will take a lot of time to go through large messages to filter out spam. Thus, making it a tiring, tedious and complex task to perform.

The second alternative is to use traditional "rule based" otherwise known as “system oriented approaches” which uses approaches such as “white lists”, “black lists” and” finger print” in coding to filter out spam in these mediums which is faster and effective method than human intervention. However, due to the ever evolving nature of spammers, this method of coding has turned out to be "dumb" in circumstances where the rule changes in pattern or a malicious modification in the network traffic behavior by spammers [4]. This means the program will never be able to filter spam except if the coding is modified by the developer to serve its purpose. Consequently, making these methods a daunting task for spam filtering. Thus, the need for ML/DL solution which without specific instructions (rule-based codes) use algorithms and statistical models to filter spam by relying on patterns and inferences in the content of message; that is, it parse the message, learn from it and decides if it is ham or spam.

This approach is economical, faster, and efficient than the other two traditional methods since it can filter huge messages in short time and the ML/DL algorithms learn from the previous messages and apply it to incoming messages.

1.2. Aim and Objectives of the Study.

The aim of this thesis is to implement ML and DL classifiers on both Email and SMS datasets and making a comparison of both classifiers on the two datasets using known classification evaluation metrics. The objectives of the study are:

i. Obtain Email and SMS datasets from UCI repository.
ii. Perform preliminary analysis in order to comprehend the datasets using open source tools such as word cloud and bar chart.
iii. Preprocess the unstructured datasets to eliminate external or insignificant data.
iv. Split the processed/structured datasets into training and testing data for modeling and evaluation.

v. Make a comparison of the model from the evaluation metrics.

1.3. Contribution of the Study.

Previous studies on spam filtering techniques mostly compared ML classifiers or both ML and DL classifiers either on Email or SMS dataset. However, this thesis contributes by implementing ML and DL classifiers both on Email and SMS dataset. As a result, presents comparison result for classification performance of ML and DL models on the two datasets using known classification evaluation metrics.

Furthermore, this thesis work will contribute to existing research knowledge on spam filtering using ML and DL classification algorithms by providing analytical insight on the methodology used.

Lastly, it will provide researchers and information security firms/practitioners the opportunity to consolidate upon this work; Similarly, produce a cutting edge optimized model for spam filtering in E-mail and SMS medium of communication which will sequentially protect users of these mediums from the harmful effect of spam.

1.4. Thesis Organization.

There are six chapters in this thesis. Chapter 1 introduce the history of communication, modern communication medium and their vulnerability, statement of the problem, aim and objectives of the study and lastly, the contribution of the study.

Chapter 2 briefly explained terms related to the study. Chapter 3 reviewed works related to Email and SMS spam filtering using ML/DL techniques.

Chapter 4 described the methodology used to implement the ML/DL techniques on Email and SMS datasets. Chapter 5 presented the result and made comparisons of the ML/DL models.

Chapter 6 summarize the study, conclude and make future recommendations.

2. OVERVIEW OF TERMINOLOGIES

2.1. Introduction

This chapter begins by discussing the concept of email and SMS spam techniques. It also discusses the various types of spam techniques such as appending spam, image spam, blank spam, and backscatter spam. The Spam filtering process such as content, header, blocklist, rule-based, and language spam filtering techniques are also explained in the chapter. Moreover, text cleaning and feature extraction data preprocessing techniques are also present in this chapter.

2.2. Overview of spam and techniques

The term spam is not an acronym for anything rather it is a reference to any unsolicited sent or received an email or short messaging service (SMS) message [5]. Most of these spam emails and SMSs sent or received are commercial, perhaps that is why sending commercial email was prohibited at the beginning of the introduction of ARPANET [6]. In this study, four types of spam techniques will be introduced as follows:

1. Appending spam: Appending spam is a situation where a marketer has a database with customer details. Then the marketer will request or sometimes make a payment to have their database details to be matched with an external database, which also contains the email or phone number of users. In most cases, the external company would have the resources to send email and SMS messages to users who did not even request the email or the SMS.
2. Image spam: As opposed to text-based messages, image spam is an image-based message that hid the text in an image sent to the user's email or phone. The main idea of hiding the text message in the image in image spam messages is to avoid the text been discovered by any spam detecting or filtering mechanism.
3. Blank spam: As the name suggested, this type of spam appears to be a blank email or SMS while the truth of the matter is that it is not. In a blank spam email or SMS, the recipients will see that the body and subject of the message are missing. For example, VBS.Davinia.B is a worm most sent via email messages that appear as blank to the email users, but this message has a hidden hypertext markup language code that will execute and download other files from the user's device.
4. Backscatter spam: In a situation where email service providers or mobile phone service provider's servers are misconfigured than they end up sending bogus messages, and this is known as backscatter spam. However, in the case where the sender's address has forged the email or SMS message will end up going to an innocent third party.

2.3. Spam filtering technique

The main task of spam filters for either email or SMS is to detect any unsolicited email or unwanted SMS message or have a virus, etc., and stop it from getting into the user's emails of mobile phones. Most email and mobile phone service providers are using spam filtering techniques to filter both outgoing and incoming spam messages from entering their networks [7]. The use of spam filters has become necessary for the service providers as the unsolicited emails and SMS have proved to be expensive for the service provider as they lead to things like loss of subscribers [8]. Even so, 100% spam filtering cannot be achieved, and globally 70% of email messages are classified to be spam emails [9], which are increasing in number because the spammers are also getting more sophisticated and creative. Although there are many spam filters out there, such growth in talents and knowledge the spammers are getting is the reason why there must be continuous updates on spam filtering techniques for both email and SMS messages [10].

2.3.1. Types of spam filters

Spam filters use algorithms that are heuristics with specific predefine rules to check the content of each email or SMS message [11]. In most cases, a threshold is defined and if the probability score of the content of the message passes the threshold then the message is classified as spam. Below are different types of spam filters for different scenarios:

1. Content spam filters: The content spam filters are the categories of spam filters that scan and analyze the words in an email or SMS message and compared them to the words that are commonly used in spam messages. The result of the comparison will be used to flag the received or sent the message as either spam or non-spam message.
2. Header spam filters: In this type of spam detection technique on the message header is considered. The spam filter will analyze the message header and look for any unwarranted information like withheld information or the spammer’s email address then it will detect and classify the status of the received message.
3. Blocklist spam filters: This spam filtering technique is used to block any incoming message from a flag list of spammers IP addresses.
4. Rules-based spam filters: Mostly this kind of spam filter is implemented at the organization levels. It is designed to exclude messages from some specific users that contain some specific wordings.
5. Language spam filters: The language spam filter is a language sensitive spam filter that categorized any email or SMS messages that is not in the language of the recipient as a spam message.

2.4. Artificial Intelligence, Machine Learning and Deep Learning

Artificial intelligence (AI) relates to systems that exhibit intelligent behavior in order to accomplish specific objectives by observing their environment and taking actions often with level of freedom. In the virtual world, AI-based systems can solely be software-based such as speech and face recognition systems or enclosed with hardware devices such as autonomous cars. Artificial Intelligence's primary purpose is to build systems which are self-reliant and can think and behave like humans [12]. By learning and problem-solving, these systems can imitate human actions and accomplish actions. To solve complex problems, the majority of AI systems simulate natural intelligence. Below are categories of Artificial Intelligence scenarios:

1. Reactive Devices - These are just responding systems. Such systems do not create memories, and they do not use any previous experiences to make new decisions.
2. Limited Memory - These systems relate to the past, and over a period of time, information is added. The information referred is brief.
3. Mind Theory - This includes systems that are capable of understanding human thoughts and how they impact decision-making. They are taught to change their actions appropriately.
4. Self-awareness - In order to be conscious of themselves, these type of systems are developed. They perceive their own inner states, anticipate the feelings of other individuals, and behave accordingly [12].

Figure 2.1 below shows the relationship of Artificial Intelligence, Machine Learning and Deep Learning.

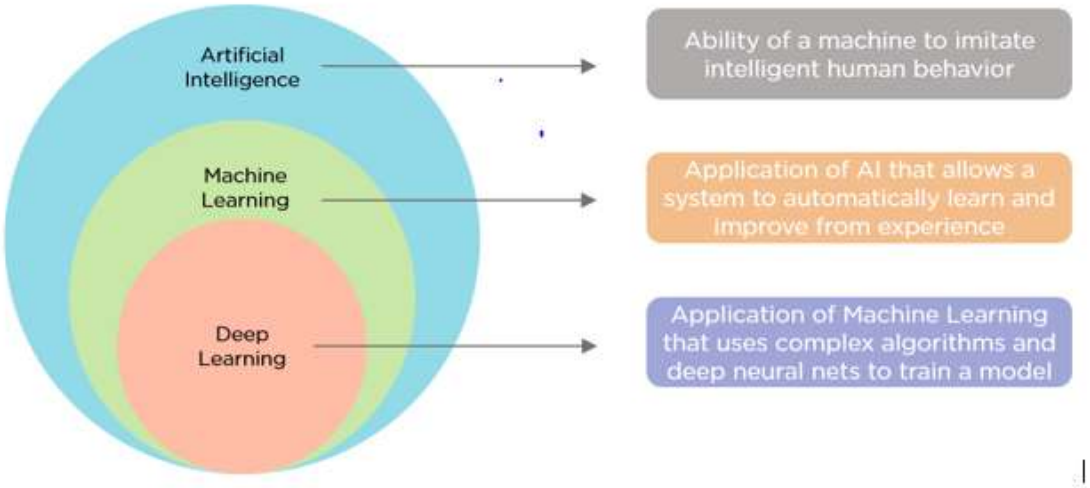

Figure 2.1. Relationship of Artificial Intelligence, Machine Learning and Deep Learning

2.4.1. Machine Learning

Machine learning is a subset and key component of artificial intelligence that uses computer algorithms and analytics to creates predictive models that can solve business problems. it also aims to provide information and inferences on global activities to computers which it uses to learn and then make decisions. The information gained enables computers to respond to changing conditions accurately. To foretell an outcome, machine learning learns from large quantities of structured and unstructured data by using several algorithms and techniques. There are three main classes of ML which are listed below [12]:

i. Supervised Learning: In supervised learning, the features in the data are labeled showing target variable. Systems can predict future results based on past data using this method of learning. This involves that the model be provided both an input and output variable for it to be trained.
ii. Unsupervised Learning: Using unlabeled input data, unsupervised learning systems are able to recognize hidden features and build models on their own from the data. The trends and similarities become more apparent until data is more understandable.
iii. Reinforcement Learning: training an agent under an unpredictable environment to complete a mission is the primary purpose of developing reinforcement learning applications. Through collecting reward and environmental observations, the agents send behavioral patterns to the environment. This incentive test how effective it was to achieve the goal of the mission.

2.4.2. Deep Learning

Shaped in the form of the human brain, deep learning is a subset of machine learning that is concerned with algorithms that works on overwhelming size of both structured and unstructured data. Thus allowing decision making in its neural networks by computers makes it the basic principle of deep learning.

The network has an input layer which accepts the data inputs and then calculate its weighted sum. To find any hidden features of the data, the hidden layer is used to transfer the sum of weights as activation function input which applies a bias, and determines whether or not the neuron should be fired. Then the output layer provides the output predicted which are contrasted with the real output. The model uses the backpropagation method after training the neural network to boost the network performance. The cost function helps to lower the rate of error. Figure 2.2 below shows the working of deep neural network [12].

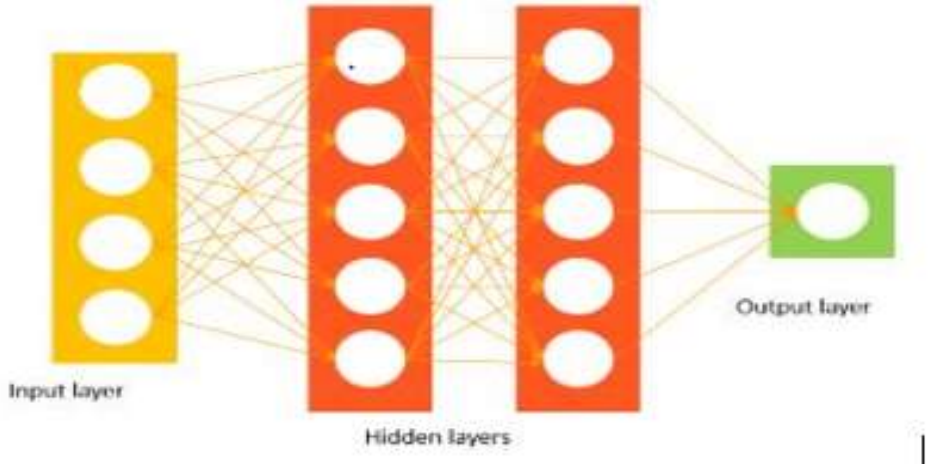

Figure 2.2. Deep Neural Network

2.5. Natural Language Processing

Natural Language Processing is the dedicated unit of Artificial Intelligence aimed at helping computers understand linguistics, statements or printed text in human language. Ease and desire to communicate with computer, brought NLP into being. A language describes a collection of rules, symbols or mixed to express or transmit information. The knowledge of human language is one features of NLP which is connected to various theories and techniques that communicates with computers. Natural Language Processing can essentially be divided into two groups. These are Natural Language Understanding and the task of developing understanding and generating the text called Natural Language Generation. Figure 2.3 below shows the division of NLP and their components [13].

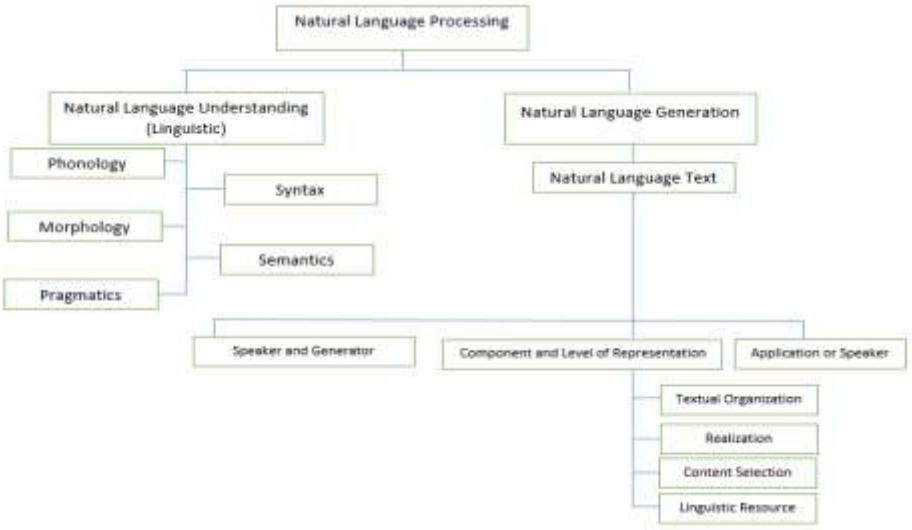

Figure 2.3. Division of NLP and their Components

2.6. Uses of Natural Language Processing

The field of NLP has experienced tremendous growth over the last decade, this is due to its application in various human endeavor and business organization. Some of the fields NLP can be used are spam detection, summarization, chatbots, machine translation, grammar checker, voice assistance, targeted advertising, search autocorrect and autocomplete, social media monitoring, sentiment analysis and many more [13].

2.7. Text cleaning

In machine learning and deep learning data preprocessing, text cleaning is an important technique of removing noise, white spaces, bogus links, and unwanted characters from the text of the email or SMS message [14]. Researchers use a different approach for text cleaning. For example, removing characters, numbers, hyperlinks, white spaces, stop words, punctuation, or even converting the alphabets from lower case to upper case or vice versa. Other techniques include word stemming and lemmatization [15]. Word stemming is a process of removing prefixes and suffixes in wording according to the language of the words. Table 2.1 shows an example of word stemming in the English language.

Table 2.1. Word stemming example

Word form	Suffix/Prefix	Stem
Running	-ing	Run
Runner	-ner	Run
Predefined	-pre	Defined
Consolidate	-ate	Consolidate
Uncover	-un	Cover
Review	-re	View

Word lemmatization is the process of changing words of a particular language into their base form in the language. For example, given in Table 2.2 is the lemmatization process of some words in the English language.

Table 2.2. Word lemmatization example

Word form	Lemma
Studies	Study
Act	Actor
Form	Formulate
Ran	Run
Norm	Normality
Ego	Egomaniac

2.8. Feature extraction

In most cases, the machine learning or deep learning algorithms are expecting numerical integer of float inputs. Therefore; this necessitates the need for a future extraction layer that will be responsible for converting the words in the email or SMS text to integers or floats [16]. The common techniques of future extraction in a machine and deep learning methods are CountVectorizer, TF-IDF-Vectorizer, and word embedding.

The CountVectorizer future extraction techniques are achieved through feeding the training data to the CountVectorizer function, which will set up a dictionary and id of each word. The techniques required the word id to keep track of the word count throughout the training period. The major issue with this future extraction is that some words that are not going to add more information to the data can be encountered too many times.

Therefore, the TF-IDF-Vectorizer is another alternative approach to feature extraction, which calculates the frequency of the words using term frequency and inverse document frequency. The term frequency is used for counting the number of times a word appears in the document while the inverse document frequency is used for downscaling the words that have too much appearance across the documents under investigation. Hence, the technique can be called word frequency scores because it learns from the document vocabulary and inverses the document frequency, which gives insight to interesting words, for example, frequent words in a particular document but not across all documents.

The word embedding future extraction process begins by converting the word into a vector and then the vector will be the word representation in a higher-dimensional space. In this regard, similar words will have a short distance-vector so that they will be closer to one another. Since the words themselves are vectors, this will allow mathematical operations to be carried out on the words. The word embedding technique is one of the most popular processes of future extraction technique that is widely implemented using the Word2Vec algorithm.

3. RELATED WORKS

3.1. Introduction

This chapter presents the literature review of previously published scholarly materials about email and SMS spam detection techniques. The scholarly written materials that are included in the chapter are collected from journal articles. The content of the articles is screened to establish if the content is relevant for inclusion. The chapter begins by explaining the various algorithms and detection techniques used for filtering email spam. It also explained the process of SMS spam detection and filtering technique.

3.2. Email spam

In the paper of Bhuiyan et al. [17], email is described as one of the most used communication mediums for web data exchange that is seeing both a tremendous increase in popularity and unsolicited data. This increase in unsolicited data presses the need for filtering techniques to reduce the amount of spam email messages. The paper surveyed and studied different machine learning algorithms used for email spam filtering. The findings of the work are presented in terms of accuracy rate, classification, evaluation, and filtering algorithm comparisons. The major conclusion of the study is that on average all the methods under review are effective but recommends that there is a need for next-generation email spam filtering algorithms, which can be applied to a large number of multimedia data to achieve more prominent email spam filtering technique.

Email is a cheap way of communication between different units of an organization that has become necessary for any business to achieve competitiveness and sustainability as highlighted by Trivedi and Dey [18]. The study also noted that there is an increase in complexity of spam emails send by the spammers which have led to difficulty in differentiating spam and non-spam emails. Considering this, the study proposed a novel method that involves combining the classifier with committee selection and president. Bayesian and Naïve Bayes probabilistic classifiers are selected as the committee while the support vector machine is selected as the president. The justification for this committee and president selection given in the study is that machine-learning classifiers had shown good performance in email spam classification, also there is a need to identify the best classifier combination. However, the study also points out that low false positive is still issued yet to be addressed. Findings of the study show that the classifier combination yields sensitive classification with reduced informative features. The results of the study are benchmark with other machine learning classifiers and it shows better performance and less false positive rate. In [19] the authors show the effect of using feature selection on machine learning classifiers in spam email detection. The findings show that concerning feature selection in spam email detection, Greedy

Stepwise Search can yield good performance. While the support vector machine yields the best performance concerning the accuracy and false positive rate among machine learning classifiers algorithms.

A comparative study of various spam email classification algorithm techniques is provided in the work Elhamayed [20]. The classification techniques used in the work are lazy, rules, function, Bayesian, and tree while the comparison was done based on some performance parameters of a selected spam dataset. Performance of the classification algorithms is calculated and recorded using a different ratio of the training and testing periods in the dataset which shows that the is an increased performance when there is an increase in the ratio of the training phase. The results show that when a filter is used on the dataset, J48, and PART classifiers yield the best performance while KNN yields the best performance for low variance filter with future selection. The study recommends the use of other future selection techniques to improve the performance of the classifier.

Another work about the ham and spam email classification techniques using machine learning methods is presented by Bassiouni et al. [21]. The work surveyed the ten most efficient classification techniques out there. This survey shows that the most widely used databases in spam email classification are Spambase data from the UCI machine learning repository and Enron. Moreover, the survey also indicates how researchers are either using filtering or machine learning techniques or the combination of the two to achieved spam detections. Taking all this information into account, the work proposed a classification method based on data preprocessing, ILFS for future selection, and data classification using the ten surveyed classification techniques and the Spambase dataset as a case study. Results of the work show that spam email classification can be achieved with Random Forest techniques having the best performance among the ten benchmarked classification techniques.

Nandhini and Marseline [22] also used the Spambase dataset from the UCI machine learning repository to evaluate the performance of five machine learning classification algorithms. The study was motivated by the need to provide a solution to the increased security threats being posed by spammers to email users. The performance evaluation results are used to train and proposed a spam detection model, which is based on the combination of various machine-learning methods. The random tree technique outperforms other classification techniques in all performance metrics from the result of the study. However, the random tree does not outperform KNN rather it is faster in yielding the results. In conclusion, because KNN is slow to build the model when compared to the Random tree, therefore, Random tree is the recommended model in this study.

The paper of Singh and Bhardwaj [23] discusses the strength and shortcomings of current spam email classification methods using the global training set as a case study. The scope of the work combined classification methods and knowledge engineering to propose a better filtering process. The highlights of the paper are (1) classification techniques require data training before

the classification job, (2) spam persist even after training of thousand samples of data because new spam email emerges almost on daily basis. The paper also makes the following recommendations: (1) data training set should be updated daily by feeding it with daily new email spams, (2) global training set collection center should be established to enable users to contribute and report new spam as soon as they see it. The recommendations of the paper can raise spam email detection efficiency if it can be successfully applied.

In another comparative study of classification algorithms used for spam email detection by Sharaff et al. [24], a comparison and discussion of four machine learning classification algorithm effectiveness is presented. The authors also discuss how spam emails cause inconveniences such as the consumption of space and bandwidth to the users. It has also been pointed out that spammers are always trying to tackle the filtering algorithms, which is making it difficult to stop spam emails. Different performance measures are used to evaluate the finding of the study. The method is applied to the Enron dataset collected from the Athens University of Economics and Business. The obtained result differ from the norms in the literature because it shows that BayesNet and J48 outperform the support vector machine classification algorithm, which indicates that data can reshape algorithm performance.

Dada et al. [25] conduct a review on approaches and open research problems on machine learning methods use for spam email detection and classification. The scope of the systematic reviews include topics such as important concepts, efficiency, and annals of trends in spam email filtering using machine learning algorithms. Using leading email service providers; Yahoo, Gmail, and Outlook as a case study, the paper discuss the strength and weakness of machine learning spam email classification techniques. Besides, issues and open research problems concerning machine learning techniques use in spam email filtering are outlined in the paper. The study recommends that to overcome the persistence of spam emails, deep learning, and deep adversarial learning machine-learning techniques should be employed to tackle the problem.

In the analysis study of the Naïve Bayes algorithm for email spam filtering across multiple datasets, Rusland et al. [26] test the algorithm using spam data and the Spambase datasets. The power of the algorithm is evaluated using the constraints of precision, F-measure, and algorithm accuracy. The findings of the work which was conducted using the WEKA tool environment show that data instances influence performance in the Naïve Bayes algorithm. Another comparative analysis of the classification algorithm for email spam detection by Abdulhamid et al. [27] compares different machine learning algorithms used in email spam detection. Findings of the work indicate that all algorithms were able to achieve spam email sorting with Rotation Forest having the best accuracy of 94.2%.

Zhang et al. [28] in their paper about the evaluation of statistical spam filtering techniques evaluates five supervised machine learning spam filtering techniques. The paper discusses how

researchers tend to ignore email headers during the spam filtering process. Results presented in the paper show that when a message that appears on the email header is classified, then a better performance can be archived when compared to the classifying techniques that only consider features from the body of the email. A novel hybrid approach for future selection optimization in spam email classification is proposed in the study of Hassani et al. [29]. The method is applied to the UCI machine learning repository Spambase dataset and it yields 97.61% accuracy one of the highest accuracy seen in this dataset. Similarly, Alauthman [30] proposed GRU-RNN with SVM for spam email classification and tested the method on the Spambase dataset from the UCI machine learning repository. The method gives statistically acceptable results with an accuracy of 98.7%.

3.3. SMS spam

According to Xia and Chen [31], most machine-learning methods such as Naïve Bayes, SVM, LSTM, CNN, etc. are developed according to the assumption that the collection of words in an SMS are unordered. In an attempt to address this, the study proposed a discrete hidden Markov model for SMS spam detection that considers word order and low term frequency. The UCI Spambase dataset is employed to test the performance of the proposed approach. In addition to the Spambase dataset, a Chinese SMS spam dataset that contains about two thousand messages is also used to further test the performance of the approach. The results of the two different datasets under investigation show that the proposed method is not language sensitive and can accurately filter SMS spam with high accuracy.

There is a rapid increase in SMS spam messages due to the rapid growth of mobile phone users, this is what motivated Almeida et al. [32] to discuss and present the work about SMS spam filtering. The work highlighted that lack of SMS datasets and contributors is an issue that must be addressed. Therefore, the authors present the public and non-encoded SMS spam collection process. After the spam SMS collection, comprehensive analysis shows that the approach does not lead to any message redundancy. Machine-learning classification techniques are applied to the collected data and their performances are evaluated. The support vector machine algorithm was found to be superior to the other benchmarked classification techniques. The authors, therefore, recommend SVM to be used as a baseline classification method for future benchmarking.

The impact of deep learning techniques in the SMS spam filtering process is investigated by Gomaa [33]. The study noted that so far many researchers depend on manual future extraction using classical machine-learning classification techniques. The study proposed the use of a deep learning method on SMS spam filtering coupled with automatic future extraction. Respectively seven deep learning techniques and six classification techniques are used in the study. The method is applied to a dataset that contains 5574 instances and the experimental result obtained from the

Random Multi-model deep learning algorithm yields a 99.26% accuracy, hence, it is the selected deep learning process for the case study data.

Chandra and Khatri [34] proposed a novel approach that employs a recurrent neural network and LSTM to detect and classify SMS spam and ham messages. The method is implemented using the Keras model in the Tensorflow python programming backend. The study test the performance of the novel method using the popular UCI machine learning repository SpamSMSCollection dataset. Before applying the method to the dataset, the authors preprocessed the dataset using TF-IDF vectorization, tokenization, and stop words exclusions. To benchmark the method, the obtained results are compared with the results of Naïve Bayes and SVM techniques. The experimental results show that the proposed method only took 13.44 seconds to build the model with an accuracy of 98%, which is the better performance when compared to the benchmark methods.

In the work of Sravya and Guntur [35], SMS messages are categorized into ham and spam. On the ham side, a textual dataset of contents of SMS messages is recoded with a flag that is showing if the SMS content is still legitimate or not. While the textual dataset of spam SMS is also using the flag, but the difference is that the flag in the spam SMS dataset is indicating junk messages. The major drawback of this process is that it is not free, it requires the users to make a payment per any received SMS. However, to address this financial constraint, the study proposed a forecasting method that is based on a machine-learning algorithm. Therefore, the forecasted messages are used for feeding the textual dataset, and the accuracy of the method is calculated using the SMS spam dataset.

A review of soft techniques used for SMS spam classification is presented in the study of Alli et al. [36]. The study aimed to classify existing spam SMS detection studies according to the constraints of artificial intelligence methods, deployed environment approach, and current SMS applications acceptability by the users. In the quest to review as many existing studies as possible the authors explored eleven databases and considered 1198 publications. After applying the underlined inclusion and exclusion criteria developed by the study, only 83 papers were screened for inclusion in the study. The study conducted a quantitative evaluation of the 83 publications and finds that machine learning algorithms are used by 49% of the papers, then statistical analysis with 39%, and evolutionary algorithms 12%.

The paper of Jain et al. [37] proposed a semantic convolutional neural network layer to detect and classify spam SMS using the Spambase dataset and Twitter dataset. The proposed method was able to give a good performance with 98.65% accuracy on the Spambase dataset and 94.40% accuracy on the Twitter dataset. Another work by Kumar et al. [38] used the combination of convolutional neural networks and the long short-term memory deep learning models to detect and class SMS messages as spam or ham. The proposed approach considers text-only data or self-extracted future sets. The method is tested on an SMS text dataset that consists of 747 and 4827

spam and ham messages respectively. The experimental result using this dataset shows that the proposed method was able to achived 99.44% accuracy.

Alzahrani and Rawat [39] present a comparative study of machine learning algorithms for SMS spam detection. The study identified logistic regression, Naïve Bayes, SVM, and neural networks as the most popular machine learning algorithms used by researchers for SMS spam detection and classification jobs. Hence, the study adopts the use of these four machine-learning methods to detect and classify spam SMS messages. The accuracy of each algorithm is recorded to establish which of the algorithms has the best performance and accuracy. The obtained results of the algorithms are compared and the findings of the comparative analysis of the study show that the neural network method has the best performance. Perhaps this is because of the train classifier future that exists in the model which will give it the advantage of filtering any incoming SMS and classifying it as either spam or non-spam message.

Another review study on SMS spam filtering techniques by Abdulhamid et al. [40] describes spam SMS as an unsolicited message sent to the users of mobile phones and its popularity is becoming a major concern for the SMS service providers. Besides, the marketing cost is also a major issue for the service providers considering the way it makes the customers upset which in turn makes the providers lose subscribers. The paper review existing approach and their challenges. It also discusses what the future direction of SMS spam filtering techniques should be and what the future direction of mobile spam SMS mitigation should also be. The fining of the study and the recommendation provided by the authors can be used by future researchers to design solutions that will be addressed the open research gap identified in the study.

In most of the published literature SMS spam detection and filtering mechanism are applied in the case of incoming messages, however, that is not always the case. Bosaeed et al. [41] put forward a spam detection tool for both incoming and outgoing SMS messages. This has become necessary as highlighted in the study because a security threat to a mobile phone can cause a spam message to be sent out. The method used three classification methods and five future extraction methods to develop a multi-machine-learning classifier. The proposed method is tested using 15 public SMS datasets that can be used either on the cloud, fog, or the edge layers. The developed method was able to detect spam SMS and alert the filters and classification techniques that can handle the spam SMS.

Nagwani [42] proposed a bi-level SMS text classification approach that gives room for spam filtering and priority messages. The object of the approach is to provide an effective way of handling and managing SMS messages. In the first classification stage, the SMS messages are classified as spam or non-spam messages using the binary classification method. While in the second stage of the classification the non-spam email is divided into priority and non-priority messages. The categorization was done using four different machine learning classification

techniques. The finding of the study shows that SVM has the best performance concerning both filtering the spam SMS and categorizing the priority SMS messages.

Using the combination of neural language processing and deep learning approach the work Poomka et al. [43] proposed a novel method for SMS spam detection that is based on English language SMS messages. Before feeding the SMS dataset to the proposed model the authors preprocess the data using tokenization, word embedding, padding, and truncating data. The output of the proposed method shows that the model has an accuracy of 98.18%.

4. MATERIAL AND METHOD

For the formal process of conducting any research study, methodology is the conceptual concept used to achieve the aim and objectives of a research [44] . Therefore, the methodology in this study is made up of the description of the dataset.

4.1. Description of the Datasets

The Email dataset is divided into three parts: spam: 500 spam messages, which were obtained from non-spam-trap sources. 2500 non-spam messages: Easy ham. These are usually very easy to separate from spam, because spam signatures sometimes do not contain any (like HTML etc.). hard ham: 250 non-spam messages that in certain cases are similar to typical spam: HTML use, irregular HTML markups, colored text, "spammish-sounding phrases" etc. Easy ham 2: 1400 messages that are not spam. A newer addition to the collection. spam 2: 1397 messages with spam. More recent again. All making a complete count of: 6047 texts, with around a spam ratio of 31 percent which is obtained from [51].

The SMS dataset is a public set of messages labeled SMS that have been collected for research into spam on mobile phones. It has set of 5,574 English, real and non-encoded text, tagged as legitimate (ham) or spam. This corpus has been gathered from research sources on the Internet for free. A list of 425 SMS spam messages from the Grumbletext website was compiled manually for the first source. This is a UK platform where mobile phone users, many without reporting the spam message they received, make public comments about SMS spam messages. Recognizing the text of spam in the messages is by careful searching of hundreds of web pages of messages which is difficult and time-consuming task. The second source, a subset of NUS SMS Corpus (NSC) 3,375 SMS randomly selected ham messages, which is a dataset of approximately 10,000 genuine messages collected by the Department of Computer Science at the National University of Singapore for study. The messages are mainly from Singaporeans mostly university students who volunteered and aware their messages are going to be accessible to the public. The third source, a collection of 450 SMS ham messages obtained from the PhD Thesis by Caroline Tag. Lastly, the SMS Spam Corpus v.0.1 Big included which has 1,002 ham messages and 322 SMS spam messages can be obtained from UCI machine learning repository at [52].

4.2. Preprocessing the Datasets

Data preprocessing is a method that prepares the raw data and makes it relevant for a model of machine learning. It is the first and critical step in the development of a model for machine learning. It is not always the case that we come across clean and formatted data while developing

a machine learning project. And it is mandatory to clean it to get rid of any noise and place it in a formatted way while doing any operation with data. For both the Email and SMS datasets, the following are the preprocessing techniques that was performed:

i. Tokenize the texts to words
ii. Change all the text in the word to lower case
iii. Get rid of stop words like the, a, with, about etc.
iv. Get rid of punctuation
v. Get rid of white space and hyperlinks
vi. Get rid of numbers
vii. Perform word stemming and lemmatization on the texts using NLTK library.

4.3. Explore the Email and SMS Datasets

The next step in the methodology is to analyze the Email and SMS datasets to better understand it. To explore the datasets, this study used WordCloud and bar chart as used in [45].. WordCloud shows the most common words in the dataset. In this study, both the Email and SMS datasets were split into two data frame for ham and spam as shown below;

Figure 4.1. Ham WordCloud for Email

Figure 4.2. Ham WordClould for SMS

From Figure 4.1 and 4.2, the most common words in Ham Email WordCloud are group, user, irish, linux etc. while for SMS are see, today sorry, will etc. The WordCloud for spam is shown below:

Figure 4.3. Spam WordCloud for Email

Figure 4.4. Spam WordCloud for SMS

From Figure 4.3 and 4.4, the most common words for Email spam wordcloud are free, offer, order, million etc. while for SMS are claim, call, free, prize etc.

Another tool the study used is bar chart. Before using the bar chart, the Email and SMS dataset were split into training and test set to properly assess them before implementing ML and DL algorithm. The datasets were split to 80% for training and 20% for testing.

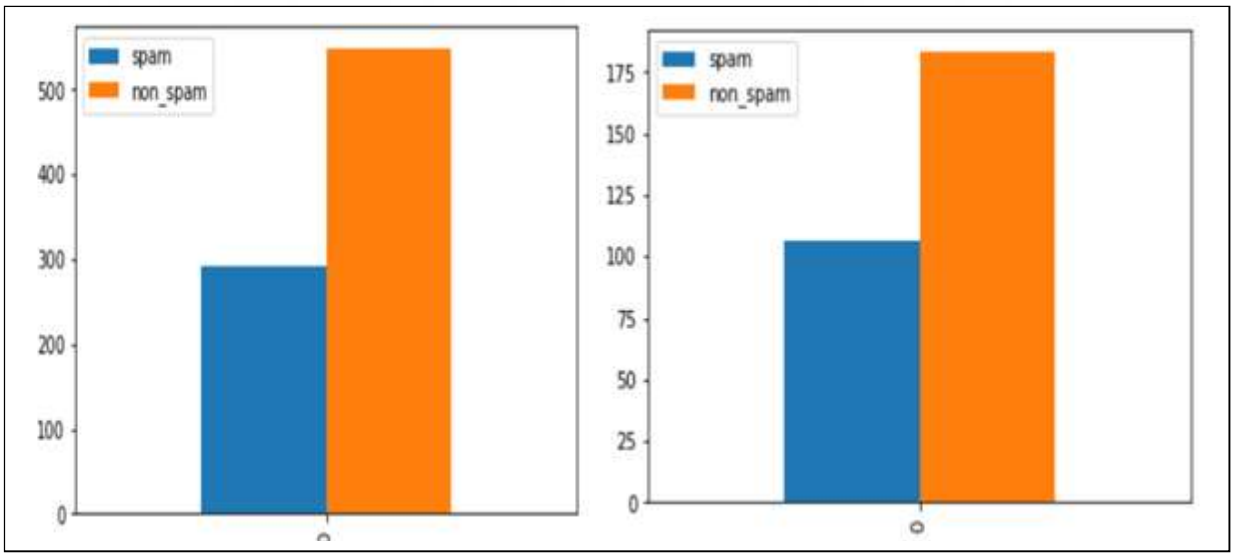

Figure 4.5. Distribution of Email Dataset for Ham and Spam in Training and Testing

The distribution for Email dataset was balance for both ham and spam in training and testing as can be seen in Figure 4.5 but for the SMS dataset, there was an imbalance for ham and spam as shown in Figure 4.6 below;

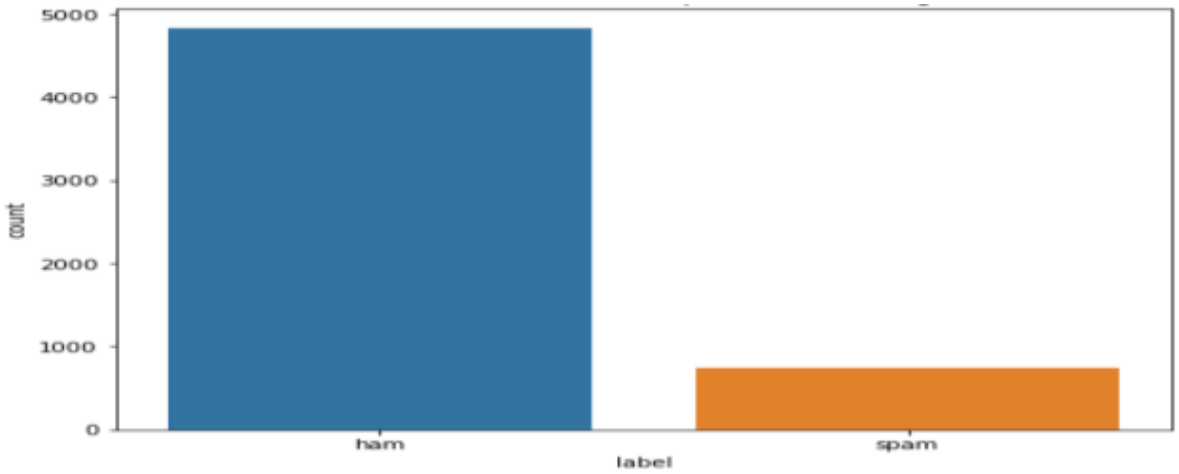

Figure 4.6. Distribution of SMS Dataset for an Imbalance Ham and Spam

The distribution above in Figure 4.6, shows an imbalance category for ham and spam in SMS dataset. Therefore, to balance the dataset, some features of the ham messages were deleted to balance up with the spam message using down sampling technique [45]. The new distribution will now have 747 classes for ham and spam respectively as shown in Figure 4.7 below;

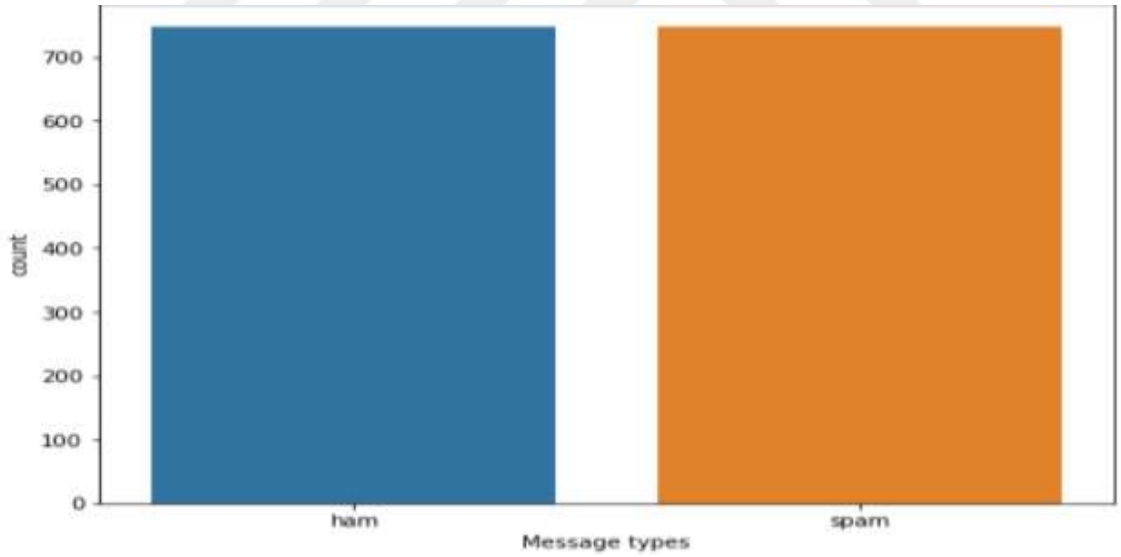

Figure 4.7. Distribution for SMS Dataset after Down sampling the Ham message to Balance with Spam Message

4.4. Vectorization

Before implementing the machine learning algorithm, the text in the datasets needs to be tokenize to words then composed to integers or floating point values. The converted values of words in the dataset is used as the input for implementing machine learning algorithm which is also known

as feature engineering [46]. This study used CountVectorizer technique to change the text in Email and SMS dataset to a “matrix of token counts” using the scikit-learn library [47].

4.5. Implementation of Machine Learning Algorithms.

Now that the texts have been tokenize to construct a corpus of known words in the two datasets using CountVectorizer, the following ML algorithm were implemented on the Email and SMS datasets using scikit- learn library [47];

4.5.1. Support Vector Machine Classifier

Support Vector machines can be described as systems that use a linear function space hypothesis space in a high-dimensional feature space trained with a learning algorithm that implements a learning bias derived from statistical learning theory from optimization theory. Support Vector Machine (SVM) is a prediction method for classification and regression that uses machine learning theory to optimize predictive accuracy while preventing over-fitting data automatically [7].

4.5.2. Random Forest Classifier

An ensemble technique used for classification is random forests. In a classification problem, random forests are used to rank the significance of variables [48].

4.5.3. Decision tree

Decision trees for classification and prediction are efficient and common instruments. Decision trees define cases or examples by beginning with the root of the tree and before a leaf node passes across it [48].

4.5.4. Multinomial Naïve Bayes classifier

The multinomial Naive Bayes classifier is ideal for discrete feature classification (e.g., word counts for text classification). Normally, integer feature counts are needed for the multinomial distribution [48].

4.5.5. Logistic Regression Classifier

Another methodology borrowed from the field of statistics by machine learning is logistic regression. It is the go-to approach for problems with binary classification (problems with two class values. Logistic regression uses an equation as a representation, input values (x) are combined

linearly to estimate an output value using weights or coefficient values (referred to as the Greek capital letter Beta) [48].

4.6. Implementing Dense Model Deep Learning

Dense neural network indicates that the neurons in a network layer are totally linked (dense) by layers. In a layer, each neuron receives an input from all the neurons present in the previous layer, so they are densely connected. In other words, a completely linked layer is the dense layer, meaning that all the neurons in a layer are linked to those in the next layer [49].

Dense model, a deep learning algorithm was applied on both Email and SMS datasets. Python Tensor flow Keras is used in this study and sequential Keras model object called which enables layers to be added to the dense model in a sequential order. The first layer in the model is the embedding layer; dimensional vector of real numbers. The embedding dimension is the size of this vector. We choose 16 in this study. This embedding layer also assigns close vectors to any two words with similar meaning_and embedding in this case also serve as input layer or first hidden layer to our model.

The input dimension is passed as our maximum padding length (500). The second layer in our model is important, this is because it helps reduce the number of parameters in the model, which in return reduces over fitting. Averaged pooling is employed here, and it is 1 dimensional. Next deuse layer added having a neutron with rely activation, followed by dropout to avoid overfitting. Another deuse layer with 19 neurons is added with a similar nature. The final layer is deuse with a single output neuron. Sigmoid activation function is used to activate each layer in the model. It returns values between 0 and 1. 0 means not active, and 1 means active. Finally, fitting the model to the data using 99 Epochs. a fraction of the training data is also used for validation during the training [45].

4.7. Evaluating the Machine and Deep Learning models

The following metrics were used in this study to evaluate both machine learning and deep learning model on Email and SMS datasets [50].

4.7.1. Accuracy

The accuracy metric calculates the ratio of accurate predictions with respect to the total number of evaluated instances. The formula is

TP + TN) / (TP + TN + FP + FN)

4.7.2. Precision

Precision is used to calculate the positive patterns that are in a positive class which are correctly predicted from the total predicted patterns.

TP / (TP + FP)

4.7.3. Recall

Used to calculate positive pattern fraction that are correctly classified.

TP/ (TP + FN)

4.7.4. F1-score

Harmonic mean between recall and precision values.

2*(Precision*Recall) / (Precision + Recall)

4.7.5. AUC/ROC

AUC is one of the common metrics of the ranking type used to create an optimized learning model and to compare learning algorithms as well. The AUC value, unlike the threshold and probability metrics, represents a classifier's overall ranking results.

$S_P – N_P (N_N+1)2 / N_P N_N$

5. RESULTS AND DISCUSSION

This chapter presents the all the result of implementing machine learning and deep learning algorithm on Email and SMS datasets. Table 5.1 below shows the result of both ML and DL classifiers on the datasets;

Table 5.1. Performance comparison of Dense Neural Network with Machine Learning classifiers on Email dataset

Classifiers	Accuracy	Precision	Recall	F1-Score	ROC-AUC
SVM	81.182%	81%	99%	89%	56.318%
Random Forest	87.014%	86%	99%	92%	69.98%
Multinomial NB	73.173%	82%	85.0 %	83.0%	57.499%
Logistic Regression	77.061%	80%	94%	87%	54.87%
Decision Tree	80.015%	88%	86%	87%	71.812%
DNN	97.512%	98%	99%	98%	95.271%

From the Table 5.1 it can be seen that the dense deep learning classifier out performs the machine learning classifiers. The model produces a 97.5% score for accuracy, 98% for precision, 99% for recall, 98% for F1 Score and 95.217% ROC.

Table 5.2. Performance comparison of Dense Neural Network with Machine Learning classifiers on SMS dataset

Classifiers	Accuracy	Precision	Recall	F1-Score	ROC-AUC
SVM	79.264%	82%	78%	80%	71.311%
Random Forest	81.271%	83%	81%	82%	81.286%
Multinomial NB	74.916%	76%	77%	76%	74.778%
Logistic Regression	77.258%	79%	78%	78%	77.22%
Decision Tree	80.935%	83%	80%	82%	80.97%
DNN	95.318%	98%	93%	95%	95.455%

Table 5.2 also shows the performance of both machine learning and deep learning classifier.

As shown in Figure 5.1 and Figure 5.2, can be seen that the dense neural network performs better than all other machine learning classifiers.

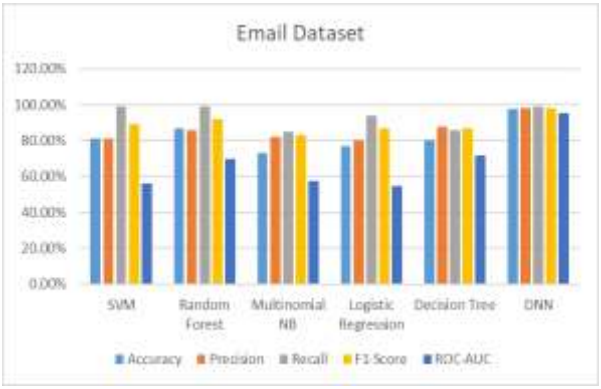

Figure 5.1. Performance comparison of DNN and Machine Learning Classifiers on Email Datasets

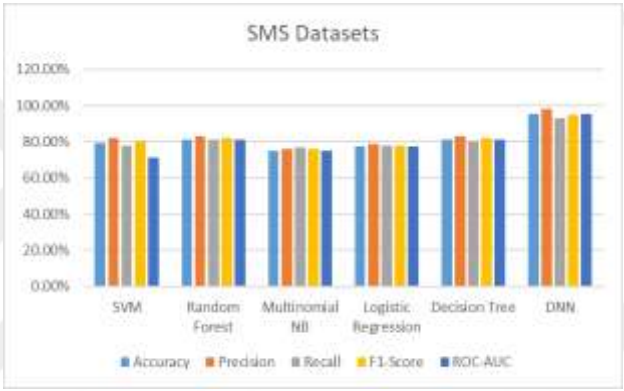

Figure 5.2. Performance comparison of DNN and Machine Learning Classifiers on Email Datasets

6. CONCLUSIONS

Social Engineering attacks such as spam is rapidly expanding in complexity, intensity and effects to digital communication medium. Telecommunication service providers, manufacturers of mobile devices and computer systems need to change their security procedures, improve end-user security and implement emerging technologies that is powered by artificial intelligence in their defense mechanism. Artificial Intelligence is driving digital transformation across business and industrial system. Machine Learning and Deep Learning which are subset of AI, are leveraged to transform and improve efficiency in system without being programmed but rather learn from large data it has been exposed to using complex algorithm that is structured and functions like human brain to build a model of an intelligent application system. This study implemented machine learning and deep learning classifiers on Email and SMS datasets. The result of the study showed that Dense Neural Network performs better in detecting spam in both datasets than machine learning classifiers.

Future works should include other communication medium such as WhatsApp, twitter, and Facebook. Other deep learning techniques such as transformer can be implemented. Finally, the machine learning classifiers can be optimized for a better result.

REFERENCES

[1] G. Le Bodic, *Mobile Messaging Technologies and Services*. 2002.

[2] K. Bahia and S. Suardi, "Connected society: the state of mobile internet connectivity 2019," *Gsma*, pp. 1–56, 2019.

[3] Salesforce, "2014 Mobile Behavior Report," *Salesforce Mark. Cloud*, pp. 11–12, 2014.

[4] T. R. Lynam, "Spam Filter Improvement Through Measurement," 2009.

[5] B. Barlowe, J. Blackbird, and W. S. Davis, "The evolution of malware and the threat landscape - a 10-year review," p. 48, 2012.

[6] W. Paper and C. C. Stacy, "Getting Started Computing at the Al Lab by Table of Contents," no. September, 1982.

[7] "Machine Learning and Security: Protecting Systems with Data and Algorithms - Clarence Chio, David Freeman - Google Books." [Online]. Available: https://books.google.iq/books?id=lyJJDwAAQBAJ&pg=PT163&lpg=PT163&dq=identify+the+same+family+of+malware&source=bl&ots=5qr1Csjak7&sig=ACfU3U36EVAg47__7uArpyJIMJ7x02IsXw&hl=ar&sa=X&ved=2ahUKEwi-lYqE79jnAhX5zMQBHWzmCNgQ6AEwDHoECAoQAQ#v=onepage&q&f=false. [Accessed: 29-Sep-2020].

[8] M. Singh, R. Pamula, and S. K. Shekhar, "Email spam classification by support vector machine," *2018 Int. Conf. Comput. Power Commun. Technol. GUCON 2018*, pp. 878–882, 2019.

[9] V. Vishagini and A. K. Rajan, "An Improved Spam Detection Method with Weighted Support Vector Machine," *2018 Int. Conf. Data Sci. Eng. ICDSE 2018*, pp. 1–5, 2018.

[10] M. Diale, T. Celik, and C. Van Der Walt, "Unsupervised feature learning for spam email filtering," *Comput. Electr. Eng.*, vol. 74, pp. 89–104, 2019.

[11] S. K. Tuteja and N. Bogiri, "Email Spam filtering using BPNN classification algorithm," *Int. Conf. Autom. Control Dyn. Optim. Tech. ICACDOT 2016*, pp. 915–919, 2017.

[12] E. Of, "I NDEPENDENT H IGH -L EVEL E XPERT G ROUP ON A RTIFICIAL I NTELLIGENCE SET UP BY THE E UROPEAN C OMMISSION A D EFINITION OF AI :," 2019.

[13] D. Khurana, A. Koli, K. Khatter, S. Singh, and M. Rachna, "Natural Language Processing : State of The Art , Current Trends and Challenges Department of Computer Science and Engineering Accendere Knowledge Management Services Pvt . Ltd ., India Abstract," no. Figure 1.

[14] A. A. Alurkar *et al.*, "A proposed data science approach for email spam classification using machine learning techniques," *Jt. 13th CTTE 10th C. Conf. Internet Things - Bus. Model. Users, Networks*, vol. 2018-Janua, pp. 1–5, 2017.

[15] --Yangyan Li, "A Brief Introduction to Deep Learning."

[16] G. Apruzzese, L. Ferretti, M. Marchetti, M. Colajanni, and A. Guido, "On the Effectiveness of Machine and Deep Learning for Cyber Security."

[17] H. Bhuiyan, A. Ashiquzzaman, T. I. Juthi, S. Biswas, and J. Ara, "A Survey of Existing E-Mail Spam Filtering Methods Considering Machine Learning Techniques," *Glob. J. Comput. Sci. Technol. Softw. Data Eng.*, vol. 1, no. 2, 2018.

[18] S. K. Trivedi and S. Dey, "A combining classifiers approach for detecting email spams," *Proc. - IEEE 30th Int. Conf. Adv. Inf. Netw. Appl. Work. WAINA 2016*, pp. 355–360, 2016.

[19] S. K. Trivedi and S. Dey, "Effect of feature selection methods on machine learning classifiers for detecting email spams," *Proc. 2013 Res. Adapt. Converg. Syst. RACS 2013*, pp. 35–40, 2013.

[20] S. H. A. Elhamayed, "Comparative Study on Different Classification Techniques for Spam Dataset," *Int. J. Comput. Commun. Eng.*, vol. 7, no. 4, pp. 189–194, 2018.

[21] M. Bassiouni, M. Ali, and E. A. El-Dahshan, "Ham and Spam E-Mails Classification Using Machine Learning Techniques," *J. Appl. Secur. Res.*, vol. 13, no. 3, pp. 315–331, 2018.

[22] Nandhini.S and .Jeen Marseline.K.S, "Performance Evaluation of Machine Learning Algorithms for Bitcoin Price Prediction," *Proc. 4th Int. Conf. Inven. Syst. Control. ICISC 2020*, pp. 110–114, 2020.

[23] V. K. Singh and S. Bhardwaj, "Spam mail detection using classification techniques and global

training set," *Adv. Intell. Syst. Comput.*, vol. 673, pp. 623–632, 2018.

[24] A. Sharaff, N. K. Nagwani, and A. Dhadse, "Emerging Research in Computing, Information, Communication and Applications," *Emerg. Res. Comput. Information, Commun. Appl.*, pp. 237–244, 2016.

[25] E. G. Dada, J. S. Bassi, H. Chiroma, S. M. Abdulhamid, A. O. Adetunmbi, and O. E. Ajibuwa, "Machine learning for email spam filtering: review, approaches and open research problems," *Heliyon*, vol. 5, no. 6, 2019.

[26] N. F. Rusland, N. Wahid, S. Kasim, and H. Hafit, "Analysis of Naïve Bayes Algorithm for Email Spam Filtering across Multiple Datasets," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 226, no. 1, 2017.

[27] S. Muhammad Abdulhamid, M. Shuaib, O. Osho, I. Ismaila, and J. K. Alhassan, "Comparative Analysis of Classification Algorithms for Email Spam Detection," *Int. J. Comput. Netw. Inf. Secur.*, vol. 10, no. 1, pp. 60–67, 2018.

[28] L. E. Zhang, J. Zhu, and T. Yao, "An evaluation of statistical spam filtering techniques," *ACM Trans. Asian Lang. Inf. Process.*, vol. 3, no. 4, pp. 243–269, 2004.

[29] Z. Hassani, V. Hajihashemi, K. Borna, and I. Sahraei Dehmajnoonie, "A classification method for e-mail spam using a hybrid approach for feature selection optimization," *J. Sci. Islam. Repub. Iran*, vol. 31, no. 2, pp. 165–173, 2020.

[30] M. Alauthman, "Botnet spam e-mail detection using deep recurrent neural network," *Int. J. Emerg. Trends Eng. Res.*, vol. 8, no. 5, pp. 1979–1986, 2020.

[31] T. Xia and X. Chen, "A discrete hidden Markov model for SMS spam detection," *Appl. Sci.*, vol. 10, no. 14, 2020.

[32] T. A. Almeida, J. M. Gomez Hidalgo, and T. P. Silva, "Towards SMS Spam Filtering : Results under a New Dataset," *Int. J. Inf. Secur. Sci. T.*, vol. 2, no. 1, pp. 1–18, 2012.

[33] W. H. Gomaa, "The impact of deep learning techniques on SMS spam filtering," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 1, pp. 544–549, 2020.

[34] A. Chandra and S. K. Khatri, "Spam SMS Filtering using Recurrent Neural Network and Long Short Term Memory," *2019 4th Int. Conf. Inf. Syst. Comput. Networks, ISCON 2019*, pp. 118–122, 2019.

[35] G. S. Sravya and G. Pradeepini, "Mobile Sms spam filter techniques using machine learning techniques," *Int. J. Sci. Technol. Res.*, vol. 9, no. 3, pp. 384–389, 2020.

[36] O. Abayomi-Alli, S. Misra, A. Abayomi-Alli, and M. Odusami, "A review of soft techniques for SMS spam classification: Methods, approaches and applications," *Eng. Appl. Artif. Intell.*, vol. 86, no. September 2018, pp. 197–212, 2019.

[37] G. Jain, M. Sharma, and B. Agarwal, "Spam Detection on Social Media Using Semantic Convolutional Neural Network," *Int. J. Knowl. Discov. Bioinforma.*, vol. 8, no. 1, pp. 12–26, 2018.

[38] P. Kumar and J. Prakash, "This is a repository copy of Deep learning to filter SMS spam . White Rose Research Online URL for this paper : Version : Accepted Version Article : Deep Learning to Filter SMS Spam," 2020.

[39] A. Alzahrani and D. B. Rawat, "Comparative Study of Machine Learning Algorithms for SMS Spam Detection," *Conf. Proc. - IEEE SOUTHEASTCON*, vol. 2019-April, pp. 5–10, 2019.

[40] S. M. Abdulhamid *et al.*, "A Review on Mobile SMS Spam Filtering Techniques," *IEEE Access*, vol. 5, pp. 15650–15666, 2017.

[41] S. Bosaeed, I. Katib, and R. Mehmood, "A Fog-Augmented Machine Learning based SMS Spam Detection and Classification System," *2020 5th Int. Conf. Fog Mob. Edge Comput. FMEC 2020*, pp. 325–330, 2020.

[42] N. K. Nagwani, "A Bi-level text classification approach for SMS spam filtering and identifying priority messages," *Int. Arab J. Inf. Technol.*, vol. 14, no. 4, pp. 473–480, 2017.

[43] P. Poomka, W. Pongsena, N. Kerdprasop, and K. Kerdprasop, "SMS Spam Detection Based on Long Short-Term Memory and Gated Recurrent Unit," *Int. J. Futur. Comput. Commun.*, vol. 8, no. 1, pp. 11–15, 2019.

[44] S. Gounder, "Chapter 3 - Research methodology and research questions," *Res. Methodol. Res. Method*, no. March 2012, pp. 84–193, 2004.

[45] A. S. Mirza Rahim Baig, Thomas V. Joseph, Nipun Sadvilkar, Mohan Kumar Silaparasetty, "Deep learning 简介 一 、 什么是 Deep Learning ？," *Nature*, vol. 29, no. 7553, pp. 1–73, 2019.

[46] J. Brownlee, “How to Encode Text Data for Machine Learning with scikit-learn,” *Deep Learn. Nat. Lang. Process.*, pp. 1–29, 2017.

[47] I. None and T. Menu, “sklearn.feature_extraction.text,” pp. 1–6, 2021.

[48] S. Bird, E. Klein, and E. Loper, *Covariance structure analysis of health-related indicators in the elderly at home with a focus on subjective health*, vol. 53, no. 9. 2015.

[49] M. Rampurawala, “Classification with TensorFlow and Dense Neural Networks,” *Hear. Beat*, pp. 1–8, 2019.

[50] H. M and S. M.N, “A Review on Evaluation Metrics for Data Classification Evaluations,” *Int. J. Data Min. Knowl. Manag. Process*, vol. 5, no. 2, pp. 01–11, 2015.

[51] https://spamassassin.apache.org/old/publiccorpus/?C=D;O=D Accessed June-2020

[52] https://archive.ics.uci.edu/ml/datasets/SMS+Spam+Collection Accessed June-2020

CURRICULUM VITAE

[redacted]

PERSONAL INFORMATIONS

RESEARCHER INFORMATION

[redacted]

EDUCATION

Master Degree **:** "Email and SMS Spam Detection Based on Deep Learning."
Firat University, Sciences Graduate School, Department of Computer Engineering, 2021
Supervisor: Prof. Dr. Mehmet Kaya

Bachelor **:** Bayero University Kano, Faculty of Information and Computer Science, Department of Computer Science, 2010

High School **:** Bayero University Staff Secondary School, Kano, Nigeria, 2000

RESEARCH EXPERIENCES

- ✓ Computer Programming languages available (C-C++, MATLAP, LABVIEW, vb.)

WORK EXPERIENCE

2012 – 2024: Class Teacher, Kano State Polytechnic Staff School Kano.
2010 – 2015: Lecturer, Kano State Institute of Information Technology, Kano, Nigeria

ACADEMIC ACTIVITIES