FRACTIONAL-EWA BASED DEEP CNN FOR PHISHING ATTACK DETECTION

Arshey M

Research Scholar, Noorul Islam Centre for Higher Education, Tamil Nadu
arshi.sm@gmail.com

Dr. K.S Angel Viji

Associate Professor, College of Engineering, Kidangoor, Kerala
ksangelviji@gmail.com

Abstract

Phishing attack is acclaimed as one of the recognized cybercrime attacks over the internet and mail users. Phishing is a form of unauthorized access of confidential information, like passwords, user names and credit card details. Detection of phishing attacks and classifying the mails still remains a challenging issue. This research presents an effective strategy by developing a newly proposed method called Fractional-EarthWorm Algorithm (EWA) based Deep Convolutional Neural Network. The Fractional-EWA is derived by inclusion of fractional calculus concept to EarthWorm Optimization. The features are extracted using the term frequency and the feature is selected using the Levenshtein distance. The DCNN is trained by exploiting the proposed Fractional-EWA. However, this algorithm achieved maximum accuracy, maximum sensitivity, and maximum specificity of 0.781, 0.782, and 0.722 respectively for chunk percentage of data and achieved the maximum accuracy, maximum sensitivity, and maximum specificity of 0.744, 0.725, and 0.723, respectively for number of features.

Keywords: **Phishing attack, Spam mail classification, Fractional Calculus, Earthworm Optimization Algorithm (EWA), Levenshtein distance, Deep Convolutional Neural Network (DCNN)**

1. Introduction

Recently, phishing has been recognized as one of the significant problems in mail services over the internet and it has a crucial influence on both the business and financial fields. The hackers send false messages requesting the web users to meet one of the websites so that the user is requested to submit their confidential data. However, these websites are only established to steal user password and data to enter the e-mail without having knowledge that the information was accessed [25] [4]. Phishing is considered as one of the social attacks to attract users to carry put adverse actions behind the attacker [1]. The main goal of these impacts is to hoax users to steal their private data. Because of this, phishers tried to copy web pages or e-mails that represent an extreme visual coincidence to authorized web pages. The web pages are then combined with one login to collect the user's personal or confidential data [1] [2]. A phishing attack is commonly processed by considering the merits of the visual coincidence between the real and fake web pages. The hacker generates a web page which is identical to that of the authentic web page. After that, the link of the phishing web pages is transmitted to millions of internet users by emails and in various forms of communication. Typically, the false email data resembles a little sense of fear or emergency or provide money and requests the user to take immediate step [7].

Electronic mail is considered as one of the significant ways of communication. Nowadays, most of the people across worldwide use e-mail for various purposes because it is the fast, inexpensive and very simple way of communication. This is the reason that the intruders generally attack such type of tools for different illegal purposes [16]. Typically, email is very susceptible to spam emails due to its wide usage, simplicity and cheap in cost and also in a single click, anyone can communicate anywhere around the world. Emails are broadly categorized into two kinds, such as spam email and non-spam email. Spam mail is always termed as junk mail or unnecessary mail, while non-spam mails are look like a real mail but it is specifically designed for an individual purpose. Spam emails are the unnecessary emails that are daily transmitted to thousands of inboxes of different users [1]. Mostly, spam emails consist of number of copies of identical messages, commercial advertisements and unrelated posts [3] [26]. The main intention of spam emails of advertisements targets at posting different products and services, including electronics, software, loans, jewelry, gambling, and phishing attempts. The major limitations of spam mails are misusing of mailbox space, and network bandwidth. In addition to this, it also consumes lot of user's valuable time as well as provoking them to detect and destroy the unnecessary messages. Thus, spam detection still remains as one of the challenging tasks for many organizations [28]. However, there have been different types of methods to alleviate the quantity of spam.

An anti-spam law has been utilized by authorizing penalty for issuing spam emails. Machine learning is another technique commonly employed for email spam detection that has the potential to identify and categorize the email data into either ham email or spam email [30] [26]. An effective e-mail spam filtering methods must reply to the recent unwanted emails, bulk emails, and spam threats. There are numerous spam email filtering techniques, like SPAMfighter, iHateSpam, SpamEater, MailWasher, and Spam buster. However, these modern techniques failed to detect the new spams effectively and failed to provide high accuracy. The effectiveness of email spam filter is enhanced by considering the efficient feature reorganization and conceptual and semantic similarity. However, spam email filtering technique is a significant application of data retrieval area [1] [25] Nevertheless such challenges, Machine learning techniques have been recognized as one of the hotspots in research areas. There is an indisputable impact of the features utilized for categorization of emails [16] [17]. The main problem associated with the feature selection is very important to construct the detection of phishing techniques that are popular in common. It is practicable in which the group of best features discriminating the phishing websites that copy a financial organization must be dissimilar from websites copying an e-commerce domain. The latest procedure in feature selection methods are commonly depends on some heuristic threshold [1]. Different learning algorithms have been utilized for email spam detection, such as Random Forest, Support Vector Machine (SVM), Naïve Bayes, Artificial Neural Networks (ANN), and K-Nearest Neighbor (K-NN) [6] [15].

The primary motive of this research is to develop a newly proposed method called Fractional-EWA based DCNN for effective spam mail classification and phishing attack detection. The developed approach mainly consists of four stages, such as pre-processing, feature extraction, feature selection, and spam mail categorization. At first, the input data that is acquired from dataset is passed through the pre-processing phase in order to eliminate the unwanted noises and external calamities. The pre-processing step consists of stop word removal and stemming technique that is essential for removing redundant words. Once the pre-processing is done, the pre-processed output is carried out to process feature extraction, where the appropriate features like term frequency are extracted. After that, feature selection is performed to choose the data using Levenshtein distance. Furthermore, the selected features are passed through the Deep DCNN to categorize the spam and non-spam mails, where network classifier is trained using developed Fractional–EWA algorithm. However, the developed Fractional-EWA algorithm is derived by incorporating the Fractional calculus and Earthworm Optimization Algorithm (EWA). Moreover, the phishing attacks of spam emails are detected using the Fractional-EWA depending on error condition and weight bounding.

2. Related Works

In this section, the literature survey of recently published papers related to spam filtering techniques are discussed and also the challenges existed in the existing methods are also presented. Mahdieh Zabihimayvan and Derek Doran developed a Fuzzy Rough Set (FRS) approach to choose the appropriate features from three benchmarked datasets. Here, three classification methods, such as a multilayer perceptron, a random forest, and SMO were utilized effectively for phishing detection. Moreover, the developed strategy reduced the detection time and it was more resistant to zero-day attacks. Besides, the method effectively handled the limitations of dimensionality for classification techniques. Hassani Z [26] *et al.* modeled a hybrid scheme for effective spam detection method. The hybridized approach was derived by integrating the Binary whale Optimization Algorithm (BWOA), and Binary Grey Wolf Optimization Algorithm (BGWO) in order to grasp the best features. Here, the subset of the features was evaluated using the two classifiers, like K-Nearest Neighbor (KNN) and Fuzzy K-Nearest Neighbor (FKNN). However, main advantage of this approach was its highest accuracy, but failed to detect the spam messages effectively. Kumaresan T [28] *et al.* introduced a spam categorization strategy called S-Cuckoo and hybrid kernel-based support vector machine (HKSVM). This method involved mainly three phases, namely textual and visual feature extraction, best feature selection, and spam classification. Initially, features were excerpted from e-mails depending on text and also image. In order to extract the textual features, TF-term frequency was utilized, whereas image-based features were extracted using the wavelet and correlogram token. Although the approach provided better accuracy in terms of detecting the spam messages, it failed to utilize optimal kernel parameter selection as it enhances the accuracy of the classifier.

Venkatraman S [31] *et al.* developed a semantic similarity approach to encounter the problems raised by means of polysemy in spam detection. The effectiveness of e-mail spam filtering was enhanced by exploiting feature detection and semantic similarity according to mail user preferences. The advantages included in this system were reduced overhead costs, better performance against zero-day attacks, and no backscatter. Moreover, it utilized conceptual concept and semantic similarity-based spam filters concept for analyzing the data to detect, recognize, and protect unwanted e-mails created using IoT systems. The major barrier exists in this approach is that it reduced the accuracy because of the complexity increases in the structure of semantic similarity in some extreme cases. Anuj Kumar Singh [8] *et al.* devised a method to recognize the best classifier for spam mail classification using Fuzzy C-Means algorithm. This technique used the membership threshold value of 0.5 for detecting the spam e-mails, but it failed to extend the system using machine learning approaches. Ali Mohammad H. Al-Ibrahim [4] introduced an electronic phishing algorithm named Sequential Minimal Optimization (SMO)

to protect the internet users from attackers while stealing the confidential information. The SMO algorithm was selected and applied through Weka program and provided better classification rate as well as accuracy. However, some misclassifications were found during phishing attack.

Ankit Kumar Jain, et al [7] developed a machine learning scheme that effectively detected the phishing charge by evaluating the links presented in the HTML source code of website. The developed strategy categorized the hyperlink features into 12 various forms in order to train the algorithms of machine learning. Moreover, the developed approach did not rely on any third-party users and thus there is a no need to wait or delay for the results. However, there was a possibility of predicting the false results since the developed approach relies on the source code of website. This method failed to recognize non-HTML websites with better performance and also it had the inability to recognize the phishing attacks in mobile surrounding. Adebowale M.A [24] et al. modeled an Adaptive Neuro-Fuzzy Inference System (ANFIS) for phishing detection and protection using the text, frames, and images features. The major contribution of this method was enhanced the detection accuracy as well as reducing the scrutiny time. However, the method failed to provide a web browser plug-in in order to prevent the users in real time applications.

2.1 *Challenges*

Some of the limitations confronted by the traditional techniques of phishing attack detection and spam mail classification are deliberated as follows:

- An Artificial Neural Network (ANN) for better classification of detecting phishing attacks developed in [12], had sufficient potential to identify phishing as well as pharming attack quite accurately. Moreover, it failed to consider the URL from g-mail directly.
- In [10], whale optimization algorithm (WOA) effectively selected significant features in the e-mail library and also utilized random forest algorithm for categorizing the e-mails as spam and non-spam ones. The major drawback of this algorithm is that it was limited to certain number of datasets for feature selection process.
- An effective clustering technique integrating the features of K-means algorithm and BIRCH algorithm modeled in [10], was a best clustering algorithm that required a single scan of whole information and thus preserving the time, but this method was not suitable for testing with different algorithms by varying the size of large datasets.
- An Artificial Immune System (AIS) introduced in [15], effectively reduced the false positive rate and generated efficient spam detectors. The major difficulty found in this system was reduction of recall rate.
- Rule-based method failed to refine the rules to enhance false positives and false negatives on newer datasets. Moreover, the deployment and testing of the system in real world applications was not accomplished [17].

3. Proposed Adaptive spam filtering using Fractional EWA-based DCNN

Spam mail classification and phishing attack detection is considered as the challenging task over the internet users and websites and still it remains a troublesome process to recognize the spam and non-spam emails. In this research, an effective strategy is designed and developed for efficient spam mail classification and phishing attack detection utilizing the proposed Fractional-EWA based DCNN. However, the developed approach is derived by integrating Fractional Calculus [13], and Earthworm Optimization Algorithm (EWO) [22]. At first, the input data acquired from dataset is fed as an input to pre-processing phase in order to eliminate the unwanted noises and external calamities. Here, the pre-processing step consists of stop word elimination and stemming technique to evacuate the redundant words in the data. Once the pre-processing is completed, the pre- processed output is subjected to feature extraction process, in which the desired and appropriate features are excerpted by applying term frequency. After that, the features are selected using the Levenshtein distance [20]. Furthermore, the selected features are passed through the DCNN [23] to classify the spam mails, where the network classifier is trained by employing the proposed Fractional-EWA based DCNN. Moreover, the phishing attack detection is detected using Fractional-EWA based DCNN depending on error condition and weight bounding. Fig. 1 represents the schematic view of the proposed approach for effective spam mail classification.

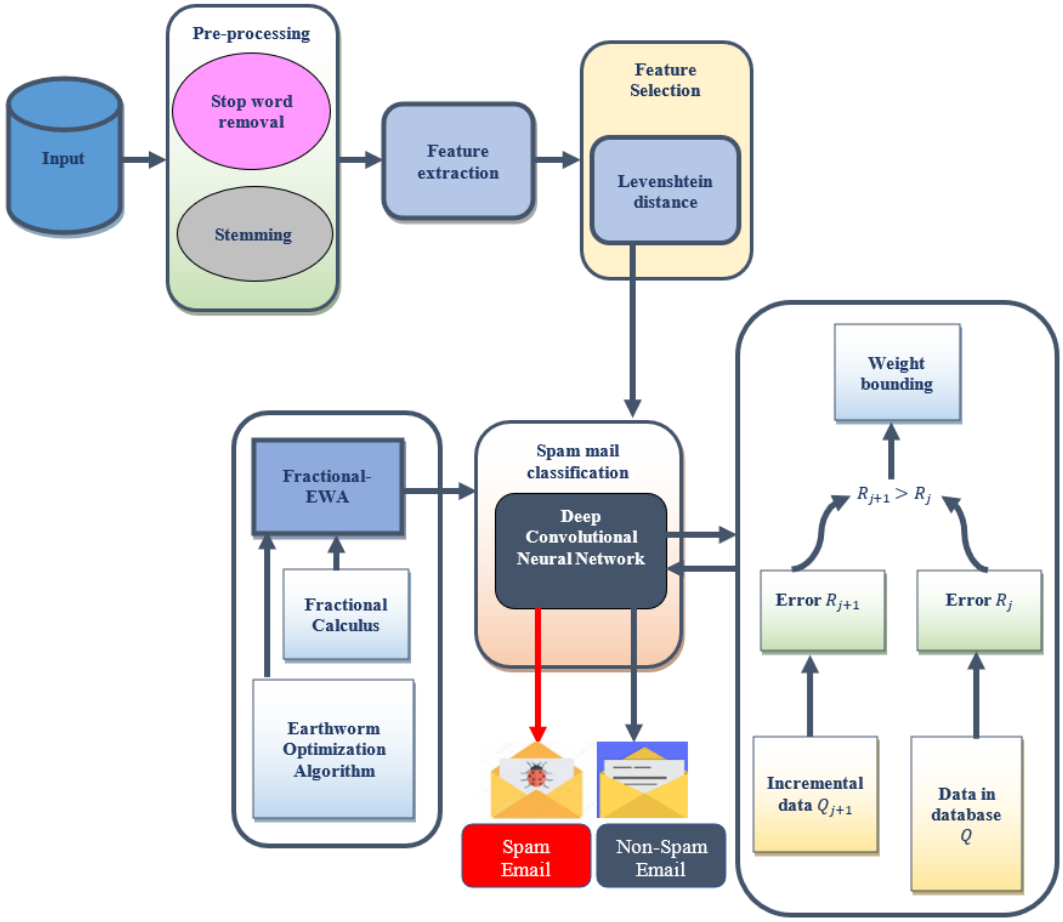

Fig. 1. Schematic view of the developed Fractional-EWA based DCNN for spam mail classification

3.1 *Acquisition of input data*

The first step in the process of spam mail classification is the acquisition of input data from the given dataset. Here, the emails from the dataset are considered as an input data for further processing. Let us consider the dataset as T and Y be the training set and it is expressed as,

$$T = \{Y_1, Y_2, \ldots, Y_i, \ldots, Y_n\} \tag{1}$$

where, *i= {1,2…n}* and *n* indicates the total number of samples with a dimension [*f* x *g*] and Y_i is considered as the input, which is passed through the pre-processing phase.

3.2 *Pre-processing using stop word removal and stemming*

The input data Y_i is passed through pre-processing phase in order to eliminate unwanted noises. Typically, pre-processing is considered as the most significant task that converts the raw information into a corrected data for further detection. In order to enable the pre-processing process, it is ensured that the dataset must be clean and noise-free. However, pre-processing is performed to make perfection in data from various errors [29]. Generally, a URL includes meaningful and meaningless words. Therefore, it is necessary to eliminate such meaningless words from the URL by means of pre-processing [14].

Pre-processing is used to differentiate authorized and spam contents. Initially, all words are transformed to letters in lower-case and the process of token distribution is carried out. In order to define the weights of the pre-processed words, the most significant [9] Bag of Words (BoW) method *tf.idf*, which is a particular weighting scheme is used. The weight of the scheme W_{uv} is calculated as below,

$$W_{uv} = (1 + log(tf_{uv})) \times log\left(\frac{M}{df_{uv}}\right) \tag{2}$$

where, M specifies the whole number of messages, tf_{uv} denotes frequency of u^{th} word in v^{th} message and df_{uv} represents the total number of messages at least one chance of occurring of the u^{th} term. Here, *tf.idf* assumes both document length and term rareness. In order to choose the most appropriate words, it is significant to rank them based on their *tf.idf* weights.

3.2.1 Stop word removal

Stop word removal is the significant step in pre-processing phase. Stop-words are the words which are generally occurred in texts without depending a specific concept, like articles, prepositions, and conjunctions. Hence, the stop-words are commonly considered as an irrelevant part in the process of text classification and it is significant to remove the stop-words before classification [21].

It is significant to eliminate the stop words that speed up the operation in text processing. This process mainly converts the initiative words of document to the mode of base root.

$$J = \{A_e, 1 \leq 0 \leq d\} \tag{3}$$

where*, d* represents the total words in a dictionary from database. Therefore, the dictionary words are obtained from pre-processing phase and further feature extraction is carried out using the acquired words from the dictionary.

3.2.2 Stemming

The ultimate goal of the stemming process is to achieve the correct stem word or root forms of computed words. If the derived words are identical to their root words, the phenomenon of words are normally determined after performing the stemming process [21]. Typically, word stemming extracts the root or stem of a misspelled or altered word, such that the efficiency of the spam filtering is enhanced. A simple rule-based stemming algorithm is especially developed for spam detection. The result obtained through the pre-processing phase is denoted as P_i, which is subjected to feature extraction module to extract the desired and appropriate features.

3.3 *Feature extraction using term frequency*

Feature extraction is significant process in email classification. The goal of feature extraction is to separate the meaningful features from emails, so that they are classified based on their source. Usually, the frequency number of the disturbed words is considered as text features. Initially, the T dataset is categorized into two groups, such as training set T^{TR} and testing set T^{TS}. The process of feature extraction is performed for both testing and training dataset. Here, the term frequency feature is extracted while performing the feature extraction process.

3.3.1 Term frequency feature

The term frequency feature (TF) [28] is defined as the number of times the term resembles in email about summation of occurrences of terms in it. The occurrence count of each feature with an e-mail is utilized as an element of 257dimensional feature vector and feature vector is regularized by total text size. The term frequency is determined for every mail to separate the textual feature. In feature extraction process, $F_{1,}\ F_{2\ldots\ldots.,}F_{257}$ denotes the frequency number of 257 spam words. These obtained outputs are employed as a text feature vector for further classification. The weights are similar to number of times that the features occur in a document for term frequency. Generally, term frequency weights contain huge information when compared with the binary weights. The term frequency is expressed as,

$$TF_{u,v} = \frac{N_{u,v}}{\sum_a N_{a,v}} \tag{4}$$

where, $N_{u,v}$ is the number of times the term $T_{u,v}$ happens in a document D**.**

3.4 *Levenshtein distance-based feature selection*

Feature selection is the most significant step and it is described as the mechanism of choosing a subset of original features in order to minimize the dimension of feature space based on some criterion. It is essential to evacuate the redundant and unrelated features, which increase the efficiency in learning tasks, thereby enhancing learning performance, such as predictive accuracy and improving the learned results. However, infinite number of features results some negative impacts on the classification accuracy that degrades the performance of the classifier. Typically, feature selection is categorized into two important types. One category chooses features that actually do not cause any impacts on classification performance. However, it alters the original features based on some concepts to establish a new group of features. Furthermore, it selects a subset of converted features and the chosen subset has less dimensionality than indigenous feature. On the other hand, the second category selects the features based on performance of the classification model [11].

3.4.1 Levenshtein distance

Levenshtein distance is utilized to choose the appropriate features. Levenshtein distance is a metric that is used for calculating the number of differences between two strings. Generally, the difference in percentage is formulated according to the levenshtein distance and it is utilized to determine the quantity of characters that must be inserted, substituted, or eliminated to achieve perfect matching. The selected feature FS is fed as an input to the DCNN for further process as shown in Fig. 2.

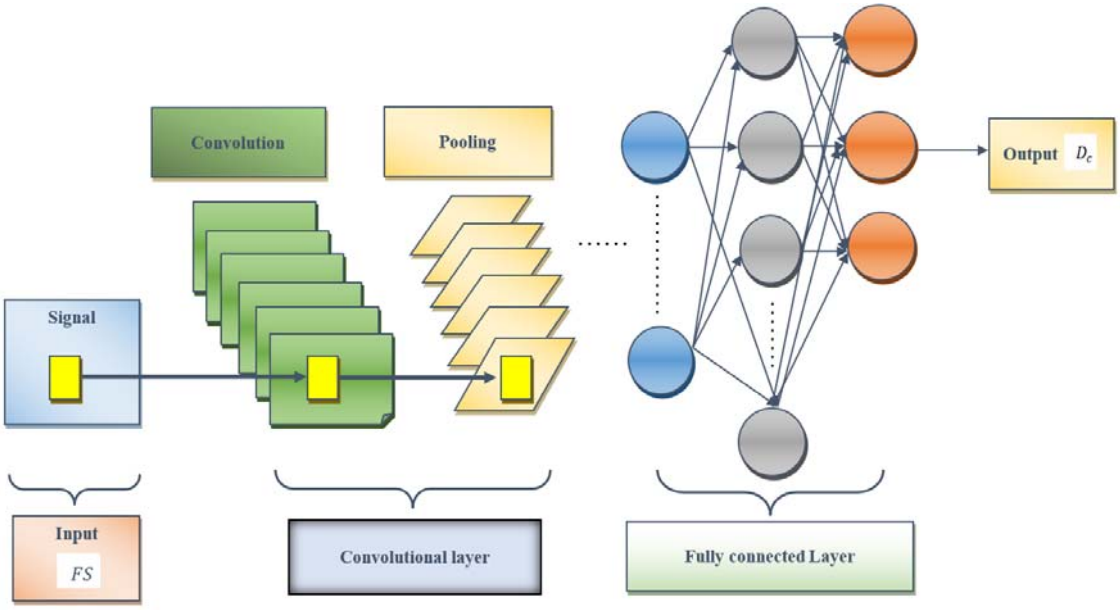

Fig. 2. Architecture of Deep CNN

3.5 *Spam mail classification using Deep Convolutional Neural Network*

The selected feature *FS* is subjected to the Deep Convolutional Neural Network (DCNN) to classify the spam mails. The major advantage of the DCNN is that it consumes less time for training process. Moreover, the network achieved high accuracy with less parameters and low complexity.

3.5.1 Structure of Deep CNN

Deep CNN plays a vital part in the analysis of compressed signals for high classification results. In Deep CNN architecture, a number of neurons are linked to the single neuron consisting in next layer. The structure of Deep CNN is comprised with mainly three layers, namely Convolutional (Conv) layers, Pooling (POOL) layers, and a Fully Connected (FC) layer. The structure of the Deep CNN is illustrated in Fig. 2. Various convolutional and pooling layers are arranged to generate a deep neural network infrastructure and such layers are responsible for extracting the informative and discriminative representations with respect to data. The layers in the deep CNN process some individual purposes, like establishment of feature maps in the conv layers, sub-sampling of the feature maps in pooling layers, and classification is performed in FC layer finally.

(1) Convolutional layers: The purpose of the convolutional layers is to attain the patterns impelled in the input compressed signal utilizing the convolutional filters that are linked with the receptive fields that provides an interlink between neurons of preceding layer with the consecutive conv layers along with trainable weights. Let us consider the input *FS*, which is obtained from the feature selection process is fed as an input to deep CNN and result of conv layer is expressed as,

$$(\omega_s{}^h)q,w = \left(O_s{}^h\right)q,w + \sum_{\alpha=1}^{x_1{}^{\alpha-1}} \sum_{\beta=x_1}^{x_1{}^{nn}} \sum_{\gamma=-x_2{}^{nn}}^{x_2{}^{nn}} \left(\delta_{s,\alpha}\right)_{\gamma,\beta} * \left(FS_\alpha{}^{nn-1}\right)_{q+\beta,w+\gamma} \quad (5)$$

where, $*$ denotes the convolutional operator, $\left(\omega_s{}^h\right)_{q,w}$ signifies the invariant feature map or result of the nn^{th} conv layers, which is positioned as *(q, w).* The result of the preceding *(nn-1)* th layer generates input to z^{th} conv layer. Let us assume the weights of conv layers be $\delta_{s,\alpha}$, which is the weights of nnth conv layer and $O^h{}_s$ be the bias of nn^{th} conv layer. Consider that there are α conv layers, $(1<=z<=\infty)$ and the symbols α, β and γ signifies the feature map that plays as the output of the individual conv filter.

$$\omega_s{}^h = Af_n\left(\omega_s{}^{h-1}\right) \quad (6)$$

The role of ReLU layer is that it provides the sufficient potential to manage with greater number of networks.

(2) Pooling layers: The pooling layer follows the convolutional layer to replace the output of the previous into a type of summary statistic of closest outputs. The maximum pooling operation generates the maximum value between a group of closest inputs and it is expressed as,

$$PP_{i_1}{}^{l_1,j_1} = \underset{ss \in SS}{max}\left(A_{c_{l_1}\times ST+SS}{}^{l_1,j_1}\right) \tag{7}$$

where, *SS* is the pooling size, *ST* denotes the pooling stride and $PP_{i_1}{}^{l_1,j_1}$ is the pooling layer output. A simple SoftMax classifier is used to recognize the activities, which is presented at the topmost layer.

(3) Fully Connected (FC) layers: The signal obtained from the pooling layer is subjected to FC layer. At the end side, the signals are converted into a single signal that indicates the classes of signal. The result of the fully connected layer is given by,

$$FC_s{}^h = \underset{\alpha-1}{\psi}\left(\omega_s{}^h\right) with \;\; \omega^h{}_s = \sum_{\alpha=1}^{x_1{}^{\alpha-1}} \sum_{\beta=x_1{}^n}^{x_1{}^{nn}} \sum_{\gamma=-x_2{}^{nn}}^{x_2{}^{nn}} \left(\delta_{s,\alpha}\right)_{\gamma,\beta}\left(FS_\alpha{}^{nn-1}\right)_{q+\beta,w+\gamma} \tag{8}$$

The weight values of the DCNN are determined employing the SGD that trains the DCNN so as to achieve the optimal weights. The output obtained from the Deep CNN is denoted as D_c.

3.5.2 Training of DCNN using proposed Fractional-EWA

Fractional-EWA integrates the benefits of the Earth Worm Optimization Algorithm (EWA) [22] and Fractional Calculus [13], in order to train the Deep CNN effectively using this Fractional-EWA in order to optimize the best solution. EWA is a bio-inspired metaheuristic optimization that is influenced by the reproducing nature of earthworms that is categorized into two main phases of reproduction and in later step, the best solutions are generated by adding weights to produce new earthworms. The Cauchy operator employed in EWA is used to enhance the searching capability of the earthworms thus resulted in reducing the avoidance of local optimal. From the two kinds of reproduction, the reproduction-1 is a category, where one offspring is produced. In other side, in reproduction-2, one or more young ones are produced by the parent earthworms. In addition to this, EWA is possible for real-world applications. Factional calculus concept is applied to enhance the accuracy level of categorization since the optimal solutions of preceding iterations are utilized for updating the solution.

- *Reproduction 1:* The individual earthworm is included in producing the young ones as it is familiar that earthworms are hermaphrodites in nature. The mathematical expression of the reproduction-1 is expressed as,

$$Z_{k,y} = Z_{max,y} + Z_{min,y} - \eta Z_{k,y} \tag{9}$$

  where, $Z_{k,y}$ is the y^{th} element of Z_k or it is denoted as the location of k^{th} earthworm and $Z_{k1,y}$ is the y^{th} component of Z_{k1}, which signifies the new location of k1th earthworm. Here, the location is limited to upper and lower bounds, $Z_{max,y}$ and $Z_{min,y}$. η denotes the similarity parameter that lies between 0 and 1 and the similarity parameter determines distance between the kth earthworm and reproduced earthworm Z_{k1}.

- *Reproduction 2:* In the second form of reproduction obtains one or more young ones. Here, three examples are assured throughout the changes in parents number P_a to obtain the young ones O_s that is not small than zero. The crossover described here is of three types, namely individual point crossover, multi-point crossover, and uniform crossover. However, the process of choosing the parents is completely depends on mechanism of roulette wheel selection.

  - *Case 1:* With $P_a = 2$, $O_s = 1$: In such case, two parents are chosen that is indicated as jj_1, and jj_2 and in terms of single point crossover, two variables are needed that is calculated as,

$$S_{p_1} = \sigma\left(jj_1(Z{:}X), jj_2(Z{:}X)\right) \tag{10}$$

$$S_{p_2} = \sigma\left(jj_2(Z{:}X), jj_2(Z{:}X)\right) \tag{11}$$

    where, σ denotes the set difference of arrays $jj_1(Z{:}X)$ and $jj_2(Z{:}X)$. The value of Z is invariant between 1 and X and reproduced young ones are expressed as,

$$Z_{12} = \left[jj_2\left(1{:}X - \left|S_{p_1}\right|\right), S_{p_1}\right] \tag{12}$$

$$Z_{22} = \left[jj_2\left(1{:}X - \left|S_{p_2}\right|\right), S_{p_2}\right] \tag{13}$$

where, $\left|S_{p_1}\right|$ and $\left|S_{p_2}\right|$are the variables length and the offspring is produced depending on the following criterion as,

$$Z_{k2} = \begin{cases} Z_{12} rand\langle 0.5 \\ Z_{22} else \end{cases} \tag{14}$$

where, *rand* denotes the random number generated. However, the multi-point crossover is relied on two random numbers Z_1 and, Z_2 where Z_2 is lies below Z_2 and produced off springs depend on the below expressions,

$$Z_{12} = [jj_1(1{:}Z_1), jj_2(Z_1 + 1{:}Z_2), jj_1(Z_1 + 1{:}X)] \tag{15}$$

$$Z_{22} = [jj_2(1{:}Z_1), jj_1(Z_1 + 1{:}Z_2), jj_2(Z_2 + 1{:}X)] \tag{16}$$

The young ones of the reproduction 2 follows Eqn. (17). In uniform crossover, the off springs are produced depending on the *rand* as,

$$Z_{12,y} = jj_{1,y}{:}Z_{22,y} = jj_{2,y}; if rand\rangle 0.5$$

$$Z_{12,y} = jj_{2,y}{:}Z_{22,y} = jj_{1,y}; otherwise \tag{17}$$

- *Case 2:* with $P_a = 2$, $O_s = 2$: Two parents are utilized and latest produced offspring is expressed as,

$$Z_{k2} = \lambda_1 Z_{12} + \lambda_2 Z_{22} \tag{18}$$

where, λ_1 and λ_2 represents the weighting factors, which is computed depending on the fitness of earthworms Z_{12} and Z_{22} calculated using the Eqn. (15), and Eqn. (16). Correspondingly, the off springs in multipoint crossover is calculated using Eqn. (17) and Eqn. (18).

- *Case 3:* With $P_a = 3$, $O_s = 3$: In this case, three parents are selected for producing three off springs. Here, three variables, such as S_{P_1}, S_{P_2}, and S_{P_3} are produced and it is expressed as,

$$S_{P_1} = \omega\left(jj_1(Z{:}X), jj_3(Z{:}X)\right) \tag{19}$$

$$S_{P_2} = \omega\left(jj_2(Z{:}X), jj_2(Z{:}X)\right) \tag{20}$$

$$S_{P_3} = \omega\left(jj_2(Z{:}X), jj_1(Z{:}X)\right) \tag{21}$$

Thus, the generated off springs are expressed as,

$$Z_{31} = \left[jj_3\left(1{:}X - \left|S_{P_1}\right|\right), S_{P_1}\right] \tag{22}$$

$$Z_{32} = \left[jj_2\left(1{:}X - \left|S_{P_2}\right|\right), S_{P_2}\right] \tag{23}$$

$$Z_{33} = \left[jj_3\left(1{:}X - \left|S_{P_3}\right|\right), S_{P_3}\right] \tag{24}$$

Once the young ones are generated, the new earthworms of type-1 reproduction is given by,

$$jj_{k2} = \sum_{y=1}^{3} \lambda_y jj_{3y} \tag{25}$$

After the completion of reproduction of offspring from both kinds of reproduction, the location of the k^{th} earthworm is represented as,

$$Z_{k,y}{}^{t^{n+1}} = \theta . Z_{k,y}{}^{1} + (1 - \theta) Z_{k,y}{}^{2} \tag{26}$$

where, the proportional factor is denoted as θ that is considerably reduced with increasing number of iterations so that the large value of θ enhances the search. It is also described that the controlling factor between the local search and global search is stimulated utilizing two constants. In order to improve the potential of search, the Cauchy function is utilized to compute the location of earthworm.

The y^{th} location of k^{th} earthworm is calculated according to the Cauchy operator and it is expressed as,

$$Z_{k,y}{}^{\tau+1} = Z_{k,y}{}^{\tau} + \lambda_y * \vartheta \tag{27}$$

where, ϑ indicates the random number created utilizing the Cauchy distribution and λ_y signifies the weight vector of y^{th} location and $Z_{k,y}{}^{\tau+1}$ represents the y^{th} position of k^{th} earthworm at instant τ. In this step, the concept of fractional calculus is applied to modify the above equation and it is expressed as,

$$Z_{k,y}{}^{\tau+1} - Z_{k,y}{}^{\tau} = \lambda_y * \vartheta \tag{28}$$

$$\partial^{\mu}\left[Z_{k,y}{}^{\tau+1}\right] = \lambda_y * \vartheta \tag{29}$$

$$\partial^{\mu}\left[Z_{k,y}{}^{\tau+1}\right] = Z_{k,y}{}^{\tau+1} - \mu.Z_{k,y}{}^{\tau} - \frac{1}{2}\mu.Z_{k,y}{}^{\tau-1} - \frac{1}{6}(1-\mu).Z_{k,y}{}^{\tau-2} - \frac{1}{24}\mu.(1-\mu)(2-\mu)Z_{k,y}{}^{\tau-3} \tag{30}$$

where, $Z_{k,y}{}^{\tau}$, $Z_{k,y}{}^{\tau-1}$, $Z_{k,y}{}^{\tau-2}$signifies the location of earthworms in preceding iterations and the fractional constant is signified as μ.

Substituting the Eqn. (30) in Eqn. (29), results the update equation of the developed Fractional-EWA algorithm.

$$Z_{k,y}{}^{\tau+1} - \mu.Z_{k,y}{}^{\tau} - \frac{1}{2}\mu.Z_{k,y}{}^{\tau-1} - \frac{1}{6}(1-\mu).Z_{k,y}{}^{\tau-2} - \frac{1}{24}\mu.(1-\mu)(2-\mu)Z_{k,y}{}^{\tau-3} = \lambda_y * \vartheta \tag{31}$$

$$Z_{k,y}{}^{\tau+1} = \mu.Z_{k,y}{}^{\tau} + \frac{1}{2}\mu Z_{k,y}{}^{\tau-1} + \frac{1}{6}(1-\mu).Z_{k,y}{}^{\tau-2} + \frac{1}{24}\mu.(1-\mu)(2-\mu)Z_{k,y}{}^{\tau-3} + \lambda_y * \vartheta \tag{32}$$

The solution of current iteration $Z_{k,y}{}^{\tau+1}$ is updated depending on solutions in earlier iteration. However, convergence of the problem is improved and there is an efficient balance between the global optima and local optima. In addition to this, the search ability of algorithm is enhanced with a global convergence. Table. 1 portrays the pseudo code of developed Fractional-EWA.

Sl. No	Pseudo code of proposed Fractional-EWA
1	Input: $Z_{k,y}$
2	Output: $Z_{k,y}{}^{\tau+1}$
3	Begin
4	Initialize the parameters
5	Set the generation counter as $\tau = 1$
6	Initialize the population $P_A\langle\rho, \rho$is the total earthworms
7	Maximum generation τ_{max}
8	Similarity factor η
9	Initialize the proportional factor
10	Compute the fitness measure
11	While $\tau\langle\tau_{max}$
12	Arrange all the earthworms using fitness measure
13	For all $P_A\langle\rho$
14	Process reproduction-1
15	Generate $Z_{k_1,y}$ using reproduction-1
16	Process reproduction-2
17	If $\rho\rangle$Number of kept earthworms
18	Define the selected number of parents and offspring earthworms
19	Select the parents based on the roulette wheel selection
20	Generate the offspring
21	Determine $Z_{k_2,y}$

22	*else*
23	Select $Z_{k_2,y}$ in random
24	*end if*
25	Update the earthworm position using Eqn. (32)
26	*end for*
27	For all unselected earthworm
28	Perform Cauchy mutation
29	*end for*
30	Determine population based on newly updated position $\tau = \tau + 1$
31	*end while*
32	Obtain the best solution
33	Terminate

Table. 1. Algorithm of pseudo code of the proposed Fractional-EWA

4. Results and Discussion

This section elaborates the results and discussion of developed Fractional-EWA based DCNN and the performance of the developed approach is evaluated with respect to the evaluation metrics.

4.1 *Experimental setup*

The implementation of developed Fractional-EWA based DCNN is done in JAVA tool and the dataset utilized for the implementation purpose are Enron [32]. Moreover, the experimentation of the developed method is carried out in personal computers with intel core i-3 processor.

4.1.1 Dataset description

The dataset, such as Enron [32], are employed for the proposed Fractional-EWA based DCNN. The raw subdirectory of Enron consists of indigenous messages and spam messages are accessible in non-Latin encodings. Moreover, the ham messages are transmitted from the mailboxes of the authorized user to themselves. The attacked contents are separated without leaving any changes.

4.2 *Evaluation metrics*

The analysis of developed Fractional-EWA based DCNN is made using evaluation metrics, like accuracy, specificity, and sensitivity.

4.2.1 Accuracy

Accuracy is termed as the degree of closest value of measurements of a quantity. In other words, accuracy is also stated as the standard of the correct or precise measurement.

$$Accuracy = \frac{\alpha_p+\alpha_n}{\alpha_p+\beta_p+\beta_n+\alpha_n} \tag{33}$$

where, α_p denotes the true positive, α_n refers the true negative. The false positive and false negative are denoted as β_p and, β_n respectively.

4.2.2 Sensitivity

Sensitivity is defined as the true positive rate that determines the proportions of positives that are perfectly identified.

$$Sensitivity = \frac{\alpha_p}{\alpha_p+\beta_n} \tag{34}$$

Where, α_p denotes the number of True positives and the number of false negatives is represented as β_n.

4.2.3 Specificity

Specificity is referred as the true negative rate that calculates the proportion of negatives that are correctly identified.

$$Specificity = \frac{\alpha_n}{\alpha_n + \beta_p} \quad (35)$$

4.3 *Comparative methods*

The performance improvement of developed Fractional-EWA based DCNN is analyzed by comparing the existing techniques, such as Naive Byes (NB) [5], Deep Belief Networks (DBN) [9], Neural Networks (NN) [20], EWA-DBN, and Fractional EWA-DBN.

4.4 *Comparative analysis*

This section deliberates the comparative analysis made by developed Fractional-EWA based DCNN using the dataset Enron.

4.5 *Analysis using dataset from Enron*

This section demonstrates the analysis of developed strategy using dataset by changing the chunk data and the number of features.

4.5.1 Analysis using chunk data

The analysis of developed Fractional-EWA based DCNN using the dataset by varying the chunk data is portrayed in figure below. Fig.3 shows the analysis of approach with respect to accuracy by varying chunk data. When the chunk data is 50%, the accuracy achieved by developed Fractional-EWA based DCNN is 0.876, whereas the conventional methods, such as NB is 0.533, DBN is 0.545, NN is 0.556, EWA-DBN is 0.571, and Fractional EWA-based DBN is 0.857. By varying the chunk percentage to 90%, the proposed Fractional-EWA based DCNN attained the accuracy of 0.781 that shows the percentage improvement of proposed with that of the conventional techniques, like NB is 31.813%, DBN is 24.383%, NN is 14.688%, EWA-DBN is 3.296%, and Fractional-EWA based DBN is 2.934%, respectively. However, the accuracy obtained by the traditional schemes, such as NB is 0.533, DBN is 0.591, NN is 0.667, EWA-DBN is0.756, and Fractional EWA-DBN is 0.759.

The analysis made by the developed Fractional-EWA based DCNN using sensitivity is portrayed in Fig.4. When the chunk percentage is 50%, the sensitivity attained by the developed Fractional-EWA based DCNN is 0.840, whereas the existing approaches attained the sensitivity for NB is 0.456, DBN is 0.563, NN is 0.704, EWA-DBN is 0.704, and Fractional-EWA DBN is 0.818. Similarly, if the chunk percentage is increased to 90%, the sensitivity obtained by the conventional techniques, such as NB is 0.489, DBN is 0.563, NN is 0.703, EWA-DBN is 0.703, and Fractional-EWA DBN is 0.757. However, the proposed Fractional-EWA based DCNN achieved a sensitivity of 0.782. The performance enhancement of proposed with that of the existing methods are 37.541% for NB, 28.095% for DBN, 10.151% for NN, 10.134 for EWA-DBN, and 3.312 for Fractional EWA-DBN.

Fig.5 illustrates the analysis using specificity by varying chunk data. If the chunk data is 50%, the specificity attained by developed Fractional-EWA based DCNN is 0.907, whereas the conventional approaches, like NB is 0.505, DBN is 0.563, NN is 0.703, EWA-DBN is 0.704, and Fractional EWA-DBN is 0.880. By varying the chunk percentage to 90%, the performance improvement of proposed approach while comparing it with the conventional techniques, like NB is 29.183%, DBN is 22.021%, NN is 12.249%, EWA-DBN is 2.576%, and Fractional EWA-DBN is 2.565%, respectively. However, the specificity obtained by the conventional schemes, such as NB is 0.511, DBN is 0.563, NN is 0.633, EWA is 0.703, and Fractional EWA-DBN is 0.703.

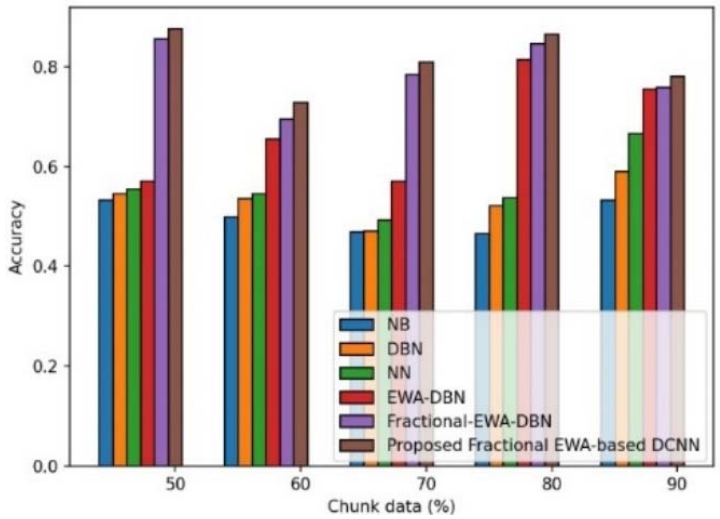

Fig.3. Analysis using dataset Enron by varying chunk data for Accuracy.

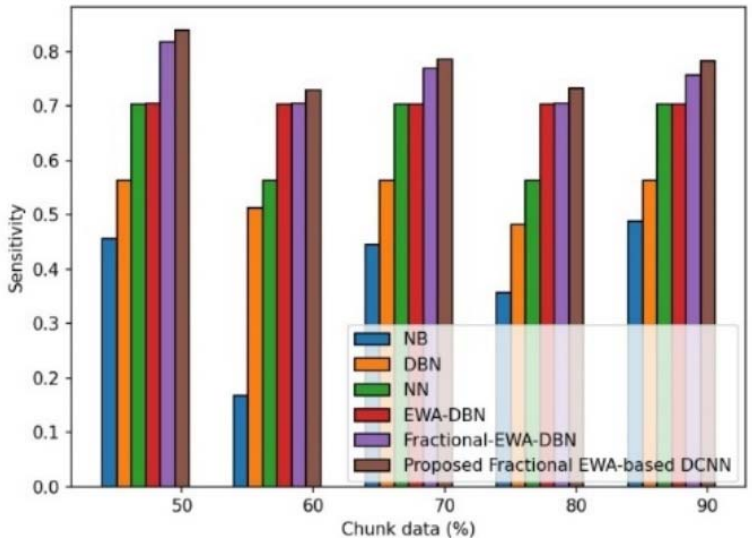

Fig.4. Analysis using dataset Enron by varying chunk data for Sensitivity.

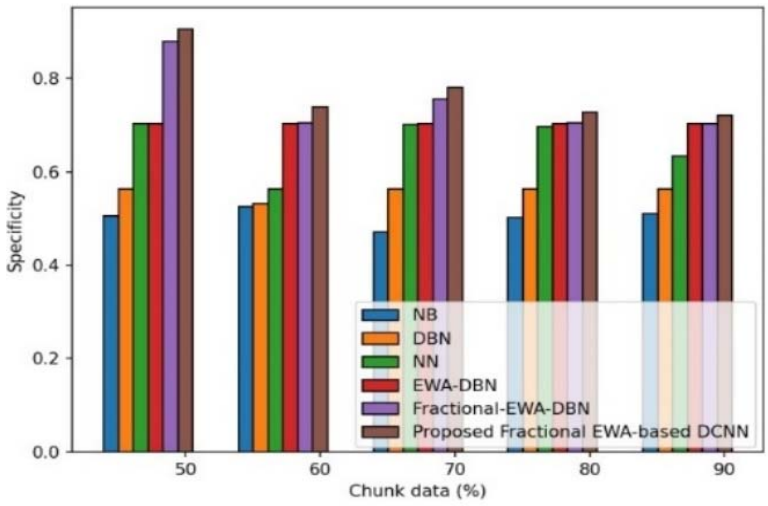

Fig.5 Analysis using dataset Enron by varying chunk data for Specificity

4.5.2 *Analysis using number of features*

The below figures portray the analysis of developed Fractional EWA-based DCNN using the dataset by changing the feature numbers. Fig.6 illustrates the analysis using accuracy. When the number of features = 50, the accuracy achieved by the developed Fractional EWA-based DCNN is 0.810, whereas the conventional techniques attained the accuracy for NB is 0.469, DBN is 0.471, NN is 0.493, EWA-DBN is 0.571, and Fractional EWA-DBN is 0.785. By varying the features to 90, the accuracy achieved by the developed Fractional-EWA based DCNN is 0.744 that outcome the performance enhancement of developed approach with that of the conventional schemes, such as NB is 52.314%, DBN is 48.662%, NN is 46.218%, EWA-DBN is 34.589%, and Fractional EWA-DBN is 2.311%. However, the accuracy achieved by the conventional techniques, like NB is 0.355, DBN is 0.382, NN is 0.400, EWA-DBN is 0.486, and Fractional EWA-DBN is 0.727, respectively.

Fig.7 portrays the analysis of proposed Fractional-EWA based DCNN using sensitivity. If the number of features is 50, the sensitivity achieved by the developed Fractional-EWA based DCNN is 0.786, whereas the conventional schemes attained the sensitivity of 0.445 for NB, 0.563 for DBN, 0.703 for NN, 0.704 for EWA-DBN, and 0.768 for Fractional EWA-DBN. Correspondingly, if the features are increased to 90 in numbers, the sensitivity achieved by the proposed Fractional-EWA based DCNN is 0.725 that shows the performance development of the developed approach with that of the existing methods, such as NB is 65.541%, DBN is 38.847%, NN is 22.258%, EWA-DBN is 3.117%, and Fractional EWA-DBN is 3.113%.

The analysis made by the proposed Fractional-EWA based DCNN using specificity is depicted in Fig. 8. When the number of features is 50%, the specificity obtained by the developed approach is 0.774, whereas the existing methods attained the specificity for NB is 0.470, DBN is 0.563, NN is 0.702, EWA-DBN is 0.703, and Fractional

EWA-DBN is 0.756. If the number of features is increased to 90, the specificity attained by the proposed approach is 0.723, whereas the conventional approaches achieved the specificity of 0.466 for NB, DBN is 0.563, NN is 0.571, EWA-DBN is 0.701, and Fractional EWA-DBN is 0.703. However, the performance improvement of proposed with that of the existing techniques, such as NB is 35.600%, DBN is 22.067%, NN is 20.949%, EWA-DBN is 2.979%, and Fractional EWA-DBN is 2.786%.

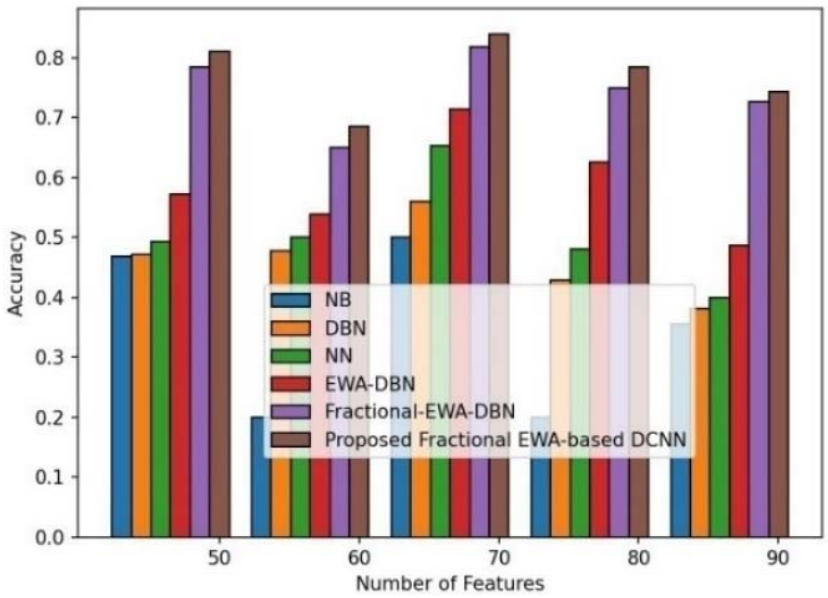

Fig.6. Analysis using dataset Enron by varying the number of features for Accuracy

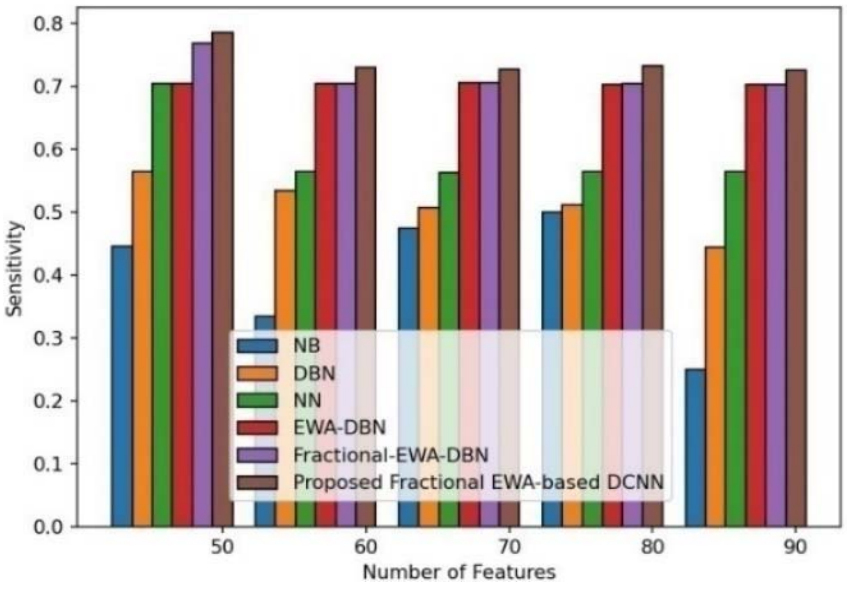

Fig.7. Analysis using dataset Enron by varying the number of features for Sensitivity

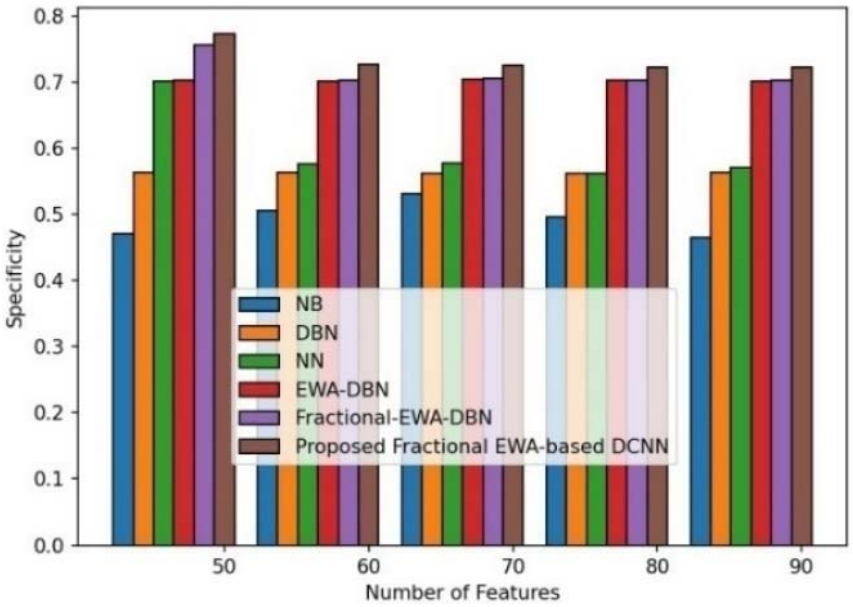

Fig.8.Analysis using dataset Enron by varying the number of features for Specificity

4.7 Comparative discussion

Table 2 portrays the comparative discussion of developed approach. The data from Enron dataset are considered here for the comparison with respect to the evaluation metrics, like accuracy, sensitivity, and specificity. For chunk percentage of 90%, the accuracy, sensitivity, and specificity achieved by developed Fractional-EWA based DCNN are 0.781, 0.782, and 0.722, respectively. The specificity achieved by the existing methods, such as NB is 0.466, DBN is 0.563, NN is 0.571, EWA-DBN is 0.701, and Fractional EWA-DBN is 0.703 by varying the number of features.

Datasets	**Metrics**		**NB**	**DBN**	**NN**	**EWA -DBN**	**Fractional EWA-DBN**	**Proposed Fractional EWA based DCNN**
Dataset	Chunk data	*Accuracy*	0.533	0.591	0.667	0.756	0.759	**0.781**
		Sensitivity	0.489	0.563	0.703	0.703	0.757	**0.782**
		Specificity	0.511	0.563	0.633	0.703	0.703	**0.722**
	Number of features	*Accuracy*	0.355	0.382	0.400	0.486	0.727	**0.744**
		Sensitivity	0.250	0.444	0.564	0.703	0.703	**0.725**
		Specificity	0.466	0.563	0.571	0.701	0.703	**0.723**

Table. 2. Comparative discussion

5. Conclusion

The unauthorized access of confidential information, like user name, passwords, and credit card details using various tricks by hackers is called as phishing attack. Recently, phishing attacks are considered as one of the frequently caused cybercrime attacks over the internet users. However, detection of such phishing attacks and classification of spam mails is still a major hurdle. To counter such issues, this research introduces a new effective strategy for adaptive spam filtering by developing a newly proposed method called Fractional-EWA based DCNN.

The developed Fractional-EWA is derived by the inclusion of fractional calculus concept to the Earthworm Optimization. The proposed strategy includes four phases, namely pre-processing, feature extraction, feature selection, and spam mail classification. The features are separated using the term frequency, whereas the selected features are derived utilizing levenshtein distance. The Deep Convolutional Neural Network is trained by employing the developed Fractional-EWA. However, the developed Fractional-EWA based DCNN attained the maximum accuracy of 0.781, sensitivity of 0.782, and specificity of 0.722 for chunk percentage and achieved the maximum accuracy of 0.744, sensitivity of 0.725, and specificity of 0.723, respectively for number of features.

References

[1] Abdulla.S, and Altyeb Altaher, "IHASS: Intelligent and Hybrid Anti-Spam System", International Journal of Computer Networks and Communications Security, vol.5, no.6, p.115, 2017.
[2] Abu-Nimeh. S, Dario Nappa, Xinlei Wang, and Suku Nair, "A comparison of machine learning techniques for phishing detection", In Proceedings of the anti-phishing working groups 2nd annual eCrime researchers' summit, pp. 60-69, October, 2007.
[3] AHazem Al Saied, Nicolas Dugue, and Jean-Charles Lamirel, "Automatic summarization of scientific publications using a feature selection approach", International Journal on Digital Libraries, vol.19, no.2, pp.203-215, 2018.
[4] Ali Mohammad H. Al-Ibrahim, "Using Sequential Minimal Optimization for Phishing Attack Detection", Modern Applied Science, vol.13, no.5, 2019.
[5] Androutsopoulos, I., Paliouras, G., Karkaletsis, V., Sakkis, G., Spyropoulos, C.D. and Stamatopoulos, P., "Learning to filter spam e-mail: A comparison of a naive Bayesian and a memory-based approach", arXiv preprint cs/0009009, 2000.
[6] Androutsopoulos.I, John Koutsias, Konstantinos V. Chandrinos, George Paliouras and Constantine D. Spyropoulos, "An evaluation of naive Bayesian anti-spam filtering", arXiv preprint cs/0006013, 2000.
[7] Ankit Kumar Jain, and B. B. Gupta, "A machine learning based approach for phishing detection using hyperlinks information", Journal of Ambient Intelligence and Humanized Computing, vol.10, no.5, pp.2015-2028, 2019.
[8] Anuj Kumar Singh, Shashi Bhushan, and Sonakshi Vij, "Filtering spam messages and mails using fuzzy C means algorithm", In IEEE 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), pp. 1-5, April 2019.
[9] Barushka. A, and Petr Hajek, "Spam filtering using integrated distribution-based balancing approach and regularized deep neural networks", Applied Intelligence, vol.48, no.10, pp.3538-3556, 2018.
[10] Basavaraju M, and R. Prabhakar, "A novel method of spam mail detection using text-based clustering approach", International Journal of Computer Applications, vol.5, no.4, pp.15-25, 2010.
[11] Behjat.A. R, Aida Mustapha, Hossein Nezamabadi-pour, Md. Nasir Sulaiman, and Norwati Mustapha, "GA-based feature subset selection in a spam/non-spam detection system", In IEEE International Conference on Computer and Communication Engineering (ICCCE), pp. 675-679, July 2012.
[12] Gajera.K, Mukul Jangid, Palash Mehta, and Jayashri Mittal, "A novel approach to detect phishing attack using artificial neural networks combined with pharming detection", In 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), pp. 196-200, 2019 June.
[13] Gorenflo, R. and Mainardi, F., "Fractional calculus: integral and differential equations of fractional order", arXiv preprint arXiv:0805.3823, 2008.
[14] Ismaila Idris, Ali Selamat, and Sigeru Omatu, "Hybrid email spam detection model with negative selection algorithm and differential evolution", Engineering Applications of Artificial Intelligence, vol.28, pp.97-110, 2014.
[15] Ismaila Idris, and Abdulhamid Shafi'i Muhammad, "An improved AIS based e-mail classification technique for spam detection", arXiv preprint arXiv:1402.1242, 2014.
[16] Jiachang Qian, Jiaxiang Yi, Jinlan Zhang, Yuansheng Cheng, and Jun Liu, "An Entropy Weight-Based Lower Confidence Bounding Optimization Approach for Engineering Product Design", Applied Sciences, vol.10, no.10, p.3554, 2020.
[17] Ram B. Basnet, Andrew H. Sung, and Quingzhong Liu, "Rule-based phishing attack detection", In Proceedings of the International Conference on Security and Management (SAM) (p. 1). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing, 2011.
[18] Ronaoo, C.A. and Cho, S.B., "Evaluation of deep convolutional neural network architectures for human activity recognition with smartphone sensors", pp.858-860, 2015.
[19] Shuaib. M, Shafii Muhammad Abdulhamid, Olawale Surajudeen Adebayo, Oluwafemi Osho, Ismaila Idris, John K. Alhassan, and Nadim Rana, "Whale optimization algorithm-based email spam feature selection method using rotation forest algorithm for classification", SN Applied Sciences, vol.1, no.5, p.390, 2019.
[20] Smadi.S, Nauman Aslam and Li Zhang, "Detection of online phishing email using dynamic evolving neural network based on reinforcement learning", Decision Support Systems, vol.107, pp.88-102, 2018.
[21] Uysal.U. K, and Serkan Gunal, "The impact of preprocessing on text classification", Information Processing & Management, vol.50, no.1, pp.104-112, 2014.
[22] Wang, G.G., Deb, S. and Coelho, L.D.S., "Earthworm optimization algorithm: a bio-inspired metaheuristic algorithm for global optimization problems", International Journal of Bio-Inspired Computation, vol.12, no.1, pp.1-22, 2018.
[23] Zabihimayvan. M and Derek Doran, "Fuzzy rough set feature selection to enhance phishing attack detection", In IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1-6, June 2019.
[24] Adebowale M A, K.T. Lwin, E. Sanchez, and M.A. Hossain, "Intelligent web-phishing detection and protection scheme using integrated features of Images", frames and text. Expert Systems with Applications, vol.115, pp.300-313, 2019.
[25] Dhamdhere. B.D, Kaushal Sudhakar Dhonde, and Rohit Gopal Chinchwade R.G., "A Hybrid Model to Detect Phishing-Sites using Clustering and Bayesian Approach", IJCSNS, vol.15, no.1, p.92, 2015.
[26] Hassani Z, V. Hajihashemi, K. Borna, and I. Sahraei Dehmajnoonie, "A Classification Method for E-mail Spam Using a Hybrid Approach for Feature Selection Optimization", Journal of Sciences, Islamic Republic of Iran, vol.31, no.2, pp.165-173, 2020.
[27] Korkmaz . M, Ozgur Koray Sahingoz, and Banu Diri, "Feature Selections for the Classification of Webpages to Detect Phishing Attacks: A Survey", In IEEE International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1-9, June 2020.
[28] Kumaresan T, S. Saravanakumar, and R. Balamurugan, "Visual and textual features-based email spam classification using S-Cuckoo search and hybrid kernel support vector machine", Cluster Computing, vol.22, no.1, pp.33-46, 2019.
[29] Renukha. D.K and P. Visalakshi, "Weighted-based multiple classifier and F-GSO algorithm for email spam classification", International Journal of Business Intelligence and Data Mining, vol.12, no.3, pp.274-298, 2017.

[30] Sahingoz.O. K, Ebubekir Buber, Onder Demir, and Banu Diri c, "Machine learning based phishing detection from URLs", Expert Systems with Applications, Vol.117, pp.345-357, 2019.
[31] Venkatraman S, B. Surendiran, and P. Arun Raj Kumar, "Spam e-mail classification for the internet of things environment using semantic similarity approach", The Journal of Supercomputing, vol.76, no.2, pp.756-776, 2020.
[32] Enron dataset taken from, "http://nlp.cs.aueb.gr/software_and_datasets/Enron-Spam/index.html" accessed on March 2021.

Authors Profile

Arshey M is currently working as Assistant Professor, Department of Computer Science and Engineering, SCMS School of Engineering and Technology, Ernakulam. She is having more than 12 years of teaching experience and 2 years of Industrial experience. She received her B.E. degree and M.E degree in Computer Science and Engineering from Anna University, Chennai. She is currently a Research Scholar from Noorul Islam Centre for Higher Education. Her areas of interest are Network Security, Cyber Security, Machine Learning and Deep Learning. She has guided projects for B. Tech and MCA degree students in the domain of Machine Learning and Security in Android. She is a member of ISTE Society.

Angel Viji K S is currently working as Associate Professor, Department of Computer Science and Engineering, College of Engineering, Kidangoor. She is having over 12 years of teaching experience and holds a doctorate in area of Medical Image Processing. She received her B.E. degree and M.E degree in Computer Science and Engineering from Anna University, Chennai. Her dedicated involvement in R & D in Medical image Processing and Network Security has been recognized through 31 publications amongst which 3 are SCI indexed, 5 are SCOPUS indexed, 11 papers in refereed, indexed Research Journals and the few in indexed conferences such as IEEE, Springer etc. In view of sharing knowledge and serve technical community, she became the professional member of IEAE and IEEE.