# Bio-inspired Metaheuristic Optimization Techniquefor the Detection of Phishing Emails

Arshey M
Department Of Computer Science and Engineering
Noorul Islam Centre for Higher Education,
Thuckalay

Dr. Angel Viji K S
Department Of Computer Science and EngineeringCollege of Engineering,
Kidangoor

***Abstract*—The most trending cybersecurity threat all over the world is Phishing, which uses the public through various media especially e-mail, to gather the individual's private particulars. This rapid rise of undesired information needs to be coped with, raising the need to develop suitable and efficient anti-phishing methods. This paper emphasizes a process to detect email phishing based on optimization algorithms using deep belief networks. At the first, the emails are subjected to pre-processing using stemming and stop word removal mechanisms are implemented to assure that the significant words are identified for further processing. Term-Frequency (TF) is used for feature extraction from the significant words, followed by the Bhattacharya distance for feature selection. The features selected are fed as input to the deep belief neural network (DBN), which is then trained using the proposed Earth Worm optimization (EWA) Algorithm. The analysis of the spam mail detection is performed using the datasets and found that the accuracy, sensitivity, and specificity of the proposed EWA DBN are found to be a maximal value of 0.671, 0.814, and 0.804, respectively.**

***Index Terms*—E-mail, Phishing, Optimization, deep learning, spam mails, Deep Belief Network(DBN)**

## I. INTRODUCTION

Technology enhancement brings out fresh criminal ways and many new types of crimes. The Web is upright for developing and refining worldwide commerce to already far-fetched statures, cultivating momentous headways in instruction, and inspiring round-the-world communication that was once seen to be constrained and exorbitant. Regardless, the Web, with its boundless measure and as of now unimaginable capacities, remembers a despairing side for that it has opened windows of effectively dark criminal openings that not in a manner of speaking test, but rather too transcend every actual limit, boundaries, and limitations to detect, rebuke and lessen what appears at being a creating social issue of overall degrees. Cybercrime is an offense to data, the public, associations, or governments. The idea of digital infringement isn't radically different from the idea of standard bad behavior. Both fuse directly whether act or prohibition, which causes a break of rules of law counterbalanced by the support of the state. [1] Computer-based wrongdoing insinuates any bad behavior that incorporates a system and an organization.

Phishing is the technique for delicate data, similar to passwords, usernames, and credit card data for noxious purposes, through dissimulating as a dependable individual in electronic correspondence. Phishing messages consist of sites linked with malware. Phishing is subsequently performed utilizing texting or email parodying, which makes the clients give their subtleties in any phony site that looks and seems indistinguishable from the real site. Phishing remains an occurrence for social designing strategies that misleads clients, and adventures helpless convenience of present security advances in the web. Phishing is a danger that forces huge negative effects on online media, similar to Twitter, Facebook, and Google+. Programmers clone a site and demand the web clients to give the individual data that is at last sent to the programmers [2]. Additionally, there are various anti-phishing procedures to perform phishing and smishing is a consolidation of Phishing assaults that use a basic instant message or Short Messaging Service (SMS) on mobiles to claim the individual credentials [3] [4]. Accordingly, it is outstanding that phishing irritated the clients as well as caused financial damage for people and associations [5].

Spam is the undesirable message of a sender sent electronically to a beneficiary, who doesn't have any relationship with the individual [6]. Email spam alludes to a subset of electronic spam that takes enormous time since the clients participate in recognizing and eliminating the undesired messages. The common issues on the web are in regards to email spam. [7]Spamming is the consistently enduring issue that is accessible from the hour of the presence of mailboxes. The methods utilized for separating are progressing with time and the level of spam messages are rising definitely with time, causing tremendous traffic in the messages. In this way, a successful spam channel is utilized for upgrading the efficiency of the client and limits the utilization of the assets related to the data innovation, similar to help work areas.

There are various spam filters utilized alongside the AI techniques, similar to decision trees, Naive Bayes classifiers, k-closest neighbor algorithm, SVM, K-means algorithm, and many more8. Machine Learning techniques consequently build up the word records alongside their weights for arranging the messages as two classes. The input messages could be either spam or not. Also, there are various strategies utilized for identifying spam.

The main aim of this research is to develop an approach for eliminating phishing by recommending an optimization algorithm. The proposed method involves four steps, which include pre-processing, feature extraction, feature selection, and classification of phishing emails. Initially, the stop word elimination and stemming of the input dataset is performed

in the pre-processing stage followed by the feature extraction process. The features are selected based on extracting the keyword frequency from the output of the pre-processing. The next step is the feature selection using Bhattacharya distance to identify the significant features for the classification stage. The selected features are subject to classification using the Deep Belief Network and trained using the proposed EWA.

## II. Related Works

The review of various methods is deliberated in this section. Smadi, S et al. [9] developed an algorithm to detect the zero-day phishing attacks using 2 techniques namely Feature Evaluation and Reduction algorithm and (DENNuRL) Dynamic Evolving Neural Network using a Reinforcement learning algorithm. As per the algorithm, the result revealed a higher performance and provided reasonable error rates. The main drawback of this technique was due to the insufficient amount of dataset chosen for classification, which was critical to group the spam mails.

Barushka. A and Hajek P [10] designed an algorithm to effectively handle the class distributions which shows the imbalance and misclassification costs with some difficult forms of text patterns. The Algorithm namely Distribution-based balancing along with the regularized deep multi-layer perceptrons NN model with rectified linear units (DBB-RDNN-ReL) can help in effectively tackling the class distributions with imbalance. The disadvantage of the method is that it makes use of numerous hidden layers and the units in the model would exhibit noise in the data, which leads to unsatisfactory performance.

Kovalluri, S.S et al. [11] designed a system based on Artificial Intelligence using LSTM. This helped in reducing the application of fake mails to sneak the data, proliferate, and made it difficult to track the victims. The disadvantage of this technique was that this model had errors during sentence generation.

Ruano-Ordás et al. [12] designed a model using the Genetic programming algorithm to be used for datasets that were large and also identified the patterns which improved the accuracy to great extent. However, it further helped in the reduction of the computational overhead related to the e-mail filter server. The main drawback of this technique was the requirement of security features to prevent False Positive errors.

Sonowal, G and Kuppusamy, K.S [13]designed an algorithm that used the SMIshing Detection based on the Correlation Algorithm (SmiDCA) that accomplished higher efficiency to confront datasets based on both the English and non-English. However, to improve the accuracy the system had to depend on deep learning technologies.

## III. DBN based spam mail Classification

An Email Spam causes affliction in the digital world and it imprints the loss of time, space, and communication bandwidth. Almost more than 40% of the mails are fake nowadays that implies that more than 15 billion emails a day, thereby increasing the price of cyberspace users. This research work focused on the spam mail classification technique using the Deep Belief Network classifier, tuned perfectly using the Earthworm Algorithm. The dataset is first pre-processed based on which the keywords are identified and followed by feature extraction. This is then followed by feature selection. The selection of features is performed using the Bhattacharya distance. The features retrieved using the Bhattacharya distance are then subjected to spam mail classification in which Deep Belief Network is used which identifies the spam mails. Fig 1. shows the proposed plan for spam mail detection.

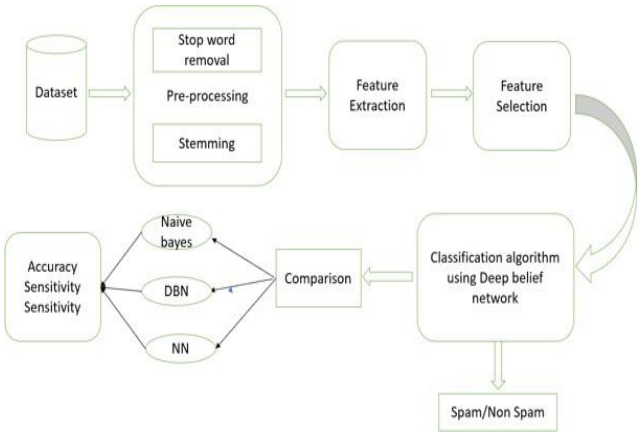


Fig. 1. Proposed plan of Email Phishing Detection

### A. Pre-processing

Pre-processing is the first step in the identification of phishing attacks. [14] This phase involves 2 processes which include the elimination of stop words and stemming. The dataset for the email is chosen from UCI and Enron and the mail has words as a sentence or a paragraph. The stop words mainly a, an, in, and so on searched are eliminated from the mail. This is followed by stemming in which certain words in the mail document are changed to their root word. The output of the pre-processing step is known as dictionary words. This in turn acts as input to the feature extraction.

### B. Feature extraction

The dictionary words are then put through the feature extraction using Term Frequency which identifies the frequency of the dictionary words used in the particular mail. TF is an arithmetic method of retrieving the significant word from a dataset. Term frequency is an efficient algorithm to extract the frequency of terms from dictionary words and also in the method of assigning word weights. Therefore Term Frequency expresses the total number of times an individual word appears in an email.

### C. Feature selection

Feature selection is the method of selecting prominent features from the identified dictionary words. The Bhattacharya distance is computed between the individual feature

and the class and the feature with the maximal Bhattacharya distance is selected as the effective features for the classification using the DBN. The Bhattacharya distance is calculated based on,

$$BD(g_k, C_l) = \frac{1}{4}\left[\ln\left[\frac{1}{4}\left[\frac{\sigma_k^2}{\sigma_l^2} + \frac{\sigma_l^2}{\sigma_k^2} + 2\right]\right] + \frac{1}{4}\left[\frac{\mu_k - \mu_l}{\sigma_k^2 + \sigma_l^2}\right]\right] \quad (1)$$

where $BD(g_k, C_l)$ refers to the Bhattacharya distance be- tween the $k^{th}$ feature and the $l^{th}$ class. The mean of the $k^{th}$ feature and the $l^{th}$ class is denoted as $\mu_k$ and $\mu_l$ , and variance of the $k^{th}$ feature and the $l^{th}$ class is denoted as $\sigma_k$ and $\sigma_l$.

*D. Classification using Deep Belief Neural Networks*

The features identified from the feature selection are classified using the DBN. The classified data are first given as input to the classifier and then trained using the proposed Earthworm Algorithm which in turn tunes the optimal weights of DBN. This helps in differentiating spam mails from relevant mails.

*1) Deep Belief Neural networks (DBN):*

Deep Belief Network is a generative network and it is implemented by stacking several layers with each middle layer consisting of the visible and hidden neurons [15]. The DBN layers include Restricted Boltzmann Machine (RBM) layers and a Multilayer perceptron (MLP) layer. Each RBM layer in turn consists of its input and hidden layers and the MLP layer comprises the input, hidden, and output layers [16]. The effectiveness of DBN is the interconnection between the hidden and the input neurons that are interlinked by a set of tunable weights. The architecture of the DBN network is shown in Fig 2.

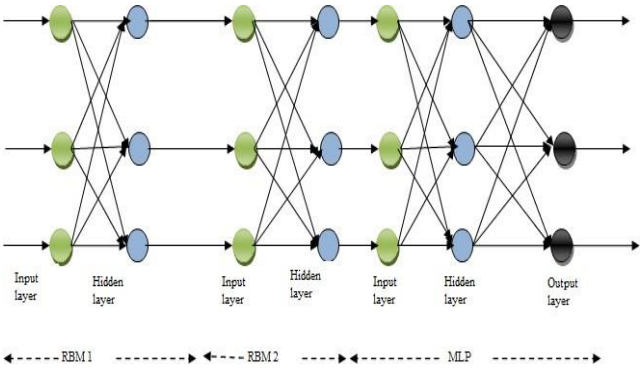


Fig. 2. Proposed plan of Email Phishing Detection

*2) Training phase of DBN:*

The DBN classifier has to be trained to acquire the correct weights and biases that help to reveal the spam mails. This phase points at fine-tuning the RBM and MLP layers, which entirely depends on the optimal weights derived using the proposed EWA algorithm.

The newly enhanced optimization algorithm helps in fine-tuning the optimal weights and biases and therefore ensures a minimal level of error values. The following steps are followed in the training of DBN are:

Step 1: Train the 2 layers RBM1 and RBM2.

Step 2: Train the MLP layer

The first step involves providing an RBM1 layer with the input data and then subjected to a probability distribution. The data is then encoded using weights to compose an output which forms the input to the RBM2 layer.

The process of training the DBN can be further repeated to retrieve the input to the MLP layer.

- Initialization of MLP weights
- Determine the output of the MLP layer
- Determine the error of the network
- Weight calculation in the MLP layer
- Termination

The following steps are repeated for a maximum number of iterations till a globally optimal solution is obtained.

*3) Determination of weights of DBN based on Optimization algorithm:* Earthworm Algorithm is a bio-inspired metaheuristic algorithm based on the reproducing pattern of the earthworms [17]. This can be viewed as 2 types of reproduction and the new obtained solutions are calculated by counting the weights for producing new earthworms. The searching tendency in EWA was enhanced by the use of Cauchy operators. In the reproductive capability of the Earthworm, the type-1 Reproduction produces only 1 offspring and the type-2 Reproduction produces 2 or more offspring.

Type-1Reproduction: In this type of reproduction, the single earthworm is involved in reproduction as earthworms are known as hermaphrodites.

Type-2 Reproduction: This type of reproduction produces two or more two offspring. Crossovers are considered as parents can be changed accordingly to produce the offspring to ensure that offspring produced is not less than zero. The crossover mentioned is single-point, multi-point, and uniform crossover. The parents selected for crossover are based on the strategy named roulette wheel selection.

Case 1: With 2 parents and 1 offspring and it follows a single-point crossover. The multipoint crossover with 2 parents is based on two random numbers generated.

Case 2: With 2 parents and 2 offspring

Case 3: With 3 parents and 3 offspring

The position of the earthworm based on the 2 types of offspring generated can be calculated as,

$$u_{p,q}^{\tau+1} = \alpha.u_{p,q}^{1} + (1-\alpha).u_{p,q}^{2} \quad (2)$$

where, up,q is the qth element of up, which is the position of pth the earthworm and α is the proportional factor.

The Cauchy operator gives the position of the earthworm based on the formula,

$$u_{p,q}^{\tau+1} = u_{p,q}^{\tau} + \omega_q * R \quad (3)$$

where, R indicates the random number obtained by performing the Cauchy distribution and $\omega_q$ denotes the weight assigned for the $q^{th}$ position, and $u_{p,q}^{\tau+1}$ determines the $q^{th}$ position of the $p^{th}$ earthworm at time $\tau$ .

## IV. EXPERIMENTAL SETUP

The proposed algorithm is implemented using JAVA and the datasets utilized for the analysis include Enron and UCI. The effectiveness of the proposed algorithm for spam mail detection is computed based on three metrics, namely accuracy, specificity, and sensitivity. The datasets like Enron and UCI have the original messages which include both ham and spam mails in non-Latin encodings.

### *A. Performance metrics*

The algorithm is analyzed based on the performance metrics, mainly accuracy, specificity, and sensitivity. The accuracy can be determined by calculating the accurate number of spam mails, Specificity is the metric to determine the negatives which are correctly detected and sensitivity determines the positives correctly identified.

$$Accuracy = \frac{Tn + Tp}{Tn + Tp + Fn + Fp} \quad (4)$$

$$Specificity = \frac{Tn}{Tn + Fp} \quad (5)$$

$$Sensitivity = \frac{Tp}{Tp + Fn} \quad (6)$$

where refers to the values as true positive, refers to true negative, denotes the false negative values, and denotes the values as false positive.

### *B. Comparative Analysis*

The proposed algorithm is being compared with the following methods namely Naive Bayes (NB) [18], Deep Belief Networks (DBN), and Neural Networks (NN). The proposed EWA-DBN algorithm is employed for the detection of email phishing and compared with the above methods to determine its effectiveness.

*1) Analysis of Dataset by varying the data percentage :* The figure below denotes the comparative analysis based on the data chosen for training. Fig 3 denotes the comparative analysis based on the accuracy of the algorithm chosen.When the percentage of the data is 50, the accuracy of the methods, NB, DBN, NN and EWA-DBN is 0.5333, 0.5455, 0.5556 and 0.5714, respectively. Fig 4 denotes the comparative analysis on the sensitivity of the algorithm chosen. When the percentage of the data is 50, the sensitivity of the methods, NB, DBN, NN and EWA-DBN, is 0.4558, 0.5531, 0.7035 and 0.7223 respectively. Fig 5 denotes the comparison based on the specificity of the algorithm chosen. When the data percentage is 50, the specificity of the methods, NB, DBN, NN and EWA-DBN, is 0.5052, 0.5631, 0.7028 and 0.7104, respectively.

### *C. Comparative discussion*

Table 1 shows the comparative results based on the various methods chosen depending on the performance metrics from the dataset ENRON [19] and UCI [20]. The accuracy value of the methods, NB, DBN, NN and EWA-DBN is 0.5233, 0.5465, 0.5568 and 0.6714. The sensitivity range of the methods, NB, DBN, NN and EWA-DBN is 0.4978, 0.5642, 0.7235 and 0.8145 respectively. Similarly, the specificity value of the methods, NB, DBN, NN and EWA-DBN is 0.5152, 0.5845, 0.7238 and 0.8040 respectively. It is evident from the comparison that the proposed new algorithm has accomplished the maximum value with regards to accuracy, sensitivity, and specificity in comparison with the already existing methods.

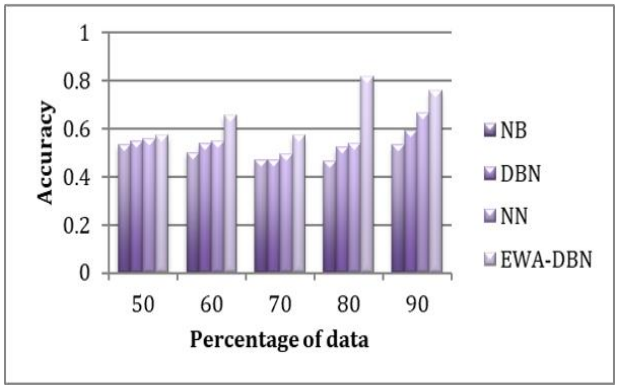


Fig. 3. Comparative analysis based on the training percentage using dataset accuracy

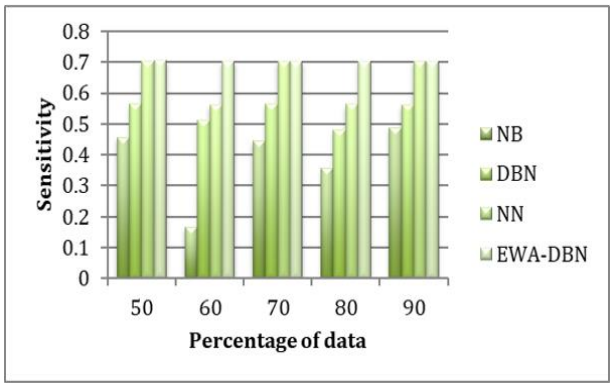


Fig. 4. Comparative analysis based on the training percentage using dataset Sensitivity

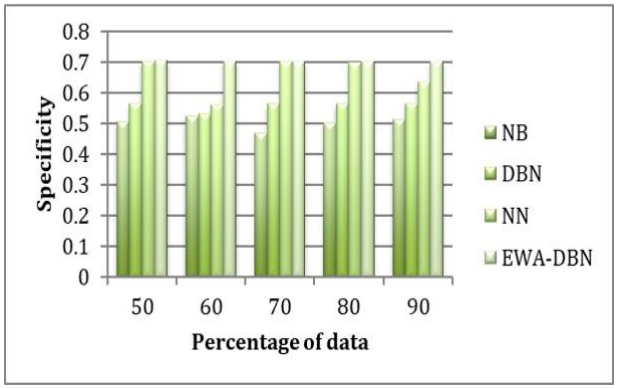


Fig. 5. Comparative analysis based on the training percentage using dataset Specificity

TABLE I
COMPARATIVE DISCUSSION

| Metrics | NN | DBN | NN | EWA-DBN |
|---|---|---|---|---|
| Accuracy | 0.5233 | 0.5465 | 0.5568 | 0.6714 |
| Sensitivity | 0.4978 | 0.5642 | 0.7235 | 0.8145 |
| Specificity | 0.5152 | 0.5845 | 0.7238 | 0.8040 |

## V. CONCLUSION

The email phishing has created a havoc among the internet users. The phishing detection is carried out using the optimization-based deep learning networks. The mail received are first pre-processed to furnish only the selected words to the next step namely feature extraction. The extracted features are then provided to feature selection using the method of Bhattacharya distance. This is in turn fed to the classification algorithm based on the deep belief neural networks. The classifier after fine tuning based on the proposed EWA aims at detecting the spam mails effectively. The comparison is performed using the datasets, UCI and Enron, which is analyzed based on the performance metrics, such as accuracy, sensitivity, and specificity, which is 0.671, 0.814, and 0.804, respectively. The research can be further extended by performing hybrid optimizations so as to enhance the phishing detection ratio.

## REFERENCES

[1] J. Ma, Y. Zhang, Z. Wang, and B. Chen, "A new fine-grain sms corpus and its corresponding classifier using probabilistic topic model.," *TIIS*, vol. 12, no. 2, pp. 604–625, 2018.
[2] P. Patil, R. Rane, and M. Bhalekar, "Detecting spam and phishing mails using svm and obfuscation url detection algorithm," in *2017 International Conference on Inventive Systems and Control (ICISC)*, pp. 1–4, IEEE, 2017.
[3] S. J. Delany, M. Buckley, and D. Greene, "Sms spam filtering: Methods and data," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9899–9908, 2012.
[4] L. Zhang, J. Zhu, and T. Yao, "An evaluation of statistical spam filtering techniques," *ACM Transactions on Asian Language Information Processing (TALIP)*, vol. 3, no. 4, pp. 243–269, 2004.
[5] G. Dalkılıc¸ and D. Sipahi, "Spam filtering with sender authentication network," *Computer Communications*, vol. 98, pp. 72–79, 2017.
[6] B. Zhou, Y. Yao, and J. Luo, "Cost-sensitive three-way email spam filtering," *Journal of Intelligent Information Systems*, vol. 42, no. 1, pp. 19–45, 2014.
[7] G. V. Cormack, "Email spam filtering: A systematic review," 2008.
[8] C. Laorden, X. Ugarte-Pedrero, I. Santos, B. Sanz, J. Nieves, and P. G. Bringas, "Study on the effectiveness of anomaly detection for spam filtering," *Information Sciences*, vol. 277, pp. 421–444, 2014.
[9] S. Smadi, N. Aslam, and L. Zhang, "Detection of online phishing email using dynamic evolving neural network based on reinforcement learning," *Decision Support Systems*, vol. 107, pp. 88–102, 2018.
[10] A. Barushka and P. Hajek, "Spam filtering using integrated distribution-based balancing approach and regularized deep neural networks," *Applied Intelligence*, vol. 48, no. 10, pp. 3538–3556, 2018.
[11] S. S. Kovalluri, A. Ashok, and H. Singanamala, "Lstm based self-defending ai chatbot providing anti-phishing," in *Proceedings of the first workshop on radical and experiential security*, pp. 49–56, 2018.
[12] D. Ruano-Ordás, F. Fdez-Riverola, and J. R. Méndez, "Using evolutionary computation for discovering spam patterns from e-mail samples," *Information Processing & Management*, vol. 54, no. 2, pp. 303–317, 2018.
[13] G. Sonowal and K. Kuppusamy, "Smidca: an anti-smishing model with machine learning approach," *The Computer Journal*, vol. 61, no. 8, pp. 1143–1157, 2018.
[14] R. M. Silva, T. C. Alberto, T. A. Almeida, and A. Yamakami, "Towards filtering undesired short text messages using an online learning approach with semantic indexing," *Expert Systems with Applications*, vol. 83, pp. 314–325, 2017.
[15] E. Benavides, W. Fuertes, S. Sanchez, and M. Sanchez, "Classification of phishing attack solutions by employing deep learning techniques: A systematic literature review," *Developments and advances in defense and security*, pp. 51–64, 2020.
[16] G. Tzortzis and A. Likas, "Deep belief networks for spam filtering," in *19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, vol. 2, pp. 306–309, IEEE, 2007.
[17] G.-G. Wang, S. Deb, and L. D. S. Coelho, "Earthworm optimisation algorithm: a bio-inspired metaheuristic algorithm for global optimisation problems," *International Journal of Bio-Inspired Computation*, vol. 12, no. 1, pp. 1–22, 2020.
[18] I. Androutsopoulos, J. Koutsias, K. V. Chandrinos, and C. D. Spyropoulos, "An experimental comparison of naive bayesian and keyword-based anti-spam filtering with personal e-mail messages," in *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 160–167, 2000.
[19] *http://nlp.cs.aueb.gr/software*$_a$*nd*$_d$*atasets/Enron – Spam/index.html,* 2021(*accessedMarch*202021).
[20] *https://archive.ics.uci.edu/ml/machine-learning-databases/00228*, 2021 (accessed March 20 2021).