Ingénierie des Systèmes d'Information
Vol. 25, No. 6, December, 2020, pp. 747-753
Journal homepage: http://iieta.org/journals/isi

Classification of Social Media Text Spam Using VAE-CNN and LSTM Model

Ayyappa Chakravarthi Metlapalli[1*], Thillaikarasi Muthusamy[1], Bhanu Prakash Battula[2]

[1] Department of CSE, Annamalai University, Chidambaram 608002, Tamil Nadu, India
[2] GVR&S College of Engineering and Technology, Guntur 522013, Andhra Pradesh, India

Corresponding Author Email: me.chakravarty@gmail.com

https://doi.org/10.18280/isi.250605

Received: 30 July 2020
Accepted: 12 November 2020

Keywords:
spam data, convolution neural network (CNN), (long short term memory networks) LSTM, deep learning, variational auto-encoder (VAE)

ABSTRACT

Presently a day's human relations are kept up by online life systems. Customary connections now days are outdated. To keep up in affiliation, sharing thoughts, trade information between we utilize web-based social networking organizing locales. Web based life organizing locales like Twitter, Facebook, LinkedIn and so forth are accessible in the correspondence condition. Through Twitter media clients share their sentiments, interests, information to others by messages. Simultaneously a portion of the client's mislead the certifiable clients. These certified clients are additionally called requested clients and the clients what misguidance's identity is called spammers. These spammers present undesirable data on the non-spam clients. The non-spammers may retweet them to other people and they follow the spammers. Generally most of the spam messages are in the form of text, images and different multimedia formats. Considering all different formats in one process may not give the best classification results. In this paper address the process and classification of text spam messages. Classification of text messages is a complex task in order to achieve this deep learning based hybrid VAE-CNN and LSTM model is proposed and evaluated the model using the performance metrics of precision, recall and F measure metrics.

1. INTRODUCTION

Spam is an undesirable correspondence expected to be conveyed to an aimless objective, legitimately or in a roundabout way, despite measures to forestall its conveyance. Spam channel is a computerized strategy to distinguish spam for the reason for forestalling its conveyance. The motivation behind spam is that the receiver is supplied with data that comprises payloads such as (for example, unprofitable, unconstitutional or non-existent advertisements for an item), ransom machine traps, advancement without a purpose, computer malware without the recipient's device being captured. As it is so modest to submit results, just a limited number of the recipients - maybe one in 10 thousand or less - have to be willing to respond to the payload such that spam is useful for their sender [1]. The fundamental qualities of spam are undesirable, unpredictable, guileful, payload bearing. Undesirable spam implies spam messages are definitely not needed by lion's share of individuals. Aimless spam implies Spam is transmitted outside of any sensible relationship or forthcoming connection between the senders what's more, the beneficiary. By and large, it is more practical for the spammer to send more spam than to be specific with respect to its target [1]. Pretentious spam infers considering the way that spam is unwanted and flighty, it must cover itself to redesign the open door that its payload will be passed on and followed up on [2]. The content of a spam message may be transparent or hidden; spam decreases can be enhanced in any case by understanding the content and the background from which spammers gain. Collectively, payloads are a repository of condensed identities, ideologies, network links, phone numbers, etc. They can either be transparent or be blurred in order to make it obvious that the human being always is aware of the Screen. Alternatively, they could be jumbling on the other side to appear like they take the human being into account and cause some damaging PC behavior. The load may include a boring term or joint such as "gouranga" or "Platypus Race," in the name of intrigue, a web look and a cumulative charge for the website of the spammer, or of a paying advance. A second type of roundabout transportation of payload is backscatter: spam message is sent by certified mail server to a non-existing customer with the return address (produced) of the true customer [3]. Email is a powerful, quick and modest correspondence way. Hence spammers want to send spam through such kind of correspondence. These days pretty much consistently client has an E-mail, and therefore they are confronted with spam issue. Email Spam is non-mentioned data sent to the E-post boxes. Spam is a major issue both for clients and for ISPs. The causes are development of estimation of electronic correspondences from one perspective and improvement of spam sending innovation then again. By spam reports of Symantec in 2013, the normal worldwide spam rate for the year was 89.1%, with an expansion of 1.4% contrasted and 2012. In social media spam data is of different types.

Labelling different types of text documents is both important and desirable. There are plenty of different situations where this is useful. Automating these tasks is therefore something of great value if the automated system performs on a par with, or better than, humans. Labelling essays with grades is an example of a task that is time consuming but also important, which is why a lot of research has been put into Automated Essay Scoring. Removing posts

from social media platforms which are against the terms of use or illegal (e.g., hate speech or threats of physical violence) is another case where automation, if done right, would be beneficial. Automated labelling of texts can also be useful in other cases. Classifying e-mails as spam, classifying reviews as either positive or negative and assigning topics to Wikipedia articles [4] are more examples of useful applications. When assigning topics to Wikipedia articles, the number of target classes is larger. The literature makes a distinction between the case when the target classes are binary (e.g., "Spam"/"Not Spam", "Positive"/"Negative") and when there are several possible target classes (e.g., different topics), where the latter problem is more complex. A different text classification task is when a document can be labelled with several labels. This is referred to as a multi-label classification task. If the label-space is very big, the task becomes an extreme multi-label classification task (XMTC) [5-7]. In this paper we are concentrating mainly on social media text data, for that we use the CNN and LSTM and auto encoder mechanism. The rest of the paper is organized as follows section-2 describes the literature, section-3 details the proposed work, section-4 and 5 describes experimental data and comparative analysis and finally section-6 concluded the paper.

2. LITERATURE SURVEY

The importance, appropriateness and ongoing creation of deep AI models in the detection of malware, disruption and spam has been evaluated by Apruzzese et al. [4]. The authors claimed that the application of UBE identification may be improved by utilizing different AI classifiers to classify specific undertakings; in all situations they rendered no extraordinary decision on deep neural models.

Basnet and Sung [8] suggested a technique to distinguish phishing communications with weighted confidence lead classificators [9]. The developers have always used the email content as highlights and have avoided the usage of clear highlights in heuristic phishing.

Bergholz et al. [10] published a groundbreaking work in the area of phishing email separation in which developers replicated a range of innovative elements, including realistic templates for email subject photos, email [11] text and external link review and study of mounted logos [12] on concealed salt.

Bhowmick and Hazarika [13] presented a broad illustration of a portion of the cutting-edge, UBE-based strategies. Their research explored several important ideas in UBE, which differentiate between the feasibility of existing ventures, and the ongoing trends in UBE organization [14, 15], based on popular AI approaches for the position of a text. Additionally, they explored the evolving concept of UBE attacks [16, 17] and examined a variety of AI [18, 19] estimates to counter such signals.

Dhanaraj et al. [20-25] have discussed and developed proposed solutions for moderate email spam. About the inventiveness of pictorial techniques, the research on the AI models or the corpus [26] used did not explain them.

Fette et al. [27] utilized a lot of orthographic highlights to accomplish a programmed bunching of phishing messages, which brought about more prominent proficiency and better execution by means of IG [28-30] with C4.5 [31, 32]. They utilized the adjusted worldwide K-implies way to deal with create the target work esteems, for chose highlight subsets, which aided acknowledgment of groups.

Gansterer and Pölz [33] proposed an arrangement of separating the approaching messages into ham, spam, and phishing, in light of FSC [34], which gave better (98%) grouping precision [35] (ternary order) than that came about because of the utilization of two parallel classifiers.

Lueg [36] introduced a short study investigating the method of applying data recovery and data separating instruments to hypothesize spam wavering in a hypothetically grounded and legitimate way. In spite of the fact that the creator planned for presenting an operationally productive spam finder, the introduced study didn't detail the re-enactment instruments, AI draws near, or the datasets used.

Wang and Cloete [37] looked into a few methodologies of 2 Note that PCA encourages include extraction as opposed to highlight determination. Immaterialness of AI in spam and phishing email… identifying spam messages, ordered spontaneous spam messages into various levelled envelopes, and encouraged programmed guideline of the errands concerning the reaction to an email. In any case, the creator didn't cover any AI draws near.

The key work in UBe recognition and agreements is provided by Chandrasekaran et al. [38] and their research identified and used subordinate email highlights such as the lavishness of the content and the amount of helpful terms (e.g. bank, loan, and payment) in order to segregate the phishing from real communications. They used SVM to differentiate between phishing messages and avoid them from entering the package of the client, thus raising any potential implementation.

Throughout both situations, those critical parts of the spam networks required to be explained. In Sanz et al. [39] concerns relevant to UBE studies, their impacts on consumers and the approaches to minimize these impacts are answered. Their exploration work explained on a few AI calculations used in UBE recognition. In any case, their work did not have a relative investigation of different substance channels.

Cormack [40] utilized a helped outfit style which depended on C5 DT, and occasion based learning group methods to rename messages that were delegated non-phishing by C5 DT [41]. They acquired a decent exactness using C5 DT [42] and sent percent review from the group.

The clear use of URLs and the quality of the web highlights offered a structured position solution to Chinese e-business pages [43]. Four classifiers like RF, SMO, strategic replication and NB have been used by the designers to test their results using Chi-squared measurements.

The importance of disclosure of anomalies in UBE splitting from the precondition of organization of UBEs was explained by Laorden et al. [44]. Our research explores a UBE sieving method focused on irregularity, which utilizes a knowledge reducing strategy that reduces pre-treatment when taking care of T concurrently. Information on account qualifications relevant to account existence from Gangavarapu et al. [30]. More recently other experiments have been designed to explore, derived from the ability of these forms to cope with research, adaptation and summarizing, the application of the numerous AI methods including KNN, SVM, NB, neural networks and others to spam and phishing communications.

Sah and Parmar [45] suggested a model for the efficient detection of spam malignant messages by effectively identifying the variable, monitored by three AI methods namely NB, SVM and MLP. A portion of research has used comprehensive learning frameworks to classify UBEs with the successful achievement of deep neural systems in various

applications.

Hassanpour et al. [46] displayed the email material as highlights in the Word2-Vec format utilizing certain deep learning tools, and developers have achieved 97 percent of general accuracy.

Vorobeychik and Kantarcioglu [47] have used unwilling AI to build email assessments and prepared a classification framework to identify the assessments they have been created.

3. PROPOSED WORK

The proposed mechanism mainly concentrates on text data classification. Here the input data is spam data it is a binary classification problem weather a message is spam or not. Here CNN and Encoder for text processing but to maintain relation among the words LSTM will came into place. The detailed proposed work is discussed below.

Word embedding is a collection of methods in Natural Language Processing for making vector representations of words where the idea is to map words into a vector space where similar words get grouped together. One of the simplest methods for word embedding is the one-hot embedding scheme wherein each word is represented with a vector of the same length as the total number of unique words in the corpus. The vector is then filled with zeros except for one position which corresponds to the position of the word in an ordered list of all unique words. The zero at this position is changed to one, hence the name "one-hot". This results in a very sparse vector with possibly thousands of zeros for a decent size corpus.

A CNN for the most part has the accompanying structure: an info layer, a convolutional layer, a pooling layer, a completely associated (likewise called thick) layer lastly a yield layer. The advantages of CNNs are the spatial invariance and programmed highlight age that originates from the utilization of scholarly channels for the convolution step and the utilization of pooling. This implies CNNs naturally creates includes in the learning procedure and that these highlights are spatially invariant, for instance a straight line in an image will be perceived regardless of whether it is moved. For a similar explanation CNNs are best utilized for input information that makes them request, for example, pixels in an image or words in a sentence. Information which can be rearranged without losing any data will not benefit from the convolution and pooling steps. The convolution layer consists of filters that will be convoluted with the input to generate feature maps. A filter is a matrix with predetermined dimensions and filled with numbers that are initialized randomly and later learned through the training. During the convolution the filter is swept with a given stride length over the input and for each step a convolution is made which results in a number that is put into the feature map. In the pooling layer a pooling window with a predetermined size and stride length is swept over the feature map. For each step the numbers inside the pooling window are condensed into a smaller set of numbers depending on the pooling strategy. The most common strategy is max-pooling where only the largest number in the pooling window is kept. The pooling step concentrates the information in the feature map and makes it less dependent on the position of the features in the original input.

Here Figure 1 shows the proposed model architecture. Initially it takes input social media data, captured by word2ec model it converts the text into vectors. Vectors of data is supplied to encoded CNN model, CNN model extracts features and give the processed features to LSTM. It takes the processed features and make classification of data.

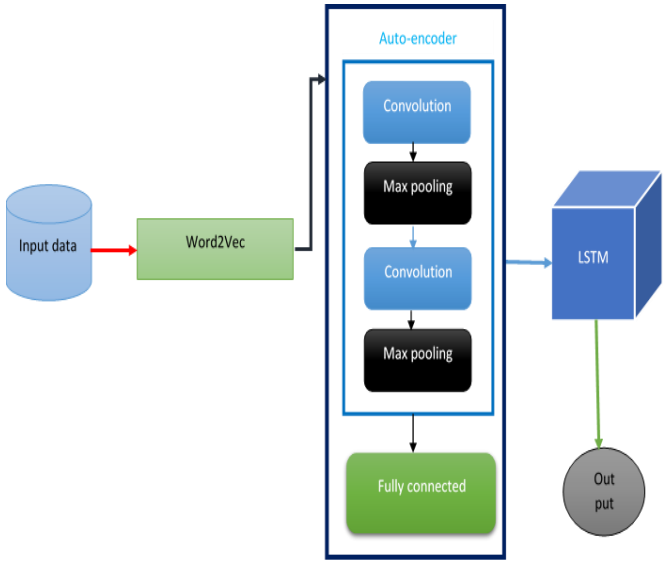

Figure 1. Proposed mechanism

Algorithm:

Input: Text messages
Output: Label: Spam or Ham
For each text message p ∈ {1, ..., n}
do
//text message is converted to word vectors of numerical form
if Word2Vec(pn) not equal to NULL then
yn = Word2Vec(Pn)
else if WordNet (Pn) not equal to NULL NULL then
pn = WordNet (Pn)
yn = Word2Vec(Pn)
else if ConceptNet (Pn) not equal to NULL NULL then
pn = ConceptNet (Pn)
yn = Word2V ec(Pn)
else
yn = Random(Pn)
end if
end for
for each channel f ∈ {1, ..., F} do/Get the most significant highlights
c = CNN (x)
end for
for each time step t ∈ {1, ..., T} do
o = LST M(c)
end for
for each sentence portrayal step o ∈ {1, ..., N} do/Get the last sentence name
name = sigmoid(o)
end for

The algorithm working is as follows, initially the input data is social media text data. And it output the spam or harm data contain by the social media data. The algorithm is as follows, convolution and pooling steps can be repeated several times but finally you will end up with a set of feature maps which are then flattened into a single vector. This vector is given as input to the dense layer, which is an ordinary Neural Network that consists of one to many hidden layers. Each number in the

input vector is passed to every neuron in the hidden layer. Every connection has an associated weight which is multiplied to the number. When all numbers have been passed to a neuron they are summed together with a bias added. They are then passed through an activation function, most commonly ReLu (rectified linear units), to capture any nonlinear relations. Here an optional dropout step can be added where every neuron has a certain chance to drop its value, where the idea is that this might help against overfitting to the training data. Every neuron in the hidden layer is finally connected to every neuron in the output layer where the final result can be determined by seeing which neuron is most activated by the input. Training of the CNN is done by comparing the output of the model to the label of the input. A loss function determines the discrepancy between the output and the label and a loss minimization method, often different variations of backpropagation, adjusts the weights, biases and filters to minimize this loss.

4. DATASET DEPICTION

The Keras 2.0 API and Tensor Stream backend utilizing Python 3 for Ubuntu 16.4.2 research platform was used to incorporate the CNN and LSTM process. We use two separate datasets: SMS spam and Twitter dataset to test the efficiency of the proposed model. At first, the data collection of SMS Spam that can be viewed via the ML UCI Api. The data collection includes 5,575 English and uncoded SMS instant messages that blend 4,826 ham and 749 spam messages previously identified with a pre-defined package. Table 1 displays the class SMS shrewd distribution. Table 2 demonstrates planning and distribution of evaluation occasions as seen by class markings.

We have used twitter messages scratched out of usable live tweets from Twitter data sets. Two notors who disregards the message, whether there is some disagreement amongst them, physically label such tweets as ham or spam. The party exploitation is displayed in Table 1 while the astute case dispersion planning and evaluation grouping is shown in Table 2.

Table 1. UCI spam messages data

S.No	Class label	Training	Testing
1	Spam	672	173
2	Harm	3406	845

We utilized Twitter's Streaming API to gather tweets with URLs. The open Streaming APIs can get 1 % of all the open tweets coursing through Twitter. While it is conceivable to utilize Twitter to send spam and different messages without utilizing URLs, most of spam and different malevolent messages on the Twitter stage contain URLs [11]. In the large number of spam tweets which were physically assessed, we discovered just a bunch of tweets without URLs which could be considered as spam. Furthermore, spammers for the most part utilize implanted URLs to make it increasingly helpful to guide casualties to their outside destinations to accomplish their objectives, for example, phishing, tricks, and malware downloading [16]. In this way, we confined this exploration to the tweets with URLs. During the assortment time frame, we gathered an aggregate of more than 4K tweets with URLs.

Table 2. Tweets text data

S.No	Class label	Training	Testing
1	Spam	722	88
2	Harm	300	60
3	normal	600	240

5. EXPERIMENTAL RESULTS

Comparison of experimental results of proposed model with the NB model, SVM model with RBF kernel and Random forest with the maximum number of features Random Forest is allowed to try in individual tree.

The aftereffects of the Figure 2 plainly exhibit that for SMS spam information characterization dependent on VAE-CNN and LSTM model gives the best outcomes as True Positive Rate and Precision are most elevated for it though False Positive Rate is least. While the aftereffects of relative chart affirm that NB and SVM gives the most noteworthy True Positive rate anyway their False Positive Rate was a lot bigger. Then again Random backwoods gives True positive rate as the NB and SVM however False Positive Rate for Random woodland is least among all the four methods so we can arrive at the resolution that for connected based highlights, proposed order is the best by demonstrating the TP rate, FP Rate and the accuracy esteem.

The aftereffects of the Figure 3 show that for Twitter spam information grouping dependent on VAE-CNN and LSTM model gives the best outcomes as True Positive Rate and Precision are most elevated for it while False Positive Rate is least. While the consequences of relative diagram affirm that NB and SVM gives the most noteworthy True Positive rate anyway their False Positive Rate was a lot bigger. Then again Random timberland gives True positive rate as the NB and SVM however False Positive Rate for Random woodland is least among all the four methods so we can arrive at the resolution that for connected based highlights, proposed grouping is the best by indicating the TP rate, FP Rate and the accuracy esteem.

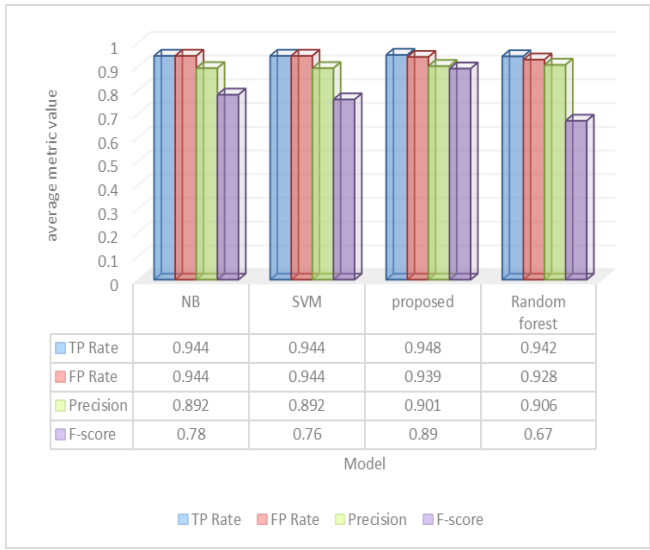

Figure 2. SMS spam data classification

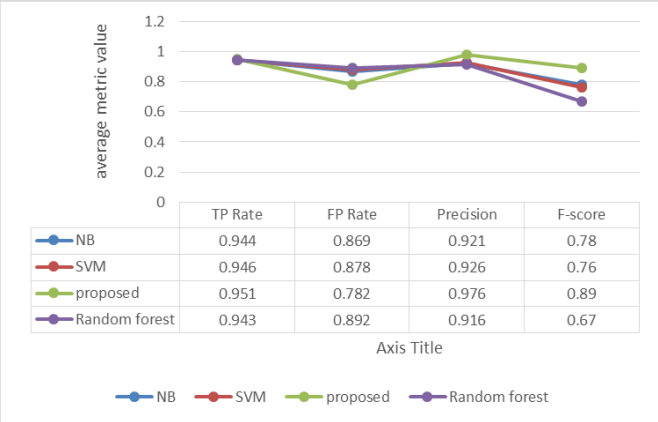

Figure 3. Twitter data classification

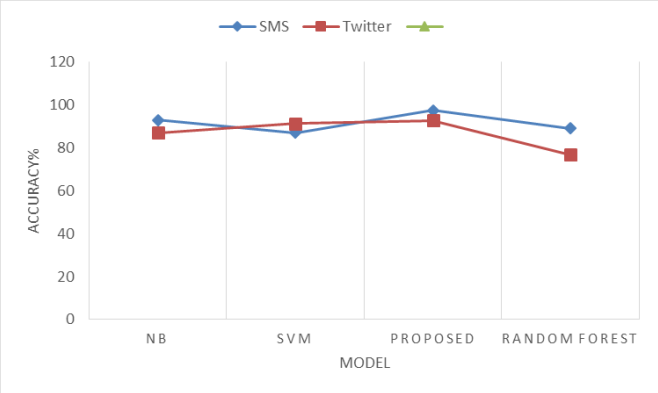

Figure 4. Accuracy of spam data classification

Here Figure 4 shows the precision correlation of proposed model just as various existing models of SVM, NB and Random backwoods. The correlation between the models depends on two informational indexes of SMS spam information and Twitter information. The proposed VAE-CNN and LSTM based model gives preferred exactness over the current models in light of the fact that the engineering of the model. CNN can take text information positively and ready to process in plainly and LSTM can make connection of the words however the current works flops in that angle.

6. CONCLUSIONS

Social networking abuse is a threat as it causes people a lot of frustration and can wasting money. Due to short text, repetitive words, and more noisy results, the detection of email spam in social media is challenging. Throughout this post, we concentrated on spam detection utilizing a deep learning approach to online networking. Although spam discovery in short loud content on Twitter is highly testing due to the lack of consistence and lack of coherence in languages that are being used through the networking media, existing effective methodologies are primarily challenged by long email messages. We suggested a thorough learning approach comprising neural structures VAE-CNN and LSTM. Explorative findings indicate that the suggested solution is focused on two databases such as SMS Dataset and Twitter Dataset and through separate methodologies.

REFERENCES

[1] Abu-Nimeh, S., Nappa, D., Wang, X., Nair, S. (2007). A comparison of machine learning techniques for phishing detection. In: Proceedings of the Anti-phishing Working Groups 2nd Annual eCrime Researchers Summit, ACM, pp. 60-69. https://doi.org/10.1145/1299015.1299021

[2] Akinyelu, A.A., Adewumi, A.O. (2014). Classifcation of phishing email using random forest machine learning technique. Journal of Applied Mathematics, 2014: 45731. https://doi.org/10.1155/2014/425731

[3] Alkaht I.J., Al-Khatib, B. (2016) Filtering spam using several stages neural networks. International Review on Computers and Software, 11(2). https://doi.org/10.15866/irecos.v11i2.8269

[4] Apruzzese, G., Colajanni, M., Ferretti, L., Guido, A., Marchetti, M. (2018). On the efectiveness of machine and deep learning for cyber security. In: 2018 10th International Conference on Cyber Confict (CyCon), pp. 371-390. https://doi.org/10.23919/CYCON.2018.8405026

[5] Gopi, A.P., Suresh Babu, E., Naga Raju, C., Ashok Kumar, S. (2015). Designing an adversarial model against reactive and proactive routing protocols in MANETS: A comparative performance study. International Journal of Electrical & Computer Engineering, 5(5). https://doi.org/10.11591/ijece.v5i5.pp1111-1118

[6] Kumar, S.A., Suresh Babu, E., Nagaraju, C., Peda Gopi, A. (2015). An Empirical critique of on-demand routing protocols against rushing attack in MANET. International Journal of Electrical and Computer Engineering, 5(5). https://doi.org/10.11591/ijece.v5i5.pp1102-1110

[7] Awad, M., Foqaha, M. (2016). Email spam classifcation using hybrid approach of rbf neural network and particle swarm optimization. International Journal of Network Security & Its Application, 8(4): 17-28. https://doi.org/10.5121/ijnsa.2016.8402

[8] Basnet, R.B., Sung, A.H. (2010). Classifying phishing emails using confdence-weighted linear classifers. In: International Conference on Information Security and Artifcial Intelligence (ISAI), pp. 108-112.

[9] Bec Scams Trends and Themes. (2019). Bec scams remain a billion-dollar enterprise, targeting 6k businesses monthly. https://www.symantec.com/blogs/threat-intelligence/bec-scams-trends-and-themes-2019, accessed on May 7, 2019.

[10] Bergholz, A., De Beer, J., Glahn, S., Moens, M.F., Paaß, G., Strobel, S. (2010). New fltering approaches for phishing email. Journal of Computer Security, 18(1): 7-35. https://doi.org/10.3233/JCS-2010-0371

[11] Bhagyashri, G., Pratap, H., Patil, D. (2013). Auto e-mails classifcation using bayesian flter. International Journal of Advanced Technology & Engineering Research, 3(4): 19-24.

[12] Yang, H.H., Moody, J. (2000). Data visualization and feature selection: New algorithms for nongaussian data. In: Advances in Neural Information Processing Systems, pp. 687-693.

[13] Bhowmick, A., Hazarika, S.M. (2016). Machine learning for e-mail spam fltering: Review, techniques and trends. arXiv preprint arXiv:1606.01042.

[14] Biggio, B., Corona, I., Fumera, G., Giacinto, G., Roli, F. (2011). Bagging classifers for fghting poisoning attacks in adversarial classifcation tasks. In: International Workshop on Multiple Classifer Systems, pp. 350-359. https://doi.org/10.1007/978-3-642-21557-5_37
[15] Bolboaca, S.D., Jäntschi, L. (2006). Pearson versus spearman, kendall tau correlation analysis on structure-activity relationships of biologic active compounds. Leonardo Journal of Science, 5(9): 179-200.
[16] Breiman, L. (2002). Manual on setting up, using, and understanding random forests v3. 1. Statistics Department University of California, Berkeley.
[17] Breiman, L. (2001). Random forests. Machine Learning, 45(1): 5-32. https://doi.org/10.1023/A:1010933404324
[18] Breiman, L. (2017). Classifcation and Regression Trees. Routledge, Abingdon. https://doi.org/10.1201/9781315139470
[19] Chandrasekaran, M., Narayanan, K., Upadhyaya, S. (2006). Phishing email detection based on structural properties. In: NYS Cyber Security Conference, Albany, New York, pp. 1-8.
[20] Chanduka, B., Gangavarapu, T., Jaidhar, C.D. (2018). A single program multiple data algorithm for feature selection. In: Abraham A, Cherukuri AK, Melin P, Gandhi N (eds), Intelligent Systems Design and Applications, Springer, Cham, 662-672. https://doi.org/10.1007/978-3-030-16657-1_62
[21] Choudhary, M., Dhaka, V. (2013). Automatic e-mails classifcation using genetic algorithm. In: Special Conference Issue: National Conference on Cloud Computing and Big Data, pp. 42-49.
[22] Christina, V., Karpagavalli, S., Suganya, G. (2010). Email spam fltering using supervised machine learning techniques. International Journal on Computer Science and Engineering, 2: 3126-3129.
[23] Cormack, G.V. (2008). Email spam fltering: A systematic review. Foundations and Trends® in Information Retrieval, 1(4): 335-455. https://doi.org/10.1561/1500000006
[24] Dhanaraj, S., Karthikeyani, V. (2013). A study on e-mail image spam fltering techniques. In: 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, Salem, India, pp. 49-55. https://doi.org/10.1109/ICPRIME.2013.6496446
[25] Dhanaraj, K.R., Palaniswami, V. (2014). Firefy and bayes classifer for email spam classifcation in a distributed environment. Aust J Basic Appl Sci., 8(17): 118-130.
[26] Díaz-Uriarte, R., De Andres, S.A. (2006). Gene selection and classifcation of microarray data using random forest. BMC Bioinform, 7(1): 3. https://doi.org/10.1186/1471-2105-7-3
[27] Fette, I., Sadeh, N., Tomasic, A. (2007). Learning to detect phishing emails. In: Proceedings of the 16th International Conference on World Wide Web, ACM, pp. 649-656. https://doi.org/10.1145/1242572.1242660
[28] Gang, S. (2017). Email overload: research and statistics [with infographic]. https://blog.sanebox.com/2016/02/18/email-overload-research-statistics-sanebox/
[29] Gangavarapu, T., Patil, N. (2019). A novel flter-wrapper hybrid greedy ensemble approach optimized using the genetic algorithm to reduce the dimensionality of high-dimensional biomedical datasets. Applied Soft Computing, 81: 105538. https://doi.org/10.1016/j.asoc.2019.105538
[30] Gangavarapu, T., Jayasimha, A., Krishnan, G.S., Kamath, S.S. (2019). TAGS: Towards automated classifcation of unstructured clinical nursing notes. In: Métais E, Meziane F, Vadera S, Sugumaran V, Saraee M (eds) Natural Language Processing and Information Systems. Springer, Cham, pp. 195-207. https://doi.org/10.1007/978-3-030-23281-8_16
[31] Gangavarapu, T., Jayasimha, A., Krishnan, G.S., Kamath, S. (2019). Predicting ICD-9 code groups with fuzzy similarity based supervised multi-label classifcation of unstructured clinical nursing notes. Knowledge-Based Systems, 190: 105321. https://doi.org/10.1016/j.knosys.2019.105321
[32] Gangavarapu, T., Krishnan, G.S., Kamath, S. (2019). Coherence-based modeling of clinical concepts inferred from heterogeneous clinical notes for ICU patient risk stratifcation. In: Proceedings of the 23rd conference on Computational Natural Language Learning (CoNLL), pp. 1012-1022. https://doi.org/10.18653/v1/K19-1095
[33] Gansterer, W.N., Pölz, D. (2009). E-mail classifcation for phishing defense. In: European Conference on Information Retrieval, Springer, pp. 449-460. https://doi.org/10.1007/978-3-642-00958-7_40
[34] Geurts, P., Ernst, D., Wehenkel, L. (2006). Extremely randomized trees. Machine Learning, 63(1): 3-42. https://doi.org/10.1007/s10994-006-6226-1
[35] Guerra, P.H.C., Guedes, D., Meira, J.W., Hoepers, C., Chaves, M., Steding-Jessen, K. (2010). Exploring the spam arms race to characterize spam evolution. In: Proceedings of the 7th Collaboration, Electronic Messaging, Anti-abuse and Spam Conference (CEAS), Redmond.
[36] Lueg, C.P. (2005). From spam filtering to information retrieval and back: Seeking conceptual foundations for spam filtering. Proceedings of the American Society for Information Science and Technology, 42(1). https://doi.org/10.1002/meet.14504201146
[37] Wang, X.L., Cloete. (2005). Learning to classify email: a survey. In: 2005 International Conference on Machine Learning and Cybernetics, Guangzhou, China. https://doi.org/10.1109/ICMLC.2005.1527956
[38] Chandrasekaran, M., Narayanan, K., Upadhyaya, S. (2006). Phishing email detection based on structural proper-ties. In: NYS Cyber Security Conference, Albany, New York, pp. 1-8.
[39] Zhong, N., Liu, J., Yao, Y., Wu, J., Lu, S., Qin, Y., Li, K., Wah, B. (2006). Spam filtering and email-mediated appli-cations. In: International Workshop on web Intelligence Meets Brain Informatics, Springer, pp. 1-31. https://doi.org/10.1007/978-3-540-77028-2_1
[40] Cormack, G.V. (2008). Email spam filtering: A systematic review. Foundations and Trends® in Information Retrieval, 1(4): 335-455. https://doi.org/10.1561/1500000006
[41] Sanz, E.P., Hidalgo, J.M.G., Pérez, J.C.C. (2008). Email spam filtering. Advances in Computers, 74: 45-114. https://doi.org/10.1016/S0065-2458(08)00603-7
[42] Toolan, F., Carthy, J. (2009). Phishing detection using classifier ensembles. In: eCrime Researchers Summit, eCRIME'09, pp. 1-9. https://doi.org/10.1109/ECRIME.2009.5342607
[43] Zhang, D., Yan, Z., Jiang, H., Kim, T. (2014). A domain-

feature enhanced classification model for the detection of chinese phishing e-business websites. Information & Management, 51(7): 845-853. https://doi.org/10.1016/j.im.2014.08.003

[44] Laorden, C., Ugarte-Pedrero, X., Santos, I., Sanz, B., Nieves, J., Bringas, P.G. (2014). Study on the effectiveness of anomaly detection for spam filtering. Information Sciences, 277: 421-444. https://doi.org/10.1016/j.ins.2014.02.114

[45] Sah, U.K., Parmar, N. (2017). An approach for malicious spam detection in email with comparison of different classifiers. International Research Journal of Engineering and Technology (IRJET), 4(8): 2238-2242.

[46] Hassanpour, R., Dogdu, E., Choupani, R., Goker, O., Nazli, N. (2018). Phishing e-mail detection by using deep learning algorithms. In: Proceedings of the ACMSE 2018 Conference, ACM. https://doi.org/10.1145/3190645.3190719

[47] Vorobeychik, Y., Kantarcioglu, M. (2018). Adversarial machine learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, 12(3): 1-169. https://doi.org/10.2200/S00861ED1V01Y201806AIM039