I.J. Intelligent Systems and Applications, **2022, 1, 24-31**
Published Online February 2022 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijisa.2022.01.02

An Intelligent Ensemble Classification Method for Spam Diagnosis in Social Networks

Ali Ahraminezhad
Department of Computer Engineering, Liyan Institute of Education, Bushehr, Iran
E-mail: ali.ahraminezhad@gmail.com

Musa Mojarad*
Department of Computer Engineering, Firoozabad Branch, Islamic Azad University, Firoozabad, Iran
E-mail: m.mojarad@iauf.ac.ir

Hassan Arfaeinia
Department of Computer Engineering, Liyan Institute of Education, Bushehr, Iran
E-mail: harfaeinia@gmail.com

Received: 10 October 2021; Revised: 09 November 2021; Accepted: 02 December 2021; Published: 08 February 2022

Abstract: In recent years, the destructive behavior of social networks spammers has seriously threatened the information security of ordinary users. To reduce this threat, many researchers have extracted the behavioral characteristics of spam and obtained good results based on machine learning algorithms to identify them. However, most of these studies use a single classification technique that often works differently for different spam data. In this paper, an intelligent ensemble classification method for social networks spam detection is introduced. The proposed heterogeneous ensemble learning framework is based on stack generalization and uses an evolutionary algorithm to improve the modeling process and reduce complexity. In particular, particle swarm optimization has been used as an evolutionary algorithm to optimize model parameters to reduce model complexity. These parameters include a subset of effective features and a subset of the most appropriate single classification techniques. The SPAM E-mail dataset used in this article contains the correct and effective features in spam prediction. Experimental results show that the proposed algorithm effectively improves the detection rate of spam and performs better than the methods used.

Index Terms: Spam detection, social networks, ensemble classification, intelligent technique.

1. Introduction

In recent years, the Internet has jumped forward and smart platforms have become increasingly popular. In this context, online social networks are an important tool for people to obtain information, disseminate information, make friends and have fun. The popularity of these networks is growing rapidly and allows users to collect a lot of information [1]. As social networks increase in popularity, users are getting a lot of information to receive. The volume of information and the ease of access to it can be the reason for attracting malicious groups. Therefore, networks are affected by spam [2]. Spam detection can facilitate the process of analyzing and monitoring social networks events as well as setting up the management of these networks. Although much work has been done to identify and solve this problem, but this issue is still an important challenge among researchers [3].

Spam detection on social networks is very important due to the rapid growth of databases. Spammers cause users' personal pages to be littered with deceptive messages, resulting in wasted users' time and high traffic load when viewing these networks. To deal with this complex issue, complex spam filters must be created. Traditional machine learning approaches such as neural networks, support vector machine, and naive bayes classification are not efficient enough to process and use the complex features of high-dimensional data [4]. In addition, traditional criteria for filtering social networks spam cannot be adapted to its various costs [5].

With the rapid development of online social networks, social networks such as Twitter has become increasingly important in real life and has become the main target of spammers. Spam detection involves a complex task to involve a wide range of features, as spam causes an unbalanced distribution of data on social networks [6]. In this regard, designing an optimal mechanism that can detect spam with acceptable accuracy and reduce the problems of these networks to some extent is complex. Therefore, having an accurate and efficient method in this field is very necessary and important. In this paper, using an ensemble classification system, a method has been proposed that can detect spam in social network.

Today, networks and the Internet have become an integral part of human life. This, in addition to the many benefits, can lead to a variety of security threats and social anomalies. Spam increases network traffic to such an extent that it sometimes prevents servers from transmitting even the smallest amount of information. This will definitely have a negative effect on user performance. Much research has been done in recent years to address this issue, but unwanted spam is still a serious problem on social media.

Various intelligent detection systems have been developed to assist technology professionals in accurately identifying spam [7]. These systems mainly use classification methods to identify a record as ordinary data or spam. Considering that there is no consensus on the various classification techniques that can perform best in any situation, recently group-based classification (group classification), which combines records with more than one classification model has been investigated. In this study, heterogeneous groups based on classical machine learning models (such as support vector machine, multilayer perceptron, decision trees, etc.) are examined, developed and evaluated by examining the effect of group members' parameter values on classification performance. In particular, the proposed method is divided into two stages of parameter optimization and ensemble classification. In the first step, the parameters of the classification models, including the features used to maximize accuracy, are optimized with evolutionary algorithms. In the second step, an intelligent ensemble classification algorithm is applied to detect social networks spam with optimized parameters.

The continuation of the article is as follows: Section 2 reviews related work. The proposed model for detecting spam on social networks is described in Section 3. In Section 4, the simulation results are presented and finally in Section 5, conclusions and future suggestions are presented.

2. Literature Review

Numerous studies have been conducted on the identification of social networks spam as well as the creation of ensemble classification models. In this section, some of the latest related research in this field is analyzed.

In [8], spam detection in social networks has been done by machine learning based methods such as neural network. This method has two layers, in the first layer the optimal feature selection is done and in the second layer the learning of artificial neural network and spam detection is done using the selected features.

In [9], spam detection was performed using artificial neural network feature selection and sine and cosine algorithms. In this method, the features are updated by the sine and cosine search algorithm and the most optimal features are selected for neural network training.

In [10], spam detection on social networks is based on peer acceptance. Peer acceptance can be achieved on the basis of common interests on several common issues. The contribution of this article is an unsupervised method of identifying spammers based on users' similar acceptance of their post content.

In [11], an in-depth learning model for identifying Twitter spam is proposed. The method is a new method based on deep learning techniques. This approach uses both tweet text and user profile data (e.g., account age, number of followers/ followers, etc.) to identify spam.

In [12], social network spam detection was performed using analysis of graph characteristics and sequence of interactions between users. In this paper, two representation models are proposed for datasets based on social interaction graphs. The first model is based on graph-based analysis, while the other model is based on sequential processing of user interactions.

In [13], the classification of spam emails is investigated using ensemble and non-ensemble machine learning algorithms. Ensemble machine learning algorithms containing non-ensemble algorithms are reinforced by voting. The results show that the non-ensemble SVM with 98.47 accuracy has the best performance and is followed by the ensemble voting algorithm with 96.80 accuracy.

The RST-SU method was introduced in [14] with the aim of reducing the size of spam classification. In this method, the net set theory (RST) and the symmetric uncertainty (SU) method are used to minimize the dimensions of the spam email data group. Comparison of the obtained results shows the efficiency of this method for classifying Spam E-mail dataset spam with 91.69% detection accuracy based on J48 model and RST reduction method.

The I2FELM method has been proposed in [1] for identifying Twitter social network spam based on an enhanced extreme learning machine. Here, Twitter network features such as user feature, content feature, activity feature and link feature are considered for spam analysis. Based on these features, the "Fuzzy Core Extreme Learning Machine (I2FELM)" algorithm for identifying Twitter social network spam has been introduced. Experience validation results show that I2FELM can efficiently identify balanced and unbalanced datasets. In addition, I2FELM detects spam more efficiently due to its fewer features.

3. Proposed Method

This section describes the proposed algorithm for spam detection in the form of a visual learning model. The proposed algorithm consists of several steps: In the first step, the data is pre-processed. In the second step, the 10-fold validation technique is used to divide the data into two sets of training and testing. The third step involves selecting a

subset of effective features as well as selecting a set of heterogeneous classification models. At this stage, heterogeneous groups based on classical machine learning models are developed and evaluated by examining the effect of group member parameter values on classification performance. These parameters are optimized based on an evolutionary algorithm. The parameters of the classification model, including the features used to maximize accuracy, are optimized. The fourth step is to create an intelligent ensemble classification model based on the stack generalization technique. The ensemble classification approach consists of k single classification models whose prediction results are taught by a stack-generalization-based meta-classification. Specifically, at this stage, the intelligent ensemble classification algorithm is applied to detect spam on social networks with optimized parameters. The fifth step involves evaluating the proposed classification model with different criteria. At this stage, the proposed model is evaluated using the experimental set data based on the most effective features selected as well as the structure of the designed heterogeneous classification model. Fig. 1 shows the flowchart of the proposed algorithm.

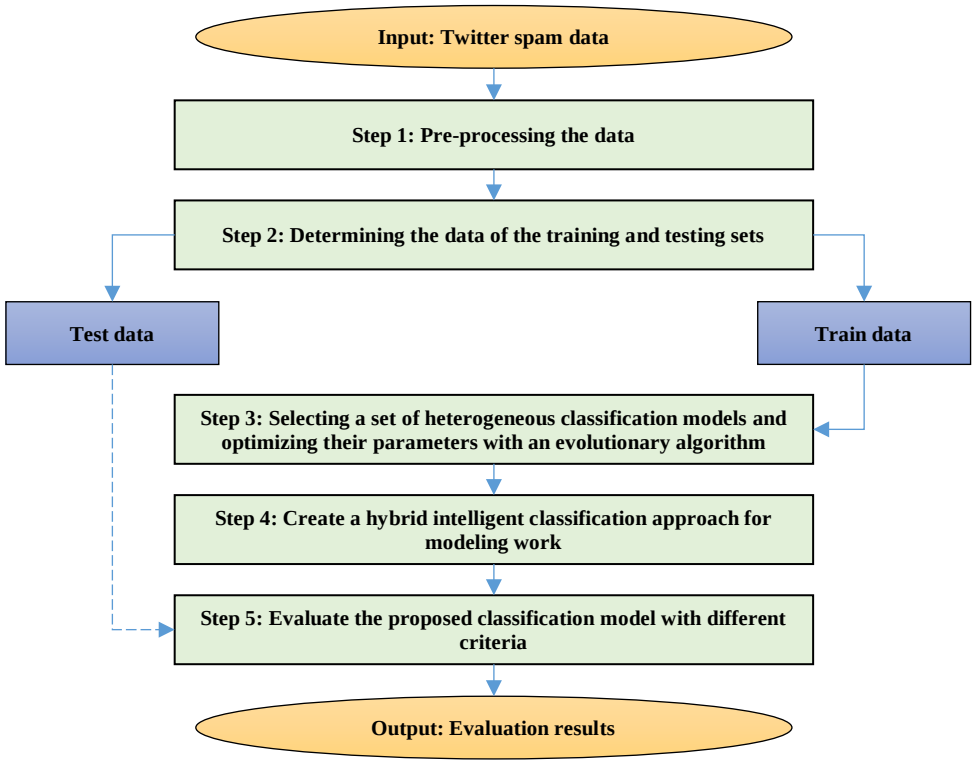

Fig.1. Flowchart of the proposed method

The popularity of social media offers a lot of convenience to people, but their rapid growth has also attracted many attackers. In recent years, the destructive behavior of social network spammers has severely threatened the information security of ordinary users. To reduce this threat, many researchers have extracted the behavioral characteristics of spammers and obtained good results by using machine learning algorithms to identify them on social media. However, current approaches to spam detection are largely based on individual classification techniques.

3.1. Pre-processing of data

Data processing is a key step in the data mining process. In this paper, the data are normalized before modeling. Normalization is used to scale the data of a feature. This paper uses the z-score normalization technique for spam data.

$$x_{i,j}^{Z-score} = \frac{x_{i,j} - \mu_j}{\sigma_j} \quad (1)$$

Where, μ_j is the mean of the samples for the j-th feature and σ_j refers to the standard deviation of the samples for this feature. In addition, $x_{i,j}$ and $x_{i,j}^{Z-score}$ are the actual value and the normalized value for the j-th feature in sample i, respectively.

3.2. Parameter optimization with particle swarm optimization

This step involves optimizing the subset of effective features as well as the subset of heterogeneous classification models. In this paper, particle swarm optimization (PSO) is used as an evolutionary algorithm to search for optimal parameter values.

Here the dimensions of a particle (solution) include two parts: feature selection and classification model selection. Therefore, the number of dimensions of each particle is equal to $m + k$. Here, m refers to the number of main features, and $F = \{f_1, f_2, \dots, f_m\}$ indicates the list of these features. Also, k refers to the number of heterogeneous classification techniques available, and $C = \{c_1, c_2, \dots, c_k\}$ is a list of these techniques. Fig. 2 shows the proposed coding in PSO.

f_1	f_2	…	f_m	c_1	c_2	…	c_k
0/1	0/1	…	0/1	0/1	0/1	…	0/1

Fig.2. Proposed binary coding for PSO

Here, the fit function has two objectives: 1) to increase the accuracy of spam prediction and 2) to reduce the complexity of the model. The weighted method is considered to solve many multi-objective problems due to its simplicity and high efficiency. By adopting the weighted sum method, the fit function can generate the final cost for the solutions in a simple and similar way to the one-objective problem. In the proposed algorithm, the weighted sum is applied to the Fit fitness function according to Eq. (2) to merge the two targets based on the weights w_1 and w_2.

$$Fit = w_1 \times Acc + w_2 \times \frac{1}{m' + k'} \tag{2}$$

Where, Acc refers to the accuracy of particle classification, and m' and k' are selected for the number of features and the number of classification techniques, respectively.

In each step, each particle is updated using the two best values, the $pbest$ and $gbest$ positions. Each particle has a position vector (solution vector) and a velocity vector of equal dimensions. Here, P_i is the position of the i-th particle, and V_i refers to the velocity of the i-th particle. After finding the best values, the velocity and location of each particle are updated according to the PSO algorithm using Eq. (3) and Eq. (4).

$$V_i^{t+1} = \omega \times V_i^t + c_1 \times r_1 \times (pbest_i - P_i^t) + c_2 \times r_2 \times (gbest - P_i^t) \tag{3}$$

$$P_i^{t+1} = \langle P_i^t + V_i^{t+1} \rangle \tag{4}$$

Where, ω is a constant weight of inertia that determines the effect of velocity on the previous iteration on the current velocity. c_1 and c_2 are fixed learning parameters and r_1 and r_2 are random numbers in the range [0, 1]. In this paper, P_i^{t+1} is rounded to the nearest integer.

3.3. Create a proposed ensemble classification model

Despite the large number of classification techniques, none of them have proven to be the best in every situation, because their performance varies from field to field. This destabilizes the performance of classification techniques. In order to deal with this issue, a new approach has been developed which is based on the ensemble method. This method uses more than one classification technique using a combination law to solve the classification problem. Since each of the classification techniques has strengths and weaknesses, the goal of the ensemble method is to reduce these weaknesses and consolidate the advantages by combining each of these techniques through the formation of a group. In general, there are two types of cumulative classification:

(1) Homogeneous sets: In this type, all classification techniques are the same.
(2) Heterogeneous sets: In this type there are at least two different classification techniques for composition. The present study is related to the heterogeneous group in spam classification.

The performance of an ensemble classification depends on the performance of its members (i.e., single classification), and it has been well established those accurate single classifications lead to accurate group classification. In addition, a variety of single techniques are needed to correct conflicting errors in the same data to improve the performance of ensemble classifications. The purpose of this section is to determine the composition rules for integrating the outputs of the classification techniques outlined in the previous step. In this research, several classical classification techniques such as NB, SVM, LDA, ID3 and KNN have been used as single classifiers. The output of these techniques is combined based on the generalized method. Figure 3 shows the structure of the proposed classification model.

In the proposed method, the subset of selected features $\{f_1, f_2, \dots, f_{m'}\}$ as well as the sample labels predicted by each single classification technique $\{p_1, p_2, \dots, p_{k'}\}$, as new data is combined. This method of composition is in accordance with the generalization of the stack. In addition, stack generalization uses a single classification technique

known as compound or meta-classification to create final predictions. Stack generalization creates a combination of predictions made as well as selected features to create a meta-classifier in a new dataset.

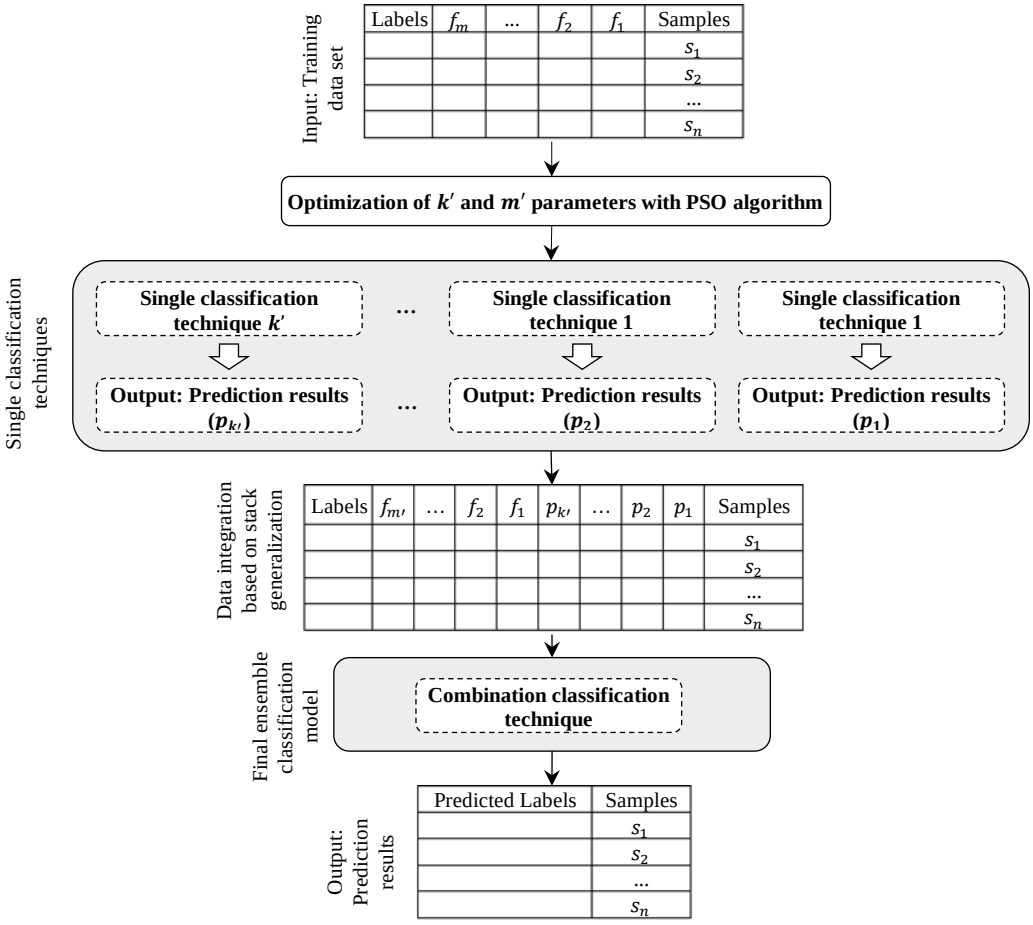

Fig.3. The structure of the proposed ensemble classification model

4. Simulation Results

In this paper, MATLAB software is used for simulation work and comparison of evaluation criteria for different methods. All tests and simulations are performed on a personal computer that has the following specifications: Intel Core i7-2670QM processor with a frequency of 2.2 GHz and 6 GB of RAM. Here the results of the proposed algorithm are compared with the well-known spam detection method such as RST-SU [1] and I2FELM [14]. Two different datasets, SPAM E-mail and UtkMl's Twitter, were used for comparison [15]. This dataset is available at https://archive.ics.uci.edu/ml/datasets/spambase [16].

Table 1 shows the detection accuracy of the proposed model in two cases with and without selecting a feature on the Spam E-mail dataset. The results show the superiority of the proposed method in the case with feature selection, where in this case only 26 features are used for modeling work and the model is less complex.

Table 1. Results of the proposed algorithm with / without selecting a feature on the Spam E-mail dataset

Simulation mode	Accuracy (%)	Number of features used
With feature selection	91.77	26
Without feature selection	97.19	57

In another experiment, the performance of the proposed algorithm against the RST-SU method on the SPAM E-mail dataset is examined. The results of this comparison are presented in Fig. 4.

The results of this experiment show that the proposed algorithm has better results than the RST-SU method in all evaluation criteria. The proposed algorithm offers the best forecast accuracy with 91.77%. Also, the superiority of the proposed algorithm compared to the RST-SU method is about 0.1% on average.

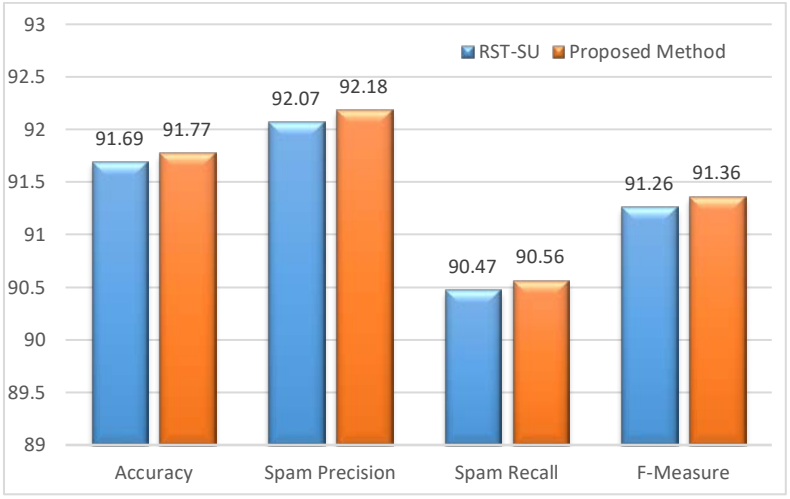

Fig.4. Comparison of the proposed algorithm with RST-SU method on SPAM E-mail dataset

In the next experiment, the performance of the proposed algorithm is compared with the RST-SU method based on the ROC curve. In this diagram, the vertical axis is TPR (true positive ratio) and the horizontal axis is FPR (false positive ratio). Fig. 5 shows this curve based on the SPAM E-mail dataset.

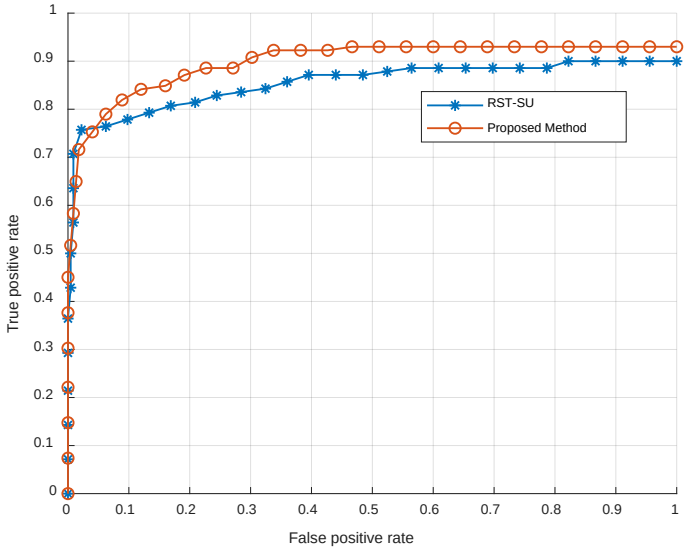

Fig.5. ROC curve of the proposed algorithm and RST-SU method based on SPAM E-mail dataset

The classification results for UtkMl's Twitter dataset are discussed below. Here the proposed algorithm is compared to the I2FELM method with different criteria. The comparison results are presented in Table 2.

Table 2. Comparison of the proposed algorithm with the I2FELM method on UtkMl's Twitter dataset

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
I2FELM algorithm	90.06	92.15	90.12	91.12
Proposed algorithm	90.14	92.23	90.16	91.18

In this comparison, both algorithms have almost the same high performance, however the proposed algorithm is more efficient than I2FELM. Here, the accuracy of the proposed algorithm is 90.14%, which has reported better results than the I2FELM method with 90.06% accuracy.

In this paper, effective features are selected using the PSO algorithm. In addition to selecting effective features, this technique also provides the right number of features. Finally, the number of effective features selected on the Spam E-mail dataset is evaluated. To do this, in the optimization process, the best calculated accuracy of any number of features is reported. Fig. 6 shows the results of this experiment. Here, for each number of different features, the accuracy of spam detection for the test suite is estimated. The best results show 91.77% accuracy with 24 features, while the total number of features is 57. These features are F1, F2, F3, F4, F7, F10, F11, F13, F14, F16, F23, F24, F25, F26, F27, F30, F31, F33, F44, F46, F53, F54, F56 and Are F57.

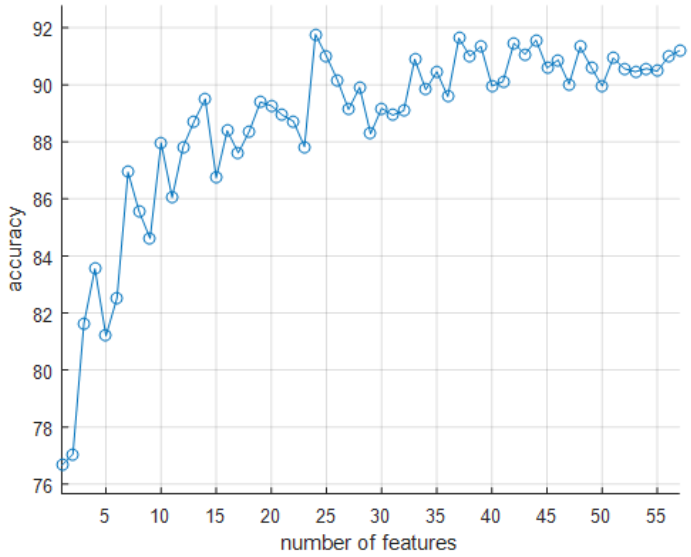

Fig.6. Accuracy criterion results for classifying Spam E-mail datasets with a number of different features

5. Conclusion

Social networks provide a way for users to connect with their friends. The growing popularity of these networks allows users to collect a great deal of personal information about their users. Unfortunately, this ease of access to user information can attract the attention of malicious groups. Most previous research on spam detection on social media has focused on feature extraction, which combines different features or extracts more features from social media accounts to teach classification. Improves, however, most of these studies use an individual classification technique that often works differently for different spam data. That is why these networks have been attacked by spammers, while much work has been done to detect and eliminate them. Due to the fact that spammers are looking for new ways to target these networks every day, continuous measures have been taken to identify malicious spammers and emails. In this paper, an intelligent ensemble classification method for social network spam detection is proposed, which has provided promising results in the simulation.

References

[1] Zhang, Z., Hou, R., & Yang, J. (2020). Detection of Social Network Spam Based on Improved Extreme Learning Machine. *IEEE Access*, 8, 112003-112014.

[2] Srinivasan, S., Ravi, V., Alazab, M., Ketha, S., Ala'M, A. Z., & Padannayil, S. K. (2021). Spam Emails Detection Based on Distributed Word Embedding with Deep Learning. In *Machine Intelligence and Big Data Analytics for Cybersecurity Applications* (pp. 161-189). Springer, Cham.

[3] Prathamesh Churi, N. T. Rao, " Teaching Cyber Security Course in the Classrooms of NMIMS University ", International Journal of Modern Education and Computer Science, Vol.13, No.4, pp. 1-15, 2021.

[4] Sedhai, S., & Sun, A. (2017). Semi-supervised spam detection in Twitter stream. *IEEE Transactions on Computational Social Systems*, 5(1), 169-175.

[5] Sohrabi, M. K., & Karimi, F. (2018). A feature selection approach to detect spam in the Facebook social network. *Arabian Journal for Science and Engineering*, 43(2), 949-958.

[6] Wu, T., Wen, S., Xiang, Y., & Zhou, W. (2018). Twitter spam detection: Survey of new approaches and comparative study. *Computers & Security*, 76, 265-284.

[7] Faris, H., Ala'M, A. Z., Heidari, A. A., Aljarah, I., Mafarja, M., Hassonah, M. A., & Fujita, H. (2019). An intelligent system for spam detection and identification of the most relevant features based on evolutionary random weight networks. *Information Fusion*, 48, 67-83.

[8] Sumathi, S., & Pugalendhi, G. K. (2021). Cognition based spam mail text analysis using combined approach of deep neural network classifier and random forest. *Journal of Ambient Intelligence and Humanized Computing*, 12(6), 5721-5731.

[9] Alauthman, M. O. H. A. M. M. A. D. (2020). Botnet spam E-mail detection using deep recurrent neural network. *Int. J*, 8(5), 1979-1986.

[10] Niranjan Koggalahewa, D., Xu, Y., & Foo, E. (2020, February). Spam Detection in Social Networks based on Peer Acceptance. In *Proceedings of the Australasian Computer Science Week Multiconference* (pp. 1-7).

[11] Alom, Z., Carminati, B., & Ferrari, E. (2020). A deep learning model for Twitter spam detection. *Online Social Networks and Media*, 18, 100079.

[12] Al-Thelaya, K. A., Al-Nethary, T. S., & Ramadan, E. Y. (2020, February). Social Networks Spam Detection Using Graph-Based Features Analysis and Sequence of Interactions Between Users. In *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)* (pp. 206-211). IEEE.

[13] Agarwal, K., Uniyal, P., Virendrasingh, S., Krishna, S., & Dutt, V. (2021). Spam Mail Classification Using Ensemble and Non-

Ensemble Machine Learning Algorithms. In *Machine Learning for Predictive Analysis* (pp. 179-189). Springer, Singapore.
[14] Khalaf, O. I., Abdulsahib, G. M., & Salman, A. D. (2018). Handling Dimensionality Reduction in Spam E-Mail Classification, Jour. of. Adv. *Research in Dynamical & Control Systems*, 10(1), 691-697.
[15] Tran Son Hai, Le Hoang Thai, Nguyen Thanh Thuy,"Facial Expression Classification Using Artificial Neural Network and K-Nearest Neighbor", International Journal of Information Technology and Computer Science, vol.7, no.3, pp.27-32, 2015.
[16] Saptarsi Goswami, Amlan Chakrabarti,"Feature Selection: A Practitioner View", International Journal of Information Technology and Computer Science, vol.6, no.11, pp.66-77, 2014.

Authors' Profiles

Ali Ahraminezhad received his B.E. degree in computer software engineering from Lian Institute, Bushehr, Iran, and Department of Computer Engineering in 2018, and has received his M.Sc. degree in computer software engineering from of Lian Bushehr Institute, Iran, in 2021. Her hobbies are Distributed Computing, Computer Communications (Networks), Computer Engineering, Computer Graphics, Cloud Computing, Edge Computing, Network Security, and E-Learning.

Mousa Mojarad received his PhD in Computer-Software Engineering in 2020. He is currently a lecturer and faculty member of the Islamic Azad University, Firoozabad Branch. His hobbies are Big Data, Face Recognition, Machine Learning, Pattern Recognition and Feature Extraction.

Hassan Arfaeinia obtained her B.E in Computer Science from Rafsanjan University, Kerman, Iran in 2009. HE received her M.S in Computer Engineering from the Amirkabir University of Technology, Tehran, Iran in 2011 and PhD in Computer Engineering from Islamic Azad University, North Tehran Branch, Iran in 2016. His main paper interests consist of Software Development, Machine Learning, Statistical Modeling and Computer Programming.

How to cite this paper: Ali Ahraminezhad, Musa Mojarad, Hassan Arfaeinia, "An Intelligent Ensemble Classification Method For Spam Diagnosis in Social Networks", International Journal of Intelligent Systems and Applications(IJISA), Vol.14, No.1, pp.24-31, 2022. DOI: 10.5815/ijisa.2022.01.02