



ISSN: 0067-2904

SMS Spam Detection Using Multiple Linear Regression and Extreme Learning Machines

Zuhair Hussein Ali^{1*}, Hayder Mahmood Salman², Alaa Hassan Harif³

¹Department of Computer Science, College of Education, Mustansiriyah University, Iraq

²Department of Computer Science, Al-Turath University College, Baghdad, Iraq

³Department of Remote Sensing and GIS, College of Sciences, University of Baghdad, Baghdad, Iraq

Received: 15/12/2022

Accepted: 13/2/2023

Published: 30/10/2023

Abstract

With the growth of the use mobile phones, people have become increasingly interested in using Short Message Services (SMS) as the most suitable communications service. The popularity of SMS has also given rise to SMS spam, which refers to any unwanted message sent to a mobile phone as a text. Spam may cause many problems, such as traffic bottlenecks or stealing important users' information. This paper, presents a new model that extracts seven features from each message before applying a Multiple Linear Regression (MLR) to assign a weight to each of the extracted features. The message features are fed into the Extreme Learning Machine (ELM) to determine whether they are spam or ham. To evaluate the proposed model, the UCI benchmark dataset was used. The proposed model produced recall, precision, F-measure, and accuracy values of 98.7%, 93.3%, 95.9%, and 98.2%, respectively.

Keywords: Spam, ham, Extreme Learning Machine, Multiple Linear Regression, SMS.

الكشف عن الرسائل القصيرة غير المرغوب فيها باستخدام الانحدار الخطي المتعدد وآلات التعلم المتطرفة

زهير حسين علي^{1*}، حيدر محمود سلمان²، علاء حسن هاريف³

¹قسم علوم الحاسوب، كلية التربية، جامعة المستنصرية، بغداد، العراق

²قسم علوم الحاسوب، كلية التراث الجامعة، بغداد، العراق

³قسم التحسس النائي، كلية العلوم، جامعة بغداد، بغداد، العراق

الخلاصة

مع زيادة استخدام الهواتف المحمولة، أصبح الناس مهتمين بشكل متزايد باستخدام خدمات الرسائل القصيرة (SMS) باعتبارها أنسب خدمة اتصالات. أدت شعبية الرسائل القصيرة أيضًا إلى ظهور الرسائل القصيرة العشوائية، والتي تشير إلى أي رسالة غير مرغوب فيها يتم إرسالها إلى الهاتف المحمول كنص. قد يتسبب البريد العشوائي في حدوث العديد من المشكلات، مثل الاختناقات المرورية أو سرقة معلومات المستخدمين المهمة. قدمت هذه الورقة نموذجًا جديدًا يستخلص سبع ميزات من كل رسالة قبل تطبيق الانحدار الخطي المتعدد (MLR) لتعيين وزن لكل من الميزات المستخرجة. أخيرًا، يتم إدخال ميزات الرسائل في Extreme Learning

*Email: zuhair72h@uomustansiriyah.edu.iq

Machine (ELM) لتحديد ما إذا كانت رسائل غير مرغوب فيها أو غير مرغوب فيها. لتقييمها، تم استخدام مجموعة البيانات المعيارية UCI. أنتج النموذج المقترح قيم استرجاع ودقة وقياس F ودقة تبلغ 98.7% و 93.3% و 95.9% و 98.2% على التوالي.

1. Introduction

The Short Message Service (SMS) has become one of the most common means of communication among millions of people as a result of the advancement of mobile communication technologies and the expansion of mobile phone capabilities, where messages must be transmitted by standard communication protocols [1]. There are numerous reasons for the widespread use of SMS messages. Most users read SMS every day, whereas emails may go unread for several days. SMS is inexpensive and, in some cases, free. Finally, the number of mobile phone users has increased dramatically in many countries, reaching several million in the United States and China [2].

Spam is defined as unsolicited and unwanted electronic messages containing potentially malicious content [3]. Email spam and SMS spam are both types of spam. Email spam is spam sent via the internet, whereas SMS spam is typically sent via mobile devices [4]. A spammer is an individual or group of individuals who collectively send unsolicited e-mails using multiple methods of communication [5]. Spam messages come in a variety of forms. Some of these are marketing communications that try to advertise something. Other sorts of spam communications can transmit malware and fool a receiver into giving personal information, which leads to stealing important data such as passwords, private photos, and credit cards' details. In addition, spam messages have a number of drawbacks, such as traffic throttling, storage space, and computational power consumption [6, 7].

With the increased use of text messaging, the problem of SMS spam is becoming more prevalent. There are a number of security methods available to combat the problem of SMS spam, but they are not yet mature [8]. Many Android apps exist on the Play Store to block spam texts; however, owing to a lack of awareness, most consumers are unaware of them. Apart from apps, the available filtering approaches are primarily focused on email spam, as email spam is a known problem. However, with the growth of mobile devices, SMS spam has become one of the most common concerns [9]. Based on the low cost of SMS, it can be considered the simplest way for attackers to phish important information from mobile devices. Attackers are trying to discover new ways to steal sensitive information from users via SMS because SMS is the most common method of communication. Phishing can be accomplished by sending a malicious link via SMS and inviting the recipient to open the link [10].

There are many reasons for the increase in SMS spam. Since there are no applications that can filter spam SMS for both senders and recipients, spam messages may be delivered unfiltered to users' mobile phones. Unlike e-mail and other online communication services, SMS doesn't need the Internet to reach the user reliably. Another reason is the availability of inexpensive bulk SMS senders that can deliver large numbers of SMS simultaneously [11]. To address the problem of SMS spam, they must be identified as soon as the message arrives.

This paper introduces a new model based on the MLR for weighting features and using the ELM to classify messages as spam or ham (legitimate messages). The following are the main contributions:

1. Determine the importance of features in detecting spam messages by assigning a weight to each feature using MLR.

2. Investigate the impact on classification performance of combining the proposed MLR-based feature weighting with ELM.

The paper is divided into five sections: Section one is an introduction. Section 2 illustrates the representations related to spam detection. Section 3 introduces the proposed spam detection model. Section 4 presents the findings attained throughout the respective experiments in this regard. Finally, Section 5 includes with the conclusion.

2. Related Work

Recently, many efforts have been made by researchers to develop a variety of approaches to identify SMS spam messages. However, many of these publicly available SMS spam message filtering methods are still in the early stages of classification and are not very mature or reliable [12]. Generally, SMS spam detection approaches have been divided into two categories: content-based approaches and collaborative-based approaches. The content-based approach uses some techniques to analyze the content of SMS text messages to decide whether the message is spam or not. Some techniques used fully extracted features to detect spam SMS, while other techniques used feature selection methods for detection purposes. The collaborative-based approach depends on usage and user experience [13].

Suleiman and Al-Naymat (2017) [14] developed a classifier based mainly on using H2O as a platform for comparison purposes between different machine learning algorithms. The proposed methods are based on using deep learning, random forests, and Naive Bayes. The UCI dataset was used in the experiment. The random forest gives the best results with precision, recall, f-measure, and accuracy of 96%, 86%, 91%, and 0.977%, respectively.

Popovac et al. (2018) [15] proposed a CNN-based architecture with one layer of convolution and pooling to filter SMS spam. They achieved an accuracy of 98.4%. In addition, Kaliyar et al. (2018) [16] proposed a model based on using different machine learning algorithms to filter messages from many languages, including Singaporean, American, and Indian English. The model attained a high degree of precision for the Indian English language.

Pumrapee et al. (2019) [17] proposed an SMS spam detection method based on long short-term memory (LSTM) and gated recurrent unit algorithms. The experimental results show that the accuracy of the model reached 98%. Sjarif (2019) [18] developed a method based on computing the term frequency-inverse document frequency (TF-IDF) and applying many machine learning algorithms to classify the messages as spam or junk. The Random Forest algorithm outperforms other algorithms with an accuracy of 97.50%.

A new method was introduced by Kumar (2020) [19] to classify a message as ham or spam by going further to identify a smishing message from spam messages. The proposed model is based on two main phases. The first is based on selecting eleven features from each message to classify the message as "ham" or "spam". The second phase is only applied to spam messages in order to separate Smishing messages from spam messages. The random forest gives the best results among many classification algorithms, with 94.9% accuracy.

Hameed and Ali (2021) [20] developed a new method based on binary particle swarm optimization to select the best fuzzy rules. A set of six features was extracted from an SMS spam dataset and introduced as input to the fuzzy system to generate a set of suitable rules for classification purposes. Finally, binary PSO is used to pick the best rules that decrease the complexity and improve the system's performance. The experimental results showed that the

recall, precision, F-measure, and accuracy of the proposed system were 98.8%, 90.8%, 94.6%, and 98.5%, respectively.

3. Proposed spam-detecting model

The details of the employed dataset, preprocessing, and feature extraction technique are first explained, followed by a suggested strategy for resolving the SMS spam problem.

3.1 Dataset Collection

This paper uses the widely adopted UCI repository dataset for performance evaluations. The dataset consists of 5,574 English SMS messages classified as spam and ham. The number of spam and ham messages is shown in Table 1. The dataset contains two columns: the message string and its class (ham or spam). These messages were collected from the National University of Singapore[21, 22].

Table 1: Number of ham and spam messages in the dataset

Message Type	Number of SMS	Percentage of SMS
Ham	4827	86.6%
Spam	747	13.4%

3.2 Preprocessing and Feature Extraction

To prepare the dataset, tokenization, stop word removal, and stemming processes were applied. Then the feature extraction process started. Feature extraction is the process of converting the actual data into numerical values that can be processed while keeping the dataset information. The feature extraction stage plays an important role in the detection of spam messages because the selection of features can have a considerable impact on the success of machine learning approaches. As a result, finding the most helpful features that can effectively classify SMS spam messages is, in most situations, a difficult task. In this paper, a set of seven features $F = \{f_1, f_2, \dots, f_7\}$ is extracted from each message as illustrated in the following.

1. Message length: The number of characters in the message, where SMS spam length tends to be longer than ham.
2. Number of words: A spam message usually contains a larger number of words than a ham message.
3. Uniform Resource Locator(URLs): The existence of URLs in the message. The value of this feature is either 1 or 0.
4. Special characters: The number of special characters in the message. Spam messages usually contain special characters since spammers use these characters for various aims. special characters such as "+, &, !, \$":
5. Capitalized words: a spam message usually consists of capital words to gain the user's attention.
6. Services words: The spammers use many words like "join" or "loan" to trick users into joining their services or helping them to get loans.
7. The existence of the dot symbol in a message is a good sign for ham messages since the dot is used to separate sentences and for chatting.

3.3 MLR-based Feature weighting

Regression analysis is a statistical model for estimating the relationship between variables in a cause-and-effect relationship. MLR helps in assigning a weight for each feature, thus

improving the system's performance. In MLR, there is one dependent variable and more than one independent variable [23, 24]. Equation (1) illustrates the formula for MLR.

$$Y = W_0 + W_1F_1 + W_2F_2 + W_3F_3 + W_4F_4 + W_5F_5 + W_6F_6 + W_7F_7 \quad (1)$$

where

[Y] is the output vector.

[W] feature weight vector.

[F] The extracted features (independent variables).

Equation (2) is used to compute the weight of the features.

$$W = (F \cdot F^t)^{-1} F^t Y \quad (2)$$

The regression model can be represented in a matrix form as follows.

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} F_{1,1} & F_{1,2} & \dots & F_{1,7} \\ F_{2,1} & F_{2,2} & \dots & F_{2,7} \\ \dots & \dots & \dots & \dots \\ F_{n,1} & F_{n,2} & \dots & F_{n,7} \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix}$$

Where n is the number of messages used to train the model . Y is the desired output, whose value is for spam and 0 for ham.. The seven extracted features were represented by (F1, F2,... F7). Finally, W represents the weight of the feature that we try to estimate. Final value features are computed as illustrated in equation (3).

$$F_1 = \sum_{i=1}^7 W_i F_i \quad (3)$$

3.4 ELM for Spam Detection

Extreme Learning Machine (ELM) was proposed by Huang in 2006 as a new learning algorithm. ELM is a training algorithm for a single hidden layer feed forward neural network (SLFN) [25]. Different from gradient-based methods, the weights between the input layer neurons and the hidden layer neurons in the ELM algorithm are randomly assigned and frozen during the training phase [26]. Generally, ELM consists of two main phases: the first one is concerned with creating the hidden layer output matrix with random hidden neurons, and the second phase is concerned with finding the output connections [27]. Compared with the BP neural algorithm, ELM has many advantages over convergence speed and local optimization [25]. The ELM can be expressed as in equation (4).

$$f_j = \sum_{i=1}^l \beta_i g(w_i x_j + b_i) \quad (4)$$

Suppose n is the number of input nodes, l is the number of hidden nodes, and m is the number of output nodes.

Where x_j is the input vector, w_i is the weight between the input nodes and hidden nodes. g is the activation function, which is usually nonlinear; b_i is the bias; and finally, β_i is the weight of the output layer of the i th hidden layer node.

4. Experimental Results

This section deals with evaluation of the proposed model and comparing the results with the state-of-the-art model.

4.1 Evaluation Metrics

The SMS spam detection system's performance can be evaluated by using well-known standard evaluation metrics, namely, recall (R), precision (P), F-measure (F), and accuracy (ACC), as calculated in Equations (5-8) [28, 29].

$$R = \frac{TP}{TP+FN} \quad (5)$$

$$P = \frac{TP}{TP+FP} \quad (6)$$

$$F - \text{measure} = \frac{2 * P * R}{P + R} \quad (7)$$

$$ACC = \frac{TP + TN}{TP + TN + FN + FP} \quad (8)$$

Where

TP is a spam message that is classified as spam.

TN is a ham message that is classified as ham.

FN is a spam message that is classified as ham.

FP is a ham message that is classified as spam.

Also, the Receiver Operating Characteristic (ROC) curve is calculated. In a ROC curve, the true positive rate (recall) is usually plotted on the y-axis, whereas the false negative rate (FNR) is plotted on the x-axis. A ROC curve is used to calculate the classifier's accuracy. Equation (9) explains how to calculate the FNR.

$$FNR = \frac{FN}{FN + TP} \quad (9)$$

4.2 Feature Analysis

The performance of the proposed model was evaluated using a number of tests. The main objective of the proposed model is to measure its ability to detect SMS spam and compare it with many other classification algorithms. Initially, a preprocessing step was performed to prepare the dataset, and then a set of seven features were extracted based on spam and ham message behaviors. Then, the MLR is applied to compute the weight of the feature. Finally, the proposed model classified the message as either spam or ham.

Figure 1 shows the distribution of the most frequent words that appear in ham and spam messages. It is clear from the figure that there is an overlap between spam and ham words, which makes distinguishing between them difficult.

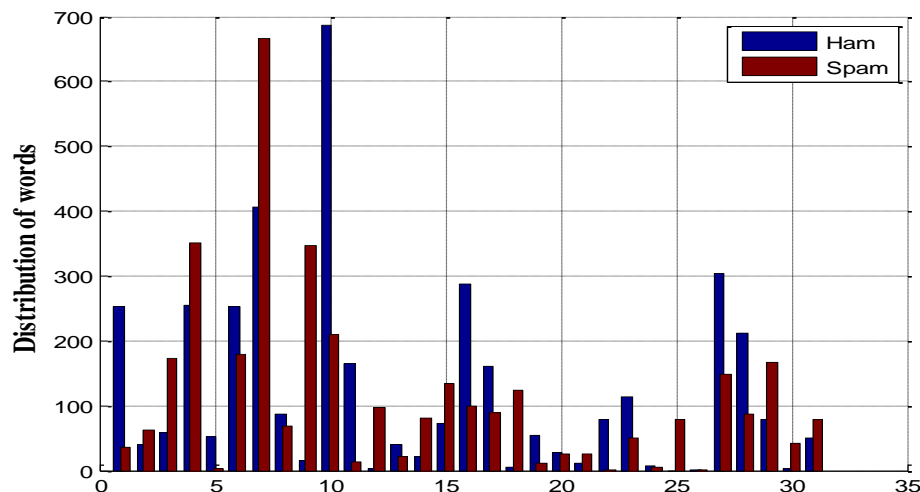


Figure1: Most Frequent words in spam and ham messages.

As is known, different features have varying degrees of significance for a given learning problem. Some features are less significant, while others are critical. Instead of using feature selection to identify the most significant features, an algorithm given the importance of the features increases its speed and accuracy. MLR was applied to satisfy two objectives: firstly, to compute feature weights, which help show the significance of each feature in the detection

model. Secondly, using these weights to assign new values to the features, as illustrated in equation 3, helps to increase the separation between spam and ham features. Figure 2 explains the weight of each feature in the proposed model.

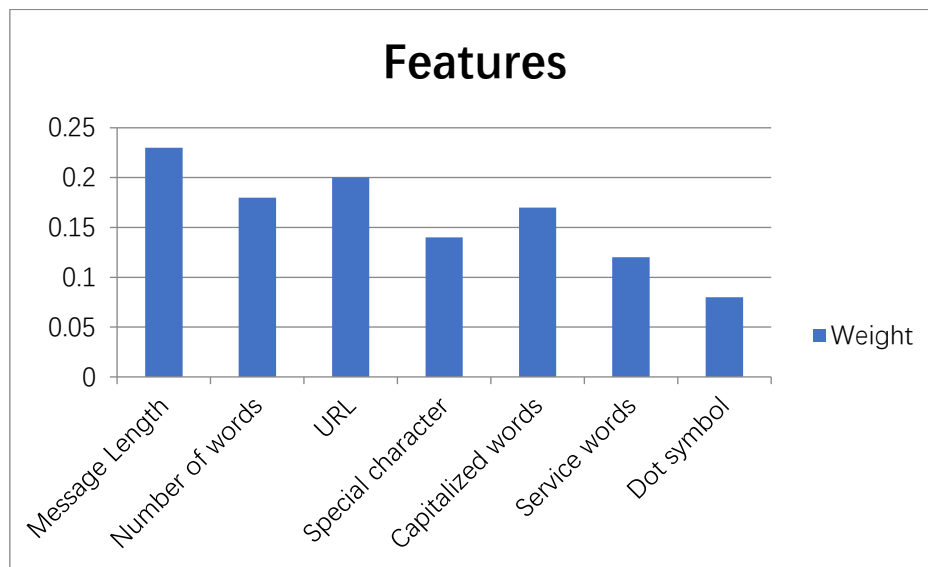


Figure 2: Features weight of the Proposed Model.

A prediction failure occurs in the classification algorithm when features of classes are close to each other. Thus, increasing the separation between features improves the accuracy of the classifier. The proposed method is based on increasing the separation between ham and spam features by multiplying each feature by its weight. As shown in Figure 2, features that identify spam messages have higher weights than features that identify ham messages. Therefore, multiplying the features by their weight increases the separation, which leads to more accurate results.

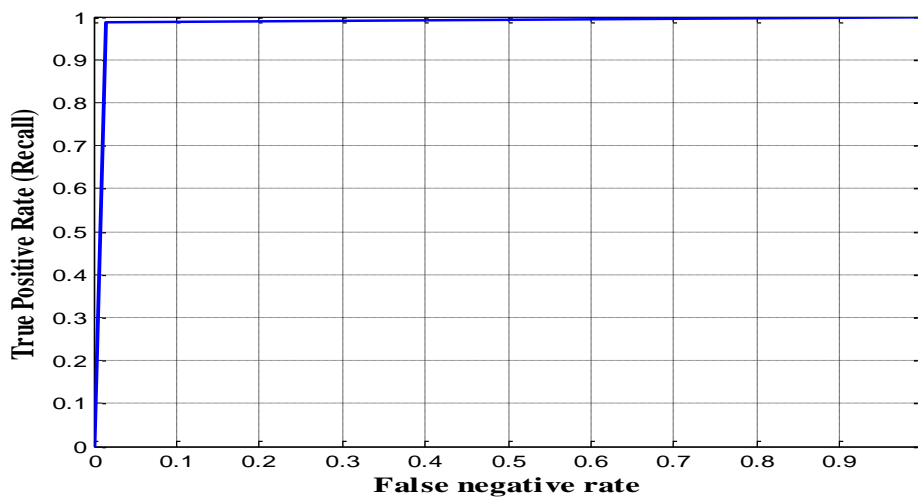


Figure 3: ROC for the proposed model.

It is clear from Figure 3 that the message length feature has the highest weight among all the features set. Other important features include the URL, the number of words, capitalized words, special characters, service words, and the dot symbol.

4.3 Comparisons with Existing Approaches

The model was evaluated using 8-fold cross-validation to determine how accurate the predictions will be in practice. After training the models, the performance of the model was evaluated using the testing dataset. In each step, all the evaluation metrics were applied. The final values were obtained by taking the mean of each step.

Figure 3 shows the ROC where the Area Under the Curve (AUC) equals 0.984. It is clear from the figure that the good performance of the classifier shows a spat between spam and ham messages.

To certify the results of the proposed model, a comparative analysis was done between the existing state-of-the-art models and the proposed model, as shown in Table 2. The proposed model outperforms other state-of-the-art models, especially in terms of accuracy. There are two main reasons behind the good results obtained: Firstly, introducing a new strategy to assign feature weight based on MLR has a direct effect on the separation of the features that distinguish between spam and ham messages, Secondly, the good classification performance of the ELM.

Table 2: Comparisons between the proposed model and other models

Model	Recall	Precision	F-Measure	Accuracy
Proposed method	98.7%	93.3%	95.9%	98.2%
Ref. [14]	86%	96%	91%	97.7%
Ref. [18]	97%	98%	97%	97.5%
Ref. [19]	97.7%	95.9%	96.7%	94.9%

As shown in Table 2, the results of the proposed model tend to be more efficient in dealing with SMS spam messages and less efficient in dealing with ham messages because most of the selected features work toward identifying spam messages.

5. Conclusion and Future Work

SMS is one of the most common ways for people to communicate. There are two categories of SMS messages: spam and ham. This paper introduces a new model to detect SMS spam messages based on feature weight and ELM. MLR is used to weight the extracted seven features, and ELM is used to determine whether SMS messages are spam or ham. According to the weighting algorithm, the message length has the highest weight among a set of features, while the dot symbol has the lowest. Furthermore, the results show that the proposed method outperforms other state-of-the-art models, particularly in terms of accuracy, which is regarded as the most important factor in determining classification algorithm performance.

A limitation of this work is that some features are dependent on the text messages being written in English only. As a result, future research should consider using a similar approach to filter spam and ham text messages written in other languages by replacing these features with others that are language independent. In addition, the dataset used in the experiment contains thousands of messages. In the future, a larger dataset can be used to prove the performance of the model.

Acknowledgment

The author would like to thank Mustansiriyah University (www.uomustansiriyah.edu.iq) in Baghdad, Iraq, for its assistance in this work.

References

- [1] P. Roy, J. Singh, and S. Banerjee, "Deep Learning to Filter SMS Spam," *Future Generation Computer Systems*, vol. 2, pp. 524–533, 2020.
- [2] O. Abayomi-Alli, S. Misra, A. Abayomi-Alli, and M. Odusami, "A review of soft techniques for SMS spam classification: Methods, approaches, and applications," *Engineering Applications of Artificial Intelligence*, vol. 86, pp. 197–212, 2019.
- [3] N. Jiang et al., "Understanding SMS Spam in Large Cellular Network: Characteristics, Strategies, and Defenses, Intelligent Systems," *IEEE*, vol. 27, no. 6, pp. 15-26, 2011.
- [4] N. Choudhary, and A. Jain, "Comparative analysis of mobile phishing detection and prevention approaches," in *International Conference on Information and Communication Technology for Intelligent Systems*, Springer, pp.349-356, 2017.
- [5] A. Aski and N. Sourati, "Proposed efficient algorithm to filter spam using machine learning techniques," *Pacific Science Review A: Natural Science and Engineering*, vol.18, no. 2, pp. 145–149, 2016.
- [6] D. Drossos, G. Giaglis, and G. Lekakos, "Determinants of Effective SMS Advertising: An experimental Study," *Journal of Interactive Advertising*, vol. 7, no. 2, pp. 16-27, 2007.
- [7] X. Liu, L. HAOYE, and A. Nayak, "A Spam Transformer Model for SMS Spam Detection," *IEEE Access*, vol.9, pp.80253–80263, 2021.
- [8] TextBlocker. <https://play.google.com/store/apps/details?id=com.thesimpleandroidguy.app.messageclient&hl=en>.
- [9] D. Puniški, R. Laurutis, R. Dirmeikis, "An artificial neural nets for spam e-mail recognition," *Elektronika Elektrotechnika*, no. 69, pp.73–76, 2006.
- [10] N. Choudhary and A. Jain, "Towards Filtering of SMS Spam Messages Using Machine Learning Based Technique," *Advanced Informatics for Computing Research*, pp.18–30, 2017.
- [11] F. Dewi, M. Fadhlurrahman, R. Rahmianto, and R. Mahendra, "Multiclass SMS message categorization: Beyond spam binary classification," *2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pp. 210-215, 2017.
- [12] S. Sheikhi, M. Kheirabadi, and A. Bazzazi, "An Effective Model for SMS Spam Detection Using Content-based Features and Averaged Neural Network," *International Journal of Engineering*, vol. 33, no. 2, pp. 221-228, 2020.
- [13] S. M. Hameed, "Differential Evolution Detection Models for SMS Spam," *International Journal of Electrical and Computer Engineering*, Vol. 11, No. 1, pp. 596-601, 2021.
- [14] D. Suleiman, and G. Al-Naymat, "SMS Spam Detection Using H2O Framework", *Procedia Computer Science*, vol. 113, pp. 154–161, 2017.
- [15] M. Popovac, M. Karanovic, S. Sladojevic, M. Arsenovic, A. Anderla, "Convolutional neural network based SMS spam detection," In: *2018 26th Telecommunications Forum (TELFOR)*, IEEE, pp. 1–4, 2018.
- [16] R. K. Kaliyar, P. Narang, and A. Goswami, "SMS Spam Filtering on Multiple Background Datasets Using Machine Learning Techniques: A Novel Approach", In: *Proc. of International Conf. On Advance Computing, Greater Noida, India*, pp. 59-65, 2018.
- [17] P. Poomka, W. Pongsena, N. Kerdprasop, K. Kerdprasop, "SMS Spam Detection Based on Long Short-Term Memory and Gated Recurrent Unit," *Int. J. Futur. Comput. Commun.*, Vol. 8, pp. 12–15, 2019.
- [18] N. N. Sjarif, N. F. Mohd Azmi, S. Chuprat, H. Sarkan, Y. Yahya, S. M. Sam Razak, "SMS Spam Message Detection using Term Frequency-Inverse Document Frequency and Random Forest Algorithm," *Procedia Computer Science*, vol. 161, pp. 509–515, 2019.
- [19] A. Kumar, "A Novel Approach to Detect Spam and Smishing SMS Using Machine Learning Techniques," *International Journal of E-Services and Mobile Applications*, vol. 12, no. 1, 2020.
- [20] S. M. Hameed and Z.H. Ali, "SMS Spam Detection Based on Fuzzy Rules and Binary Particle Swarm Optimization," *International Journal of Intelligent Engineering & Systems*, vol. 14, no. 2, pp. 314-322, 2021.
- [21] T. A. Almeida, G. Hidalgo, and J. M., Yamakami, "Contributions to the Study of SMS Spam filtering: New Collection and Results," In: *Proc. of ACM Symposium on Document Engineering*, pp. 259-262, 2011.
- [22] T. A., Almeida, J. M., Gómez Hidalgo, T. P. Silva, "Towards SMS Spam Filtering: Results under a New Dataset," *International Journal of Information Security Science*, vol. 2, no. 1, pp. 1-18, 2013.

- [23] S. Malallah and Z. H. . Ali, "TOPSIS with Multiple Linear Regression for Multi-Document Text Summarization," *Iraqi Journal of Science*, vol. 58, no. , pp. 1298–1307, Dec. 2021.
- [24] M. J. Gatea and S. M. Hameed, "An Internet of Things Botnet Detection Model Using Regression Analysis and Linear Discrimination Analysis," *Iraqi Journal of Science*, vol. 63, no.10, pp. 4534-4546, 2022.
- [25] G.B. Huang, Q.Y. Zhu and C. K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489–501, 2006.
- [26] L. Yang, Y. Li and Z. Li., "Improved-ELM method for detecting false data attack in smart grid," *International Journal of Electrical Power and Energy Systems*, vol. 91, pp. 183–191, 2017.
- [27] L. Yang, J. Zhang, X. Wang, Z. Li and Y. He., "An improved ELM-based and data preprocessing integrated approach for phishing detection considering comprehensive features," *Expert Systems with Applications*, vol.165, 2021.
- [28] A. Iqbq, S. Aftab, I. Ullah, M. S. Bashir and M.A. Saeed, "A feature selection based ensemble classification framework for software defect prediction," *International Journal of Modern Education and Computer Science*, vol.11, no.9, pp. 54-64, 2019.
- [29] Husam A. Abdul Mohsen and Alaa Harif, "Agent-based grid computing load balancing at application level," *Iraqi Journal of Science*, vol. 53, no. 4, pp. 899-902, 2012.