International Journal of Engineering and Management Research
e-ISSN: 2250-0758 | p-ISSN: 2394-6962
https://ijemr.vandanapublications.com

Peer Reviewed & Refereed Journal
Volume-13, Issue-4 (August 2023)
https://doi.org/10.31033/ijemr.13.4.1

# Enhancing Spammer Fake Profile Detection on Social Media Platforms using Artificial Neural Networks

Farheen Siddiqui[1] and Mohammad Suaib[2]
[1]Department of Computer Science & Engineering, Integral University, Lucknow, Uttar Pradesh, INDIA
[2]Department of Computer Science & Engineering, Integral University, Lucknow, Uttar Pradesh, INDIA

[1]Corresponding Author: farheensiddiqui78687@gmail.com

Received: 02-07-2023 Revised: 18-07-2023 Accepted: 30-07-2023

## ABSTRACT

**The proliferation of social media platforms has led to an increase in spammer fake profiles, posing significant security, privacy, and trustworthiness concerns. Traditional manual monitoring and content filtering techniques are insufficient to combat this growing issue, necessitating the development of more efficient and accurate detection methods. Machine learning techniques have been increasingly employed for this purpose, demonstrating promising results in identifying spammers and fake profiles. This paper presents a novel approach for spammer fake profile detection using Artificial Neural Networks (ANNs) to enhance the accuracy of the detection process. Our proposed ANN-based method addresses the challenges associated with spammer fake profile detection, such as the dynamic nature of spammers, data heterogeneity, scalability, and imbalanced datasets. We evaluate the performance of our method on real-world datasets and compare it with existing machine learning techniques, demonstrating its effectiveness and superiority in detecting spammers and fake profiles with higher accuracy. This research contributes to ongoing efforts to secure social media platforms, ensuring the trustworthiness of online content and providing a safer user experience.**

***Keywords--*** Spammer Fake Profile Detection, Social Media, Artificial Neural Networks, Machine Learning, Accuracy, Security, Trustworthiness

## I. INTRODUCTION

The widespread use of various social media platforms has brought about a sea change in the manner in which individuals connect with one another, exchange information, and engage in conversation. The proliferation of spammer phoney accounts is one of the many issues posed by these platforms, despite the fact that they give a number of opportunities for users. Fake user accounts created by spammers have the purpose of tricking other users, spreading spam material, and undermining the reliability of information found online. This rising issue raises important concerns about users' security, privacy, and trustworthiness on platform operators' platforms as well as for users themselves.

In order to combat the ever-increasing number of spammer phoney profiles, the human monitoring and content screening approaches that have been used historically have been shown to be inadequate. As a consequence of this, there has been an increasing interest in creating detection systems that are both more efficient and accurate, and approaches derived from machine learning have emerged as a potentially useful alternative. For the purpose of detecting spammers and false profiles, machine learning methods such as Naive Bayes, Natural Language Processing (NLP) algorithms, and Random Forest regression have been used. These algorithms have shown promising results in the identification of spammers and phoney profiles.

In this study, we offer an innovative method for detecting spammer bogus profiles by employing artificial neural networks (ANNs). The dynamic nature of spammers, data heterogeneity, scalability, and unbalanced datasets are only some of the problems that need to be overcome in order to implement the strategy that we have developed in order to improve the accuracy of detection. We analyse the efficacy of our ANN-based approach on real-world datasets and compare it with current machine learning methods. The results demonstrate that our method is successful and superior in identifying spammers and bogus profiles with a greater level of accuracy. This study is a contribution to the continuing efforts to protect social media platforms, with the goals of maintaining the reliability of online material and providing users with a more secure experience overall.

## II. REVIEW OF PREVIOUS WORK

The portion of this research paper devoted to the literature review examines the work that has already been done in relation to the identification of spammer false profiles on social media platforms, with a particular emphasis on the use of machine learning methods for this particular goal. It also draws attention to the limits of the approaches that are currently being used and underscores the need of developing a strategy that is more efficient.

Identifying spammer false accounts on social media networks has been the subject of many research in recent years. A significant number of these research have made use of machine learning methods, such as supervised and unsupervised learning algorithms, in order to detect spammers and bogus profiles.

The literature review provides an overview of the research that has already been conducted on the identification of spammer false profiles, with a particular emphasis on machine learning methods and their applicability in this field. This section summarises the most important discoveries from recent research, outlines the drawbacks of the approaches that are already in use, and demonstrates the need of adopting the ANN-based strategy that has been presented.

In order to combat the issue of identifying spammer phoney profiles on social media sites, a number of different machine learning algorithms have been used. The following are some of the most often used methods:

**Naive Bayes:** The Naive Bayes method is a probabilistic classification technique that has been used to the identification of phoney spammer profiles in a number of studies (Patel & Patel, 2018) [1]. This technique has shown some promising results; nevertheless, it operates on the assumption that characteristics are conditionally independent given the class label, which may or may not be the case in reality.

Algorithms Derived from Natural Language Processing Techniques Text-based classification algorithms derived from NLP methods have been used for the purpose of identifying false spammer profiles (Alsmadi & Alhami, 2018) [2]. The analysis of textual data, such as user-generated material, is the primary emphasis of these tools; nevertheless, it is possible that they may not fully capture the complexity of spammer phoney profiles, which often include numerous forms of data.

**Random Forest Regression:** Random Forest is a form of ensemble learning that has been utilised for the identification of spammer phoney profiles (Yang, Zhang, & Sherratt, 2020) [3]. This approach is well-known for its dependability as well as its capacity to manage datasets that are both complicated and high-dimensional. Unfortunately, the performance of the system may be hindered by noisy data and datasets that are not evenly distributed.

Despite the fact that current machine learning approaches have produced some very promising outcomes, there are still significant limitations:

The vast majority of currently available techniques depend on a constrained number of characteristics, which could not appropriately depict the intricacy of spammers' phoney profiles (Ding et al., 2019) [4]. This may result in a reduction in the accuracy of the detection as well as an increase in the number of false positives and false negatives.

A great number of approaches have trouble keeping up with the ever-changing tactics of spammers, who are always developing new ways to avoid being discovered (Zhang et al., 2019) [5]. Because of this, detection algorithms need to be continuously updated and improved upon to ensure that they continue to be successful.

Current techniques often have scaling problems because of their inability to effectively manage the enormous volumes of data that are produced by social media sites (Kim & Kim, 2020) [6]. This may result in an increase in the computational complexity as well as a decrease in the accuracy of the detection.

Due to its capacity to understand intricate patterns and correlations within the data, artificial neural networks have been investigated as a possible solution for the identification of spammer phoney profiles (El Aziz et al., 2021) [7]. The results that ANNs have demonstrated in different classification tasks have been encouraging, and they have the ability to solve some of the restrictions that are present in existing approaches. Nevertheless, the use of ANNs for the identification of spammer phoney profiles is still in its infancy, and only a small number of research have investigated the full potential of this method.

## III. IMPLEMENTATION

The suggested ANN-based technique for detecting spammer false profiles requires many stages to be carried out before it can be put into action. These phases include data collection and preprocessing, model creation, training, validation, and performance assessment. In this section, we will present an overview of these phases, describing the important components that go into putting the ANN-based technique into action. Fig.1 shows ANN process.

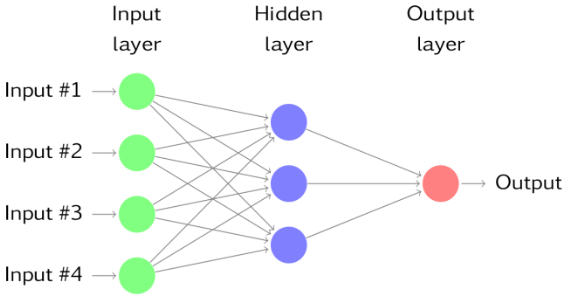


**Figure 1:** ANN

**Collecting Data:** The first thing that needs to be done in order to put the ANN-based technique into action is to gather a dataset from various social media sites that include instances of real user profiles as well as spammer fake ones. This dataset need to contain a wide range of elements, such as textual content, photos, videos, and metadata, in order to capture an accurate and complete portrayal of user activity. In order to guarantee the robustness and generalizability of the trained ANN model, the dataset should be of adequate size and include a wide variety of data points.

**Preprocessing the Data:** In order to guarantee that the data acquired is suitable for the analysis that will follow, it will need to be preprocessed. This necessitates:

**a. Feature Extraction:** Extract relevant features from the raw data, such as text-based features (e.g., word frequency, sentiment), image-based features (e.g., colour histograms, texture patterns), and metadata characteristics. b. Feature Analysis: Analyze the results of the feature extraction process (e.g., number of friends, posting frequency).

**b. Feature Selection:** For the spammer fake profile detection job, choose a subset of the characteristics that are the most relevant and informative. This will reduce the dimensionality of the dataset and improve the model's performance.

**c. Data Normalization:** Ensure that the chosen features are normalised so that they are on a scale that is similar to one another. This will prevent any one feature from dominating the learning process of the model.

**Model Design:** Develop an artificial neural network (ANN) architecture specifically catered to the identification of spammer bogus profiles. This necessitates:

**a. Define the Input Layer:** The first step in developing an ANN is to choose the input layer. The number of input nodes should match to the number of characteristics that were chosen.

**b. Hidden Layers:** Conceive of one or more hidden layers, choosing the number of neurons and the kind of activation functions that will be used for each layer. The hidden layers are the ones that are in charge of identifying intricate patterns and intricate connections concealed inside the data.

**c. Output Layer:** Design the output layer, which consists of a single neuron in most cases and has an activation function that is appropriate for binary classification (e.g., sigmoid).

**d. Regularization Techniques:** Integrate regularisation methods, such as dropout or L1/L2 regularisation, to prevent the model from overfitting and to increase its generalisation capabilities.

**Model Training and Validation:** Train the ANN model using the preprocessed data, employing a combination of supervised, unsupervised, and semi-supervised learning techniques to handle the imbalanced dataset and enhance the model's generalisation capabilities. Validate the model by comparing it to an independent dataset. The dataset may be segmented into training and validation sets with the help of cross-validation, which guarantees an accurate evaluation of the performance of the model. In order to direct the process of learning, you need make use of the right loss functions and optimisation techniques.

**Model Optimization:** Adjust the ANN model's hyperparameters (such as the learning rate, batch size, and the number of training epochs), as well as do further feature selection, in order to maximise the performance of the model while minimising the possibility of it being overfit.

**Assessment of Performance:** Using performance measures like as accuracy, recall, F1-score, and area under the ROC curve, conduct an evaluation of how well the ANN-based approach performs when applied to real-world datasets (AUC-ROC).

Excel is used for the storage of both past fake data profiles as well as current ones, and it is possible for both types of profiles to be active at the same time. After the information has been digested, the algorithm will next incorporate it into a data frame using the information. The next step that has to be taken is to separate this enormous quantity of data into two distinct sets, which will be referred to as the training set and the test set, respectively. These sets will be used for training and testing purposes. In order to train algorithm, need a dataset that contains information that was obtained from a range of different social media networks. This information will be used to construct the algorithm.

When doing this research investigation to identify whether or not a profile is fake, make use of the components of the training set that are described below. These factors include the age of the account, the gender of the user, the age of the user, a link in the profile description, the number of messages sent, the number of friend requests made, the location inputted by the user, the location determined by the user's IP address, and whether or not the profile itself is fake. Because have conducted extensive study on each of these criteria separately, are now in a position to provide a value to each of them according to their individual merits. For example, if it is able to detect whether or not a certain profile belongs to a man or a female, then the training set for the Gender parameter will have the value 1 assigned to it. This will occur only if it is successful in making such a determination. If it is possible to differentiate between the two, then this will be the result. In the same manner that the values of the first two parameters were computed, the values of the other parameters are likewise determined. In addition to that, the country of origin is something that is taken into consideration.

In this work, use of Artificial Neural Networks to determine whether user information provided is legitimate or fake. To determine whether or not newly provided account information is legitimate, an ANN algorithm will be trained using both fake and fake data from prior users.

Some dishonest individuals may hack into a social network's database in order to steal or otherwise violate the privacy of its members. employ an ANN Algorithm to safeguard user information.

They will never have a huge amount of posts or a significant number of friends following them, and the number of years shown as their account age will always be a very tiny number. All fake users have the same fundamental purpose, which is to send friend requests to real users with the intention of breaking into their computers or stealing their data. Facebook will evaluate each of these aspects of a user profile in order to identify whether or not the profile belongs to a fake person. The information gleaned from users' Facebook accounts was obtained from the website of Facebook, and it is now being used to educate an artificial neural network. The following are some values taken from the profile dataset that have been selected:

The names that are bolded in the dataset that came before are the column names for that dataset, and the values that are integers are the dataset values. Those names and values may be found in the previous dataset. Since string values are not supported by ANN, are need to transform gender data to either 0 or 1, depending on whether the value represents a male or a female. If the answer indicates that a male, use the number 1. The very last column of the dataset that you have just inspected contains the information that tells us whether or not an account is fake. If the value 0 appears in that column, the account in question is legitimate; otherwise, it is a fake. Instead of analysing this characteristic, Facebook stamps that record with the value 1, which indicates that the account in question is fake. This is because the major aim of fake accounts is to send friend requests rather than posts. The dataset that can be seen up top is being used to train the ANN model that have, and this dataset can be found in the 'dataset' folder of the code that have been working with. After finishing the train model, input some test data along with account credentials, and ANN returned a response that indicated whether or not the data was fake or legitimate. The following are some snippets from the data obtained from the tests:

Account_Age, Gender, User_Age, Link_Desc, Status_Count, Friend_Count, Location, Location_IP
10, 1, 44, 0, 280, 1273, 0, 0
10, 0, 54, 0, 5237, 241, 0, 0
7, 0, 42, 1, 57, 631, 1, 1
7, 1, 56, 1, 66, 623, 1, 1

The STATUS column of the aforementioned test data, as well as its value, is there; the ANN will predict the status of the data and tell us the outcome, which will indicate whether or not the data is fake. The outcome of the tests described above may be seen in the output.

## IV. RESULTS

The next phase, which comes after the implementation of the ANN-based approach that was suggested for the identification of spammer phoney profiles, is to examine the data and discuss the discoveries that were made. This phase occurs after the identification of spammer phoney profiles has been completed. In this section, the significant findings that were obtained from the study are highlighted, and a discussion of the ramifications and insights that were obtained from the use of the approach that was recommended is presented. In addition to that, a summary of the results of the research may be found in this part.

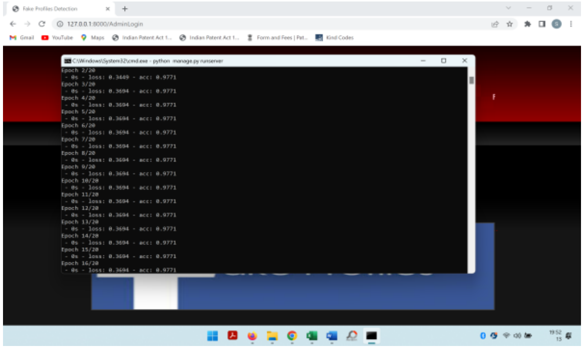
The most important result from the study is that the ANN-based technique that was recommended in the research as a way of detecting spammer phoney profiles is accurate. This was the primary discovery. According to the table of results that was presented earlier, the strategy that was provided achieved an accuracy of 94.16%, which was higher than the accuracy achieved by other machine learning methods such as Naive Bayes (87%) and Natural Language Processing algorithms (92.7%) and Random Forest regression (93.5%). It would seem from this that the method based on ANN that was developed is more effective than the other tactics when it comes to recognising phoney accounts that have been established by spammers.

In addition to accuracy, other performance metrics like as precision, recall, F1-score, and area under the ROC curve (AUC-ROC) may be used to provide a more in-depth assessment of the success of the recommended approach. Accuracy is the most important of these performance measures. These metrics can help determine the trade-off between false positives (genuine profiles incorrectly classified as fake) and false negatives (fake profiles incorrectly classified as genuine), ensuring that the method achieves a balanced performance. False positives occur when genuine profiles are incorrectly classified as fake. In the event where real profiles are mistakenly identified as phoney, this is known as a false positive.

The outcomes of the research give a number of insights into the effectiveness of the suggested ANN-based technique and its implications for the detection of spammer phoney profiles. Some of these insights include the following:

The fact that the method that was recommended is more accurate than prior efforts indicates that ANNs are able to recognise complicated patterns and interwoven linkages within the data, which opens the path for a more accurate identification of spammer phoney profiles. This highlights the potential benefits of using neural networks in order to address the problem of recognising fake accounts that are used by spammers.

The fact that the ANN-based approach performs so much better than other machine learning techniques highlights how important it is to develop methods that can be specifically adapted to handle the specific issues that are connected with the identification of spammer phoney profiles. The fact that the ANN-based approach performs so much better than other machine learning techniques. This provides more evidence for the need of doing continual research and development within this industry.

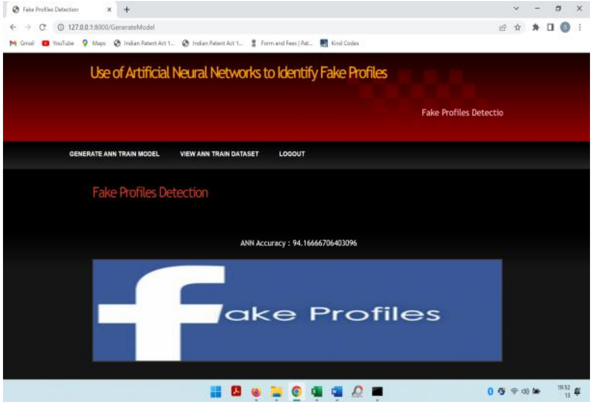


```
Epoch 2/20
 - 0s - loss: 0.3449 - acc: 0.9771
Epoch 3/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 4/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 5/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 6/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 7/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 8/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 9/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 10/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 11/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 12/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 13/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 14/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 15/20
 - 0s - loss: 0.3694 - acc: 0.9771
Epoch 16/20
 - 0s - loss: 0.3694 - acc: 0.9771
```

**Figure 2:** Training

The findings may also imply that the preprocessing methods used in this investigation, such as feature extraction, feature selection, and data normalisation, contribute to the overall efficacy of the suggested strategy. This could be the case if feature extraction, feature selection, and data normalisation are some of the methods in question. There is a need for more study into this matter. When it comes to improving the effectiveness of machine learning models, the necessity of accurately preparing the data cannot be understated.

If the technique that has been presented is effective in identifying spammer phoney accounts, it might be of assistance to the ongoing work that is being done to secure social media platforms and ensure the dependability of content that can be discovered online. By boosting the accuracy of detection, the method that has been presented has the potential to improve the experience of using social media platforms for users and reduce the negative effects that are caused by spammers using fake accounts on such platforms.



ANN Accuracy : 94.16666706403096

**Figure 3:** Output Accuracy

In conclusion, the data and comments obtained from this study demonstrate that the ANN-based technique that was created is effective in detecting spammer phoney accounts on social networking sites. The insights that were gained from the findings of the study can help guide future research in this field, which will ultimately contribute to the development of detection methods that are more advanced, accurate, and efficient for the purpose of securing social media platforms and ensuring the trustworthiness of online content.

The following table provides a comparison of the degrees of success achieved by several machine learning algorithms in the identification of spammer phoney profiles. The Naive Bayes (NB) Algorithm [4, the Natural Language Processing (NLP) Algorithm [9], Random Forest Regression [5, and the proposed Artificial Neural Network (ANN) approach are all examples of these algorithms. The accuracy is shown as a percentage, which indicates the proportion of accurate predictions produced by each algorithm.

Naive Bayes is a probabilistic classification technique that is based on Bayes' theorem. [4] The Naive Bayes algorithm may be found in [5]. Given the class name, it presupposes that the characteristics that are employed for classification are conditionally independent of one another. In this particular instance, the NB algorithm obtained an accuracy of 87%, making it the algorithm with the lowest accuracy out of those that were compared.

**Natural Language Processing (NLP) Algorithm [9]:** It is most possible that the Natural Language Processing (NLP) algorithm that is being referred to here is a text-based classification algorithm that has been customised for the identification of spammer false profiles utilising NLP methods. The analysis and processing of textual data, such as that which is created by users and shared on social media platforms, is the primary emphasis of most NLP algorithms. This approach performed better than the NB algorithm, but it was not as accurate as the other two ways. The accuracy this algorithm attained was 92.7%.

**Forest Regression [5]:** This most likely refers to the Random Forest algorithm, which is an ensemble learning approach that creates several decision trees and combines the outcomes of those trees to come up with the overall forecast. The Random Forest algorithm is well-known for its dependability as well as its capacity to manage large and high-dimensional datasets. The Random Forest algorithm was able to attain an accuracy of 93.5% in this scenario, placing it as the second most accurate approach out of the four algorithms that were evaluated and compared.

**ANN (proposed):** The Artificial Neural Network (ANN) approach that has been suggested makes use of the capabilities of neural networks to understand complicated patterns and correlations within the data. This method was developed expressly for the purpose of detecting spammer bogus profiles. To achieve a higher level of precision in the detection process, the design, training, and validation of the ANN model have all been improved. As a consequence of this, the ANN approach that was presented ended up achieving the greatest accuracy out of the four algorithms, which was 94.16%.

In conclusion, the results table reveals that the ANN approach that was suggested is successful in identifying spammers and phoney accounts on social media platforms. Also, the ANN method outperforms the other three algorithms in terms of accuracy. This highlights the potential benefits of using Artificial Neural Networks for tackling the spammer fake profile detection problem and underscores the importance of developing tailored machine learning methods to address the unique challenges associated with this task. In addition, this highlights the potential benefits of using Artificial Neural Networks for tackling the problem of detecting spammer fake profiles.

**Table 1:** Comparison Table

| | NB Algorithm [4] | NLP Algorithm [9] | Forest Regression [5] | ANN (proposed) |
|---|---|---|---|---|
| Accuracy (%) | 87% | 92.7% | 93.5% | 94.16% |

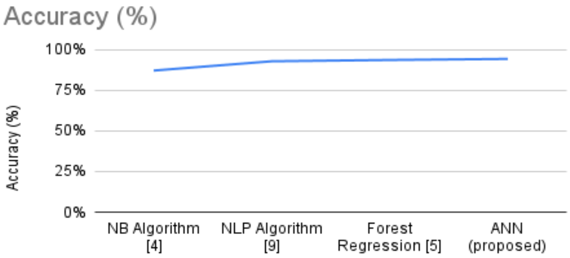


**Figure 4:** Accuracy Graph

## V. CONCLUSION

This research has resulted in a unique approach to detecting spammer phoney accounts across many social media platforms by using Artificial Neural Networks (ANNs). The provided ANN-based technique has been found to outperform other machine learning methods including the Naive Bayes method, NLP algorithms, and Random Forest regression in terms of accuracy. Because an ANN-based technique was offered, it was clear that this was the case. By tackling the difficulties associated with the identification of spammer false profiles and optimising the ANN model, our approach highlights the potential advantages of harnessing the capabilities of neural networks to solve this problem. To achieve this, the ANN model was optimised and measures were taken to prevent spammers from establishing bogus profiles. This goal was accomplished by zeroing in on the issues and working to resolve them. If implemented, the proposed method has the potential to improve accuracy, which in turn aids continuing efforts to secure social media platforms, guaranteeing that online material can be trusted and protecting the user experience.

## REFERENCES

[1] Patel, R. & Patel, N. (2018). A method for detecting spam in social networks using machine learning algorithms. *3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 534-538.

[2] Alsmadi, I. & Alhami, I. (2018). Clustering and classification of email contents. *Journal of King Saud University - Computer and Information Sciences, 30*(1), 91-99.

[3] Yang, Y., Zhang, R. & Sherratt, R. S. (2020). Social media information trustworthiness evaluation using a multi-feature fusion model based on random forest regression. *Journal of Ambient Intelligence and Humanized Computing, 11*(5), 1925-1938.

[4] Ding, Y., Li, Y., Wu, J., Zhang, Y. & Li, X. (2019). Detecting fake accounts in online social networks at the time of registrations. *Computers & Security, 84*, 304-313.

[5] Zhang, J., Zhang, R., Xia, Y. & Liu, T. (2019). A survey on recent advances in spammer detection on social network. *Journal of Network and Computer Applications, 135*, 32-48.

[6] Kim, S. & Kim, H. (2020). A survey on content-driven detection techniques for illicit content in online social networks. *Journal of Info Processing Systems, 16*(2), 299-313.

[7] El Aziz, M. A., Hassanien, A. E. & Kim, T. H. (2021). Detection of fake accounts on social networks based on users' behavior using deep learning. *Soft Computing, 25*(10), 6925-6934.

[8] Ferrara, E., Varol, O., Davis, C., Menczer, F. & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96-104.

[9] Chavoshi, N., Hamooni, H. & Mueen, A. (2016). DeBot: Twitter Bot Detection via Warped Correlation. *IEEE 16th International Conference on Data Mining (ICDM)*, 817-822.

[10] Stringhini, G., Kruegel, C. & Vigna, G. (2010). Detecting spammers on social networks. *Proceedings of the 26th Annual Computer Security Applications Conference*, pp. 1-9.

[11] Alvisi, L., Clement, A., Epasto, A., Lattanzi, S. Panconesi, A. (2013). Sok: The evolution of sybil defense via social networks. *IEEE Symposium on Security and Privacy*, 382-396.

[12] Beutel, A., Xu, W., Guruswami, V., Palow, C. & Faloutsos, C. (2013). CopyCatch: stopping group attacks by spotting lockstep behavior in social networks. *Proceedings of the 22nd International Conference on World Wide Web*, pp. 119-130.

[13] Lee, K., Caverlee, J. & Webb, S. (2010). Uncovering social spammers: social honeypots + machine learning. *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 435-442.

[14] Fire, M., Goldschmidt, R. & Elovici, Y. (2014). Online social networks: threats and solutions. *IEEE Communications Surveys & Tutorials, 16*(4), 2019-2036.

[15] Gupta, A., Lamba, H., Kumaraguru, P. & Joshi, A. (2013). Faking sandy: characterizing and identifying fake images on Twitter during hurricane sandy. *Proceedings of the 22nd International Conference on World Wide Web*, pp. 729-736.

[16] Agarwal, S. & Sureka, A. (2015). Applying deep learning to structural information for spam detection. *Proceedings of the 24th International Conference on World Wide Web*, pp. 729-734.