F2PMSMD: Design of a Fusion Model to Identify Fake Profiles from Multimodal Social Media Datasets

Bhrugumalla L. V. S. Aditya[1](✉), Gnanajeyaraman Rajaram[2], Shreyas Rajendra Hole[3], and Sachi Nandan Mohanty[1]

[1] School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh, India
aditya.22phd7023@vitap.ac.in

[2] Department of AI & ML, Saveetha School of Engineering, SIMATS, Chennai, India

[3] School of Electronics Engineering (SENSE), VIT-AP University, Amaravati, Andhra Pradesh, India
rajendra.20phd7078@vitap.ac.in

Abstract. Modern-day social media is one of the most used platforms by millennials for sharing personal, and professional events, thoughts & other entities. These entities include photos, texts, videos, locations, meta data about other users, etc. Thus, securing this content from fake-users is of utmost importance, due to which a wide variety of techniques are proposed by researchers that includes but is not limited to, deep learning models, high density feature processing models, bioinspired models, etc. But these models are either highly complex, or require large user-specific datasets in order to improve their detection capabilities. Moreover, most of these models are inflexible, and cannot be scaled for large social networks with multiple parameter sets. To overcome these issues, this text proposes the design of a novel fusion model to identify fake profiles from multimodal social media datasets. The proposed model initially collects multimodal information about users that includes the presence of profile pic, username length ratios, number of words in the full name, length of their personal description, use of external URLs, account type, number of posts, number of followers & following users, etc. These information sets are pre-processed via a Genetic Algorithm (GA) based feature selection model, which assists in the identification of highly variant feature sets. The selected feature sets are classified via a fusion of Naïve Bayes (NB) Multilayer Perceptron (MLP), Logistic Regression (LR), Support Vector Machine (SVM), and Deep Forest (DF) classifiers. Due to a combination of these classifiers, the proposed model is capable of showcasing an accuracy of 98.5%, precision of 94.3%, recall of 94.9%, and F-measure score of 94.7% across multiple datasets. Due to such a high performance, the proposed model is capable of deployment for a wide variety of social media platforms to detect fake profiles.

Keywords: Social · Media · Fake · Profile · Genetic · Algorithm · NB · MLP · LR · SVM · DF · Precision · Recall · Accuracy · F-measure · Fusion · Features

S. Nandan Mohanty et al. (Eds.): ICISML 2022, LNICST 471, pp. 13–23, 2023.
https://doi.org/10.1007/978-3-031-35081-8_2

1 Introduction

Identification of fake profiles in social media is a multidomain task that involves user tracing, context analysis, analysis of user posts, their login patterns, personal profile update patterns, image metadata patterns, etc. A typical social network analysis model (Krishnan et al. 2020) that uses network activity logs, profile attributes, graph analysis, ranking models, network information sets, etc. is depicted in Fig. 1, wherein a threshold-based ranking of profiles is done for identification of normal profiles from fakes. In this model, researchers were able to combine multiple data sources, and extract their nodes, edges, weights, and other parameters for improving classification performance under different scenarios. But these models showcase moderate accuracy, and cannot be scaled for larger network deployments. To overcome these issues, researchers have proposed multiple deep learning and machine learning based classification models, that aim at dynamically modifying their performance based on contextual requirements. A survey of such models (Latha et al. 2022), (Harris et al. 2021), (Patel et al. 2020) in terms of their functional nuances, contextual advantages, application-specific drawbacks, and deployment-specific future scopes is discussed in the next section of this text.

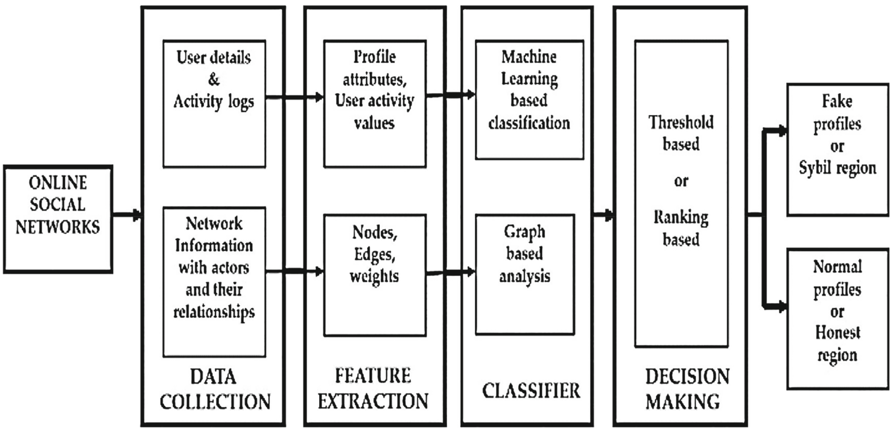

Fig. 1. Design of a threshold-based model for identification of fake profiles

Based on this discussion, it was observed that these models are either highly complex, or require large user-specific datasets in order to improve their detection capabilities. Moreover, most of these models are inflexible, and cannot be scaled for large social networks with multiple parameter sets. To overcome these issues, Sect. 3 of this text proposes design of a novel fusion model to identify fake profiles from multimodal social media datasets. The proposed model was validated in Sect. 4, wherein its performance was compared with other models in terms of accuracy, recall, F-measure, and precision metrics. Finally, this text is concluded with some contextual observations about the proposed model, and also recommends methods to further improve its performance under different network types.

2 Related Work

A wide variety of models are proposed for analysis of social media profiles, and each of them varies in terms of their internal performance characteristics. For instance, work in (Siva Rama Krishna et al. 2022), (Parihar et al. 2021) proposes use of C4.5 Decision Tree (CDT), and Collaborative Filtering that assists in identification of fake profiles for Twitter, but cannot be extended to other networks. To overcome this limitation, work in (Kulkarni et al. 2022) proposes use of Artificial Neural Networks (ANNs) that can be extended to multiple network types. Similar models are discussed in (Bhattacharya et al. 2021), (Ajesh et al. 2021), (Antonio Theophilo, Rafael Padilha, Fernanda A. Andal, 2022) which propose use of Deep Learning, ensemble of Deep Forest, Support Vector Machine (SVM) with Optimized Naïve Bayes (ESDN), and Authorship Attribution sets, which aim at incorporating multiple parameters for improved social media analysis. Extensions to these models are discussed in (Rathod, 2022), (Ekosputra et al. 2021), (Hosseini Moghaddam & Abbaspour, 2022) which propose use of Author Profiling, ANN, and botnet-based detection, which assists in automating the detection process under different network types. Similar methods that use Random Walk (Le et al. 2020), Adversarial Model of Network Disruption (AMND) (Chen and Racz, 2022), and Particle Swarm Optimization with Deep Reinforcement Learning (PSO DRL) (Lingam et al. 2021) that aims at incorporating multiple deep learning models with extended feature processing techniques for identification of fake profiles are also discussed and applied for different use cases. Based on this discussion, it can be observed that these models are either highly complex, or require large user-specific datasets in order to improve their detection capabilities. Moreover, most of these models are inflexible, and cannot be scaled for large social networks with multiple parameter sets. To overcome these issues, next section proposes design of a novel fusion model to identify fake profiles from multimodal social media datasets. The proposed model was also validated on multiple datasets, and compared with various state-of-the-art methods under different scenarios.

3 Data Set Description

The proposed model uses a combination of multimodal information sets along with Genetic Algorithm for feature selection, and ensemble classifier for identification of fake profiles. To validate its performance, the model was tested on the following datasets,

- Social Network Fake Account Dataset, which is available at https://www.kaggle.com/datasets/bitandatom/social-network-fake-account-dataset
- Instagram fake spammer genuine accounts Dataset, which is available at https://www.kaggle.com/datasets/free4ever1/instagram-fake-spammer-genuine-accounts

4 Design of the Proposed Fusion Model to Identify Fake Profiles from Multimodal Social Media Datasets

Following a review of the existing models, it can be observed that these models are either very complex or call for sizable user-specific datasets to enhance their detection capabilities. In addition, the majority of these models lack flexibility and are incapable of scaling to large social networks with multiple parameter sets. The design of a novel fusion model to identify fake profiles from multimodal social media datasets is suggested in this section as a solution to these problems. The proposed model's process is shown in Fig. 2, where it can be seen that it begins by gathering multimodal data about users, such as whether they have profile pictures, how long their usernames are on average, how many words are in their full names, how long their personal descriptions are, whether they use external URLs, what kind of accounts they have, how many posts they have made, how many followers and users they follow, etc. These data sets are pre-processed using a feature selection model based on a Genetic Algorithm (GA), which helps identify feature sets with a high degree of variation. A combination of Naive Bayes (NB), Multilayer Perceptron (MLP), Logistic Regression (LR), Support Vector Machine (SVM), and Deep Forest (DF) classifiers is used to categorize the chosen feature sets.

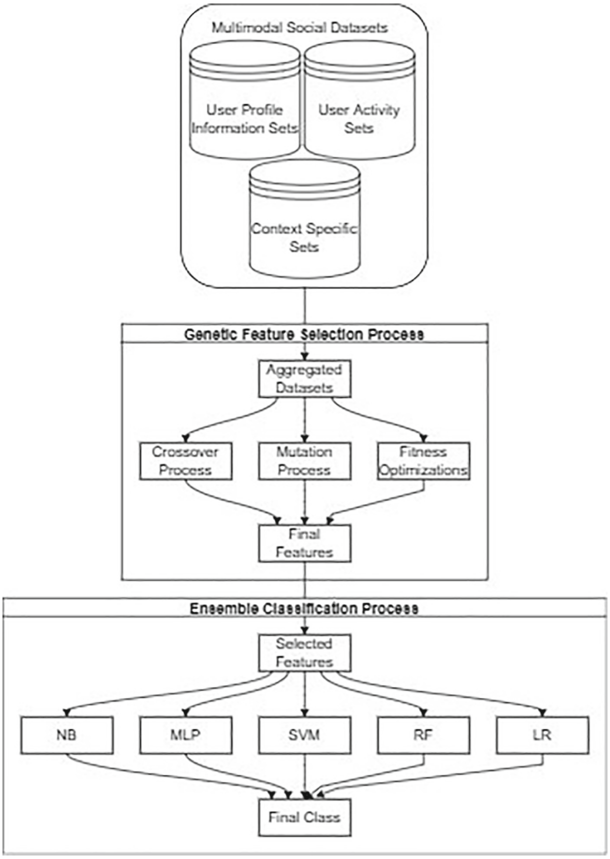

Fig. 2. Overall flow of the proposed fake-profile identification process

The process initially collects multimodal information sets from different sources, and applies a Genetic Algorithm (GA) for the identification of other f optimal feature sets. These feature sets are obtained via the following process,

4.1 Initially, setup the following optimization parameters,
  4.1.1 Iterations needed for GA optimization (*Ni*)
  4.1.2 Solutions to be generated for optimization (*Ns*)
  4.1.3 Rate at which the model will learn (*Lr*)
  4.1.4 Number of features extracted from the multimodal datasets (*Nf*)
4.2 To start the process, generate *Ns* solutions as follows,
  4.2.1 Stochastically select *N* features from the dataset via Eq. 1,

$$\mathrm{N} = \mathrm{STOCH}(\mathrm{Lr} * \mathrm{Nf}, \mathrm{Nf}) \tag{1}$$

where, *STOCH* is a stochastic Markovian process, that generates value sets between given ranges.

4.3 Based on the selected features, identify class-level variance (or fitness of solution) via Eq. 2,

$$f = \sqrt{\frac{\sum_{a=1}^{m}\left(fa - \frac{\sum_{i=1}^{m}\sqrt{\frac{\sum_{j=1}^{n}\left(f_j - \frac{\sum_{a=1}^{n} fa}{n}\right)^2}{n-1}}}{m}\right)^2}{m-1}} \tag{2}$$

where, m, n represents feature values in current class and other classes respectively such that $N = m + n$, while fa represents selected features.

4.4 Once all solutions are generated, then identify optimization fitness threshold via Eq. 3,

$$f_{th} = \sum_{i=1}^{N_s} f_i * \frac{L_r}{N_s} \tag{3}$$

4.5 Check fitness of all solutions, and mark solutions with $f > f_{\mathrm{th}}$ as 'Crossover', while mark others as 'Mutate'
4.6 Once the distinction is done, then scan all solutions for *Ni* iterations, and regenerate 'Mutate' solutions via Eq. 1 and 2
4.7 Repeat this process for *Ni* iterations

At the end of final iteration, select solution with maximum fitness, and use its features for classification process. This classification process is performed via a combination of Naive Bayes (NB), Multilayer Perceptron (MLP), Logistic Regression (LR), Support Vector Machine (SVM), and Deep Forest (DF) classifiers. The parameters used for these classifiers are indicated in Table 1 as follows,

Table 1. Parameters used for different classifiers

Classifier	Parameters
NB	Priors = Interclass Fitness Levels Smoothing = 1e-9
LR	Normalization = True
MLP	Hidden Layers = Number of classes Solver = Stochastic Gradient Descent Learning Rate = *Lr*
SVM	Regularization parameter = 0.5 Kernel = Sigmoid Class Weights = Balanced
DF	Number of trees = 100*Number of classes Maximum Forest Depth = 10

These parameters were selected via 'hit & try' method, which assisted in identification of classification parameters for optimal accuracy levels. Classification results of these classifiers are combined via Eq. 4,

$$C_{out} = C(NB) * A(NB) + C(LR) * A(LR) + C(MLP) * A(MLP) + C(SVM) * A(SVM) + C(DF) * A(DF) \tag{4}$$

where, C & A represents output class, and testing accuracy for given classifiers. Based on this process, classification is performed on different social media sets. Results of this classification, can be observed from the next section of this text.

5 Results and Discussion

The sets were combined to form a total of 2000 entries, out of which 1500 were used for training, while 250 for testing and 250 for validation purposes. The model's performance was compared with CDT (Siva Rama Krishna et al., 2022), ESDN, and PSO DRL (Lingam et al. 2021)in order to identify its efficiency over standard implementations. Based on this strategy, accuracy (A) of fake profile detection w.r.t. Number of Test entries (NT) can be observed from Table 2,

Table 2. Accuracy evaluation for different fake profile detection models

NT	100	200	300	400	500	600	750	900	1000	1200	1500	2000
CDT	87.7	88.31	88.84	89.22	89.45	89.57	89.61	89.65	89.69	89.78	89.95	90.15
ESDN	86.87	87.76	88.32	88.73	88.96	89.2	89.49	89.8	90.23	90.6	90.93	91.25
PSO DRL	86.09	86.83	87.36	87.76	87.99	88.17	88.32	88.5	88.73	88.95	89.18	89.44
F2PM SMD	95.58	96.4	96.99	97.43	97.68	97.89	98.05	98.25	98.52	98.76	99.02	99.31

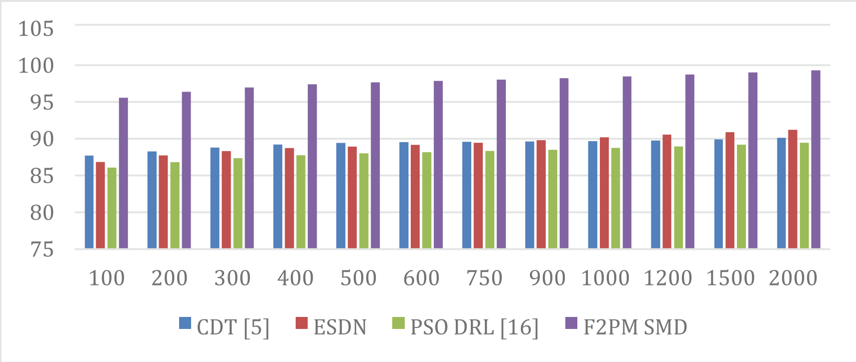

Fig. 3. Accuracy evaluation for different fake profile detection models

Based on this evaluation and Fig. 3, it can be observed that the proposed model showcases 9.2% better accuracy than CDT (Siva Rama Krishna et al. 2022), 8.3% better accuracy than ESDN and 9.5% better accuracy than PSO DRL (Lingam et al. 2021), which makes it highly useful for a wide variety of fake profile detection use cases. This is due to selection of optimal feature sets, that aim at improving classification performance for different scenarios. Similarly, precision (P) of evaluation can be observed from Table 3 as follows,

Table 3. Precision evaluation for different fake profile detection model

NT	100	200	300	400	500	600	750	900	1000	1200	1500	2000
CDT(Siva Rama Krishna et al., 2022)	87.7	88.31	88.84	89.22	89.5	89.57	89.61	89.7	89.7	89.78	89.95	90.2
ESDN	86.87	87.76	88.32	88.73	89	89.2	89.49	89.8	90.2	90.6	90.93	91.3
PSO DRL(Lingam et al., 2021)	86.09	86.83	87.36	87.76	88	88.17	88.32	88.5	88.7	88.95	89.18	89.4
F2PM SMD	95.58	96.4	96.99	97.43	97.7	97.89	98.05	98.3	98.5	98.76	99.02	99.3

Based on this evaluation and Fig. 4, it can be observed that the proposed model showcases 8.5% better precision than CDT (Siva Rama Krishna et al. 2022), 8.3% better precision than ESDN and 9.4% better precision than PSO DRL (Lingam et al. 2021), which makes it highly useful for a wide variety of fake profile detection use cases. This is due to selection of optimal feature sets and use of ensemble learning, that aim at improving classification performance for different scenarios. Similarly, recall (R) of evaluation can be observed from Table 4 as follows,

Based on this evaluation and Fig. 5, it can be observed that the a) of evaluation can be observed from Table 5 as follows,

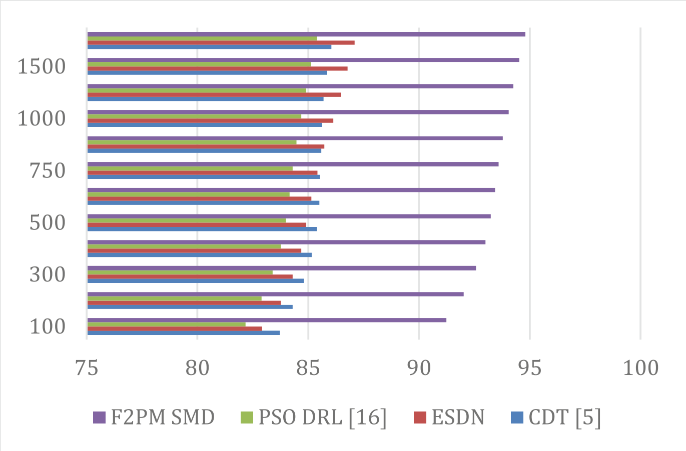

Fig. 4. Precision evaluation for different fake profile detection models

Table 4. Recall evaluation for different fake profile detection models

NT%		100	200	300	400	500	600	750	900	1000	1200	1500	2000
R	**CDT** (Siva Rama Krishna et al., 2022)	85.14	85.72	86.24	86.61	86.83	86.95	86.98	87.03	87.07	87.15	87.31	87.51
R	**ESDN**	84.32	85.19	85.73	86.13	86.35	86.59	86.87	87.17	87.59	87.94	88.26	88.57
R	**PSO DRL** (Lingam et al., 2021)	83.57	84.29	84.8	85.19	85.41	85.58	85.73	85.9	86.13	86.34	86.57	86.82
R	**F2PM SMD**	92.78	93.58	94.15	94.57	94.82	95.02	95.18	95.38	95.63	95.87	96.12	96.4

Table 5. F-measure evaluation for different fake profile detection models

NT(%)		100	200	300	400	500	600	750	900	1000	1200	1500	2000
F	**CDT** (Siva Rama Krishna et al., 2022)	84.42	85	85.51	85.88	86.1	86.22	86.25	86.3	86.34	86.4	86.6	86.77
F	**ESDN**	83.62	84.47	85.01	85.4	85.63	85.86	86.14	86.4	86.86	87.2	87.5	87.83
F	**PSO DRL** (Linga m et al., 2021)	82.86	83.58	84.09	84.47	84.7	84.86	85.01	85.2	85.4	85.6	85.8	86.09
F	**F2PM SMD**	92.15	92.8	93.36	93.78	94.02	94.22	94.38	94.6	94.83	95.1	95.3	95.59

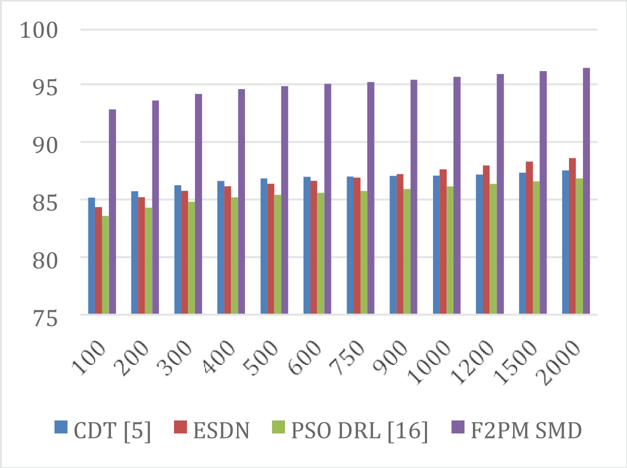

Fig. 5. Recall evaluation for different fake profile detection models

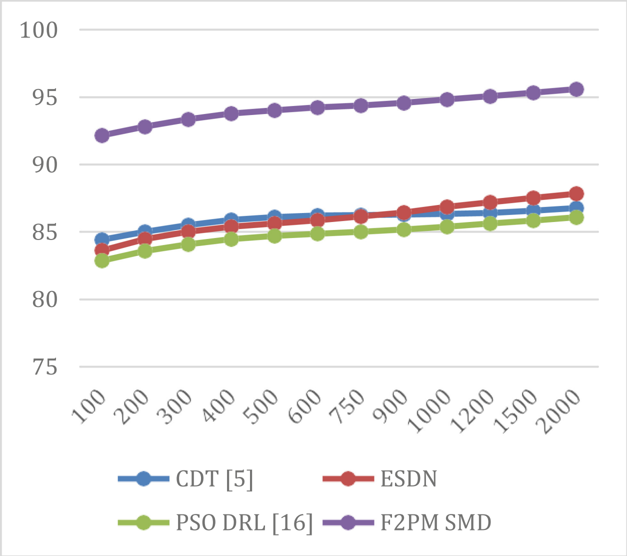

Fig. 6. F-measure evaluation for different fake profile detection models

Based on this evaluation and Fig. 6, it can be observed that the proposed model showcases 9.8% better F-measure than CDT (Siva Rama Krishna et al. 2022), 8.5% better F-measure than ESDN and 9.5% better F-measure than PSO DRL (Lingam et al. 2021), which makes it highly useful for a wide variety of fake profile detection use cases. This is due to use of high efficiency classifiers & selection of optimal feature sets, that aim at improving classification performance for different scenarios. Due to this performance enhancements, the proposed model is capable of use for a wide variety of real- time social media fake profile detection scenarios.

6 Conclusion and Future Scope

The suggested model combines multiple multimodal data sets, genetic algorithms for feature selection and ensemble classifiers for fake profile detection. Because of this, the model was able to demonstrate accuracy that was 9.2% better than CDT (Siva Rama Krishna et al. 2022), 8.3% better than ESDN, and 9.5% better than PSO DRL (Lingam et al. 2021), making it extremely useful for a variety of fake profile detection use cases. This is because the best feature sets were chosen, which aims to enhance classification performance for various scenarios. It was also able to achieve precision levels that were 8.5% better than CDT (Siva Rama Krishna et al. 2022), 8.3% better than ESDN, and 9.4% better than PSO DRL (Lingam et al. 2021), making it extremely beneficial for a variety of fake profile detection use cases. The model also demonstrated recall improvements of 8.3% over CDT (Siva Rama Krishna et al. 2022), 7.5% over ESDN, and 9.5% over PSO DRL (Lingam et al. 2021), making it extremely useful for a variety of fake profile detection use cases. This is because ensemble classifiers are used and the best feature sets are chosen with the intention of enhancing classification performance in various scenarios. The model demonstrated an F-measure that was 9.8% better than CDT (Siva Rama Krishna et al. 2022), 8.5% better than ESDN, and 9.5% better than PSO DRL (Lingam et al. 2021), making it extremely useful for a variety of fake profile detection use cases. This is a result of the use of highly effective classifiers and the choice of the best feature sets, which aims to enhance classification performance for various scenarios. The proposed model can be used for a wide range of real-time social media fake profile detection scenarios as a result of these performance improvements. In future, researchers can extend the model's performance by integration of multiple bioinspired techniques for feature extraction & classification processes. The model's performance must be validated on larger network sizes, and can be enhanced via use of hybrid deep learning techniques like Q-Learning, Generative Adversarial Networks (GANs), etc. which will assist in accuracy enhancement under different use cases.

References

Ajesh, F., Aswathy, S.U., Philip, F.M., Jeyakrishnan, V.: A hybrid method for fake profile detection in social networkusing artificial intelligence. Security Issues and Privacy Concerns in Industry 4.0 Applications, 89–112 (2021). https://doi.org/10.1002/9781119776529.ch5

Theophilo, A., Padilha, R., Fernanda, A.A., Rocha, R.: Explainable Artificial Intelligence For Authorship Attribution On Social Media Artificial Intelligence Lab . (Recod . ai) Institute of Computing , University of Campinas , Br. ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2909–2913 (2022)

Bhattacharya, A., Bathla, R., Rana, A., Arora, G.: Application of machine learning techniques in detecting fake profiles on social media. In: 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2021, pp. 1–8 (2021). https://doi.org/10.1109/ICRITO51393.2021.9596373

Chen, M.F., Racz, M.Z.: An adversarial model of network disruption: maximizing disagreement and polarization in social networks. IEEE Trans. Netw. Sci. Eng. **9**(2), 728–739 (2022). https://doi.org/10.1109/TNSE.2021.3131416

Ekosputra, M.J., Susanto, A., Haryanto, F., Suhartono, D.: Supervised machine learning algorithms to detect instagram fake accounts. In: 2021 4th International Seminar on Research of Information Technology and Intelligent Systems. ISRITI 2021, pp. 396–400 (2021). https://doi.org/10.1109/ISRITI54043.2021.9702833

Harris, P., Gojal, J., Chitra, R., Anithra, S.: Fake instagram profile identification and classification using machine learning. In: 2021 2nd Global Conference for Advancement in Technology, GCAT 2021, pp. 1–5 (2021). https://doi.org/10.1109/GCAT52182.2021.9587858

Hosseini Moghaddam, S., Abbaspour, M.: Friendship preference: scalable and robust category of features for social bot detection. IEEE Trans. Depend. Sec. Comput. **5971**(c), 1–14 (2022). https://doi.org/10.1109/TDSC.2022.3159007

Krishnan, P., John Aravindhar, D., Reddy, P.B.P.: Finite automata for fake profile identification in online social networks. In: Proceedings of the International Conference on Intelligent Computing and Control Systems (2020)

ICICCS 2020, Iciccs, 1301–1305. https://doi.org/10.1109/ICICCS48265.2020.9121086

Kulkarni, V., Aashritha Reddy, D., Sreevani, P., Teja, R.N.: Fake profile identification using ANN, pp. 375–380 (2022). https://doi.org/10.1049/icp.2022.0372

Latha, P., Sumitra, V., Sasikala, V., Arunarasi, J., Rajini, A. R., Nithiya, N. Fake profile identification in social network using machine learning and NLP. In: 2022 International Conference on Communication, Computing and Internet of Things, IC3IoT 2022 - Proceedings, pp. 20–23 (2022). https://doi.org/10.1109/IC3IOT53935.2022.9767958

Le, N.C., Dao, M.T., Nguyen, H.L., Nguyen, T.N., Vu, H.: An application of random walk on fake account detection problem: a hybrid approach. Proceedings - 2020 RIVF International Conference on Computing and Communication Technologies, RIVF 2020 (2020). https://doi.org/10.1109/RIVF48685.2020.9140749

Lingam, G., Rout, R.R., Somayajulu, D.V.L.N., Ghosh, S.K.: Particle swarm optimization on deep reinforcement learning for detecting social spam bots and spam-influential users in twitter network. IEEE Syst. J. **15**(2), 2281–2292 (2021). https://doi.org/10.1109/JSYST.2020.3034416

Parihar, P., Devanand, Kumar, N.: Fake profile detection from the social dataset for movie promotion. In: Proceedings of the IEEE International Conference Image Information Processing, 2021-Novem, 495–498 (2021). https://doi.org/10.1109/ICIIP53038.2021.9702684

Patel, K., Agrahari, S., Srivastava, S. Survey on fake profile detection on social sites by using machine learning algorithm. In: ICRITO 2020 - IEEE 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), 1236–1240(2020). https://doi.org/10.1109/ICRITO48877.2020.9197935

Rathod, S.: Exploring Author Profiling for Fake News Detection, pp. 1614–1619 (2022). https://doi.org/10.1109/compsac54236.2022.00256

Siva Rama Krishna, S., Umakanth, K., Anji Reddy, T., Saiteja, A., Sumanjali, R.: Detection of fake and clone accounts in twitter using classification and distance measure algorithms. In: Reddy, A.N.R., Marla, D., Favorskaya, M.N., Satapathy, S.C. (eds.) Intelligent Manufacturing and Energy Sustainability. SIST, vol. 265, pp. 391–399. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-6482-3_39