International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 04 April 2025
Page No: 1514 - 1520
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0216

An Ensemble ML Classifier-based Fake Profile Identification Framework in Online Social Network

R Anieshma[1], S Mahalakshmi[2], S Lakshmi Subhagini[3], S Jegathambigai[4], K Sabitha Shree[5]
[1]PG Scholar, Dept. of CSE Jai Shriram Engineering College, Tirupur, Tamil Nadu, India.
[2,3,4,5]UG Scholar, Dept. of CSE Jai Shriram Engineering College, Tirupur, Tamil Nadu, India.
Emails: *anieshmaramesh@gmail.com[1], mahashankarraj2004@gmail.com[2], subhasenapathi@gmail.com[3], jegathambigaisivakumar@gmail.com[4], shreesabitha24@gmail.com[5].*

Abstract

Typically, online social networks include huge range of people all around the world and this becomes a huge part of their life. People exploit social networks for sharing their feelings so as to make friends, for setting up new businesses, thus to connect with family, friends & so on. Online social network offers huge benefits to individuals in varied ways but it too suffers with few limitations. There exist several people who exploit these networks for making harm to others on creating fake accounts on these networks. To detect such fake and genuine users, machine learning (ML) algorithm is used in this work. The ML models are employed for predicting and classifying datasets over three ensemble ML approaches employed. It becomes complex sometimes to differentiate among the outcomes of varied schemes and for this reason, an ensemble ML model is employed to make the task much easier and to enhance the accuracy. In this work, three different classification algorithms termed Random forest (RF), Logistic Regression (LR), and Support Vector Machine (SVM) are used by estimating their accuracy on Twitter dataset acquired from Kaggle repository. The results of these models attained are compared and is concluded that the best accuracy is attained from RF classifier model on comparing other two classifier employed.

Keywords: *Online social network, fake profile identification, twitter dataset, ML algorithm, Logistic regression, Random forest, Support vector machine classifier.*

1. Introduction

Social media is growing exponentially in a daily basis. Every day, millions of users are joining social media platforms. Varied people are exploiting social media for varied purpose for instance some people exploit them for business purposes. Social network becomes very usual as we known, since the platforms are exploited by around 90% of people worldwide by using information retrieval system [1]. Since, varied peoples are using them for varied purposes, there were some hackers or unauthorized people exploit these platforms for wrong purposes using fake profiles with the fake pictures using another person for spreading false news, scamming peoples, and for several other wrong purposes [2]. Several thousand cases are reported on internet in a daily basis. Varied platforms work on these issues for stopping abuse or unauthorized activities on implementing varied security examinations. Likewise, it is a serious issue since it is an attack on personal property of someone else [3]. It was recently reported that hackers from cyber attacker scientists using fake women pictures.

Since, it is known that privacy in internet is a challenging issue nowadays, and misuse of some other profile is increasing exponentially [4]. Likewise, by the social media profiles, this becomes easy for hackers for creating and to theft some other personal information which includes email address, pictures and so on. Fake or false profiles are using false credentials so as to create its profiles [5]. Mostly, these profiles are involving widely in malicious actions and undesirable actions which affect directly by creating issue for the social community customers [6]. Varied existing techniques are working on it already for controlling these false profiles. Facebook have created a system of security for giving privacy to their user from spamming, phishing, and other hackers as it is the largest social media platform having billions of users [7]. This makes them a challenging aspect to offer security for their users. So as to provide security for their user, a security system named Facebook Immune system was created [8]. However, it is not properly implemented yet and is not working to observe false profiles and it seems to be useless at this condition. Since, the existing fake profile detection methods are not that much successful, there is a need to create solution based on big data by covering new technologies and algorithms to predict and recognize fake and real profiles. For this purpose, the proposed model is implemented.

1.1 Research gap

The modeling of fake profile identification system is considered an existing issue because of several challenges likes reduced accuracy, lack of extracting relevant features, and use of single ML algorithm and due to this reason, this issue still exists with lot of research gaps which has been identified and needs to be resolved upon.

1.2 Research Objective

The major intention of proposed work is to detect such fake and genuine users using ensemble machine learning (ML) algorithm. The ML models are employed for predicting and classifying datasets over three ensemble ML approaches employed like SVM, LR, and RF. To compare the proposed model's accuracy for validating the efficiency of proposed algorithm.

2. Related Works

In the work [9], a spam detection artificial intelligence model was suggested for Twitter social network. In this model, a model of SVM, a neural network, and random forest scheme was implemented for building a scheme. The outcomes of this model indicate that on comparing RF and ANN models, SVM is offering enhanced outcomes on metrics like recall, F-measure, and precision. The article presented in [10] aims at summarizing recent advancements in the detection methodology of fake account on the social networking website. A primary focus of detecting fake account was to spread rumor, spam contents, and other unauthorized information's on platform. Therefore, it is essential to filter fake accounts, however it has several challenges. A technology of fake account detection was summarized in this article by discussing limitations and challenges of traditional schemes. An automatic fake news detection model in chrome environment was proposed in the work [11] through which it could detect fake news on Facebook. More specifically, multiple features were employed associated with the Facebook account with some features of news content over deep learning model. An experimental outcome in real-world information denotes that proposed model attains high accuracy on comparing previous works. The main intention of the work [12] was to detect fake account on analyzing varied characteristic which spreads malicious contents in real-world environment. In this work, chrome extension dependent framework was discussed which detects fake accounts in the twitter environment on analyzing varied features. The work suggested in [13] offers a solution for existing problems and is a Spam ML dependent work which predicts fake profiles of high accuracy than other current models employed for profile detection. This model comprises of Spark ML libraries which includes RF classifier and other such plotting tool. The model diagram was given and results were depicted in graphical form such as confusion matrix, learning curve, ROC plot for better knowledge. A deep learning based fake news detection was suggested in the work [14] which focus on survey concerning specified topics for multimodal fake news detection on social media. The survey offers, an explanation and review on relevant features like DL model employed, their accessed data, and fusion strategy employed. In the work [15], the author aims at checking and detecting fake and trusted profiles that were created in OSN over ML dependent models. On employing SVM, DNN, and RF to treat missing values, variable identification, validation and cleaning of data was carried on entire dataset. In this python language was employed for improving the code efficiency. By using this, millions of fake profiles could be automatically detected.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 04 April 2025
Page No: 1514 - 1520
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0216

3. Proposed Work

The proposed model is described briefly in this section. Initially, preprocessing was carried to clean the data. The feature extraction process is then carried for extracting significant and relevant features so as to enhance the accuracy of classifier. The features extracted are fed to classifier for attaining outcome as predicted real or fake profiles. The ensemble classification algorithms termed SVM, RF, and LR were used for prediction purpose. The entire working module is depicted in figure 1.

3.1 Input data Processing

As an initial step, preprocessing is carried for attaining better outcome from any such ML model and this data processing was used for cleaning data from raw data and to import libraries that are usually numpy, panda, scikit-learn. For this purpose, preprocessing is carried to import and clean data. Once the data is cleaned, filtered and imported, the next step is to carry feature extraction process.

3.2 Feature Extraction

The next step of the proposed framework is to extract features from the preprocessed data. The columns like 'statuses count', 'count of followers', 'count of friends', 'listed counts', 'favorites count', and 'gender' are retrieved in the process of feature extraction. It will be utilized for determining the accuracy of profile data. The feature extraction and on describing columns might give the output regarding the total counting of every column in the data, standard deviation, mean, maximum and minimum number.

The process of feature extraction is employed for determining the optimal range of features to create the model on eliminating inappropriate and redundant features thus concentrating simply on needed features. The feature's language code is of string type which could be converted to integer. Once the extract feature function is called, it prints the features extracted in summary on printing quartile, mean, std, min, max, count, and so on. The intention of this is to minimize the training time for the detection scheme on reducing over-processing thus to enhance the generalizability of the model and to aid researchers in interpreting the scheme.

3.3 Ensemble Machine Learning Classifier

For the purpose of classification ensemble machine learning algorithm is employed. In this, Logistic regression (LR), SVM (support vector machine), and RF (random forest) are ensembled to predict the fake profile in online social network.

3.3.1 SVM Classifier

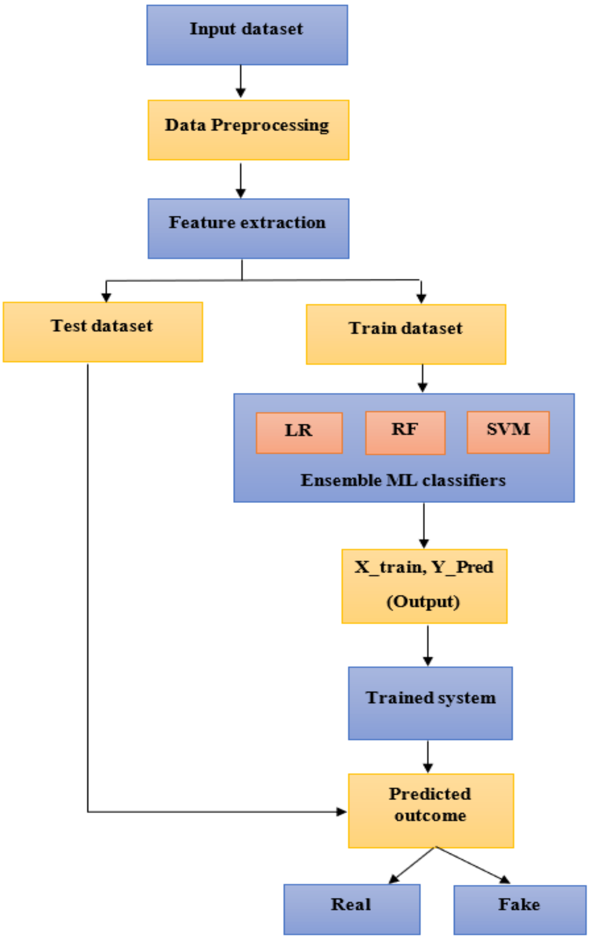


Figure 1 Working Flow of Proposed Method

3.3.2 RF Classifier

RF is considered a powerful ML approach which performs classification and regression of tasks. A basic building block of RF is derived from decision tree. The scheme is attained on considering data as bootstrapping samples based on number of trees which needs to be performed, on building a simple detection scheme in every section thus combining the outputs depending on ensemble learning model for getting final output.

3.3.3 Logistic Regression (LR) Classifier

Algorithm 1: Ensemble ML classifier algorithm

```
Input: data from CSV file

Output: predicted outcome, accuracy

Step 1: Read dataset

Step 2: Feature extarction: Covert non-integer features to integer. Store & overwrite the features listed in X.

Return X

Step 3: Split the data as training and testing and store them in x_train, x_test, y_train, y_test.

Step 4: Scale X data (x_train, x_test)

Step 5: Employ ensemble classifier SVM, RF, and LR

Step 6: Store outcome in y_pred vaiable and

Return y_pred

Step 7: Repeat step 3 with y_pred and x_test

Step 8: Output attained from step 3 is given to classifier and then store outcome in y_pred

Step 9: Testing

Step 10: Print classification accuracy on testing data

Step 11: Exit
```

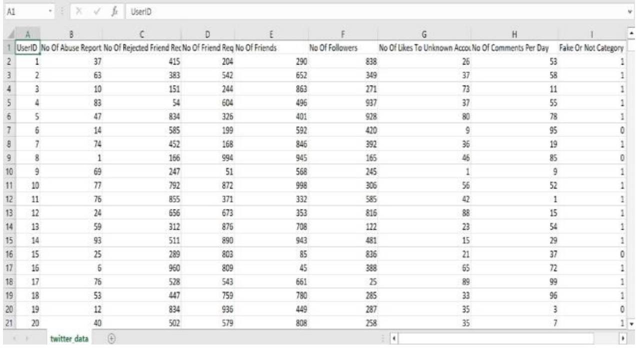
LR is one such ML model employed for binary classification purpose. This is simple and typically employed model which estimates the relationship among one and several variable dependent variables which needs to be predicted. Since it employs logistic function for estimating probabilities it must be converted to binary values, termed as sigmoid function task. This sigmoid function is a curve in S form which could take any number of real values and thus placing them in range among one and zero. In suggested model, initial algorithm level includes RF, SVM, and LR are trained on training set, and is validated by collecting expected values from entire level of algorithms for using them as input meta classifier. The entire steps are employed for generating predictions on the test set. The accounts available in test suite are thus classified as fake and real accounts depending on training suit offered. The data is then divided as testing and training group this choosing and splitting 75% for training and 25% for testing group so as to ensure the equal division and maintaining simple classes proportion. Default classifier parameters were exploited, by changing random state parameter to (one) for each of the train-test split and for RF model for having acceptable and steady results. The prediction of each of the three employed models is carried using dataset. The classifiers are employed with 10-fold validation so that each time the classifier is trained for nine parts and is being tested on tenth part. The process of training and testing are then replicated ten times. After this, predictions are fed to LR classifier for creating prediction class as fake or real.

4. Performance Analysis

The performance assessment of the proposed scheme is given in this section.

4.1 Dataset Description

The data collection is carried on collecting them from twitter platforms and is created by means of Crawler. The datasets are collected over online from well-known Kaggle websites. The model is tested on dataset collected from Kaggle from which two CSV files are used which corresponds to genuine and fake users. Figure 2 given here shows the CSV file sample.



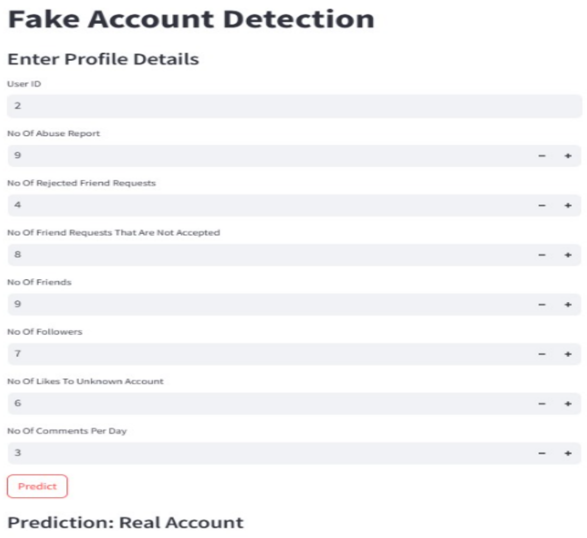
A1 UserID

UserID	No Of Abuse Report	No Of Rejected Friend Rec	No Of Friend Req	No Of Friends	No Of Followers	No Of Likes To Unknown Accou	No Of Comments Per Day	Fake Or Not Category
1	37	415	204	290	838	26	53	1
2	63	383	542	652	349	37	58	1
3	10	151	244	863	271	73	11	1
4	83	54	604	496	937	37	55	1
5	47	834	326	401	928	80	78	1
6	14	585	199	592	420	9	95	0
7	74	452	168	846	392	36	19	1
8	1	166	994	945	165	46	85	0
9	69	247	51	568	245	1	9	1
10	77	792	872	998	306	56	52	1
11	76	855	371	332	585	42	1	1
12	24	656	673	353	816	88	15	1
13	59	312	876	708	122	23	54	1
14	93	511	890	943	481	15	29	1
15	25	289	803	85	836	21	37	0
16	6	960	809	45	388	65	72	1
17	76	528	543	661	25	89	99	1
18	53	447	759	780	285	33	96	1
19	12	834	936	449	287	35	3	0
20	40	502	579	808	258	35	7	1

Figure 2 Sample of CSV File

4.2 Performance Analysis of Proposed Strategy

The performance estimation carried for proposed model is given and is projected in this section. The results attained on evaluating with considered dataset with their corresponding output are given.



Fake Account Detection

Enter Profile Details

User ID
2

No Of Abuse Report
9

No Of Rejected Friend Requests
4

No Of Friend Requests That Are Not Accepted
8

No Of Friends
9

No Of Followers
7

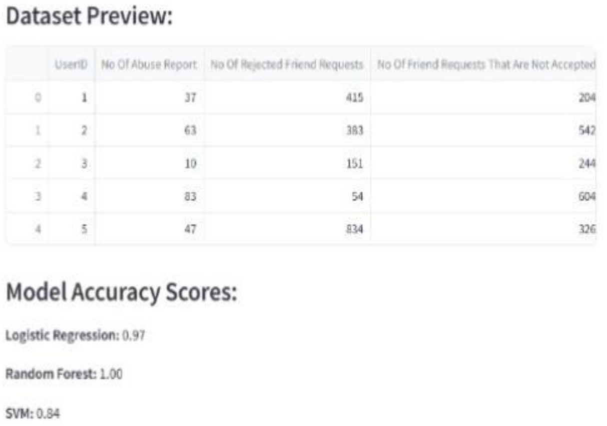
No Of Likes To Unknown Account
6

No Of Comments Per Day
3

Prediction: Real Account

Figure 3 Attained Outcome on Manual Input

Figure 3 is the attained result on employing manual input. The input is given manually for columns denoting user id, number of abuse report, rejected friend requests, requests that are not accepted, number of friends, number of followers, number of likes to unknown accounts, number of comments per day. The predicted outcome will be attained as real or fake.



Dataset Preview:

	UserID	No Of Abuse Report	No Of Rejected Friend Requests	No Of Friend Requests That Are Not Accepted
0	1	37	415	204
1	2	63	383	542
2	3	10	151	244
3	4	83	54	604
4	5	47	834	326

Model Accuracy Scores:

Logistic Regression: 0.97

Random Forest: 1.00

SVM: 0.84

Figure 4 Outcome Attained on CSV File Upload

Figure 4 is the results attained on uploading CSV file and predicting the fake or real accounts thus estimating the accuracy of each algorithm attained. The accuracy of LR is attained as 0.97, RF (1.00), and 0.84 for SVM model.

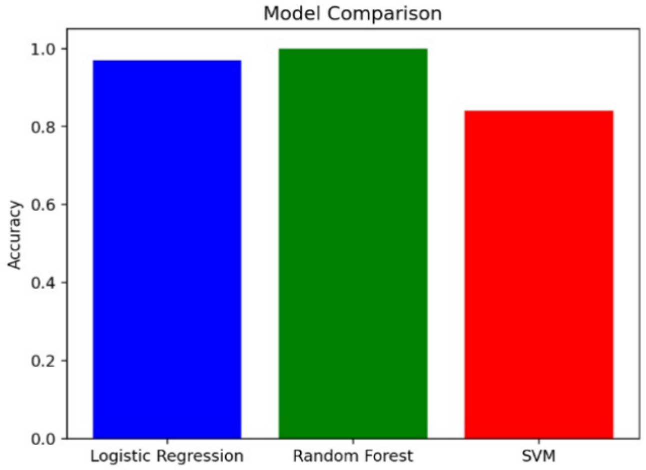


Figure 5 Accuracy Output Comparison on Three ML Algorithms Used

Figure 5 is the output comparison of three ML algorithms employed. The results attained for LR, RF, and SVM is graphically presented and the outcome shows that the RF classifier is better in offering enhanced accuracy than other two classifier employed. From the overall analysis, it is obvious that the proposed scheme offers better accuracy and predicting process is effective. Among three algorithms employed, it is observed that RF is giving enhanced outcomes than other two schemes.

Conclusion

In this paper, fake profile identification on online social network using twitter dataset was employed with the use of ensemble ML approaches like RF, SVM, and LR. Initially, preprocessing was carried to clean the data. The feature extraction process is then carried for extracting significant and relevant features so as to enhance the accuracy of classifier. The features extracted are fed to classifier for attaining outcome as predicted real or fake profiles. The ensemble classification algorithms termed SVM, RF, and LR were used for prediction purpose and efficiency of this algorithms are tested by conducting experimentation of input twitter dataset acquired from Kaggle repository. From the overall analysis, it is obvious that the proposed scheme offers better accuracy and predicting process is effective. Among three algorithms employed, it is observed that RF is giving enhanced outcomes than other two schemes

References

[1]. Shreya, K., Kothapelly, A., & Shanmugasundaram, H. (2022, December). Identification of Fake accounts in social media using machine learning. In 2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT) (pp. 1-4). IEEE.

[2]. Jain, V., Kumar, V., Pal, V., & Vishwakarma, D. K. (2021, April). Detection of cyberbullying on social media using machine learning. In 2021 5th international conference on computing methodologies and communication (ICCMC) (pp. 1091-1096). IEEE.

[3]. Xu, T., Goossen, G., Cevahir, H. K., Khodeir, S., Jin, Y., Li, F., ... & Pearce, P. (2021). Deep entity classification: Abusive account detection for online social networks. In 30th {USENIX} Security Symposium ({USENIX} Security 21).

[4]. Kanagavalli, N., & Priya, S. B. (2022, March). Design of hyperparameter tuned deep learning based automated fake news detection in social networking data. In 2022 6th International Conference on Computing Methodologies and Communication (ICCMC) (pp. 958-963). IEEE.

[5]. Heidari, M., James Jr, H., & Uzuner, O. (2021, April). An empirical study of machine learning algorithms for social media bot detection. In 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS) (pp. 1-5). IEEE.

[6]. Goyal, B., Gill, N. S., Gulia, P., Prakash, O., Priyadarshini, I., Sharma, R., ... & Yadav, K. (2023). Detection of fake accounts on social media using multimodal data with deep learning. IEEE Transactions on Computational Social Systems.

[7]. Bharti, N. S. G., & Gulia, P. (2023). Exploring machine learning techniques for fake profile detection in online social networks. International Journal of Electrical and Computer Engineering (IJECE), 13(3), 2962-2971.

[8]. Boahen, E. K., Bouya-Moko, B. E., Qamar, F., & Wang, C. (2022). A deep learning approach to online social network account compromisation. IEEE Transactions on Computational Social Systems, 10(6), 3204-3216.

[9]. Prabhu Kavin, B., Karki, S., Hemalatha, S., Singh, D., Vijayalakshmi, R., Thangamani, M., ... & Adigo, A. G. (2022). Machine learning-based secure data acquisition for fake accounts detection in future mobile communication networks. Wireless Communications and Mobile Computing, 2022(1), 6356152.

[10]. Roy, P. K., & Chahar, S. (2021). Fake profile detection on social networking websites: a comprehensive review. IEEE Transactions on Artificial Intelligence, 1(3), 271-285.

[11]. Sahoo, S. R., & Gupta, B. B. (2021). Multiple features based approach for automatic fake news detection on social networks using deep learning. Applied Soft Computing, 100, 106983.

[12]. Sahoo, S. R., & Gupta, B. B. (2021). Real-time detection of fake account in twitter using machine-learning approach. In Advances in Computational Intelligence and Communication Technology: Proceedings of CICT 2019 (pp. 149-159). Springer Singapore.

[13]. Awan, M. J., Khan, M. A., Ansari, Z. K., Yasin, A., & Shehzad, H. M. F. (2022). Fake profile recognition using big data analytics in social media platforms. International Journal of Computer Applications in Technology, 68(3), 215-222.

[14]. Comito, C., Caroprese, L., & Zumpano, E. (2023). Multimodal fake news detection on social media: a survey of deep learning techniques. Social Network Analysis and Mining, 13(1), 101.

[15]. Soumya, T. R., Manohar, S. S., Ganapathy, N. B. S., Nelson, L., Mohan, A., & Pandian, M. T. (2022, September). Profile similarity recognition in online social network using machine learning approach. In 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 805-809). IEEE.