International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET2310264

Fake Profile Detection Using Machine Learning

K. Harish, R. Naveen Kumar, Dr. J. Briso Becky Bell
Department of Information Technology, Kings Engineering College, Sriperumbudur, Tamilnadu, India

ARTICLEINFO

Article History:
Accepted: 10 April 2023
Published: 29 April 2023

Publication Issue
Volume 10, Issue 2
March-April-2023

Page Number
719-725

ABSTRACT

Platforms for social media like Facebook, Twitter, Instagram, and others have a big impact on our lives. All across the world, people are actively engaged in it. But, it also needs to address the problem of false profiles. Fake accounts are regularly made by people, software, or machines. They are employed in the spread of rumors and illegal actions like phishing and identity theft. This project uses several machine learning techniques to discriminate between fake and authentic Twitter profiles based on characteristics such as follower and friend counts, status changes, and more. Twitter profile dataset, classifying genuine accounts as TFP and E13 and fake accounts as INT, TWT, and FSF. In this section, the author talks about neural networks, LSTM, XG Boost, and Random Forest. The important traits are picked to judge the veracity of a social media page. The architecture and hyperparameters are also discussed. Lastly, after the models have been trained, results are generated. As a result, the output is 0 for true profiles and 1 for fake profiles. It is possible to disable or delete a fake profile when it is found, preventing cyber security issues.

Keywords: Social Media - World - Rumours - Fake Profiles - Detection

I. INTRODUCTION

In the current generation, everyone's social life is now entwined with online social networks. It has been simpler to add new friends and stay in touch with them and their updates. Online social networks influence a variety of fields, including research, education, community activism, employment, and business. These online social networks have been the subject of research to see how they affect people. Instructors may quickly contact their students using this, creating a welcoming atmosphere for them to learn. Teachers are becoming more familiar with these sites and using them to provide online classroom pages, assign assignments, hold conversations, and other activities that greatly enhance learning. Employers may utilize these social networking sites to find and hire skilled individuals who are enthusiastic about their jobs.

In addition to all of this, propaganda spread via social media is a problem. False accounts sharing incorrect and unsuitable information cause conflicts. The following are the main goals of this research project: These false profiles are likewise created to get

followers. More people are harmed by phony profiles than by other internet crimes. Thus, now that the user is aware, it's crucial to recognize a fake profile. False accounts on the website are mostly used to spread spam, information, and other misleading information. The technological work that has been done and what is being done right now to identify false profiles is examined in this paper.

The majority of fraudulent profiles are created with the intention of spamming, phishing, and getting more followers. Fake accounts have all the tools necessary to conduct online crimes. Identity theft and data breaches are substantial risks posed by fake accounts. All user information is transferred to faraway servers when users view the URLs supplied by these fake accounts, where it may be utilized against them. False accounts that claim to have been made on behalf of organizations or individuals can also harm their reputation and reduce the number of followers and likes they receive.

II. LITERATURE SURVEY

Several techniques were used to categorize profiles based on account activity, the number of requests that were answered, the number of messages that were sent, and other factors. The models are based on graph-based systems. Others have made an effort to differentiate between robots and cyborgs using particular methods. A list of some prior investigations is provided below. If particular terms are contained in a message, it is considered spam. This hypothesis has been used to spot fake social media profiles. These terms were located on social media using pattern-matching techniques. This criterion, however, suffers greatly from the frequent invention and usage of new terminology.

Sybil Guard [3] was developed in 2008 to lessen the negative effects of Sybil's attacks on social media. The frequency of walk-random encounters was constrained, and the dataset was the random walk of each node in Kleinberg's artificial social network. A different tactic known as the Sybil limitation was developed at about the same time as the Sybil guard. The Sybil guard functions on the same principle, except the zone outside of Sybil, quickly combines.

Each node employed a strategy that utilized several random factors to make it function. In addition, the ranking was based on how often walk intersection tails occurred. Sybil-infer was developed in 2009. It employs techniques such as model-based sampling, greedy algorithms, and Bayesian networks under the assumption that randomized walks and the non-Sybil region are swiftly integrated. A probability-based selection technique is threshold selection. Mislove's algorithm from 2010 used greedy search to choose Facebook dataset profiles based on metric-adjusted conductivity.

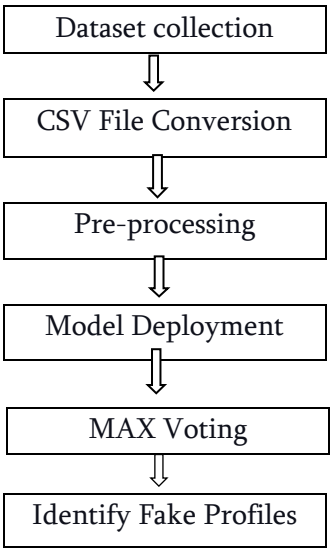
Facebook utilizes an algorithm to identify bots based on how many of your friends may have tags or connection histories. The aforementioned guidelines can be used to identify bot accounts, however, they are insufficient to identify fraudulent accounts that have been created by humans. Machine learning without supervision is used to detect bots. Information was compiled using this technical technique based on proximity rather than tagging. Co-attributes allowed grouping functions to effectively separate the bots. The Sybil rank regression method [3] [4] was developed in 2012. The profiles are arranged by interaction, tagging, and wall postings True accounts have a better ranking than false accounts, which are ranked lower. Yet this method was unreliable.

Because occasionally a genuine profile, even when it was fantastic, would get a bad rating. The Sybil frame was the second model to be produced. It employed several levels of categorization. It worked in two stages, initially using a technique based on conventions and then switching to one based on structures. These techniques have been applied in some recent studies on this topic. One of the earlier research authors [5]

created a blacklist that can distinguish between artificial characteristics and bogus accounts. Research [6] proposes Deep Profile, a technique that uses a supervised learning algorithm to identify bogus accounts, using dynamic CNN as the framework. To identify the OSN phony account, the authors of research [7] integrated the SVM, RF, and Adaboost. Regression analysis and the random forest classifiers method were specially employed in research [8] to identify fake Instagram accounts. Many authors produce a variety of interconnected works [9]–[18].

III. METHODOLOGY

To create this model, XG Boost, a random forest [19] technique, and observable features from a multi-layered neural network focusing on profiles. The retrieved features that were recorded in a CSV file may be read by the model with ease. The model's training, testing, and analysis will ultimately reveal if a profile is real or not. Researchers picked Google Collab to create models since Google offers free GPU usage. The Google Colab NVIDIA Tesla K80 GPU has a 12-gigabyte (GB) capacity and can operate nonstop for 12 hours. This method works well in spotting phony profiles. This model's accuracy after training might be higher than in past studies of a similar nature. Its design also stresses a framework that is appealing to the eye. a picture of the architecture of the system



DATASET COLLECTION:

Here dataset used is a MIB dataset [20]. The data set consisted of 3474 real profiles and 3351 fake ones. The data set used TWT, INT, and FSF for fraudulent accounts whereas E13 and TFP were used for authentic ones. The information is kept in CSV format for machine extraction.

CSV FILE CONVERSION:

The large dataset in Microsoft Excel because of its size. Exported these Excel files into CSV format so that planned software could exploit this data. After converting the Excel file to CSV, the following actions were taken:

Initiate the import of the file. Edit text

This activity may be completed using Notepad as well as other spreadsheet tools like Microsoft Excel and Google Sheets.

Instead, you might use Excel (Windows).

- Choose File.
- Choose "Save As" from the menu.
- If the author so desires, he or she may rename the file and select the.csv suffix (comma-delimited).
- Choose the Save option.
- Following the completion of these processes, the author discovered a final dataset for our suggested system, which was kept in a CSV file.

PRE-PROCESSING:

Before getting to the models in this instance, the author adds one more stage of pre-processing. The data collection is pre-processed before being given to a model. This method aims to identify if a profile is real or fake based on how it appears. The specifics have now all been settled. Just the numerical information remains after the category components have been removed. Following the combination of an accurate and unreliable user data set, each profile is given the extra label "fake," a Boolean variable. Following that, the response pertinent to profile X is recorded in the Y variable. Lastly, any blank or NAN entries are replaced with zeros.

RANDOM FOREST:

One illustration of this sort of technique is the ensemble learning strategy known as random forest (or random-decision forest).

This method is used in machine learning because it can be easily used for both classification and regression problems. Similar to the Figure, the random forest takes forecasts from each tree and predicts the outcome based on the votes of the majority of projections rather than relying just on one decision tree. Yet, compared to the decision tree approach, random-forest generates far more choice trees, and the outcome seems to be the sum of almost all of the decision trees generated.

The random forest approach to detect profiles. The model takes in data and produces pertinent outcomes. The trees (FB) are fitted to the sample for the provided set of 1 2, nX x x= and 1 2, NY y y= responses using the bootstrap aggregating process. A random sample is chosen (B times) at regular intervals.

The method utilized to calculate the outcomes for a given sample(x') after training is as follows:

$$f = \frac{1}{B}\sum_{b-1}^{B} f_b(x') \quad \ldots. (1)$$

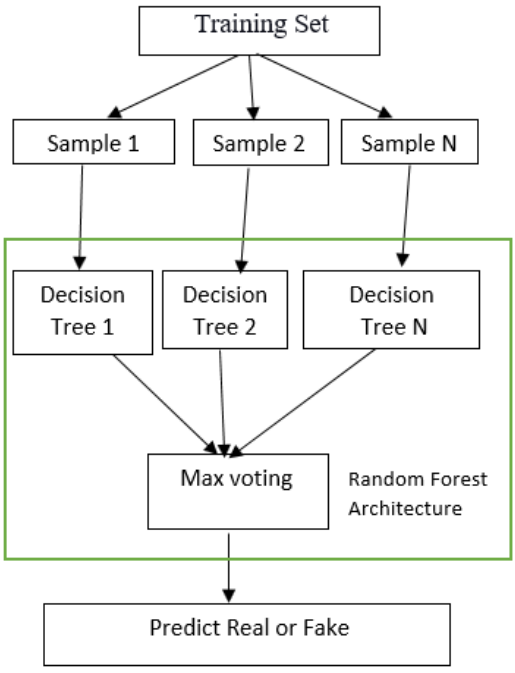


Figure 1 : Random Forest architecture

Max voting(Majority voting)

This voting procedure is typically employed for categorization issues. In this method, each model's forecast of the data is treated as a vote and numerous models are utilized to predict each piece of information. The final forecast uses the bulk of the model.

EXTREME GRADIENT BOOST:

XG Boost is a different ensemble learning approach for regression. This approach implements subsampling of various stochastic gradient boosting settings. The drawback of random forest is that it performs best with complete inputs or without any missing data. To get around this, the author uses a gradient-boosting strategy. According to the boosting method, $f_0(x)$ is first initialized in equation (2).

$$f_0(x) = \arg\min_{\gamma} \sum_{i-1}^{n} L(y_i, \gamma) \quad \ldots. (2)$$

After that, the loss function's gradient is determined repeatedly.

$$\gamma_{im} = -\alpha\left[\frac{\delta L(y_i, F(x_i))}{F(x_i)}\right] \quad \ldots. (3)$$

The boosted model F m is defined at the end (X) in equation (4).

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad \ldots. (4)$$

LSTM:

- LSTM means long short-term memory.
- An LSTM-based framework was created to evaluate a profile's trustworthiness using tweets.
- The author managed the filter that is used to remove the identification strings from each tweet when training an LSTM on this website and tweets.

- All tokens are written in lowercase, and stop words are no longer permitted in tweets.
- The words from these blockchain-enabled tweets were then built into vector representations by the author using an embedding layer.
- The single 32-dimensional vector output of the LSTM is then advanced within layers that are activated by sigmoid functions to produce the output.

OUTCOMES OF THE EXPERIMENT

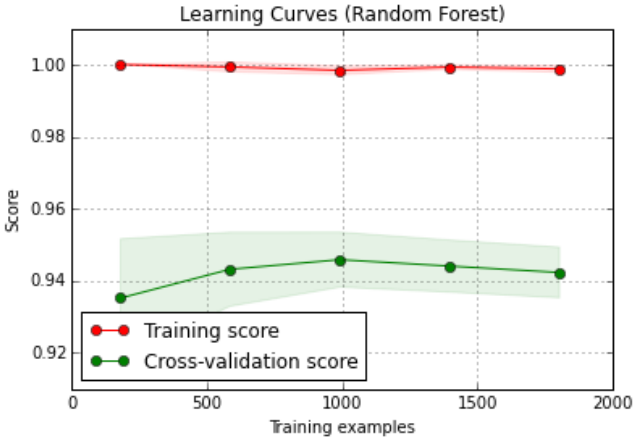
The results of training and testing for each model are listed below.

The ROC curves of the stochastic forest, XG boost, and other approaches are provided for the LSTM neural network, along with model comparisons, loss versus eras graphs, and model accuracy comparisons. Outcomes of the experiment and discussion.

DISCUSSION:

Fake accounts have the power to alter ideas like influence and popularity, which might have an impact on the nation's economy, political system, and social structure. They pose a threat to social media platforms. As stated by the authors in the introduction, this study employs several algorithms to identify bogus profiles, ensuring that consumers won't be misled or harmed by harmful individuals. A blacklist was created by the researchers of a prior study to effectively separate phony characteristics from fraudulent accounts. This research compares several machine learning algorithms to illustrate which one gives the greatest results (XGBoost 99.6%), even if the latter method's results were higher (94.9%) than those of the former (91.1%). Deep Profile was presented as a technique that uses a supervised learning algorithm to anticipate fraudulent accounts in research that made use of dynamic CNN. Another research [23] used another intriguing method to identify Sybil traits based mostly on registration time. The authors of the report claim that many respectable people were mistakenly classified as false positives because they shared IP phone numbers and addresses with Sybil. There were rates of false positives of 7%, 3%, and 21% in variously sized towns. An astounding 95% accuracy rate was achieved by the study's authors. The SVM-NN classification method obtained the greatest performance of 98.3% in predicting synthetic profiles in a research [24] that employed feature extraction using fictitious profiles.

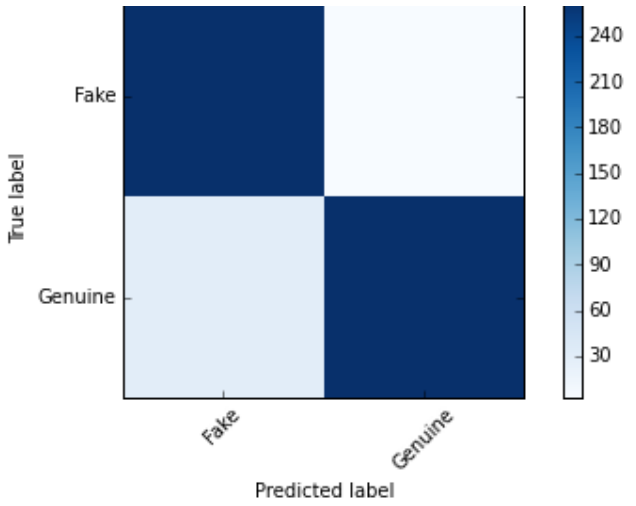
RANDOM FOREST (LEARNING CURVES)



CONFUSION MATRIX:

A confusion matrix aids in visualizing the results of a classification task by providing a table arrangement of the various outcomes of the prediction and findings. By giving a table arrangement of the different outcomes of the prediction and discoveries, a confusion matrix assists in visualizing the consequences of a categorization problem.

The real, fake, fake genuine, or genuine fake profile might be suggested using a confusion matrix.



IV. CONCLUSION

Based on the information that is easily accessible, they employed the CNN Model, Random Forests, and XG Boost supervised learning approaches in this architecture to train the system on how to identify fake accounts.

This project's primary drawbacks are that it only utilizes visible data and lacks real-time applications. Further jobs may be completed by running a CNN on the numerical, categorical, and profile photo data. Better outcomes could also occur from the addition of new parameters, the fusion of different models, and the creation of a real-time model. Depending on their size or specific significance in the identification process, the areas in the model and data may be given varying degrees of prominence. It would be simpler to discover areas where exceedingly complicated issues, such as those that infrequently appear and the latter, must be found using this technique, for example. While complicated, these hybrid models should produce better results. Occasionally, though, combining these strategies could not make a big difference in the outcome. Following that, the model will be ready for more social media platforms like LinkedIn, Snapchat, WeChat, QQ, etc.

V. REFERENCES

[1]. Using Machine Learning to Identify False Identities: Bots vs. Humans. Van Der Walt, E. and Eloff, J. 2018.6540–6549 in IEEE Access. https://doi.org/10.1109/ACCESS.2018.2796018

[2]. Ferrara and Kudugunta, S. (2018) Bot detection using deep neural networks. 312–322 in Information Sciences, 467. https://doi.org/10.1016/j.ins.2018.08.019

[3]. Fake Profile Detection Methods in Large-Scale Online Social Networks: A Complete Study, D. Ramalingam and V. Chinnaiah, 2018.165–177 in Computers & Electrical Engineering, vol. 65. https://doi.org/10.1016/j.compeleceng.2017.05.020

[4]. Y. Minoso, G. Hajdu, R. Lopez, M. Acosta, and A. Elleithy (2019)Artificial Neural Networks are used to spot fake profiles. Farmingdale, New York, 3–4 May 2019: 2019 IEEE Long Island Systems, Applications, and Technologies Conference(LISAT). https://doi.org/10.1109/LISAT.2019.8817330

[5]. Using a black-list, Swe, M.M., and Myo, N.N. (2018) detected fake accounts on Twitter. Singapore, 6–8 June 2018, IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS), 562–566. https://doi.org/10.1109/ICIS.2018.8466499

[6]. Jie, H.J., and Wanda, P. (2020) DeepProfile: Utilizing Dynamic CNN to Detect Fake Profiles in Internet Social Networks.52, Article ID 102465 in Journal of Information Security and Applications. https://doi.org/10.1016/j.jisa.2020.102465

[7]. Identification of Fake Accounts on Twitter Using Hybrid SVM Algorithm, Kodati, S., Reddy, K.P., Mekala, S., Murthy, P.S., and Reddy, P.C.S., 2021.Article No. 01046 of E3 S Web of Conferences, page 309. https://doi.org/10.1051/e3sconf/202130901046

[8]. Automated Fake Profile Detection Using Machine Learning on Instagram, Meshram, E.P., Bhambulkar, R., Pokale, P., Kharbikar, K., and Awachat, 2021. Journal of Scientific Research in Science and Technology: International, 8, 117–127. https://doi.org/10.32628/IJSRST218330

[9]. Utilizing face recognition, Chakraborty, P., Muzammel, C.S., Khatun, M., Islam, S.F., and Rahman, S. (2020) developed an automatic student attendance system. IJEAT, 9, 93–99. https://doi.org/10.35940/ijeat.B4207.029320

[10]. Evaluation of Eye-ball Movement and Head Movement Detection Based on Reading, Sayeed, S., Sultana, F., Chakraborty, P., and Yousuf, M.A. 2021.

[11]. Current Developments in Signal and Image Processing, edited by S. Bhattacharyya, L. Mri, M. Brkljai, J. V. Kureethara, and M. Koeppen, Springer, Singapore, 95–103. https://doi.org/10.1007/978-981-33-6966-5 10

[12]. Forecasting Degree of Visual Focus of Human Attention Using Machine Learning Techniques by P. Chakraborty, M. A. Yousuf, and S. Rahman. In: Proceedings of the International Conference on Trends in Computational and Cognitive Engineering, Springer, Singapore, 683-694. Shamim Kaiser, Bandyopadhyay, Mahmud, and Raym, editors. https://doi.org/10.1007/978-981-33-4673-4 56

[13]. Zero-Shot Learning to Identify Item Instances from Unknown Picture Sources. Muzammel, C.S., Chakraborty, P., Akram, M.N., Ahammad, K., and Mohibullah, M. (2020).IJITEE, 9, 988–991. International Journal of Innovative Technology and Exploring Engineering. https://doi.org/10.35940/ijitee.C8893.029420

[14]. Utilizing Template and Hog Feature Matching, Sultana, T. Ahmed, P. Chakraborty, M. Khatun, M. Hasan, and M. S. Uddin published Object Detection in 2020. IJACSA, 11, 233-238. International Journal of Advanced Computer Science and Applications. https://doi.org/10.14569/IJACSA.2020.0110730

[15]. Using machine learning techniques, Faruque, M.A., Rahman, S., Chakraborty, P., Choudhury, T., Uh, J.S., and Singh, T.P. (2021) ascertain the polarity of the public's opinions on cricket in Bangladesh.30, https://doi.org/10.1007/s41324-021-00403-8

[16]. Using a machine learning approach based on Twitter data, Sarker, Chakraborty, Sha, Khatun, Hasan, M.R., and Banerjee (2020) developed an improvised technique for data analysis and terrorist attack detection.8, 50–62, Journal of Computing and Communications. https://doi.org/10.4236/jcc.2020.87005

[17]. Identifying Fake Accounts on Social Media, S. Khaled, N. El-Tazi, and H.M. Mokhtar, 2018. Big Data 2018 IEEE International Conference, Seattle, 10–13 December 2018, 3672–3681. https://doi.org/10.1109/BigData.2018.8621913

[18]. Social Networks Fake Profiles Detection Using Machine Learning Algorithms, Y. and Z. Elyusufi, 2019. In: Innovations in Smart Cities Applications Edition 3, Springer, Cham, 30-40. Eds. Ahmed, M.B., Boudhir, A.A., Santos, D., El-Aroussi, M., and Karas, R. https://doi.org/10.1007/978-3-030-37629-1 3

[19]. Fake Social Media Profile Detection, Joshi, U.D., Singh, A.P., Pahuja, T.R., Naval, S., and Singal, G., 2021.Machine Learning Algorithms and Applications, edited by Srinivas, Sucharitha, G., Matta, and P., Scrivener Publishing LLC, Beverly, MA, pp. 193-209. https://doi.org/10.1002/9781119769262.ch11

[20]. Identifying Fake Accounts in Online Social Networks at the Time of Registrations by Yuan, D., Miao, Y., Gong, N. Z., Yang, Z., Li, Q., Song, D., Wang, D., and Liang, X. (2019).1423–1438 are included in the proceedings of the 2019 ACM SIGSAC Symposium on Computer and Communications Security, held in London from November 11–15. https://doi.org/10.1145/3319535.3363198

[21]. Fake Profile Detection on Social Networking Websites: A Complete Study, Roy, P.K. and Chahar, S. 2020. pp. 271-285 of IEEE Transactions on Artificial Intelligence. https://doi.org/10.1109/TAI.2021.3064901

Cite this article as :

K. Harish, R. Naveen Kumar, Dr. J. Briso Becky Bell, "Fake Profile Detection Using Machine Learning", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 2, pp. 719-725, March-April 2023. Available at doi : https://doi.org/10.32628/IJSRSET2310264
Journal URL : https://ijsrset.com/IJSRSET2310264