IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 3, March 2024

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

||Volume 13, Issue 3, March 2024||

|DOI:10.15680/IJIRSET.2024.1303207|

Revolutionary Machine Learning Approaches for Identifying Deceptive Online Profiles

Dr. C. Siva Balaji Yadav [1], A. Charvee Sanjana [2], C. Laasya [3], T. Sasisree [4], T. Swaroop [5], C. Dushyanth [6]

[1] Head & Professor, Dept. of AI, Annamacharya Institute of Technology and Sciences (AITS), Karakambadi, Tirupati, Andhra Pradesh, India

[2, 3,4,5,6] Dept. of AI, Annamacharya Institute of Technology and Sciences (AITS), Karakambadi, Tirupati, Andhra Pradesh, India

ABSTRACT: In the modern age, the pervasive influence of social media on individuals' lives cannot be overstated. With a significant portion of the population spending considerable amounts of time on various platforms, it has become imperative to address the growing concern of deceptive activities within these networks. While social media offers numerous benefits, it has also become a hotspot for fraudulent behavior. Traditionally, identifying fake accounts on social media has relied on various classification methods. However, the need for more accurate detection techniques has become increasingly apparent. To address this challenge, our project leverages advanced Machine Learning (ML) technologies and Natural Language Processing (NLP) methodologies to enhance the precision of fake account detection and we have adopted the Support Vector Machine (SVM) algorithm as the cornerstone of our approach. We do not rely only on individual features, but instead consider multiple factors to make better decisions. We use machine learning and natural language processing techniques to create a framework that can adapt to new deceptive tactics and reduce the number of fake accounts on social media, making the online environment safer and more reliable for users.

KEYWORDS: Natural Language Processing, Support Vector Machine, Proliferation, Machine Learning

I.INTRODUCTION

This study focuses on using SVM and machine learning techniques to identify fake social media profiles by analyzing language patterns and posting behavior. It requires expertise in misleading strategies, access to social media data, and proficiency in ML and NLP. Traditional methods are not accurate enough, so new approaches are needed. The study includes creating new ML algorithms, collecting and preparing social media data, developing features, conducting tests, and documenting findings with recommendations for further research.

[1] shows that Facebook users have a 9 percent chance of encountering spam posts and a 0.3 percent chance of encountering malicious links. Spam posts typically contain common domains, while malicious links tend to have unusual domains.

[2] suggests using spatio-temporal analysis on social networks to detect groups of users engaged in malicious activities. We propose using spatio-temporal co-occurrence to infer connections between users and found that the results closely matched actual associations within the network.

[3] discusses the progress made in detecting spam and identifying fake users on Twitter, but acknowledges that there are still many challenges to overcome in advancing these strategies.

[4] proposes a method that can detect 89% of spam tweets.

Our ability to recognize phony social media profiles is increased when we combine these techniques

II.RECENT WORKS

Researchers have developed algorithms to detect abnormal behavior and inconsistent profile information in online social networks. Although previous studies have examined certain aspects of fake profiles and user interactions, distinguishing between real and fake profiles remains challenging. Overcoming these difficulties is crucial for maintaining trust and security in online social networks.

In [3], Akshata and Veena addressed the pressing issue of detecting fake profiles and recognized their detrimental impact on platform integrity and user experience. The K-Means algorithm was used to examine methods to identify and

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

||Volume 13, Issue 3, March 2024||

|DOI:10.15680/IJIRSET.2024.1303207|

contain these malicious actors and highlight their potential use in disinformation campaigns, phishing, and the distribution of malicious content.

Eshraqi, Jalali, and Moattar recognized the difficulty in identifying fake profiles on Twitter and pointed out the limitations of traditional methods. [4] focused on creating a specialized algorithm that could cluster data from Twitter in real-time, in order to better detect spam and fake profiles. Their aim was to enhance the speed and accuracy of identifying these profiles, which is crucial for maintaining security and trust on social media platforms.

[5] have developed a method to identify Fake and Clone profiles on Twitter by considering factors such as the frequency of abuse reports, daily comments, and rejected friend requests. This method aims to detect individuals using fake accounts.

[6] implemented various machine learning methods like Neural Network, Random Forest, and XG Boost to identify fake Twitter profiles using visible data. After training and testing on the MIB dataset, the XG Boost method showed the highest accuracy of 99.46%, followed by Neural Network and Random Forest according to the result analysis.

[7] introduces a framework for identifying fake profiles on Facebook by analyzing their features and extracted data. The random forest classification yielded good results. The features collected through a chrome extension were found to enhance the performance of detecting fake accounts on social networks.

[8] conducted a survey that examines the techniques used to detect fake profiles in online social networks. It covers historical approaches and current state-of-the-art methods for detecting fake or Sybil accounts. The survey compares different approaches, their network types, and dataset statistics.

[9] utilized a unique CNN with a generic pooling function to classify fake profiles. This method creates a dynamic CNN structure by incorporating the WalkPool pooling layer to enhance accuracy and optimize computation.

[10] developed a framework that can detect fake profiles in online social networks with a 95% accuracy using a Random Forest Classifier. The identification of fake profiles can be enhanced by applying Natural Language Processing techniques and Neural Networks to analyze the posts and profiles.

[11] introduced a novel classification algorithm that enhances the identification of fake accounts on social networks. The algorithm utilizes decision values from SVM trained model to train Neural Network model, and decision values from SVM testing to test the NN model.

[12] offers a solution for detecting fake profiles by utilizing specific features. The classifier developed in the training phase successfully identified fake profiles with 97.9% accuracy using the Random Forest algorithm.

III.PROPOSED WORK

For this project, we offered a system that combines natural language processing and machine learning to detect phony profiles on social media websites. In addition, we are incorporating the Naïve Bayes method and SVM classifier to improve the accuracy of the false profile detection.

Support Vector Machine (SVM)

Support Vector Machine (SVM) is an important tool in identifying deceptive social media profiles. SVM is a supervised learning algorithm that can effectively differentiate between real and fake profiles by using a set of characteristics extracted from the profiles. SVMs are effective in handling the complex nature of social media data and can generalize well to new data. They can also handle both linear and non-linear relationships in the data, allowing for flexibility in modeling patterns.

Naïve Bayes

The Naïve Bayes method determines the likelihood that an object with specific qualities will belong to a particular crew or category. For example, let's say we want to identify fake profiles based on time, geo-location, language, and posts' date of publication. All of these characteristics, in my opinion, increase the likelihood that the false profile will exist if the other elements are present

IV.RESULTS AND DISCUSSION

This section provides an overview of the study's results, including the effectiveness of the detection model and the achieved metrics. It evaluates how various features and algorithms

contributed to identifying fake profiles and compares the new approach to existing methods. It also discusses any limitations encountered and suggests potential areas for future research. Overall, it offers a detailed analysis of the study's findings and their implications for social media security.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

||Volume 13, Issue 3, March 2024||

|DOI:10.15680/IJIRSET.2024.1303207|

SVM Metrics:

Confusion Matrix:

	Predicted Fake	Predicted Real
Actual Fake	62	1
Actual Real	8	45

Classification Report:

	precision	recall	f1-score	support
0	0.89	0.98	0.93	63
1	0.98	0.85	0.91	53
accuracy			0.92	116
macro avg	0.93	0.92	0.92	116
weighted avg	0.93	0.92	0.92	116

Figure 1: SVM Evaluation Metrics

Figure 1 contains the confusion matrix and evaluation metrics for SVM Classification. These metrics provide a comprehensive view of the performance of the SVM model for identifying fake profiles. Accuracy tells you the overall correctness of predictions, while precision, recall, and F1 score give insights into the model's performance concerning positive class identification. Specificity provides information about the model's ability to correctly identify negative cases. The SVM model achieves 92% accuracy in detecting deceptive online profiles.

Naive Bayes Metrics:

Confusion Matrix:

	Predicted Fake	Predicted Real
Actual Fake	20	43
Actual Real	1	52

Classification Report:

	precision	recall	f1-score	support
0	0.95	0.32	0.48	63
1	0.55	0.98	0.70	53
accuracy			0.62	116
macro avg	0.75	0.65	0.59	116
weighted avg	0.77	0.62	0.58	116

Figure 2: Naive Bayes Evaluation Metrics

Figure 2 contains the confusion matrix and evaluation metrics for Naive Bayes Classification. It shows the assessment measures that are essential for determining how well algorithms can detect fake profiles on social media or online platforms. These measures quantify how accurately the model can differentiate between real and fake profiles. It is observed that the Naive Bayes Classifier is 62% accurate in identifying fake profiles.

V.CONCLUSION

It is important to detect fake profiles on social media to preserve trust, safety, and genuineness in online communities. Machine learning algorithms like SVM, Naive Bayes are effective in identifying fraudulent accounts by examining different characteristics and trends in user information. SVM can effectively differentiate between genuine and fake profiles by utilizing kernel methods to handle complex data and relationships. Naive Bayes is efficient and straightforward, using probabilistic models to determine the likelihood of a profile being fake based on its attributes, making it suitable for large datasets. Researchers have made significant progress in detecting and exposing deceptive

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

||Volume 13, Issue 3, March 2024||

|DOI:10.15680/IJIRSET.2024.1303207|

profiles that manipulate users and spread false information through innovative techniques. The study emphasizes the need for collaboration between computer science, psychology, and sociology to tackle the complex issue of online deception. The effectiveness of machine learning algorithms in identifying fake profiles depends on the quality and relevance of the features used for training. By including factors like activity patterns, content analysis, and network structure, the algorithms can better detect and adapt to the changing strategies used by malicious individuals. In summary, SVM, Naive Bayes can be useful in detecting fake profiles, but their effectiveness relies on adjusting parameters, selecting features, and understanding the dataset. By combining these algorithms with advanced feature engineering methods, social media platforms can better address the issue of fake profiles and create a safer online space for users. Researchers can create better strategies for identifying and reducing deceptive behaviours on social media platforms by combining knowledge from different fields.

REFERENCES

[1] Saeed Abu-Nimeh, T. M. Chen, and O. Alzubi, “Malicious and Spam Posts in Online Social Networks,” Computer, vol.44, no.9, IEEE2011, pp.23– 28.

[2] Z. Halim, M. Gul, N. ul Hassan, R. Baig, S. Rehman, and F. Naz, “Malicious users’ circle detection in social network based on spatiotemporal co-occurrence,” in Computer Networks and Information Technology (ICCNIT),2011 International Conference on, July, pp. 35–390.

[3] Akshata, Veena,‘‘ Machine Learning Framework for Detecting Spammer and Fake Users on Twitter,’’ in International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Published by, www.ijert.org NCETESFT - 2020 Conference Proceedings

[4] N. Eshraqi, M. Jalali, and M. H. Moattar, ‘‘Detecting spam tweets in Twitter using a data stream clustering algorithm,’’ in Proc. Int. Congr. Technol., Commun. Knowl. (ICTCK), Nov. 2015, pp. 347–351.

[5] Ajith, ‘‘Fake Accounts and Clone Profiles Identification on Social Media Using Machine Learning Algorithms’’, 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/IJSRSET2293158

[6] Joshi, S., Nagariya, H.G., Dhanotiya, N., Jain, S. (2020). Identifying Fake Profile in Online Social Network: An Overview and Survey. In: Bhattacharjee, A., Borgohain, S., Soni, B., Verma, G., Gao, XZ. (eds) Machine Learning, Image Processing, Network Security and Data Sciences. MIND 2020. Communications in Computer and Information Science, vol 1240. Springer, Singapore. https://doi.org/10.1007/978-981-15-6315-7_2

[7] Sahoo, S. R., & Gupta, B. B. (2020). Fake profile detection in multimedia big data on online social networks. International Journal of Information and Computer Security, 12(2/3), 303. doi:10.1504/ijics.2020.105181

[8] D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in largescale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

[9] Wanda, P., & Jie, H. J. (2020). DeepProfile: Finding fake profile in online social network using dynamic CNN. Journal of Information Security and Applications, 52, 102465. doi:10.1016/j.jisa.2020.102465

[10] Reddy, Samala Durga Prasad. "Fake profile identification using machine learning." International Research Journal of Engineering and Technology (IRJET) 6.12 (2019): 1145-1150.

[11] Khaled, S., El-Tazi, N., & Mokhtar, H. M. O. (2018). Detecting Fake Accounts on Social Media. 2018 IEEE International Conference on Big Data (Big Data). doi:10.1109/bigdata.2018.8621913

[12] Kaubiyal, J., & Jain, A. K. (2019). A Feature Based Approach to Detect Fake Profiles in Twitter. Proceedings of the 3rd International Conference on Big Data and Internet of Things - BDIOT 2019. doi:10.1145/3361758.3361784

.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details