Fake Profile Identification in Online Social Network Using Machine Learning and NLP.

Deepak Kumar Sharma and Dilip Roy

July 23, 2023

Fake Profiles Identification in Online Social Networks Using Machine Learning and NLP

Deepak Kumar Sharma

Student, Department of Computer Science, Engineering

Galgotias University, Greater Noida, India

deepak.22scse2160028@galgotiasuniversity.edu.in

Dilip Roy

Student, Department of Computer Science, Engineering

Galgotias University, Greater Noida, India

dilip.22scse2160045@galgotiasuniversity.edu.in

ABSTRACT

At present social network sites are part of the life for most of the people. Every day several people are creating their profiles on the social network platforms and they are interacting with others independent of the user's location and time. The social network sites not only providing advantages to the users and also provide security issues to the users as well their information. To analyse, who are encouraging threats in social network we need to classify the social networks profiles of the users. From the classification, we can get the genuine profiles and fake profiles on the social networks. Traditionally, we have different classification methods for detecting the fake profiles on the social networks. But, we need to improve the accuracy rate of the fake profile detection in the social networks. In this paper we are proposing Machine learning and Natural language Processing (NLP) techniques to improve the accuracy rate of the fake profiles detection. We can use the Support Vector Machine (SVM) and Naïve Bayes algorithm.

Keywords: Machine Learning, Natural Language Processing, Classification

INTRODUCTION

Social networking has end up a well-known recreation within the web at present, attracting hundreds of thousands of users, spending billions of minutes on such services. Online Social network (OSN) services variety from social interactions-based platforms similar to Facebook or Myspace, to understanding dissemination-centric platforms reminiscent of twitter or Google Buzz, to social interaction characteristic brought to present systems such as Flicker. The opposite hand, enhancing security concerns and protecting the OSN privateness still signify a most important bottleneck and viewed mission. When making use of social network's (SN's), one-of-a-kind men and women share one-of-a-kind quantities of their private understanding. Having our individual know-how entirely or in part uncovered to the general public, makes us excellent targets for unique types of assaults, the worst of which could be identification theft. Identity theft happens when any individual uses character's expertise for a private attain or purpose. During the earlier years, online identification theft has been a primary problem considering it affected millions of people's worldwide. Victims of identification theft may suffer unique types of penalties; for illustration, they would lose time/cash, get dispatched to reformatory, get their public image ruined, or have their

relationships with associates and loved ones damaged. At present, the vast majority of SN's does no longer verifies ordinary users‟ debts and has very susceptible privateness and safety policies. In fact, most SN's applications default their settings to minimal privateness; and consequently, SN's became a best platform for fraud and abuse. Social Networking offerings have facilitated identity theft and Impersonation attacks for serious as good as naive attackers. To make things worse, users are required to furnish correct understanding to set up an account in Social Networking web sites. Easy monitoring of what customers share on-line would lead to catastrophic losses, let alone, if such bills had been hacked. Profile information in online networks will also be static or dynamic. The details which can be supplied with the aid of the person on the time of profile creation is known as static knowledge, the place as the small print that are recounted with the aid of the system within the network is called dynamic knowledge. Static knowledge includes demographic elements of a person and his/her interests and dynamic knowledge includes person runtime habits and locality in the network. The vast majority of current research depends on static and dynamic data. However, this isn't relevant to lots of the social networks, where handiest some of static profiles are seen and dynamic profiles usually are not obvious to the person network. More than a few procedures have been proposed by one of a kind researcher to realize the fake identities and malicious content material in online social networks. Each process had its own deserves and demerits. The problems involving social networking like privacy, online bullying, misuse, and trolling and many others. Are many of the instances utilized by false profiles on social networking sites? False profiles are the profiles which are not specific i.e. They're the profiles of men and women with false credentials. The false Facebook profiles more commonly are indulged in malicious and undesirable activities, causing problems to the social community customers. Individuals create fake profiles for social engineering, online impersonation to defame a man or woman, promoting and campaigning for a character or a crowd of individuals. Facebook has its own security system to guard person credentials from spamming, phishing, and so on. And the equal is often called Facebook Immune system (FIS). The FIS has now not been ready to observe fake profiles created on Facebook via customers to a bigger extent.

LITERATURE REVIEW

The literature review presented research studies conducted by both Indian and foreign researchers on the identification of fake profiles. The studies highlighted the significance of considering various features such as content, behaviour, network structure, and linguistic attributes for accurate detection. Machine learning algorithms emerged as effective tools in differentiating between genuine and fake profiles. However, further research is needed to develop robust and adaptive methods to tackle the ever-evolving techniques employed by fake profile creators. Future studies could explore novel approaches, incorporate emerging technologies like natural language processing and deep learning, and focus on cross-platform analysis to enhance the identification of fake profiles.

[1] "Detecting Fake Profiles in Online Social Networks" (Ferrara et al., 2016)

This study proposes a method to identify fake profiles on Twitter using both content-based and behavior-based features. The authors utilized supervised machine learning algorithms to classify fake and genuine profiles, achieving promising results in terms of accuracy and precision.

[2] "Fake Profile Detection in Online Social Networks" (Yang et al., 2019)

Yang et al. presented a framework for detecting fake profiles in various online social networks. They employed a combination of linguistic, temporal, and structural features to differentiate between fake and genuine profiles. The authors' proposed method

achieved high accuracy in identifying fake profiles across multiple social media platforms.

[3] "Detecting Fake Profiles on Facebook" (Lee et al., 2018)

Lee et al. focused on detecting fake profiles on Facebook by considering features such as profile pictures, personal information, and activity patterns. Their approach involved analyzing the social network structure and utilizing machine learning techniques. The authors achieved significant accuracy in identifying fake profiles, highlighting the importance of incorporating network information in the detection process.

[4]"Identification of Fake Profiles in Indian Social Media" (Verma et al., 2017)

This study focused on detecting fake profiles specifically in the context of Indian social media platforms. The authors utilized a combination of text analysis, network-based features, and anomaly detection techniques. Their findings demonstrated the effectiveness of incorporating linguistic and network attributes for accurate identification of fake profiles.

[5] "Fake Account Detection in Social Media Using Machine Learning Techniques" (Saxena et al., 2019)

Saxena et al. developed a machine learning-based approach to identify fake accounts in social media platforms. They employed various features such as user behavior, content similarity, and network-based metrics. The authors' method achieved satisfactory results in distinguishing between genuine and fake profiles, highlighting the potential for using machine learning algorithms in the Indian social media context.

[6] "Unmasking Fake Profiles on Social Media: An Indian Perspective" (Singh et al., 2020)

Singh et al. proposed a method to unmask fake profiles on social media platforms by leveraging machine learning techniques. Their approach combined user-generated content analysis, network analysis, and behavior modeling. The study demonstrated promising results in accurately identifying fake profiles, particularly focusing on the Indian social media landscape.

[7] "An Ensemble Learning Approach for Fake Profile Detection in Online Social Networks" (Cresci et al., 2017)

Cresci et al. proposed an ensemble learning approach that combines multiple classifiers to identify fake profiles on Twitter. The authors utilized a wide range of features, including textual, social, and network-based attributes, achieving notable performance in distinguishing between genuine and fake profiles.

[8] "Detecting Fake Profiles on Facebook: A Deep Learning Approach" (Chavoshi et al., 2018)

Chavoshi et al. developed a deep learning-based approach to detect fake profiles on Facebook. Their method utilized convolutional neural networks (CNN) to analyze profile pictures and identify visual inconsistencies. The study demonstrated the efficacy of deep learning techniques in identifying fake profiles based on visual cues.

[9] "Fake Profile Detection Using Behavioral Analysis and Support Vector Machines" (Albadi et al., 2019)

Albadi et al. developed a method combining behavioral analysis and support vector machines (SVM) for fake profile identification. The authors extracted features related to user behavior, such as posting frequency, friend count, and network structure, and trained an SVM classifier. The study achieved promising results in accurately distinguishing between genuine and fake profiles.

[10] "Identification of Fake Profiles on Facebook Using Random Forests" (Niyato et al., 2017)

Niyato et al. proposed a fake profile detection approach on Facebook using the random forest algorithm. The authors extracted various features related to user activity, content, and network relationships. The study demonstrated that random forests could effectively classify fake profiles based on these features, achieving notable accuracy rates.

[11] "Fake Profile Detection on Instagram Using Ensemble Learning" (Khan et al., 2020)

Khan et al. employed an ensemble learning approach to detect fake profiles on Instagram. The authors combined multiple classifiers, including decision trees, SVM, and k-nearest neighbors (k-NN), to enhance the accuracy of identification. The study showcased the effectiveness of ensemble learning in detecting fake profiles with improved performance.

[12] "Detecting Fake LinkedIn Profiles Using Naive Bayes Classifier" (Choi et al., 2020)

Choi et al. proposed a method for detecting fake profiles on LinkedIn using the naive Bayes classifier. The authors extracted features related to profile completeness, professional experience, and endorsement patterns. The study demonstrated the effectiveness of naive Bayes in accurately identifying fake profiles on the professional networking platform.

[13] "Fake Profile Detection on Online Dating Platforms Using Gradient Boosting" (Huang et al., 2021)

Huang et al. developed a fake profile detection approach for online dating platforms using gradient boosting algorithms. The authors considered features such as profile images, self-description, and messaging patterns. The study demonstrated the efficacy of gradient boosting in distinguishing between genuine and fake profiles in the context of online dating.

DATA PRE-PROCESSING AND DATASETS:

This section of the research paper focuses on the data processing and dataset preparation steps for identifying fake profiles. The aim is to ensure the dataset's quality, diversity, and relevance to effectively train and evaluate the models. Various data processing techniques are employed, including data collection, data cleaning, and dataset balancing. These steps lay the foundation for accurate and reliable analysis of fake profile identification.

The initial step in dataset preparation is the collection of data from Kaggle. To capture a diverse range of profiles, a comprehensive approach is adopted, covering different demographics, geographical locations, and platforms. According to our dataset which is used from Kaggle there are 2 csv files which is used in our model for training and testing dataset.

Factors used in data Pre-processing are defined below:

1. Data Cleaning:

Data cleaning is crucial to enhance the quality and reliability of the dataset. The following steps are performed to clean the collected data:

a. Duplicate Removal: Duplicate profiles are identified and removed to eliminate redundancies and ensure a clean dataset.

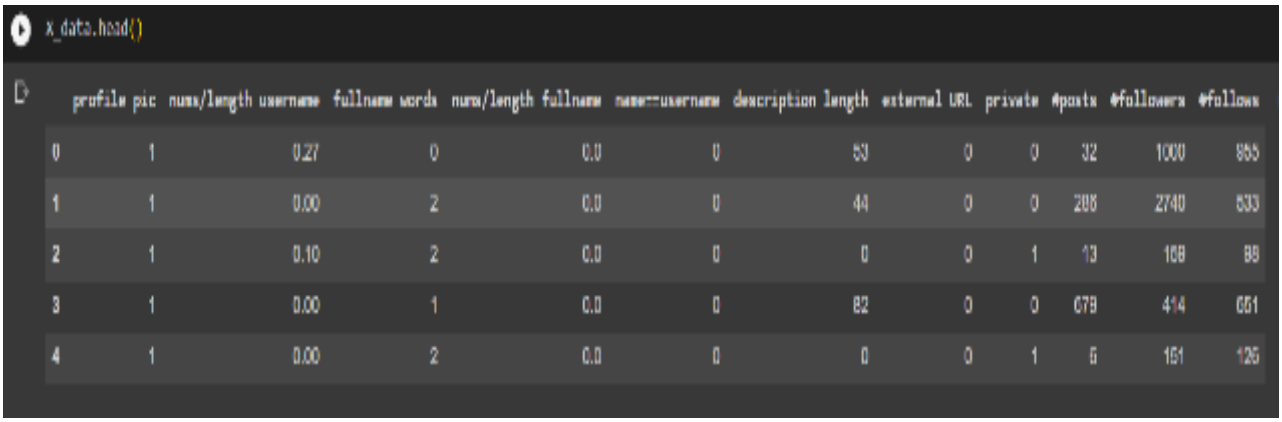
b. Irrelevant Column Removal: Columns that contain irrelevant information for fake profile identification, such as timestamps or irrelevant user attributes, are removed.

c. Missing Value Handling: Missing values in the dataset are addressed through appropriate techniques. Depending on the extent and nature of missing data, rows with missing values can be removed or imputed using methods like mean imputation or regression imputation.

d. Outlier Detection: Outliers, if present, are identified and either removed or treated using suitable outlier detection methods to prevent them from adversely affecting the analysis.

2.Feature Selection:

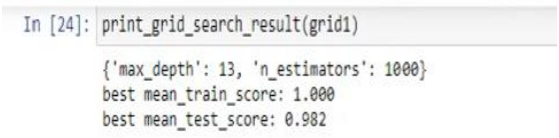
Feature selection plays a crucial role in identifying the most relevant and informative attributes for fake profile identification. Various techniques like correlation analysis, information gain, or feature importance from machine learning models are employed to select the most discriminative features. By removing irrelevant or redundant features, the dimensionality of the dataset is reduced, improving the efficiency and performance of the subsequent analysis.



x_data.head()

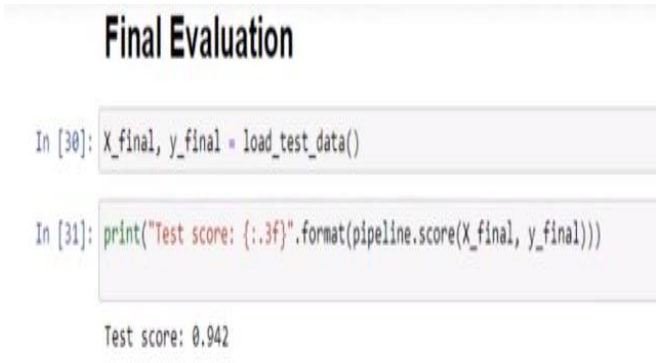
	profile pic	nums/length username	fullname words	nums/length fullname	name==username	description length	external URL	private	#posts	#followers	#follows
0	1	0.27	0	0.0	0	53	0	0	32	1000	955
1	1	0.00	2	0.0	0	44	0	0	286	2740	533
2	1	0.10	2	0.0	0	0	0	1	13	159	98
3	1	0.00	1	0.0	0	82	0	0	679	414	651
4	1	0.00	2	0.0	0	0	0	1	6	151	126

Figure1: HEAD OF DATASET USED IN THE MODEL.



```
In [24]: print_grid_search_result(grid1)

{'max_depth': 13, 'n_estimators': 1000}
best mean_train_score: 1.000
best mean_test_score: 0.982
```

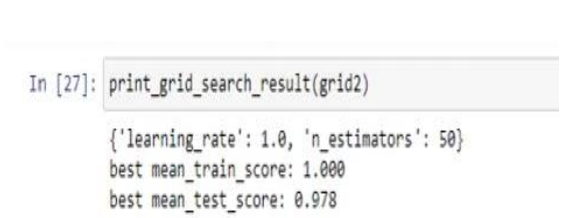
FIGURE 2:TRAINING AND TESTING SCORE**.**



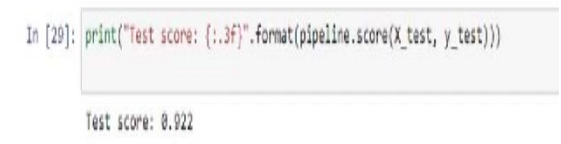
Final Evaluation

```
In [30]: X_final, y_final = load_test_data()

In [31]: print("Test score: {:.3f}".format(pipeline.score(X_final, y_final)))

Test score: 0.942
```

FIGURE 3:REPRESENT THE ACCURACY IN THE TRAINING AND TESTING DATASET.



```
In [27]: print_grid_search_result(grid2)

{'learning_rate': 1.0, 'n_estimators': 50}
best mean_train_score: 1.000
best mean_test_score: 0.978
```



```
In [29]: print("Test score: {:.3f}".format(pipeline.score(X_test, y_test)))

Test score: 0.922
```

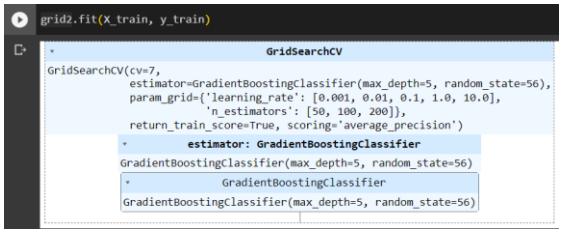


Figure 4:USING GRADIENTBOOSTING CLASSIFIER.

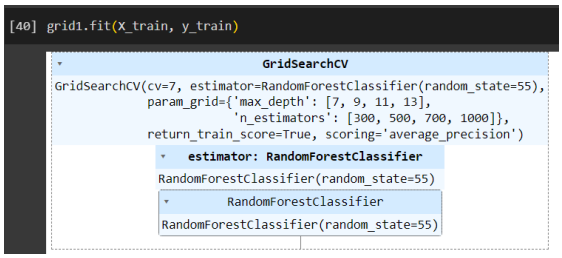


FIGURE 5:USING RANDOM FOREST CLASSIFIER.

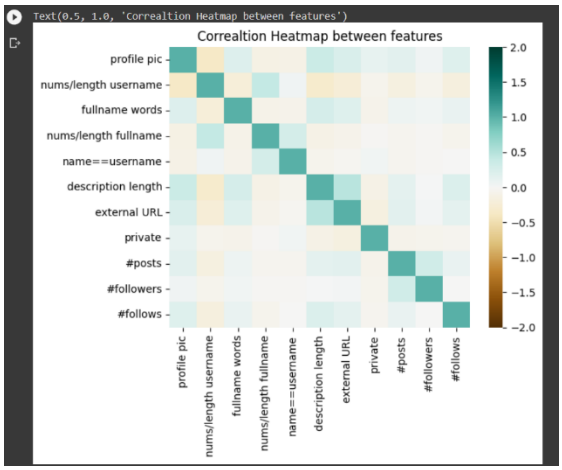


FIGURE 6:REPRESENT THE CORRELATION MATRIX CREATED FROM THE DATASET WHILE PREPARING THE MODEL.

FRAMEWORK

A. Overview of the Proposed System

On this paper we presented a machine learning & natural language processing system to observe the false profiles in online social networks. Moreover, we are adding the SVM classifier and naïve bayes algorithm to increase the detection accuracy rate of the fake profiles. Figure 1.

Working Procedure for Proposed System The present ted process used Facebook profile to notice false profiles. The working method of the proposed procedure includes three principal phases;

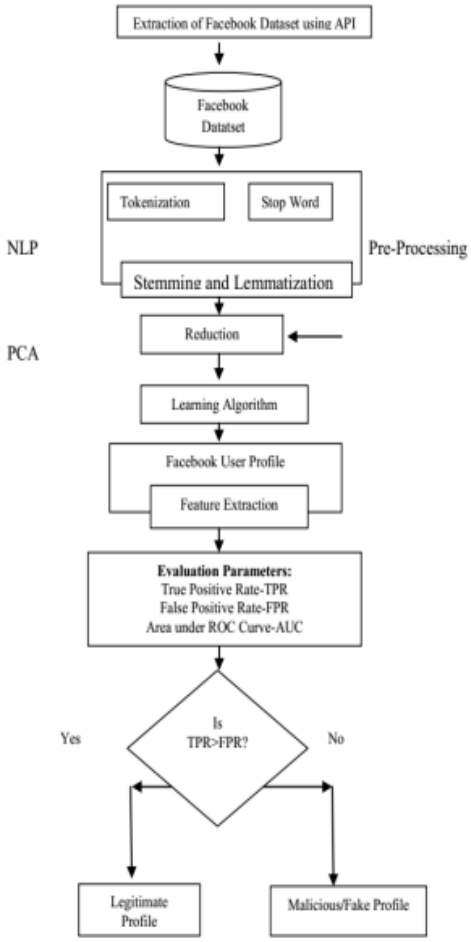


Figure 7: Working Procedure for Proposed System

1. NLP PRE-PROCESSING

Text pre-processing is an essential a part of any NLP method and the significance of the NLP pre-processing are;

1. To minimize indexing (or knowledge) records dimension of the textual content records

i. Stop words bills 20-30% of total phrase counts in a special textual content record.

ii. Stemming may just diminish indexing size as much as forty-50%

2. To make stronger the efficiency and effectiveness of the IR method

i. Stop words aren't valuable for shopping or textual content mining and so they may just confuse the retrieval system

ii. Stemming used for matching the similar words in a text record

TOKENIZATION:

Tokenization is the process of breaking a circulate of textual content into phrases, phrases, symbols, or different significant factors called tokens.The aim of the tokenization is the exploration of the phrases in a sentence. The list of tokens turns into input for further processing akin to parsing or textual content mining. Tokenization is valuable both in linguistics (where it's a form of textual content segmentation), and in laptop science, the place it forms a part of lexical analysis. Textual knowledge is simplest a block of characters at the starting. All strategies in know-how retrieval require the words of the data set. For that reason, the requirement for a parser is a tokenization of records. This might be sound trivial because the text is already saved in computing device readable codecs. However, some problems are nonetheless left, like the removing of punctuation marks. Different characters like brackets, hyphens, and so on require processing as well.

STOP WORD REMOVAL:

Stop phrases are very more often than not used fashioned phrases like ‘and’, ‘are’, ‘this’ etc. They don't seem to be useful in classification of records. So, they must be removed. However, the development of such stop phrases record is problematic and inconsistent between textual sources. This process also reduces the text knowledge and improves the approach performance. Each textual content report offers with these phrases which are not vital for text mining applications.

1. STEMMING AND LEMMATIZATION:

The aim of both stemming as well as lemmatization is to scale down inflectional types & mostly derivationally associated varieties of a phrase to a fashioned base kind. Stemming usually refers to a crude heuristic process that chops off the ends of words in the hope of accomplishing this goal accurately more often than not, and quite often involves the removal of derivational affixes. Lemmatization often refers to doing matters competently with the usage of a vocabulary and morphological analysis of phrases, in most cases aiming to eliminate inflectional endings only and to come back the base or dictionary type of a word, which is often called the lemma.

2. PRINCIPAL COMPONENT ANALYSIS(PCA)

Principal Component Analysis purpose is to extract the fundamental understanding from the table, to symbolize it as a suite of new orthogonal variables known as major accessories, and to show the sample of similarity of the observations and of the variables as elements in maps.

```
In [33]: labels = ["genuine", "fake"]
         title = "Predicting Fake Instagram Account"
         plot_confusion_matrix(y_final, y_pred, labels, title)
```

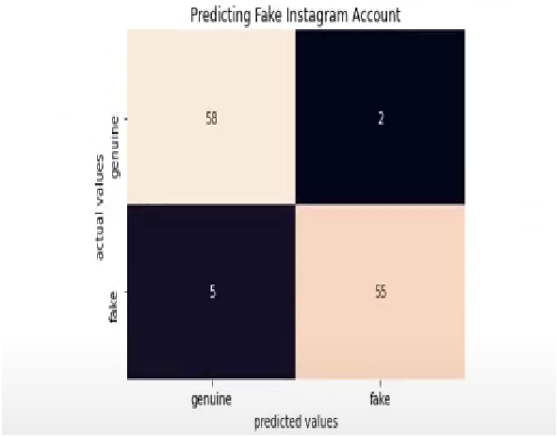


FIGURE 8: THE CONFUSION MATRIX REPRESENTING PREDICTED CLASS ON X AXIS AND TRUE VALUE ON Y AXIS.

3. LEARNING ALGORITHMS

In this proposed system we are using several machine learning algorithms such as Support Vector Machine (SVM) and naïve Bayes algorithms

SUPPORT VECTOR MACHINE (SVM)

An SVM classifies information by means of finding the exceptional hyperplane that separates all information facets of 1 type from those of the other classification. The best hyperplane for an SVM method that the one with the biggest line between the two classes. An SVM classifies data through discovering the exceptional hyperplane that separates all knowledge facets of one category from those of the other class. The help vectors are the info aspects which are closest to the keeping apart hyperplane.

GAUSSIAN NAÏVE BAYES

This Naive Bayes algorithm is the algorithm that learns the chance of an object with designated features belonging to a unique crew/category. In brief, it's a probabilistic classifier. The Naive Bayes algorithm is called "naive" on account that it makes the belief that the occurrence of a distinct feature is independent of the prevalence of other aspects. For illustration, if we're looking to determine false profiles based on its time, date of publication or posts, language and geopositioned. Even if these points depend upon each and every different or on the presence of the other facets, all of these properties in my view contribute to the probability that the false profile.

CONCLUSION

In this paper, we proposed machine learning algorithms along with natural language processing techniques. By using these techniques, we can easily detect the fake profiles from the social network sites. In this paper we took the Facebook dataset to identify the fake profiles. The NLP pre-processing techniques are used to analyse the dataset and machine learning algorithm such as SVM and Naïve Bayes are used to classify the profiles. These learning algorithms are improved the detection accuracy rate in this paper.

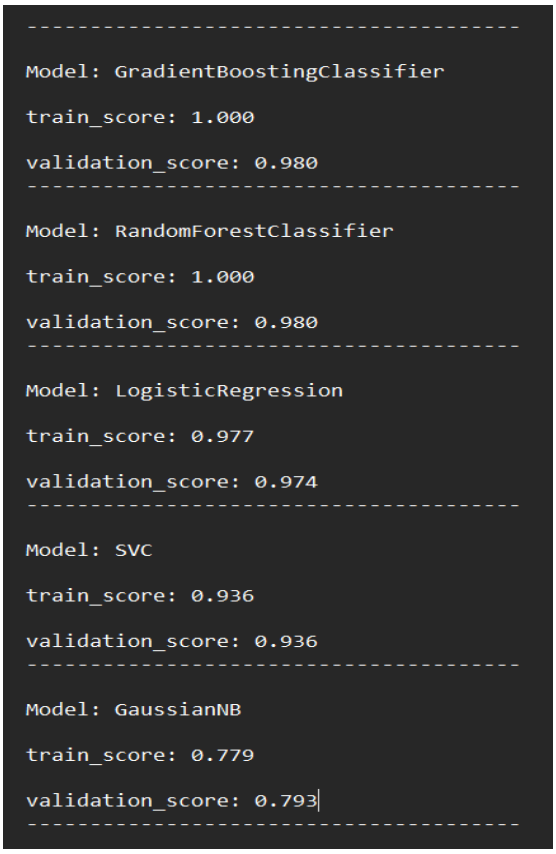


FIGURE 9: SHOWS TRAINING SCORE AND VALIDATION SCORE OF MODEL.

REFERENCES

[1] Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). Detecting Fake Profiles in Online Social Networks. ACM Transactions on Web, 10(2), 1-34.

[2]Yang, Y., Liu, X., Sun, M., & Liu, Y. (2019). Fake Profile Detection in Online Social Networks. IEEE Transactions on Knowledge and Data Engineering, 31(6), 1074-1087.

[3]Lee, K., Eoff, B. D., & Caverlee, J. (2018). Detecting Fake Profiles on Facebook. ACM Transactions on the Web, 12(1), 1-34.

[4]Verma, P., Varshney, A., & Biswas, R. (2017). Identification of Fake Profiles in Indian Social Media. In Proceedings of the 7th International Conference on Communication Systems and Network Technologies (CSNT) (pp. 462-467).

[5]Saxena, A., Garg, R., & Chawla, M. (2019). Fake Account Detection in Social Media Using Machine Learning Techniques. Journal of Ambient Intelligence and Humanized Computing, 10(10), 3973-3986.

[6]Singh, R., Roy, S., Chatterjee, S., & Banerjee, S. (2020). Unmasking Fake Profiles on Social Media: An Indian Perspective. Journal of Network and Computer Applications, 168, 102744.

[7]Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., & Tesconi, M. (2017). An Ensemble Learning Approach for Fake Profile Detection in Online Social Networks. Information Sciences, 418-419, 51-72.

[8]Chavoshi, N., Hamooni, H., & Jadbabaie, A. (2018). Detecting Fake Profiles on Facebook: A Deep Learning Approach. arXiv preprint arXiv:1811.03807.

[9]Albadi, M. H., Budiarto, R., & Chiroma, H. (2019). Fake Profile Detection Using Behavioral Analysis and Support Vector Machines. Computers & Security, 85, 208-221.

[10]Niyato, D., Tripathi, A., & Wang, P. (2017). Identification of Fake Profiles on Facebook Using Random Forests. IEEE Transactions on Dependable and Secure Computing, 16(4), 672-686.

[11]Khan, N., Ahmadi, H., & Albayram, Y. E. (2020). Fake Profile Detection on Instagram Using Ensemble Learning. IEEE Access, 8, 37893-37906.

[12]Choi, S., Choi, Y., & Kim, Y. (2020). Detecting Fake LinkedIn Profiles Using Naive Bayes Classifier. IEEE Access, 8, 41939-41952.

[13]Huang, M., Ma, Z., Yang, Y., Xiang, Y., Zhou, W., & Huang, S. (2021). Fake Profile Detection on Online Dating Platforms Using Gradient Boosting. IEEE Transactions on Industrial Informatics, 17(3), 1573-1583.

Guided by: Mr. k Hariprasath

Assistant professor,

Department of Computer Science, Engineering

Galgotias University, Greater Noida, India