IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

FAKE ACCOUNT DETECTION IN ONLINE SOCIAL NETWORKS USING MACHINE LEARNING AND NLP

N.Divya Sruthi, [1], CH.Vasavi [2]

[1]Assistant Professor,[2]Assistant Professor,

Department of CSE, Geethanjali Institute of Science & Technology,

Nellore, A.P. , India

Abstract: An online social network (OSN) is a virtual platform where users can create profiles, connect with other users, and share information, such as photos, videos, and text-based content. Examples of popular OSNs include Facebook, Twitter, Instagram, LinkedIn, and TikTok. OSNs have become an essential part of people's daily lives, enabling them to stay connected with friends, family, and colleagues, share experiences and ideas, and discover new content. However, OSNs can also present challenges, such as the risk of cyber bullying, the spread of misinformation, and the presence of fake accounts. Fake account detection in online social networks (OSNs) refers to the process of identifying and removing fraudulent accounts that are created with the intention of engaging in malicious activities, such as spreading false information, conducting scams, or engaging in cyber bullying. Traditionally we have different methods to classify fake and genuine profiles. In this project we are proposing Machine Learning and Natural Language Processing techniques to detect the fake accounts in Online Social Networks. We are using the Random Forest, Support Vector Machine and Naïve Bayes Algorithms

Index Terms— OSN, Cyber Bullying, RF, SVM, Fake Accounts

I. INTRODUCTION

In recent years, the widespread use of online social networks has created an environment in which individuals can easily create and maintain multiple profiles. However, some of these profiles may be created with the intention of deceiving others or spreading false information, known as "fake profiles." Detecting fake profiles is a crucial task to maintain the authenticity and trustworthiness of social networks. Machine learning (ML) algorithms and natural language processing (NLP) techniques have emerged as effective tools for detecting fake profiles in online social networks. In this project, we aim to explore the use of ML algorithms and NLP techniques for detecting fake profiles in online social networks. Our focus will be on identifying the key features that distinguish fake profiles from real ones and building models that can accurately classify profiles as either real or fake.

The outcome of this paper can help social networks to improve their security and user trust by eliminating fake profiles. ML algorithms and NLP techniques is multifaceted. Firstly, the presence of fake profiles can negatively impact the user experience of social networks by spreading false information, promoting malicious activities, and compromising user privacy. Therefore, it is essential to develop robust methods to detect fake profiles to maintain the authenticity and credibility of social networks.

Secondly, the identification of fake profiles can be challenging because they often exhibit subtle behavioral differences compared to real profiles. Thus, the application of ML algorithms and NLP techniques can improve the accuracy and efficiency of fake profile detection by identifying patterns and features that are difficult to detect manually. Finally, detecting fake profiles is an active area of research that can significantly contribute to the field of ML and NLP. This paper can help develop new methods and techniques for detecting fake profiles, which can be applied to other areas such as detecting fake news, spam, and fraud.

II. LITERATURE SURVEY

[1] Michael Fire et al. (2012). "Strangers intrusion detection-detecting spammers and fake profiles in social networks based on topology anomalies."

The paper by Michael Fire et al. (2012) titled "Strangers intrusion detection-detecting spammers and fake profiles in social networks based on topology anomalies" focuses on detecting spammers and fake profiles in social networks by analysing topology anomalies. The authors propose a novel approach called "Strangers Intrusion Detection" (SID) that is based on a set of features extracted from the topology of the social network.

The approach assumes that spammers and fake profiles tend to behave differently than real users in terms of their network topology. The authors propose a set of features that capture the differences in network topology between spammers/fake profiles and real users. The features include measures such as degree distribution, clustering coefficient, and betweenness centrality.

The authors train a neural network using the features extracted from the social network topology to classify users as spammers/fake profiles or real users. The results show that the approach is effective in detecting spammers and fake profiles with high accuracy.

[2] Dr. S. Kannan, Vairaprakash Gurusamy, "Preprocessing Techniques for Text Mining", 05 March 2015

The paper by Dr. S. Kannan and Vairaprakash Gurusamy titled "Preprocessing Techniques for Text Mining" discusses various preprocessing techniques that are commonly used in text mining. Text mining is a process of extracting useful information from unstructured text data, and pre-processing is an essential step in this process.The authors provide a comprehensive overview of various preprocessing techniques, including text normalization, stop word removal, stemming, and lemmatization. Text normalization involves converting all text to a standard format, such as lower case, removing punctuation, and expanding contractions. Stop word removal involves removing common words such as "the," "a," and "an" that do not carry significant meaning. Stemming involves reducing words to their root form, while lemmatization involves reducing words to their base form. The authors also discuss the importance of feature selection in text mining, which involves selecting the most relevant features or words for analysis. They provide an overview of various feature selection techniques, including term frequency-inverse document frequency (TF-IDF) and chi-squared. In addition to discussing pre-processing techniques and feature selection, the authors also provide examples of how these techniques can be applied in real-world scenarios, such as sentiment analysis and document classification. Overall, the paper provides a comprehensive overview of various pre-processing techniques used in text mining and their importance in extracting useful information from unstructured text data. The paper is a valuable resource for researchers and practitioners in the field of text mining.

[3] Shalinda Adikari and Kaushik Dutta, Identifying Fake Profiles in LinkedIn, PACIS 2014 Proceedings, AISeL

The paper by Shalinda Adikari and Kaushik Dutta titled "Identifying Fake Profiles in LinkedIn" discusses the problem of identifying fake profiles on LinkedIn, a popular professional networking platform. The authors propose a methodology that involves analysing various features of LinkedIn profiles, such as profile completeness, work experience, education, and endorsements, to identify fake profiles. The authors also use machine learning algorithms, including support vector machines and random forests, to classify profiles as fake or real. The results of the study show that the proposed methodology is effective in identifying fake profiles on LinkedIn with high accuracy. The paper provides valuable insights into the problem of fake profiles on social networking platforms and presents a practical approach for identifying them using machine learning techniques.

[4] 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

The paper by Z. Halim et al. titled "Malicious users' circle detection in social network based on spatiotemporal co-occurrence" discusses the problem of detecting malicious users' circles in social networks. The authors propose a methodology that involves analysing the spatiotemporal co-occurrence patterns of user interactions in social networks to identify groups of users that are likely to be involved in malicious activities. The authors use machine learning techniques, including k-means clustering and decision trees, to identify and classify these groups of users. The results of the study show that the proposed methodology is effective in detecting malicious users'circles in social networks with high accuracy. The paper provides valuable insights into the problem of malicious activities in social networks and presents a practical approach for detecting and mitigating these activities using machine learning techniques.

III. METHODOLOGY

The methodology for fake account detection in Online Social Networks (OSNs) using Random Forest, Naive Bayes, SVM, and NLP techniques involves the following steps: First, collect a large dataset of user accounts in the OSN, including both genuine and fake accounts. Next, extract a set of features from each account, including user-generated content and linguistic patterns, using NLP techniques. Then, use Random Forest, Naive Bayes, and SVM algorithms to select the most important features and train the classifiers on the labeled dataset. After that, deploy the trained classifiers to detect fake accounts in the OSN using the selected features and linguistic patterns. Finally, monitor the performance of the system over time and update the classifiers as needed to adapt to new types of fake accounts and evolving attack strategies. By integrating NLP techniques with machine learning algorithms, the proposed methodology can provide a more accurate and robust solution for fake account detection in OSNs.

REGISTER AND LOGIN,
PREDICT PROFILE IDENTIFICATION STATUS,
VIEW YOUR PROFILE.

Fig 4.1 Architecture Diagram

DATASET

The dataset we used for the project consists of 19 attributes. These 19 attributes are required to classify the fake and genuine profiles.

DATA PREPROCESSING

Data pre-processing is an important step for the creation of a machine learning
model. Initially, data may not be clean or in the required format for the model which can cause misleading outcomes. In pre-processing of data, we transform data into our required format. It is used to deal with noises, duplicates, and missing values of the dataset. Data pre-processing has
the activities like importing datasets, splitting datasets, attribute scaling, etc. Pre-processing of data is required for improving the accuracy of the model. For achieving better results from the applied model in Machine Learning, the format of data in a proper manner. Kaggle provides you pre-processed dataset. But how this data is pre-processed is discussed below.

Natural Language Processing (NLP) pre-processing techniques are used to transform raw text data into a format that can be easily understood and analysed by machine learning
algorithms. Here are some NLP pre-processing techniques used in project.

TOKENIZATION

Tokenization is the process of breaking down a piece of text into smaller units, called tokens. These tokens could be words, phrases, or even individual characters, depending on the requirements of the application.

Tokenization is an essential step in NLP because it enables us to process and analyse text data more effectively. By breaking down the text into smaller units, we can identify patterns, relationships, and meaning that might be difficult to detect in the raw text.

STOP WORD REMOVAL

Stop word removal is a common pre-processing technique in natural language processing (NLP) that involves removing common words that do not carry much meaning in the context of the text. These words, called stop words, are typically short and frequently occurring words such as "the," "and," "a," "an," etc.

The purpose of stop word removal is to reduce the size of the text and focus on the words that carry the most meaning. This can help improve the accuracy of analysis and reduce the noise in the data.

STEMMING

Stemming is a natural language processing technique used to reduce words to their base or root form, called a stem. It involves removing the suffixes or prefixes from words to obtain their base form.

The purpose of stemming is to reduce the variations of words that have the same meaning but are written in different forms. For example, the words "jump," "jumped," "jumps," and "jumping" have the same meaning and can be reduced to the stem "jump." By doing so, stemming reduces the size of the text and simplifies the analysis.

LEMMATIZATION

Lemmatization is a natural language processing technique that involves reducing words to their base form, called a lemma, by considering their morphological features, such as their part of speech, tense, or number. Unlike stemming, which uses a set of rules to remove suffixes and prefixes from words to obtain their root form, lemmatization relies on a dictionary or vocabulary to convert words to their base form.

The purpose of lemmatization is to obtain a canonical form of words that are semantically related to each other, which can improve the accuracy of text analysis, information retrieval, and other NLP applications.

PART OF SPEECH TAGGING

POS tagging is a natural language processing technique that involves labeling each word in a sentence with its corresponding part of speech, such as noun, verb, adjective, adverb, preposition, conjunction, interjection, etc.

The purpose of POS tagging is to understand the syntactic structure of a sentence and to identify the relationships between words. This information can be used for various NLP tasks, such as information extraction, machine translation, sentiment analysis, and more.

ALGORITHMS

Machine learning algorithms are essential for detecting fake accounts on social media platforms, as they can process vast amounts of data and identify patterns that are difficult for human moderators to detect. They can analyse large amounts of data quickly and accurately, identify patterns that may be difficult for humans to detect, and adapt to new tactics used by malicious users. The algorithms used the system are explained below.

RANDOM FOREST ALGORITHM

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the mean or average prediction of the individual trees is returned. Random decision forests correct for decision trees' habit of over fitting to their training set. Random forests generally outperform decision trees, but their accuracy is lower than gradient boosted trees. However, data characteristics can affect their performance. Random forests are frequently used as "blackbox" models in businesses, as they generate reasonable predictions across a wide range of data while requiring little configuration.

SUPPORT VECTOR MACHINE

SVM (Support Vector Machines) is a popular machine learning algorithm that is commonly used for classification, regression, and outlier detection tasks. It works by finding the hyper plane that best separates the data into different classes. In the case of binary classification, the hyper plane separates the data into two classes, while in the case of multi-class classification, multiple hyperplanes are used to separate the data into multiple classes.

NAÏVE BAYES

Naive Bayes is a probabilistic machine learning algorithm that is commonly used for classification tasks. The algorithm is based on Bayes' theorem, which describes the probability of a hypothesis given evidence. In the context of classification, Naive Bayes works by calculating the probability that a data point belongs to a particular class based on its features. The algorithm assumes that each feature is independent of the others, which is why it is called "naive." This simplifies the calculations and makes the algorithm faster and more efficient.

There are several types of Naive Bayes classifiers, including:

Gaussian Naive Bayes: This classifier assumes that the features are normally distributed.

Multinomial Naive Bayes: This classifier is commonly used for text classification tasks and assumes that the features are counts of occurrences of words.

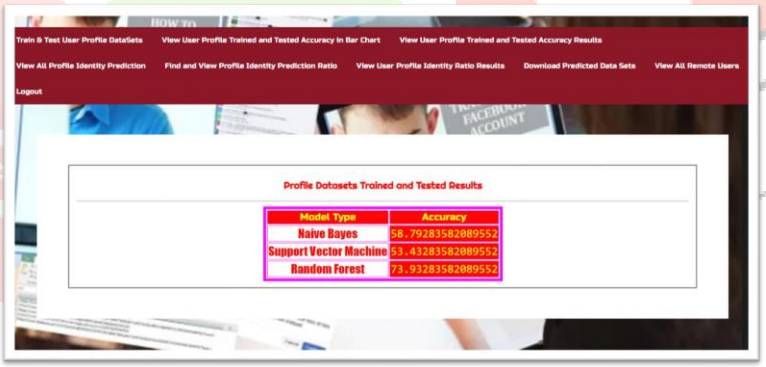
Bernoulli Naive Bayes: This classifier is also commonly used for text classification tasks and assumes that the features are binary (e.g., whether a word is present or not).

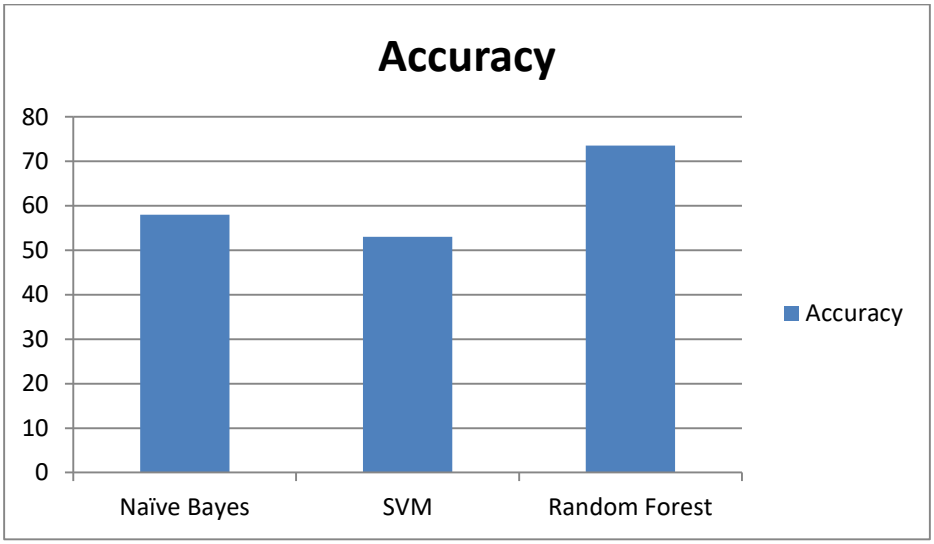
IV. RESULTS AND SCREENSHOTS

The screenshots of the proposed system are as shown in the figure below.

The system was tested on AWS Free Tier EC2 instance. A T3.nano computational engine is used and test with various users accros different platforms. The system peformed as per the expectation and could be a case for practical real time implementation

Algorithm	Accuracy
Naïve Bayes	58
SVM	53
Random Forest	73.5





V. Conclusion

In our Paper, we proposed machine learning algorithms along with natural language processing techniques to classify the fake and genuine profiles. By using these techniques, we can easily detect the fake profiles from the social network sites. In our project we took the Face book Data set to identify the fake profiles. The NLP pre-processing techniques are used to analyse the dataset and machine learning algorithms such as Random Forest, SVM and Naïve Bayes are used to classify the profiles. These learning algorithms are improved the detection accuracy rate in our project. The accuracy of present proposed system is 73.9%. The future scope of fake account detection in OSNs using machine learning algorithms such as random forest, SVM, Naive Bayes, and NLP is quite vast. Some of the potential areas of development are: Integration with deep learning: Deep learning techniques such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) can be used to improve the performance of fake account detection algorithms. Sentiment analysis: Sentiment analysis can be used to analyse the sentiment of user-generated content and metadata to detect fake accounts. Graph-based techniques: Graph-based techniques can be used to analyse the network structure of OSNs to detect fake accounts. For example, centrality measures can be used to identify accounts that are highly connected to other fake accounts. Real-time detection: Real-time detection of fake accounts can be achieved using streaming algorithms that process data in real-time and adapt to changing data distributions.

Multi-modal data analysis: Multi-modal data analysis can be used to analyse different types of data such as text, images, and videos to detect fake accounts. Transfer learning: Transfer learning can be used to transfer knowledge learned from one OSN to another OSN, improving the accuracy of fake account detection algorithms in new OSNs. Overall, the future scope of fake account detection in OSNs using machine learning algorithms such as random forest, SVM, Naive Bayes, and NLP is promising and offers many opportunities for further research and development.

REFERENCES

[1] Michael Fire et al. (2012). "Strangers intrusion detection-detecting spammers and fake profiles in social networks based on topology anomalies." Human Journal 1(1): 26-39.Günther, F. and S.

[2] Fritsch (2010). "neuralnet: Training of neural networks." The R Journal 2(1): 30-38

[3] Dr. S. Kannan, Vairaprakash Gurusamy, “Pre-processing Techniques for Text Mining”, 05

[4] Shalinda Adikari and Kaushik Dutta, Identifying Fake Profiles in LinkedIn, PACIS 2014 Proceedings, AISeL

[5] 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.

[6] Liu Y, Gummadi K, Krishnamurthy B, Mislove A,” Analyzing Facebook privacy settings: User expectations vs. reality”, in: Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference,ACM,pp.61–70.

[7] Mahmood S, Desmedt Y,” Poster: preliminary analysis of google?’s privacy. In: Proceedings of the 18th ACM conference on computer and communications security”,ACM 2011, pp.809–812.

[8] Stein T, Chen E, Mangla K,” Facebook immune system. In: Proceedings of the 4th workshop on social network systems”, ACM 2011.

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

[10] J. Jiang, C. Wilson, X. Wang, P. Huang, W. Sha, Y. Dai, B. Zhao, Understanding latent interactions in online social networks, in: Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, ACM, 2010, pp. 369–382.

[11] Kazienko, P. and K. Musiał (2006). Social capital in online social networks. Knowledge- Based Intelligent Information and Engineering Systems, Springer.