ISSN PRINT 2319 1775 Online 2320 7876

Research paper **Volume 11, Iss 11, Nov 2022**

MACHINE LEARNING AND NLP APPROACHES FOR ACCURATE DETECTION OF FAKE PROFILES IN SOCIAL NETWORKS

G. Sai Sireesha, M. Tech Student, Department of CSE, K.S.R.M. COLLEGE OF ENGINEERING (UGC-AUTONOMOUS) Kadapa, Andhra Pradesh, India- 516005 Approved by AICTE, New Delhi & Affiliated to JNTUA, Ananthapuramu**.**

Dr. Nageswara RaoSirisala, Associate Professor, Department of CSE, K.S.R.M. COLLEGE OF ENGINEERING (UGC-AUTONOMOUS) Kadapa, Andhra Pradesh, India- 516005 Approved by AICTE, New Delhi & Affiliated to JNTUA, Ananthapuramu

S. JAFFER HUSSAIN, Assistant Professor, Department of CSE, K.S.R.M. COLLEGE OF ENGINEERING (UGC-AUTONOMOUS) Kadapa, Andhra Pradesh, India- 516005 Approved by AICTE, New Delhi & Affiliated to JNTUA, Ananthapuramu.

Abstract

Now a days Most of the people using social networking sites regularly. Users may create accounts and engage with others at any time and from any location, making social networking services increasingly popular. There are several benefits of using social media, but disclosing personal information onlinecarries hazards as well. There are also risks associated with sharing personal information online. We need to categorize the user identities of online communities in order to analyze who is promoting dangers in these platforms. The categorization allows us to determine which social media profiles are authentic and which are fraudulent. Typically, Different categorization approaches have been used to recognize fake identities on social networking platforms. However, we need to enhance the reliability of fraudulent profile identification in online communities. In order to boost the trustworthiness of the false profile identification, in our paper we are going with Machine Learning (ML) technologies and Natural Language Processing (NLP) methods for this, we opted KNN, Naive Bayes and Support Vector Machine (SVM) algorithms in our work.

KEY WORDS: Support Vector Machine, Naïve Bayes, Natural language processing, KNN, Fake profiles, Machine Learning

1.INTRODUCTION

Currently, social networking has become one of the most popular activities, on the internet, engaging millions of people and consuming billions of minutes every day. social-interaction based platforms such as Face Book or My space are examples of online social
network (OSN) services [1]. The focus should be on understanding dissemination-centric platforms similar to Face Book or Twitter and google buzz, as well as social interaction attributes as exhibited by existing platforms like Flickr [2]. A major bottle neck and mission of online social network (OSN) is to protect the privateness of its users. On the other hand, security concerns and securing them still represent a critical bottleneck [3].

Social networks (SNS) allow people from different walks of life to share unique amounts of personal information. The fact that our personal information is completely exposed or partially exposed to the public, renders us excellent objectives for different types of assaults [4]. The identity theft occurs if someone uses the knowledge or abilities of another person for their own benefits or gain [5]. Millions of

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 11, Nov 2022

people around the world have been affected by online identification theft during the earlier years. Victims of fraud can experience a numerous of outcomes, including wasting duration and currency, being imprisoned, having the reputation harmed and having their relationships with friends, family were impacted [6]. In truth the majority of social networking program soften configure their privacy preferences to the lowest possible level; consequently, social networks have grown in to an ideal environment for misuse and fraud [7]. Identity fraud and imitation assaults have become more accessible by online networking features suitable for skilled and unskilled hackers [8]. To blend the issue user must provide an accurate information when creating the account in social networking sites. It becomes easy for others to closely observe everything what users share online; the potential losses can be catastrophic [9]. This is particularly true if such confidential data were to be get hacked. Profile information on the online networks can either remains fixed or change over the time [10]. Dynamically the details provided by the users during profile creation is referred to as a static information, whereas the details found by the system with in the network are called dynamic information [11].

Static knowledge comprises of person's education, employment, income, age, means (demo graphic elements and interest), while dynamic knowledge consists of a person real time behavior and their location within the network [12]. In the current state of research depends on either unchanging or variable information; but this approach does not applicable to various online networking sites. users can have access to only see a part of someone's fixed profile and they can't view profiles that change, since they remain within the user networks [13]. Many researches have come up with different ways to identify fake identities and harmful information in online community platforms. Every method exhibits distinctive strengths and weakness [14].The problems facing on the social networks like privacy, online bullying, the wrongful use of information, misusing content or data, trolling etc... are often raised by the make use of fabricated personas with in social network platforms. False identities mean the accounts which are not specific i.e. they are the personas of men and women with fraudulent credentials or accounts that are stated specific. Malicious accounts on Facebook are more frequently utilized to engage in harmful and unwanted actions producing issues for the customers who are using social networks. Individuals create false profiles for the purpose of social engineering, imitation on the internet to harm a person or a group of people and promotion and campaigning [15]. Face book has its own security mechanism to protect personal information submitted by users against spamming, phishing and other forms of fraud, and this resembles is frequently referred to as a face book immune system (FIS). The FIS has not yet been able to more closely monitor fraudulent accounts that users have created on Facebook [16]. To tackle these unanswered questions, the authors offer "a novel trust evaluation model in online networking sites using soft computing methods (TRMSC)" for Twitter. In this context, we compute direct and indirect trust for known and unknown users [17]. Social actions, such as the usage of a fuzzy inference model, are used to assess a user's direct trust in a system. Indirect trust is calculated for users who are not currently interacting with one another. The proposed approach, TMFPN, represents a social system as a fuzzy petri net and collects network data via trust propagation. In this model, people are locations, and their interactions represent changes in those locations [18].The main goal of this work is to determine the authenticity of a user through data analysis. The dataset was obtained from Kaggle, and the work involved the utilization of machine learning methodologies including naïve bayes and support vector machine, along with employing NLP techniques to improve accuracy levels.

2. LITERATURE SURVEY

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 11, Nov 2022

In [1]. A new strategy was suggested by sarod and Mishra to identify false profiles, which involves a series of processes. They gained access to several profiles using the Facebook graph API tool and created a code to retrieve the found data. This gathered data is then used to create the characteristics that the classification method will employ in their algorithm. primarily, the data is encoded in JSON structure, andit is then processed to a well-organized format (CSV) that machine learning algorithms can understand more easily. The comma delimited values will contribute to enhancing the classifier's subsequent operational efficiency. The researchers experimented with both self-learning and guided machine learning approaches. With in this situation, guided machine learning approaches produced a nearly 98% outcome. They divide the data into instructional and evaluation subsets for supervised machine learning. The classifier underwent an instructional phase, acquainting itself with 80% of the dataset, followed by validation phase of remaining 20%. Feedback is sent to the profile after the algorithm has run, requesting it to provide identification to demonstrate that it is not a false identity.

In [2], Profile details of fake Twitter accounts are determined based on activity-based pattern. it is noticed that the followers that a person or company has on online social networks (OSNs) is a key way to tell how popular they are. This metric has significant finance and/or administrative consequences. companies can modify their goods or messages based on facts about their target audience, such as age, geography, and so on. However, the availability of false profiles on social networks might skew such customizing. In this work, 62 million publicly accessible Twitter profiles were analyzed, followed by the development of an automated system designed to identify artificially created false identities was developed. Some very trust worthy subset of fraudulent accounts was discovered by combining of screen name pattern recognition mechanism and a detailed examination of posted tweets timings. The analysis of these fraudulent accounts' profile creation timings and URLs indicated unique performance of the fraudulent accounts in comparison to a reliable data set. The integration of this approach and proven social network analysis will facilitate the prompt identification of false accounts in OSNs in a timely manner. A substantial Twitter account database comprising of 62 million accounts were gathered and studied using a crawler to better understand the features of false account creation. By categorizing user accounts based on match various profile characteristics, analysis of screen name patterns, and implementation of a filter that disperses update timings, a very trustworthy malicious accounts collection was produced. Based on their Tweet behavior, a portion of the accounts that our system had flagged as phony were personally examined and confirmed to be completely fake. The false accounts were nearly continuously produced in groups and at duration of under 40 seconds, according to an analysis of the bogus accounts set's features. These accounts were produced more frequently on certain weekdays and at specific moments during the day, which suggests a non-automated component in their establishment and maintenance. When compared to a trustworthy dataset of comparable size, the false profile set that we were able to identify had substantially different creation time characteristics. The phony profile set's URLs were found to be less diverse than the real photos. Without doing an in-depth investigation of the tweets, it is possible to identify probable spammers using our activity-based profile-pattern identification approach. Our method has the drawback of only reliably identifying a tiny fraction of false accounts. The produced fraudulent profiles, however, are expected to have a minimal number of false positives, which makes it a perfect seed database to utilize with social graph approaches for effective spam identification. Twitter will be able to sustain an online community that is stocked with actual consumers and, as a result, be an effective instrument for collecting and distribution if spam-detection technologies, like that of ours, are used effectively.

In [3], M. Egele et al. introduces the COMPA social network account breach detection method. This approach is based on how people behave on social media. Normal user behavior is constant, while

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 11, Nov 2022

COMPA has shown that hacked accounts exhibit more erratic behavior. A behavioral profile is created by COMPA based on the last communication the account sent.

In [4].to extract features, profiles are processed in huge quantities. The categorization of false profiles utilizes the neural network technique resilient back propagation in conjunction with support vector machines. Sybil Frame classifies at several levels. There are both content-based and structure-based methodologies. The dataset is analyzed using a content-based technique, and data is extracted that is then utilized to derive prior knowledge about nodes and edges. Using a Markov random field and loopy belief propagation, which uses prior knowledge, the structure-based technique correlates nodes. The first step of the Sybil Frame method uses a content-based approach, and the second stage uses a structure-based approach. In stage I, clickstreams and friend suggestions are studied. Vote Trust makes use of a voting-based system that pulls user activity and uses the total number of votes cast worldwide and votes assigned based on trust to identify fraudulent profiles. Due to restrictions, such as the selling of actual accounts that have previously been hacked, it is regarded as the first line of protection.

In [5]. The extensive utilization of web browsing and the quick expansion of social networking platforms for news (such as Facebook (FB), Twitter, and Instagram) have inaugurated an unprecedented era of information dissemination that have never before been seen in human annals. Owing to the advent of digital networking sites, individuals are generating and circulating grater amount of information than they did previously. However, much of it is false and irrelevant to the conversation. It's challenging to use an algorithm to classify a piece of content as deceptive or false information. Before determining whether or not anything is true, even a subject matter expert must take into account a number of variables. Researchers advise employing a machine learning classification strategy for finding false news. Our study examines many texts related qualities that can be employed to differentiate between fake and authentic data. We employ a variety of integral methodologies to instruct a collection of unique machine learning models, and we use those attributes to assess the algorithms' performance on real-world datasets. Individual learners get surpassed by our suggested ensemble learner technique.

In [6]. One of the major issues with online social networks, which are run by automated operators and sometimes utilized for malicious reasons, is the frequent appearance of fake accounts or social bots. The researchers have made several attempts to detect these things in social media. The most popular used method is a feature-based classifier that uses machine learning, and feature selection is the key step in this process. The multi-objective hybrid feature selection technique used in the current study to detect bogus accounts facilitates feature set selection with the best classification performance. The Minimum Redundancy - Maximum Relevance (MRMR) approach was used to choose the candidate feature set based on its relationship to the target class and its least features that are redundant. The final feature set for the detection operations is then chosen, which has the stable feature set with the fewest features and can achieve optimal performance. The suggested method is evaluated on two datasets from the social network of Twitter, and the outcomes are contrasted with those of effective approaches already in use. The results demonstrate that the suggested classifier strategy performs better than the currently used techniques. The process of selecting relevant features in classification methods is really important because the system needs to operate on vast real-time data to detect bogus profiles on social media. Additionally, for the purpose of comprehending the activity of false accounts, researchers frequently study and explore the traits that have been detected. It implies that another priority that should be taken into account is the stability of the feature selection process. Additionally, stability is frequently paired with classification performance since it is not seen to be an appropriate measure for evaluating the selected characteristics on its own. In this study, a multi-objective hybrid approach is employed to determine the most useful feature set for Twitter fake account identification. The results of studies conducted on two Twitter datasets

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 11, Nov 2022

showed that, when compared to other current approaches, the suggested methodology might provide more optimum and balanced performance. With a slight modification to the feature set for fake account identification, the suggested technique may be used for a variety of social networks, which might serve as the focus of future research.

In [7], D. Freeman et al. rather than focusing on a single false account, concentrated on recognizing groups of them. Based on the information supplied at the time of registration, such as the registration IP address and registration date, this method established a cluster. The model is trained using Random Forest, SVM, Logistic Regression, and SVM is used to determine if a cluster of accounts is false or not. A study on reverse engineering mobile apps was published in Arlington, USA. It makes advantage of a method to automatically reverse-engineer user interfaces for mobile applications.

In [8]. Krishna B Kansara et al. A Sybil node detection approach based on the social network was suggested in this study. By include user behavioral factors as latent transactions and friendship rejection, this technique gets over the drawbacks of the earlier graph-based systems. Sybil node identification (SNI) and Sybil node identification using behavioral analysis (SNI-B) are the two components of the suggested design.

In [9]. Social media news consumption is a growing trend in today's society. Social media's essential advantages of quick distribution, low cost, and simple access benefit consumers. However, the standard of the news is seen as being poorer than that of traditional news sources, leading to a significant quantity of false news. Due to the negative impacts on people and society, spotting false news is becoming increasingly vital and popular. It is recommended to use social media user activities as supplementary data to enhance the identification of false information since the effectiveness of finding false information alone from data is typically unsatisfactory. Thus, a thorough knowledge of the relationship among online user accounts and deceptive information is required. In this study, we build real-world datasets that evaluate users' trust in fake news, and we choose indicative sets of both "seasoned" and "naive" consumers—the former of which can identify inaccurate information in fake news—and the latter of which are more inclined to believe it.

We conduct a comparison study of known and unknown profile characteristics across various user demographics, revealing their ability to distinguish bogus information. The outcomes of this report establish the groundwork for forthcoming research on automated detection of false news. How people interact on online platforms have the potential might assist in the identification of fraudulent news, we investigate the connection between individual profiles and bogus vs legitimate news. Experiment findings using authentic datasets show that: i) some consumers are more prone to believe false content than actual content; and ii) these consumers exhibit distinct characteristics from individual who are more prone to believe authentic news. These insights make it easier to build features for false news identification profiles. There are various intriguing future possibilities. In order to investigate additional user credentials data, such as ideological leaning and consumer trustworthiness, to determine whether these customer traits may be utilized to detect bogus news. Second, we establish a collection of plausible individual identities for false information identification in this study. To develop false news identification, we aim to look into how these features may be combined into bogus information identification models. Third, investigation has revealed that bots have extensively propagated fake news, and we will add automated software programs detection algorithms to distinguish bots from legitimate consumers in order to better use user profile information to detect fake news.

2.1 In the existing methods, the following limitations are identified.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 11, Nov 2022

1) The system is incapable of determining fake accounts because of the lack of attribute resemblance identification.
2) It is not determined if two profiles are comparable based on the degree to which their attributes are alike.
3) The Facebook dataset is small and privacy considerations prevent the release of a lot of data.
4) Rule-based approaches will have decreasing accuracy as the number of methods for creating malicious accounts continues to rise fast.
5) A high rate of false positives: The current system often makes the mistake of labelling real people as having phony profiles, which can lead to dissatisfaction among the users
6) The rule-based approach may not be scalable enough to manage social networks with millions of users.

3. DETECTION OF FAKE PROFILES IN SOCIAL NETWORKS

Existing techniques for spotting bogus accounts frequently rely on a small set of variables, which can result in less precise predictions. In some circumstances, when some inputs are inadequate, the accuracy of these systems is impacted. In response, the main goal of our research is to enhance fraudulent profile detection in social networks. We use machine learning methods like SVM, Naive Bayes, KNN, and NLP to improve accuracy and overcome the drawbacks of existing methods. It is necessary to create models that utilize machine learnings for foretelling and spotting fraudulent identities in face book. deceptive accounts using a set of guidelines that allows for a clear distinguish between counterfeit and authentic personas. By leveraging machine learning and a massive quantity of data, the system can eventually adjust to new patterns in the creation of bogus accounts. By utilizing this tool, we want to lessen the burdens of using more traditional or already used methods. The primary goal of this study was to develop a system that could quickly, accurately, and reliably determine if an account was false.

Start by obtaining the "facebook.csv" dataset from Kaggle, which contains information about various features related to Facebook. Utilize a data manipulation library such as NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn etc... in python, to load the dataset in to your machine-learning environment. This step converts dataset in to your structured format that can be easily worked with. The suggested method's steps are depicted in detail in figure1.

3.1 Step by Step Procedure for Proposed System:

*i) Data collection:*collect a sizable Facebook dataset consisting of actual and fictitious social network accounts. User profiles, posts, interactions, friend connections, and other relevant information may be included in the data collected.

ii) Data pre-processing and NLP techniques: Data pre-processing means dealing with null values and missing values like is.na () to check them. Using functions to get familiar with data like info (), describe (), value_counts (), Means these are used in used in data exploration (exploratory data processing is a part of data preprocessing) and Visualizing data using graphs like histogram, count-plot, heatmap, etc.

Scaling performance is done in data preprocessing means to scale the range of independent variables or features of data, and performing normalization is the process of translating data in to the range of 0to1

Tuning and applying algorithm: In the tuning we are using grid search algorithm to tune the data and we are applying SVM, Naive Bayes, KNN and NLP to create machine learning models which will detect the

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 11, Nov 2022

fake profile in social media. Cleaning, converting, and organizing the obtained data into a framework suitable for analysis using computational intelligence (ML) algorithms constitutes data preparation.

*NLP Techniques:*Tokenizing, deleting stop words, removing null values, punctuation marks, stemming, and lemmatizing are all examples of NLP techniques used to extract useful characteristics from textual input in preparation for further processing, such as analyzing sentiment, sorting text, term entity identification, topic modelling, etc. build a numerical representation of the text using TF-IDF or word embeddings.

*iii) Feature Extraction:*Extract the features/attributes from the profiles which are essential to solve the problem means that helps to distinguish real profiles from fake ones. Personal data may be mined for additional features like the quantity of friends, the date the account was created, posts the user's age, the frequency with which updates are made, etc... For text-based data characteristics such as sentiment scores, TF-IDF representations etc... the TF-IDF is used to convert text in to numerical vectors.

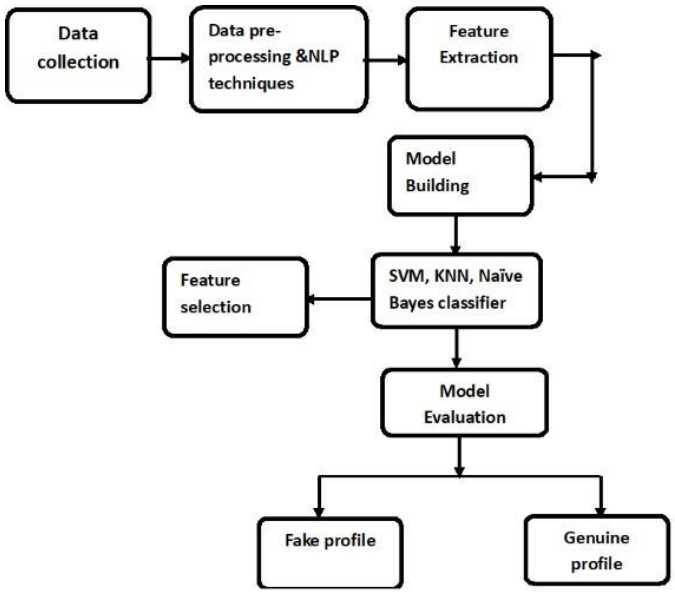

Figure1: Fake Profile Detection System of Social Networks

iv) Model building: Before a model can be trained to produce predictions, its structure, characteristics, and configuration must be designed.

*v) Machine learning algorithm:*Applying the SVM, KNN and naive bayes algorithms to the cleaned and chosen features for training. Hyperparameters may be tuned using methods like cross validation to improve performance.

vi) Model selection and training: To facilitate the process of developing ML models, the data is split into two categories: learning data and validation data. In this way, we may evaluate the model's ability to generalize by training it on a small sample of data. (After training, apply the trained models to the testing data to generate predictions means to ensure that we are getting expected result).

vii) Model evaluation: To understand how well the given model is performed by Using performance evaluation metrics like accuracy, precision, recall, and F1 score, and support. testing and refining models on a distinct testing data set is an efficient method of gauging a model's efficacy from this we can find the

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 11, Nov 2022

genuine profile and fake profiles, and by constantly monitoring means Using the feedback the model will be trained again and again i.e. In a repetitive fashion so it will be trained again and again such that accuracy prediction is good.

3.2. Fake profile Detection Using ML Algorithms

In this section different Machine learning algorithms are discussed for identification of malicious accounts of users based on their malicious activities.

3.2.1 Naïve Bayes Algorithm:

The Naïve Bayes techniques are categorized as a supervised learning approach for addressing classification issues that is based on Bayes theorem. The most common application of this method is it uses a high-dimensional training dataset for text categorization. A methodology for classification of data guided by Bayes theorem, incorporating the notion of feature independence. In machine learning models, Naïve Bayes classifier is one of the most effective and simple algorithms. It helps to create highly accurate prediction and fast learning structures. As a probabilistic classifier it predicts objects based on their probability.

Some popular instances of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and content categorization. The Naïve Bayes algorithm is consisting of two words Naïve and Bayes which can be described as:

Naïve: It is called Naïve because it presumes that existence of one attribute does not depend on existence of another attribute. For instance, apples if red, spherical and sweet are recognized as apples based on color, shape and taste. It is therefore possible to recognize an apple by each of its features without relying on others.

Bayes Theorem: It is called bayes because it relies on the principle of Bayes Theorem. Bayes' theorem is employed to ascertain the likelihood of hypothesis in the presence of prior information. This principle relies on the concept of conditional probability. The formula for Bayes theorem is given as:

$$P(A|B) = P(B|A)\ P(A)/P(B) \qquad (1)$$

In eq(1)
P(A|B) is Posterior probability.
P(B|A) is Likelihood probability.
P(A)is Prior Probability.
P(B)is Marginal Probability.

3.2.2 K-Nearest Neighbor Algorithm:

KNN serves as a classification approach that doesn't rely on parameters and falls within the category of supervised machine learning techniques. As a result, having data with predefined class or group of labels is a necessity. It is a type of instance-based learning technique, because it doesn't create a model in comparison with most other classification methods. Instead, it predicts directly based on the training set. The algorithm can be continuously updated with new training set. The working principle of k-NN is explained through the figure2.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 11, Nov 2022

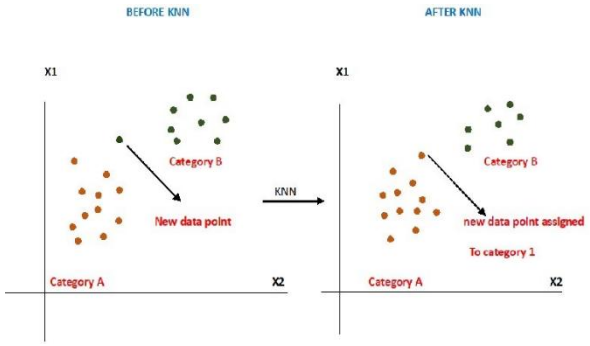

Figure2: Classification of new user profile using KNN

The methodology through which KNN operates can be understood by referring to the algorithm outlined here.

i. Choose fixed number of neighbors, denoted as 'k'.
ii. Compute the Euclidean distance between them.
iii. Euclidean Distance= $d = \sqrt{[(x_2 - x_1)^2 + (y_2 - y_1)^2]}$.
iv. collect the K Neighbors with the least Euclidean distance as determined
v. The number of data points in every segment must be counted within this group of K neighbors.
vi. Allocate the data points to the category containing the largest neighbor count.

Let's assume the above figure shows two segments. specifically, A and B and we have a new datapoint X1, so we need to determine for which category it belongs to. To tackle this sort of issue, we require a K-NN techniques. The KNN algorithm allows us to easily categorize an individual dataset using their classes.

3.2.3 Support Vector Machine

One of the most well-known SMLA, SVM is employed in a variety of Machine Learning contexts, including Classification, Outlier Detection, and Regression. Binary categorization is where it really shines. This method is among the most reliable algorithms for text categorization. SVM is used to differentiate between two data sets with similar categorization.

The hyperplanes that are produced by this data analysis technique are margins that split groups in accordance with certain patterns.

Support Vectors: Support Vector refers to the datapoints or vectors which are most near to the optimal hyperplane, and have impact on its position. SVM performs classification by finding a decision boundary. This decision boundary is also called as "hyper plane". Best hyperplane will have maximum distance from support vectors.

The main aim of SVM is to find best hyperplane with larger distance between two classes in N-Dimensional space. Text and picture classification, hand -writing recognition, face identification, and bio analysis of sequences are only few of the complicated challenges that support vector machine is applied. It's also commonly used for sentiment analysis, document categorization (news, emails, articles, websites, etc.), and other forms of categorization of texts. The working principle of SVM is shown in figure3.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, **Volume 11, Iss 11, Nov 2022**

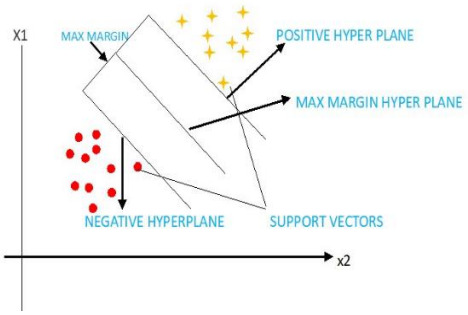

Figure3: Support Vector Machine in profile classification

3.2.4 NLP Techniques:

NLP is a technology utilized by machines to comprehend, examine, manage, analyze, translate human communication. It handles with text data. NLP is inspired to design the technology that understands the human language with greater efficiency.in figure 4 the working principle of NLP preprocessing is described.

Pre-Processing Steps: For pre-processing, this system removes unwanted samples using the removal of stop words

method. At this stage, the stop words from the sentences in the emotion data set are removed by the system. The words "for," "into," "is," "there" etc. are regarded as stop words because they have no bearing on the statement's content. Prior to performing the string transformation, this system removes the stop words.

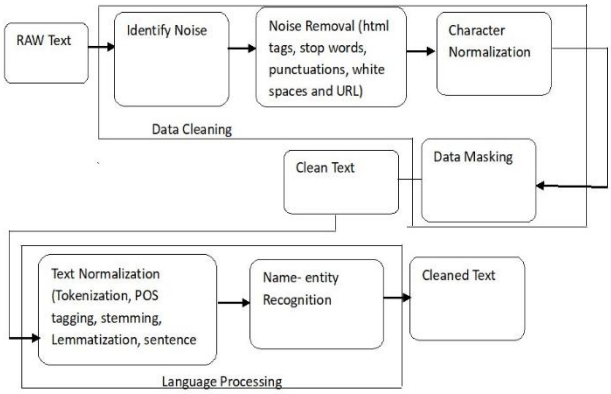

Figure 4: Preprocessing of social network dataset using NLP

Tokenization: By tokenizing statement words, single words or sets of words can be translated in to single words. N-grams can be used to implement this technique. The type of N-gram that issued based on the size of the tokens. The N-gram is frequently defined in one of three ways: unigram=1, bi-gram N=2, and tri-gram=3.

Transformation: Creating numerical tables from text data for classification is the process of transformation. The text-based data is not understood by the categorization algorithms. There are two different types of transformation strategies for text data: count vectorizers, which depend on word

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 11, Nov 2022

frequency, and TF-IDF vectorizers, which depend on the text's weight by computing the TF-TDF (term frequency-inverse document frequency).

4. EXPERIMENTAL RESULTS

Here the data set that is used for training and evaluation of models is described. The performance of ML techniques like SVM, K-Nearest neighbor, NLP, Naïve Bayes and are implemented and their performance is analyzed in tabular and graphical formats.

4.1 Dataset

Results are taken with reference of Facebook dataset collected from Kaggle. As shown in table1, the dataset consists of following features

4.2 Model Evaluation

In the following sections model's performance is evaluated in terms of precision, Recall, F1score, and Support, these are the presentation metrics commonly exist to estimate the performance of classification models. They are calculated using the confusion matrix: "Compares true value with predicted value

Table:1 social network data set

S NO	Feature_name
1	Id
2	Name
3	Screen_name
4	Fav_num
5	Statues_count
6	Followers_count
7	Friends_count
8	Favourites_count
9	Listed_count
10	Created_at
11	Location
12	Default_profile
13	Profile_img_url
14	Profile_banner_url
15	Profile_use_background_image
16	Profile_use_background_image_url
17	Profile_text_color
18	Profile_image_url_https
19	Profile_sidebar_border color
20	Profile_background_title
21	Profile_sidebar_fill_color
22	Profile_background_image_url
23	Profile_background_color
24	Profile_link_color
25	Description
26	Updated
27	Label

*Precision:*The proportion of correctly categorized positive samples (True Positive) to the total number of correctly or mistakenly classified positive samples is known as precision. the percentage of precise forecasts is calculated. as shown in eq(2).

$$\text{Precision}=TP/(TP+FP) \quad (2)$$

*Recall:*The recall is determined as the proportion of Positive samples that were properly identified as Positive to the total number of Positive samples. The recall of the model assesses its ability to recognize positive samples. The more positive samples identified, the higher the recall..as shown in eq (3).

$$\text{Recall}= TP/ (TP+FN) \quad (3)$$

ISSN PRINT 2319 1775 Online 2320 7876

Research paper Volume 11, Iss 11, Nov 2022

F1 score: F1 score is a machine learning evaluation metric that measures a model's accuracy f1 score becomes high only when both precision and recall are high. F1score is the harmonic mean of precision and recall and is a better measure than accuracyIt is derived as shown in eq (4).

F1 score=2* (precision*recall)/ (precision + recall) (4)

*Support:*The Weighted average of accuracy, recall, and f1 score are all derived from the total number of samples, which is illustrated by the support. It's useful for learning how the different types of information in the dataset are dispersed.

*Macro avg:*These studies provide a comprehensive summary of how model capabilities differ between groups without prejudice. The inconsistencies in the dataset are ignored.

*Weighted avg:*It takes in to consideration for the class probability in the dataset without making any distinctions between classes, unweighted avg makes no such allowances. A weighted average is calculated by counting instances of each class independently and then dividing the total by the total number of classes. This comes in handy when there is a disparity across social groups.

Accuracy: a well-known frequently used metric that counts the proportion of correctly labelled illustrations relative to all occurrences in the dataset, is a popular and widely used statistic that works well when classes are distributed fairly, through out a dataset. however, accuracy yet might be deceptive if there are social inequalities, as shown in eq (5).

Accuracy= (TP+TN) / Total Instances (5)

0 represents negative class (genuine profile) or “no” class and 1 indicates to the positive class or “yes” class (malicious account).

4.2.1 Naïve bayes Classification

In the below table2, it is measured the Naïve bayes accuracy:93.7442502299908, naïve bayes is best suited for text classification, it could handle and classify well user’s profiles based on their activities.

Table2: Classification Report of Naïve bayes

	precision	Recall	F1-score	Support
0	0.96	0.93	0.95	627
1	0.91	0.94	0.93	460
Accuracy			0.94	1087
Macro average	0.93	0.94	0.94	1087
Weighted average	0.94	0.94	0.94	1087

4.2.2 SVM Classification Report:

In the below table3, it is measured the SVM accuracy:94.48022079116836, SVM performance is depends on the only support vector data points and can classify fake profiles even when the data set have few instances.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper Volume 11, Iss 11, Nov 2022

Table3: Classification Report of SVM

	precision	Recall	F1-score	Support
0	0.94	0.96	0.95	627
1	0.95	0.92	0.93	460
Accuracy			0.94	1087
Macro average	0.95	0.94	0.94	1087
Weighted average	0.94	0.94	0.94	1087

4.2.3KNNClassificationReport

In the below table4, it is measured the accuracy of KNN is 93.65225390984361, since KNN is more influenced by the out-layer data points, comparatively it measured less performance than other models.

Table4: Classification Report of KNN

	precision	recall	f1score	support
0	0.93	0.96	0.95	627
1	0.94	0.91	0.92	460
Accuracy			0.94	1087
macro avg	0.94	0.93	0.93	1087
weighted avg	0.94	0.94	0.94	1087

4.2.4NLP techniques

NLP is technologies exist by systems to recognize, examine, manage, examine, translate human's languages. It handles with text data. NLP is inspired to design the technology that understands the human language with greater efficiency. In table 5, it is shown the accuracy of NLP is 95.76816927322908.

Table5: Classification Report of NLP

ISSN PRINT 2319 1775 Online 2320 7876
Research paper Volume 11, Iss 11, Nov 2022

	precision	recall	f1score	support
0	0.96	0.96	0.96	627
1	0.95	0.95	0.95	460
Accuracy			0.96	1087
macro avg	0.96	0.96	0.96	1087
weighted avg	0.96	0.96	0.96	1087

4.2.5 Model performance: In the figure5, SVM, NB and NLP could attain better performance over KNN. SVM, NB and NLP can suppress the impact of out layers, where KNN is unable to handle the out layers.

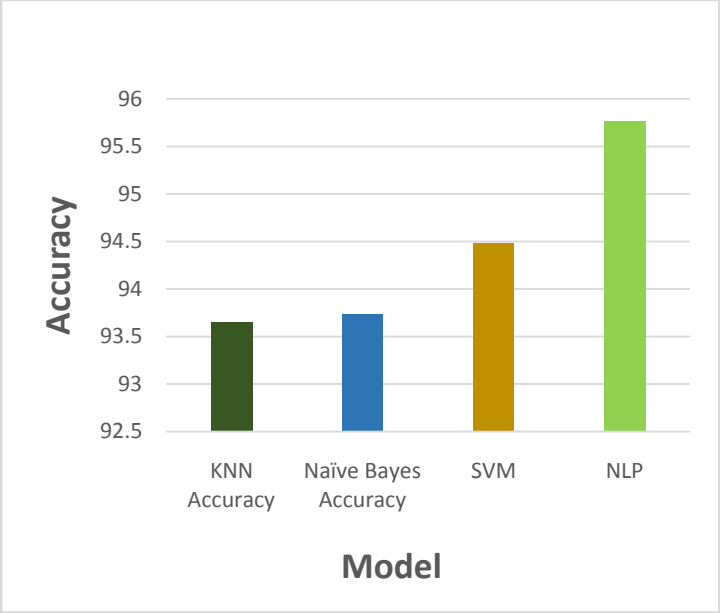

Figure5: comparison of different models in detecting fake Profiles

5. CONCLUSION AND FUTUREWORK

In this case, in addition to natural language processing approaches, ML models were used to develop this concept. These techniques can easily be implemented on social media sites in order to detect fraudulent profiles. For the purpose of highlighting the fraudulent profiles in this paper, we have analyzed the Facebook dataset. Our methods for analyzing the dataset involves using standard NLP preprocessing techniques &ML models such as the SVM, Naïve Bayes and KNN to organize the profiles. In this paper, we have improved detection accuracy with the help of learning algorithms.

One significant issue arises from the fact that someone can have multiple Facebook accounts, which gives those who create fake accounts and profiles on social media websites an added advantage. As a part of signup process, means whenever you create an account, you'll be able to enter your 12-digit Aadhar card number. Consequently, there is no possibility of fraudulent profiles in social media platforms because a single account can be set up for a single user.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper Volume 11, Iss 11, Nov 2022

REFERENCES

[1]. Akshay J. Sarode and Arun Mishra."Audit and Analysis of Impostors: An experimental approach to detect fake profile in an online social network. In Proceedings of the Sixth International Conference on Computer and Communication Technology 2015 (ICCCT '15). ACM, New York, NY, USA, 1-8.

[2]. Gurajala S., White, J.S., Hudson, B. and Matthews, J.N.,"Fake Twitter accounts: profile characteristics obtained using an activity-based pattern detection approach". International Conference on social media & Society2015, pp. 1-7.

[3]. M. Egele, G. Stringhini, C. Kruegel and G. Vigna, "Towards Detecting Compromised Accounts on Social Networks," in *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 4, pp. 447-460, 1 July-Aug. 2017,

[4]. Devakunchari Ramalingam, ValliyammaiChinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review. Computers & Electrical Engineering, Volume 65, 2018, Pages 165-177,

[5]. Ozbay, F.A. and Alates, B., "Fake news detection within online social media using supervised artificial intelligence algorithms". 2020. Physica A: Statistical Mechanics and its Applications, 540, p.123174.

[6]. Kaubiyal, Jyoti, and Ankit Kumar Jain. "A feature-based approach to detect fake profiles in Twitter." In Proceedings of the 3rd International Conference on Big Data and Internet of Things, pp. 135-139. 2019.

[7]. D. M. Freeman and T. Hwa, "Detecting Clusters of Fake Accounts in Online Social Networks Categories and Subject Descriptors," Artificial Intelligence and Security-AISec, 2015.

[8]. K.B.kansara, "security against sybil attack in social network", ICICES.no.icicles,2016, pp.1-5.

[9]. K. Shu, S. Wang and H. Liu, "Understanding User Profiles on Social Media for Fake News Detection," *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, Miami, FL, USA, 2018, pp. 430-435, Doi:10.1109/MIPR2018.00092.

[10]. Sai Pooja, G., Rajarajeswari, P., YaminiRadha, V., NavyaKrishna.G., Naga Sri Ram.B,Recognition of fake currency note using convolutional neural networks",International Journal of Innovative Technology and Exploring Engineering,(2016). pp 58-63.

[11]. Kodati, S.,Reddy, k.p., Mekala, S.,Murthy, p.s and Reddy ," detection of fake profiles on twitter using hybrid SVM algorithm". 3rd International Conference on Design and Manufacturing Aspects for Sustainable Energy (ICMED-ICMPC 2021).

[12]. Meshram, E.P., Bhambulkar, R., pokale, p., Kharbikar, k. and Awachat, A. Automatic Detection of fake profile using machine learning on Instagram. International Journal of scientific Research in science and technology, (2021) Vol 8, pp 117-127.

[13]. Adikari, Shalinda and Dutta, Kaushik, "Identifying Fake Profiles in Linkedin" (Pacific Asia Conference on Information Systems (PACIS) 2014 Proceedings.

[14]. Reddy, A. V. N., &Phanikrishna, C. "Contour tracking based knowledge extraction and object recognition using deep learning neural networks",2nd International Conference on Next Generation Computing Technologies in NGCT 2016, 352-354.

[15]. D. M. Freeman and T. Hwa, "Detecting Clusters of Fake Accounts in Online Social Networks Categories and Subject Descriptors," Artificial Intelligence and Security-AISec, 2015.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **Volume 11, Iss 11, Nov 2022**

[16]. Romanov, Aleksei, Alexander Semenov, Oleksiy Mazhelis, and Jari Veijalainenin, "Detection of fake profiles in social media-Literature review." International Conference on Web Information Systems and Technologies, vol. 2, pp. 363-369. SCITEPRESS, 2018.

[17] NageswararaoSirisala, Anitha Yarava, YCA PadbanabhaReddy, P. Veeresh "A Novel Trust Recommendation Model in Online Social Networks Using Soft Computing Methods", in Concurrency and Computations: Practice and Experience, Vol34(22),
pp1-17.

[18] Anitha Yarava, C. Shoba Bindu,"An efficient trust inference model in online social networks using fuzzy petrinets", Concurrency and Computations: Practice and Experience, Vol35(6).