TOOL TO DETECT FAKE ACCOUNTS INTWITTER - A CASE STUDY USING MACHINE LEARNING ALGORTHIMS

Madhwaraj Kango Gopal
Department of MCA
New Horizon College of Engineering(VTU)
Bangalore, India
dr.madhwaraj@newhorizonindia.edu

V. Asha
Department of MCA
New Horizon College of Engineering(VTU)
Bangalore, India
asha.gurudath@gmail.com

Binju Saju
Department of MCA
New Horizon College of Engineering(VTU)
Bangalore, India
binjusaj@gmail.com

Anju Shree R
Department of MCA
New Horizon College of Engineering(VTU)
Bangalore, India
anjurshree@gmail.com

Aishwarya R
Department of MCA
New Horizon College of Engineering(VTU)
Bangalore, India
aishwaryar1407@gmail.com

Abstract— **On the internet social media platforms are growing in popularity, and that popularity has raised increasing questions about security and privacy. Users of social networks face serious security risks due to fake and copied profiles. One major issue is the cloning of user profiles, when duplicate profiles are created using stolen user data and then utilized to harm the original profile owner. Threats like phishing, stalking, spamming, and others are also used to accomplish a variety of goals. A fake profile is one that is made on a social networking site using the name of an organization or person that does not exist and participates in destructive activities. In this study, a new tool is created that uses machine learning methods to verify user identification. People who utilize fake accounts may be recognized as having fake profiles in one of three ways: the number of abuse reports, daily comments, or rejected friend requests. Data from Twitter was used in a case study. The Random Forest algorithm and the Support vector machine approach offered a greater projected accuracy when detecting whether a user was a fraudulent or genuine user compared to other machine learning methods.**

Keywords— *Fake Profiles, Twitter datasets, social media, supervised.*

I. INTRODUCTION

Social networking platforms are those that help users to communicate with each other. Networks have grown, there has been an increasing trend in the security of such phone accounts. Users are creating false accounts on the network and exploiting them to carry out serious cybercrimes, steal data, and carry out other actions. Social networking websites appeal to a wide range of people since users can share their data and daily activities on them. One of the most popular social networking services is Twitter. Users can add friends and exchange a variety of information, such as personal, social, political, and corporate information, on online social networking sites.

The document is structured as shown below. The Literature Survey is covered in Chapter 2. The suggested technique used in this work is highlighted in Chapter 3. Chapter 4 talks about the algorithms that were used in this study. Chapter 5 highlights the results and discussion. Chapter 6 concludes the study with future work and conclusions.

II. LITERATURE REVIEW

Today, a few machine learning algorithms are employed because of the advantages they provide, including prediction, analysis, model development, etc. Machine learning approaches were utilized by Madhwaraj & Amirthavalli [3] to forecast the maintainability of free software. To estimate the number of deaths brought on by dengue fever, Viswanath et al. [1] conducted a case study. An innovative machine-learning concept was found by Madhwaraj et al. [2] to forecast the sale of washing machines inside an organization. There are several uses for machine learning algorithms.

User with false profile identification as a way to address challenges with social networking, such as privacy, online bullying, abuse, trolling, and other issues, also known as NLP and machine learning, was presented. [4].

Adikari and Dutta's [5] description on how to spot bogus LinkedIn profiles. Using just minimal profile data as input, the article demonstrates that phony with 84% accuracy and 2.44% false negatives, profiles may be found.

The approaches for spam identification in Facebook and Twitter are described by Stringhini et al [6]. The writers set up 900 fake profiles on social media and collected incoming messages and friend requests continuously for a whole year. Following the collection and analysis of user data from those who made these requests, around 16000 spam accounts were found. The authors looked deeper into how machine learning may be used to detectspamming profiles.

Aleksei et al.'s study, "Detection of Fake Profiles in Social Media," [7] was published to raise awareness of an entity related to a person but distinct from them. Using social media platforms to impersonate someone or construct a false identity in order to gain the target's confidence is a common situation for the use of false identities. This trust is then abused to acquire data for spear phishing attacks, launch spear phishing assaults, or engage in direct communication with the target to learn their information.

A. K. Visvam Devadoss et al. (eds.), *Proceedings of the 6th International Conference on Intelligent Computing (ICIC-6 2023)*, Advances in Computer Science Research 107,
https://doi.org/10.2991/978-94-6463-250-7_29

Based on spatiotemporal co-occurrence, "Malicious Users" is a circle detection seen in a social network, [8] which was related to a typical situation forutilizing Stein et al. discussed several privacy options.
[9] to safeguard user personal information in networks. Facebook users are sharing a lot of private information about themselves and their friends [10].

Spam and malicious posts on online social networks, Data Mining: Concepts and Techniques by Maclean et al. Numerous fake account detection algorithms were released in 1999 [11] to pinpoint the characteristics—or a combination of characteristics—that aid in differentiating between legitimate and fraudulent accounts. These tactics are based on the analysis of unique social network profiles.

Research on viral memes was examined by Kinuthia [12], who also developed a machine learning method to recognize campaigns. She used hashtags and a million postings to show that correct recognition could be achieved up to 95% of the time.

A mixed strategy based on machine learning and detection methods was investigated to find fake accounts [13]. This experimental approach abundantly illustrated the power of the provided strategy in terms of recognizing fake accounts from the testing dataset with high accuracy labels.

III. PROPOSED METHODOLOGY

The profiles of users were taken from the Twitter dataset and used in this research project. Several machine learning techniques were used on the dataset. Support Vector Machines, Logistic Regression Decision Trees, Random Forest, XGBoosts, and Nearest Neighbours Algorithm were some of the machine learning algorithms [14]. These algorithms were applied to determine if a certain user was real or a scam.

3.1 OVERVIEW OF THE PROPOSED SYSTEM

The gadget used in the proposed system has the crucial capability of user detection. Fraudulent or false accounts are discovered using techniques that might tell fake ones from real ones. Most fraudulent accounts lack a profile name and an image, which is one of the criteria used to identify them. They make no reference to the account. Although they frequently tweet frequently, there are times when their profiles may be entirely vacant. A detection technique exists to identify cloned and fake Twitter accounts. The frequency of abuse reports, daily remarks, or declined friend requests are three indicators that someone is using a false identity. Ten characteristics make up the suggested system. The proposed system includes the nine numerical data-type properties listed in Table 3.1, and the category feature—which determines whether a profile is fake or not—is a categorical variable.

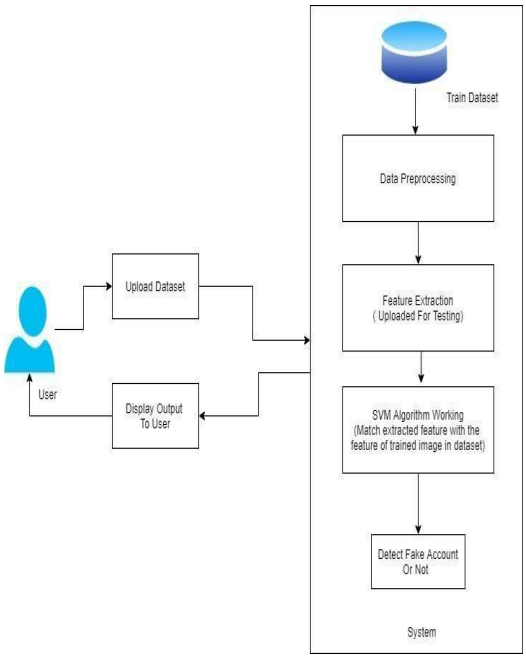

Figure 3.1 System Architecture

Feature	Data type	Description
User ID	Numerical	Id of the user
Number of Abuse Report	Numerical	Number of abuse report in social media
Number of rejected friend requests	Numerical	Number of rejected friend requests in social media
Number of friend requests that are not accepted	Numerical	Number of friend requests that are not accepted in social media
Number of friends	Numerical	Number of friends in social media
Number of followers	Numerical	Number of followers in social media
Number of likes to Unknown account	Numerical	Number of likes to Unknown account in social media
Number of comments per day	Numerical	Number of comments per day in social media
Category	Categorical	Based on above reports, classified account as genuine category or not.

Table 3.1 TEN Features and Descriptions

Feature	Description
User ID	Id of the user
Number of Abuse Report	Number of abuse report in social media
Number of rejected friend requests	Number of rejected friend requests in social media
Number of friends	Number of friends in social media
Number of followers	Number of followers in social media
Number of likes to Unknown account	Number of likes to Unknown account in social media
Number of comments per day	Number of comments per day in social media

Table 3.2 SEVEN Features and Descriptions

3.2 MACHINE LEARNING PROCESS

An object's contains attributes, which can be categorized under the word "label." The traits along with the label make up the training data. The training data presented above is used to train the model. After the model has been successfully trained, test data should be utilized to assess it. Only the features are present in the test data (given as input), not the labels. The test data will be used to evaluate the train model, which was trained using the training data. If it correctly evaluates the data, the trained model is prepared for prediction. Figure 3.2 depicts the pipeline flow diagram for machine learning.

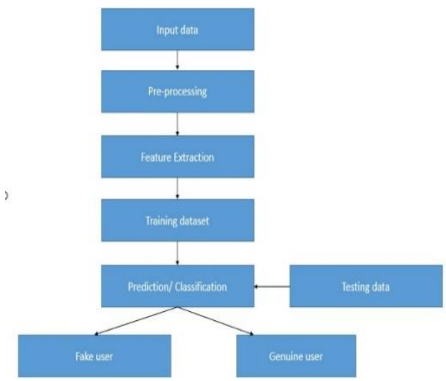

Figure 3.2. Machine learning f1o w diagram

3.3 DATA COLLECTION

The act of collecting data and information from numerous sources for use in analysis, research, or other objectives is known as data collection. When it comes to machine learning, gathering data is an essential stage in creating models that can produce precise predictions or classifications. The data may be collected using a variety of methods, including manual interventions and online scraping. Detection of Fake accounts in Twitter dataset taken from Kaggle.

3.4 DATA PRE-PROCESSING

The collection of data is the first significant stage in the construction of a machine-learning model. The quality and quantity of the data we gather at this step will ultimately decide how useful the model is; the more data we can get, the better. Web scraping and other manual interventions are examples of data-collecting techniques. There are 692 rows and 10 columns in the dataset. This module will change the data. by removing any columns and missing data. We will start by listing the column names that we wish to preserve orretain. The remaining columns are dropped or removed after that, leaving only the ones we want to maintain. Finally, we remove the data set's rows with missing values. Since there are no missing values in the dataset being used, models have already been built using this data.

IV. MODEL SELECTION

Two datasets are required when creating a machine learning model: one for training and one for testing. But for now, there is just one left. Use the 80:20 ratio to divide this into two equal halves. The data frame will also receive a label column and a feature column. The sklearn train_test_split function was imported in this place. After that, split the dataset using it. A test size of 0.2 further divides the dataset into 20% of test data and 80% of train data. The dataset is divided by a random number generator that is seeded by the random_state option. Four datasets are returned by the procedure. They were given the names test_x, test_y, train_x, and train_y, respectively. The divide is apparent when we examine the dataset's form. The data will be fitted to different decision trees using the Random Forest Classifier. By supplying the train_x and train_y parameters to the fit function, I train the model. It is necessary to test the model after it has been trained. Test_x will be sent to the predict function to make that happen.

Model Type	Parameters for Fitting the Model
Logistic Regression	Solver=lbfgs,'muli_class="auto', max_iter=2000
Random Forest	RandomForestClassifier(n_jobs=-1, random_state=123, criterion="gini', max_depth=3,)
KNN	KNeighborsClassifier(n_neighbors—7
SVM	Svm.SVC(Kernel='rbf',gamma='auto') # Linear Kernel
XGBOOST	Xgb.XGBClassifier(max_depth=10, learning_rate=0.1,n_estimatord=100, seed=10)
DECISION TREE Gini	Decision TrecClassifer(criterion = "gini", random_state = 100,max_depth=7, min_samples_leaf=5)
DECISION TREE - Entropy	DecisionTrecClassifer(criterion = "entropy", random state = 100,'max_depth=7, min_samples_leaf=5)

Table 4.1 The Model's Fitting Parameters

4.1 UNIVARIATE ANALYSIS

Analyzing just one variable Data visualization is a descriptive style of analysis that just uses one variable and makes no assumptions about how that variable is connected to other variables.

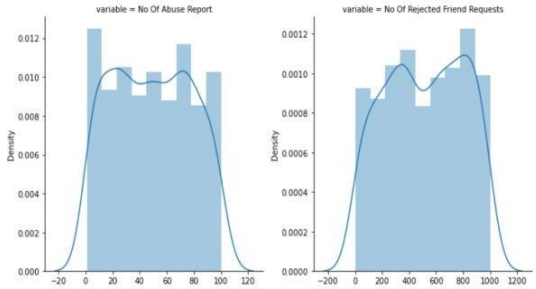

Figure 4.1 Counting the amount of abuse reports and friend requests that were declined

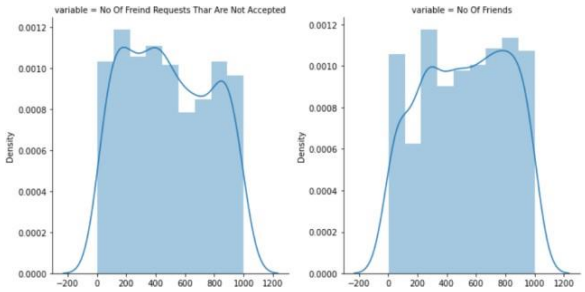

Figure 4.2 Counting the amount of friends and rejected friend requests.

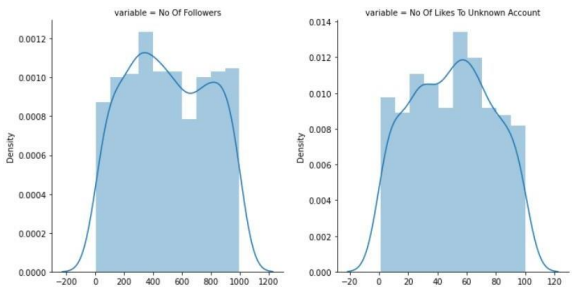

Figure 4.3 Examining the amount of likes and followers for an unidentified account

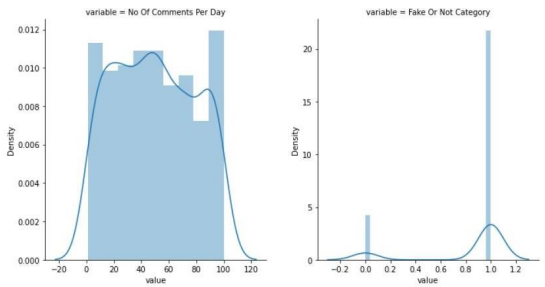

Figure 4.4 Examining the phoney or not category and daily comment count.

4.2 CORRELATION OF FEATURES

A bivariate study known as correlation [14] assesses the degree of association and the direction of a link between two variables. The correlation will range from +1 to 1.

There is a strong 60% link between the number of reports of abuse and the fake or not category. High degree of association (40%) between the category "Fake or not" and the number of likes to unknown accounts. There is a moderate (20%) association between the number of friend requests rejected and the fake or not category. Low level link between the false or not category (10%) and the numberof comments every day.

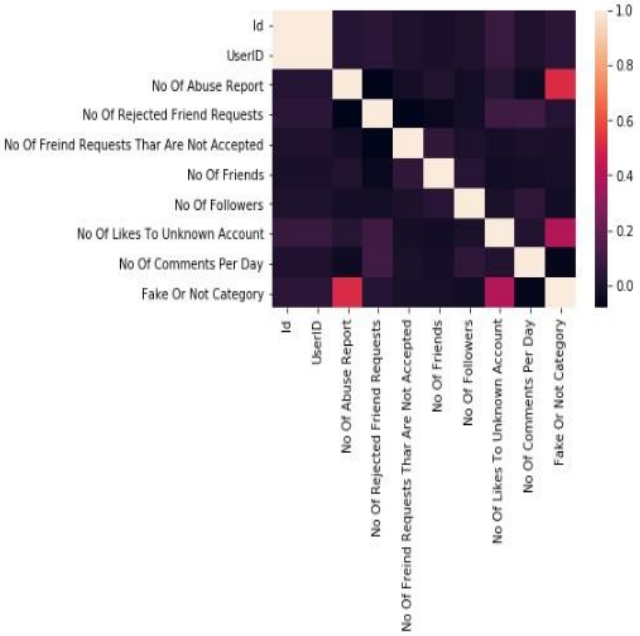

Figure 4.5 Pearson Correlation

4.1 ANALYZE AND PREDICTION

We only selected 7 characteristics from the actual dataset:

1) UserID: Twitter id
2) No Of Abuse Reports: The number of Abuse Reports
3) No Of Rejected Friend Requests: The number of Rejected Friend Requests Followers in the Twitter amount.
4) No Of Friends: The number of people friends on theTwitter amount
5) No Of Followers: The number of people Followers in the Twitter amount
6) No Of Likes to Unknown Account: The number of Likes to Unknown Account
7) No Of Comments Per Day: The number of Comments Per Day.

4.2 ACCURACY IN TESTING MODEL

Accuracy on the test would be calculated as per the dataset given.

4.3 SAVING THE TRAINED MODEL

First, use a library like pickle to store your trained and tested model in an a.h5 or pkl file if you are confident using it in a production-ready environment. Check to make sure Pickle is configured for your environment. At this point, the model will be loaded into the module and exported as a.pkl file.

V. EXPERIMENTAL RESULTS

A common tool for running and sharing data analysis experiments is Jupyter Notebook. You can create and run code in a web browser using Jupyter Notebook, and you can see the results in real-time.

Data visualization: A variety of libraries are available in Jupyter Notebook to help you produce appealing and instructive. Machine learning: You can create and test machine learning models with Jupyter Notebook. Scikit-learn, TensorFlow, and Python are just a few of the machine-learning packages that are available in Python.

Statistical Analysis: Using libraries like Pandas, NumPy, and SciPy,Jupyter Notebook enables you to do statistical analysis and hypothesis testing. Data exploration and manipulation, as well as the computation of descriptive and inferential statistics, are all possible. Text Analysis: Natural language processing (NLP) activities including sentiment analysis, text categorization, and text summarization may be carried out using Jupyter Notebook. Python has several NLP libraries.

Algorithm	Accuracy score (%)
Logistic Regression	94.23
Random Forest	94.71
KNN	82.21
SVM	82.21
XGBOOST	98
Decision Tree (Gini)	99
Decision Tree (Entropy)	98.5

Table 5.1 Experimental Results

The current user interface, Jupiter, incorporates all the well-known features of the original Jupiter Notebook, such as a notepad, terminal, text editor, file browser, rich outputs, and so on into a user experience that is both adaptable and powerful. After some time, Jupiter will supplant the ongoing Jupiter Note pad. Installation. Condo or Pip can be used to install Jupiter. Jupiter Lab 1.1.4 and Anaconda3-2019.10-Windows-x86-64 are both introduced on this machine. The following is a list of the libraries that are used in the implementation. Pandas is the most widely used Python data analysis library. It offers significantly enhanced performance due to the use of only C or Python for the back-end source code. SciPy is a Python-based open-source library used in scientific, engineering, and technical computing.

Data visualizations: To produce interactive plots, charts, and graphs, you may use Matplotlib, Seaborn, Plotly, or Bokeh It provides a sophisticated drawing instrument for creating statistical graphics that are captivating and instructive. A variety of physical formats and cross-platform interactive environments can be used in conjunction with the Python 2D plotting package Matplotlib, four graphical user interface toolkits, the Python and I Python shells, the Jupiter notebook, web application servers, and Python scripts to produce graphics of publication quality. A versatile array management library is NumPy. In addition to the capability to interact with multidimensional arrays, it provides an extremely quick array object.

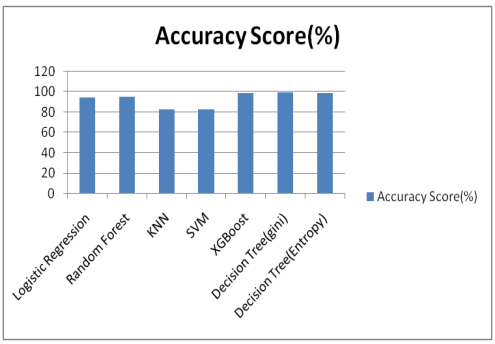

Figure 5.1 Accuracy Score

From the chart, we can see that the Decision Tree (Gini) algorithm has the highest accuracy score of 99, followed closely by Random Forest with a score of 94.71. Logistic Regression has a slightly lower score of 94.23. XGBoost has an accuracy score of 98, which is the second highest score among the algorithms. On the other hand, KNN and SVM have the lowest accuracy scores of 82.21, indicating that they may not be the best algorithms for the given dataset. Overall, the chart makes it easy to compare the accuracy scores of each algorithm and identify the best performing ones. From the chart, we can see that the Decision Tree (Gini) algorithm has the highest accuracy score of 99, followed closely by Random Forest with a score of 94.71. Logistic Regression has a slightly lower score of 94.23. XGBoost has an accuracy score of 98, which is the second highest score among the algorithms. On the other hand, KNN and SVM have the lowest accuracy scores of 82.21, indicating that they may not be the best algorithms for the given dataset. Overall, the chart makes it easy to compare the accuracy scores of each algorithm and identify the best performing ones.

VI. CONCLUSIONS AND FUTURE WORK

Finding false social media profiles is the major goal of this research. Machine learning algorithms are being used to raise customers are aware of false consumers. Administrators in this project oversee setting up the model for processing as well as administering the complete the programmed. Any extra training that the model needs will be handled by the administrator. The computation using this programmed was successful. In online social networks, fake and duplicate profiles have grown to be a very serious issue. These characteristics might occasionally manifest as threats in daily life. As a result, a detection technique that can spot both false Twitter profiles has been suggested. The issue may be expanded to include Natural Language Processing Methods.

REFERENCES

[1] Viswanath Bellie, Madhwaraj Kango Gopal, Govindaraj Venugopal “Using machine learning techniques towards predicting the number of dengue deaths in India – A case study (2014)”, International Journal of Applied Engineering Research ISSN: 2231-5381.

[2] Gopal, M. K., Bellie, V., & Venugopal, G. (2020). A novel machine learning technique towards predicting the sale of washing machines in a small organization. *Int J Psychosoc Rehab*, *24*(5), 6969-6976.

[3] Gopal, M. K., & Amirthavalli, M. (2019). Applying machine learning techniques to predict the maintainability of open-source software. *International Journal of Engineering and Advanced Technology*, *8*(5S3).

[4] Dr. Narsimha.G, Dr. JayadevGyani, P. Srinivas Rao,"Fake Profiles Identification in Online Social Networks Using Machine Learning and NLP (2018)", International Journal of Applied Engineering Research ISSN 0973-4562, Number 6, Volume 13.

[5] Adikari, S., Dutta, K., 2014. Identifying Fake Profiles in Linkedin, in: PACIS 2014 Proceedings. Presented at the Pacific Asia Conference on Information Systems.

[6] Stringhini, G., Kruegel, C., Vigna, G., 2010. Detecting Spammers on Social Networks, in: Proceedings of the 26th Annual Computer Security Applications Conference, ACSAC ’10. ACM, New York, NY, USA, pp. 1–9. doi:10.1145/1920261.1920263.

[7] Aleksei Romanov, Alexander Semenov, Oleksiy Mazhelis and Jari Veijalainen, Detection of Fake Profiles in Social Media Literature Review, University of Jyvaskyl, Finland.

[8] 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. Fritsch (2010).

[9] Stein T, Chen E, Mangla K,” Facebook immune system. In: Proceedings of the 4th workshop on social network systems”, ACM 2011, pp.1-8.

[10] Saeed Abu-Nimeh, T. M. Chen, and O. Alzubi, “Malicious and Spam Posts in Online Social Networks,” Computer,

[11] Maclean, J. Melton, J. Melton, A. R. Simon, and M. Chisholm, Data Mining: Concepts and Techniques. 1999.

[12] S. Kiruthiga, “Detecting Cloning Attack in Social Networks Using Classification and Clustering Techniques,” IEEE, 2014.

[13] N., Smruthi.M.,"A Hybrid Scheme for Detecting FakeAccounts in Facebook" ISSN: 2277- 3878, (IJRTE)International Journal of Recent Technology and Engineering (2019), Issue-5S3, Volume-7. \

[14] https://www.statisticssolutions.com/correlationpearson-kendall-spearman/.

[15] https://www.cs.princeton.edu/~schapire/talks/picasso-minicourse.pdf.