Twitter Fake Account Detection using Machine Learning

Nor Anatassia Risqah Binti Mokhtar, Dr.Shahrinaz Ismail
Department of Information System
University Kuala Lumpur, Malaysia Institute of Information
Technology Email address: nanatassia.mokhtar@s.unikl.edu.my

ABSTRACT- *Online social networks (OSNs) have grown in popularity in recent years, and people's social lives have become more intertwined with these sites. The fast proliferation of OSNs and the vast quantity of fake profiles on social media platforms such as Twitter are on the rise. These fake accounts are a popular way for criminal users to perform numerous cybercrimes. Hence, their detection and removal have become necessary to protect normal users and preserve trustworthiness. This project discusses how to recognize a fake Twitter account. This project uses Machine Learning to improve the prediction of fake accounts based on specified features such as name, status count, friends count, follower count, and favorite count. A neural network technique was employed to determine if an account was fake or legitimate.*

KEYWORDS-*Machine Learning; Fake Account Detection; Online Social Network; cyber crimes*

I. INTRODUCTION

The number of active Twitter users in Malaysia is projected to reach 2.63 million, up from 1.5 million in 2014. However, they suffer from expanding fake accounts that have been created by social media sites. The statistics were released by the Malaysian Department of Communications.

Machine learning is a massive part of artificial intelligence (AI) of a system to recognize patterns based on existing algorithms and datasets. Machine learning techniques have been used to classify the Twitter account as legitimate or not. A few techniques and algorithm that have been created to solve these issues.

The purpose of this research project is to develop a Twitter Fake Account Detection system that recognizes and detects those user accounts that are real or fake by using Machine Learning. To identify the fake accounts based on various features such as user profiles, malicious comments, or posting about political issues. Viewing the user's profile photo to identify whether the account is fake or not.

II. LITERATURE REVIEW

Researcher [5] describes the analysis of individual social network profiles relies on a variety of fake account identification methods to recognize the characteristics or a combination of them that help differentiate between genuine and fake accounts. Fake accounts are detected using the gradient boosting algorithm and with the generated decision trees. The paper concludes due to this gradient boosting was able to get a highly accurate result.

Based on research [1] identify the fake account in RenRen and Twitter. This paper shows that a fake account can be detected with 98% accuracy and 7% false-negative. The fake account was detected using two datasets of real Twitter followers. A new technique called support vector machine-neural network (SVM-NN) was used to detect fake accounts.

Researcher [6] describe used neuro-linguistic programming and text classification to identify fake accounts on Facebook and Twitter. The result showed 98% accuracy of SVM to identify the fake account and 97% accuracy that implemented in a social platform like Facebook or Twitter. In the Complement Naïve Bayes (CNB) classifier, researcher use bag-of-word (BOW) and must incorporate this on a platform like WEKA.

Researchers [7] found that Random Forest and Deep Convolutional Neural Networks give the best performance for the detection of fake vs real profiles on social media. To train the classifiers, experiment with 4-fold,10-Fold, and 12-fold cross-validation techniques.

III. PROPOSED METHODOLOGY

A new approach for detecting fake profiles on social networks has been proposed. The plan is to use publicly available information to identify the identity of the fake profile. This might be utilized by social media platforms like Facebook, Twitter, or Google. It tries to make it easier for individuals to distinguish between who they consider being real and who they consider being fake.

a. System Architecture

The proposed framework in figure 1 shows the sequence of procedures that must be followed for continuous identification of fake profiles with active learning from the feedback of the classification algorithm's results. This architecture is simple to adopt for social networking firms such as Twitter. The figure shows the steps that need to be taken by the algorithm to identify false profile names [4].

- Import the datasets that available online. There two datasets which are real users and fake users.
- After that, suitable features are selected for the classification algorithm to implement.
- After the features extracted are passed. Split into two, 30% for testing and 70% for training the datasets.
- Next, build the machine learning model and train the model.
- In this project, neural network classification has been chosen to detect fake accounts.
- Lastly, the classifier determines the system whether the profile is fake or real.

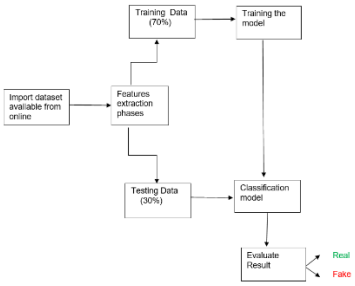

Fig 1: System Architecture Machine Learning

b. Flowchart

A neural network model will train the data users to enter and the result will come out. User needs to enter manually Name, Status_name, Follower_name Friend_count, Favourite_count. After that, the neural network will train a machine-learning algorithm to identify the fake account. The results will be shown in the flowchart below.

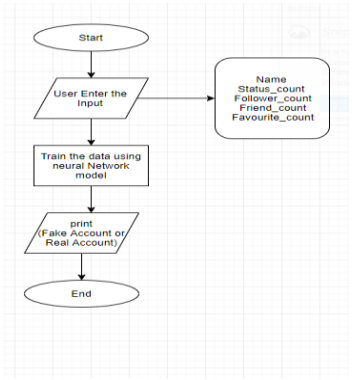

Fig 2: Flowchart for the user.

IV. RESULT AND DISCUSSION

After the implementation of a new system, testing should be performed to check that the application is functioning properly. These tests are also designed to confirm that all setup goals have been met. In this case, it is important to make sure that the system is working properly and that all set up goals are being met. This testing the functionality consists machine learning model. gender detection, and account prediction.

a. Machine Learning Model

Several neural network techniques are used to train models and forecast results based on previously learned models. The feed-forward back propagation technique was chosen as the basic algorithm. The expected results were compared to the actual genuine values (that is, whether the account is real or fake) and the prediction accuracy was computed [7].

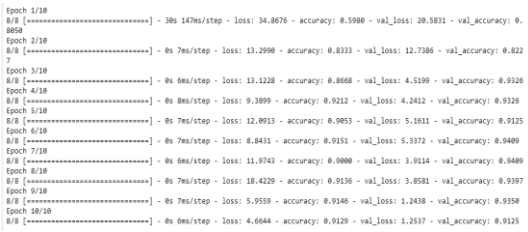

Fig 3: Accuracy training machine learning model

Based on figure 3 above, a machine learning algorithm has been able to predict the accuracy of a -time training dataset. The data was used to create a model that could be used to predict how accurate it would be after ten times. The results show that the algorithm is very good at predicting the accuracy, but not as good as it could have been in the past.

b. Gender Detection

Natural Language Toolkit (NLTK) is a framework for developing text analysis systems. Names that finish in a, e, or I am more likely to be female than those that end in k, o, r, s, or t. Names ending in the letter a or e are more often male than those ending in s. The classifier was created to better model these differences between male and female names.

```
featuresets = [(gender_features(n), g) for (n,g) in names]

train_set, test_set = featuresets[500:], featuresets[:500]

classifier = nltk.NaiveBayesClassifier.train(train_set)

print (classifier.classify(gender_features('Anat')))
```

```
male
```

Fig 4: gender detection

Based on the figure above shows, depicts the coding to configure for training gender characteristics, whether male or female. The above result is based on testing for the name Anat. Anat's name indicates that he is a male.

c. Account Prediction

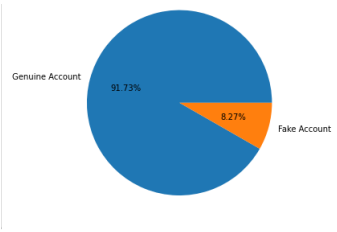

Fig 5: Genuine Account

In figure 5 shows a pie chart of the real account after users enter the data.

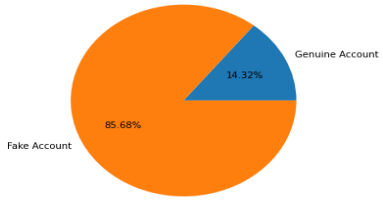

Fig 6: Fake Account

Based on figure 6 shows the pie chart of prediction fake accounts.

V. CONCLUSION

This project used supervised machine learning to develop fake account detection. The name "multi-layer perceptron" comes from the fact that a neural network is made up of multiple perceptron layers. These layers are also known as hidden layers of dense layers and are composed of many perceptron neutrons. They are the basic building blocks that make up a perceptron layer.

The project aims to connect the machine with Twitter API to retrieve the real user profile. It also aims to improve accuracy by using different ML algorithms to improve accuracy. To ensure this project executed as planned and progressively completed, recommendations and betterment is needed.

REFERENCE

[1] S. Khaled, N. El-Tazi, and H. Mokhtar, “Detecting Fake Accounts on Social Media,” 2018, pp. 3672–3681. doi: 10.1109/BigData.2018.8621913.

[2] B. Ersahin, Ö. Aktaş, D. Kilinç, and C. Akyol, “Twitter fake account detection,” 2017, pp. 388–392. doi: 10.1109/UBMK.2017.8093420.

[3] F. Masood, G. Ammad, A.Almogren, A.Abbas,H.Khattak, “Spammer Detection and Fake User Identification on Social Networks,” IEEE Access, vol. 7, pp. 68140–68152, 2019, doi: 10.1109/ACCESS.2019.2918196.

[4] N. Kumar and R. N. Reddy, “Automatic Detection of Fake Profiles in Online Social Networks,” 2012.

[5] S. P. Maniraj, Harie Krishnan G, Surya T, and Pranav R. 2019. “Fake Account Detection using Machine Learning and Data Science,” *IJITEE*, vol. 9, no. 1, pp. 583–585, Nov. 2019, doi: 10.35940/ijitee.A4437.119119.

[6] R. Raturi, “Machine Learning Implementation for Identifying Fake Accounts in Social Network,” 2018.

[7] P. Shahane and D. Gore, “Identification of Fake vs. Real Identities on Social Media using Random Forest and Deep Convolutional Neural Network,” 2019.