**Fake Social Media Profile Detection**

Satyam Rai

10/12/2023

Abstract:

In this report, I have proposed the idea of using Fake Social Media Detection algorithm which uses various parameters to increase your security over various social media platforms and also reports fake profiles protecting your data, helping you maintain your privacy, saving you from internet scams and much more.

Problem Statement:

The social life of everyone has become associated with online social networks. These sites have made a drastic change in the way we pursue our social life. Making friends and keeping in contact with them and their updates has become easier. But with their rapid growth, many problems like fake profiles, online impersonation have also grown. Fake profiles often spam legitimate users, posting inappropriate or illegal content. Several signs can help you spot a social media fake who might be trying to scam your business. Identifying fake social media profiles and taking corrective measures. Expected Output: An Application software that detects the fake social media profile Users: Crime branch and other investigative agencies.

Market/Customer/Business need Assessment:

In the rising age of social media, scam rates are increasing even more rapidly. Many phishing websites and scam profiles use social media as a medium to connect with people. Every year millions of people are attacked by Cyber Crime and many fall victim to these, losing them lots of money or important data which can cause a lot of harm. Using fake social media detection algorithm, we aim to identify these scammers before they try to attack you online and report these profiles to the authorities.

Target Specification:

The proposed system will allow the social media users to identify if a profile is fake or not, alerting them even before the scam happens ensuring they have a safe experience over the social media, thus protecting their data.

External Search:  
Sources I have used for this algorithm and the dataset used to train the model are taken from Kaggle and are attached below.

Business Model:

We will develop a browser extension which will predict if the social media account is fake or not and will automatically take preventive measures if the profile turns out to be fake.

This feature can be sold to many cyber security companies which provide protection over social media.

It can be implemented as an annual subscription to the users for the extension which runs directly in their browser and takes the input automatically through web scraping and only alerting them if the profile turns out to be a fake one. The subscription annually charges them and makes it very convenient for both the user and for us to maintain the systems and provide the services.

This feature can also be sold to many cyber security companies which provide protection over the internet such as Norton and Quick Heal.

Concept Development(Brief summary of Product that will be developed): The Final product developed will be a browser extension which will take the input data with the help of web scrapping and pass the inputs in the required format to the ANN model which will then predict if the profile is fake or not and, if fake, will report the profile and also report the user, keeping him safe.

Final Product Prototype with schematic diagram:

It will be a browser extension which automatically takes the input data whenever you open any profile of all the major social media platforms like twitter, Instagram, reddit, etc. We use web scrapping to take the input directly from the browser background and direct this input in the required form to the model. Then our ANN model predicts if the profile is fake or not. If it is not fake, the user continues without getting affected. But if the profile turns out to be fake, it will immediately be reported and the user will be prompted an alert making him aware of the scammer saving his data over the internet.

A white paper with text on it

Description automatically generated

Product details:

*How does it work*: It uses an Artificial Neural Network model with an input layer, 3 hidden layers and an output layer. It is a classification model which outputs in either 0 or 1 depending on the profile being fake or not. It uses Adam optimizer and ReLU activation function.

*Data Sources*:



The datasheet used for EDA and for training the model is inserted above. They are taken from Kaggle.

*Algorithms*: Three machine learning models were used to compare the performance and the model with the highest Precision and Recall was chosen to train the model. Three models used were Linear Regression, Random Forest and ANN model.

*Team*: Members from following domain would be required

Web Developer

Deep Learning Engineer

Machine Learning Engineer

Data Scientist

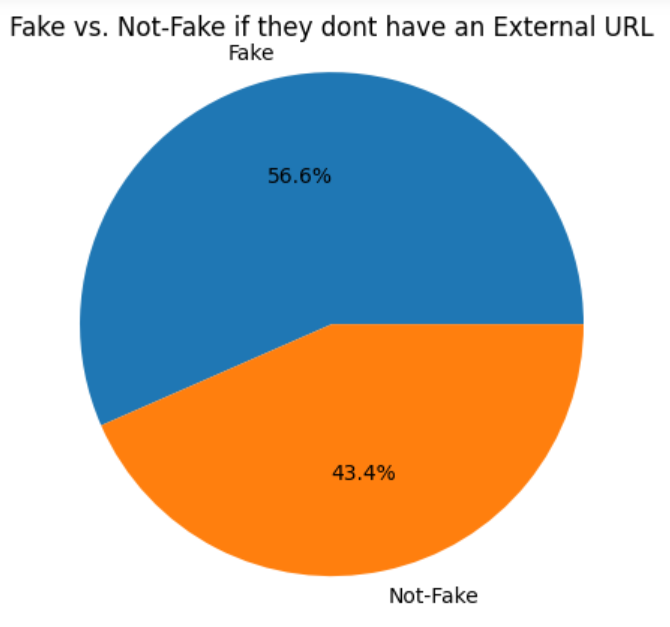
DevOps Engineer

Code Implementation/EDA:

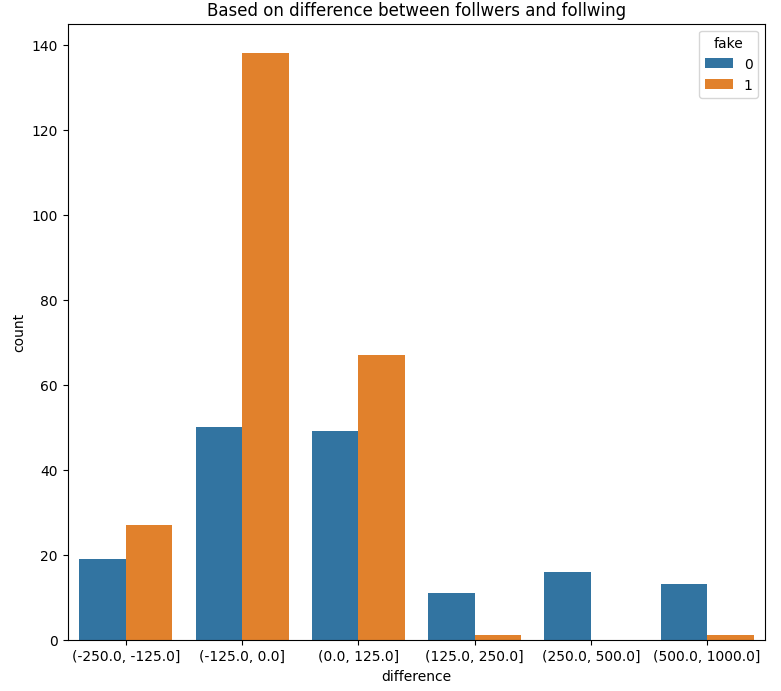


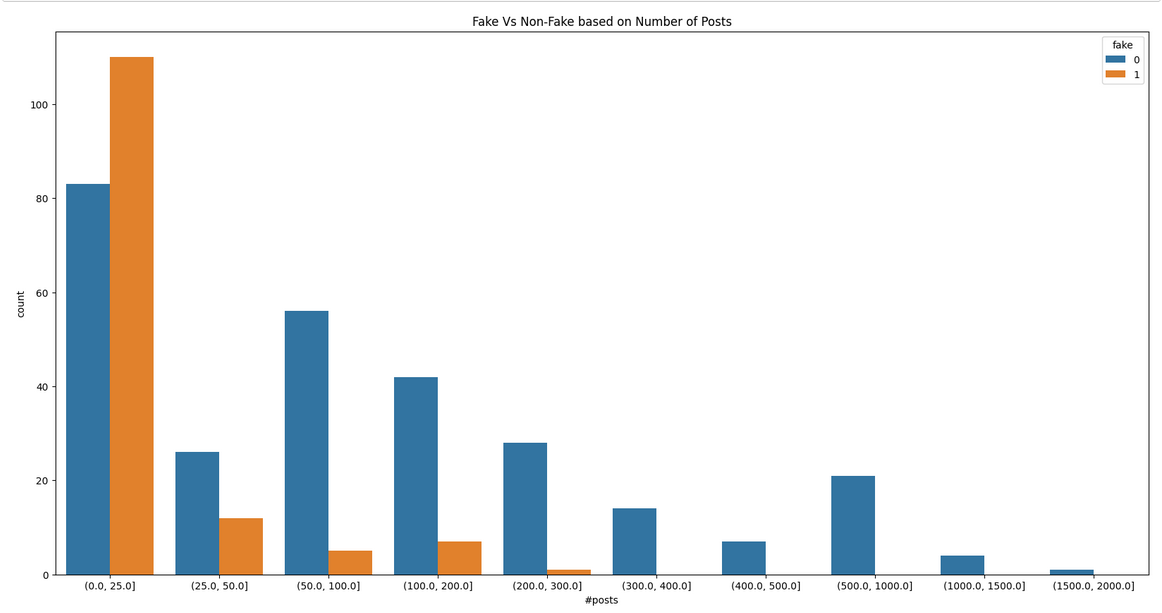
This is the notebook used for EDA.

This pie chart shows the relation between fake profiles vs not-fake profiles if an external URL is not mentioned.

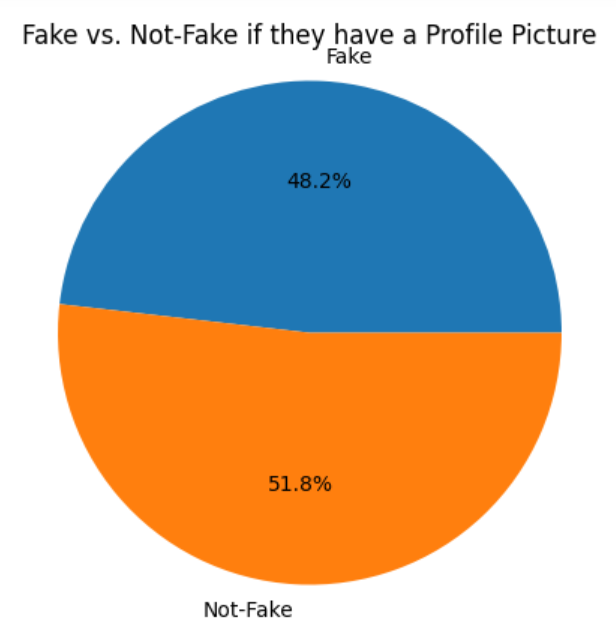


This countplot shows that if your followers count is more, then chances of the account being fake drastically decreases.

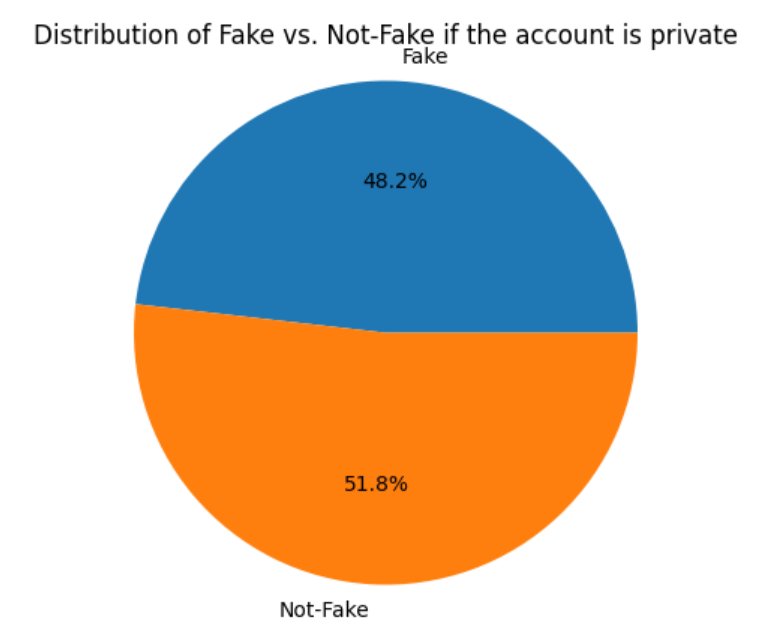


This shows that as the number of posts increases, chances of account being fake decreases drastically.

This pie chart shows the relationship between fake and not-fake profiles if they have a profile picture.



This pie chart shows the relationship between fake and not-fake profiles if the account is set to private.



Conclusion:

Viewing the current trends, it is safe to say that social media is going to continue to grow rapidly, and the rising rate of cybercrime is an important concern which is required to be handled appropriately. With the help of this project, I intend to provide a solution to reduce internet scams and protect the users.