Fake Profile Identification on Social Media

D Naga Bhushanam[1], A Venkatesh[2], G Swetha[3], G Vyshnavi[4] , V Yaswanth Sai[5]
Assistant Professor, Student, Department of CSE, AITS, Tirupati

Abstract

In the current generation, everyone's social life is now entwined with online social networks. The way we pursue our social lives has drastically changed as a result of these websites. It has gotten simpler to make friends and stay in touch with them and their updates. But along with their rapid expansion, several issues including phone profiles and online impersonation—have gotten worse. There are no practical methods available to handle these issues. I developed a methodology in this research that makes it easy and effective to identify fraudulent profiles automatically. To categorize the profiles into false or real classes, this system uses classification algorithms like CNNs or CNN-like neural networks. This automatic detection approach can be simply used by internet users because it is automatic.

Keywords: social networks, Facebook, Convolutional Neural Networks, Classification, Framework, and Dataset.

Introduction

Each user of a social networking site has a profile and can communicate with friends, post updates, and connect with new individuals who share their interests. These Online Social Networks (OSN) make use of web 2.0 technologies to enable user interaction. Social networking services are expanding quickly and altering how individuals communicate with one another. People with similar interests are brought together by online groups, making it simpler for users to establish new connections.

History

The first of these social networking sites, http://www.sixdegrees.com, appeared in 1997, followed by http://www.makeoutclub.com in 2000. Sixdegrees.com couldn't last long and shut down quickly, but new websites like Myspace, LinkedIn, and Bebo became successful. Facebook was introduced in 2004 and is currently the biggest social networking site in the world.

Social Impact

In the current generation, everyone's social life is now entwined with online social networks. The way we pursue our social lives has drastically changed as a result of these websites. It has gotten simpler to add new friends and stay in touch with them and their updates. These online social networks have been the subject of research to see how they affect people. Instructors may easily instruct students through this, creating a welcoming atmosphere for them to learn. Teachers are becoming more familiar with these sites nowadays, introducing online classroom pages, assigning assignments, holding conversations, etc., which greatly enhances education

1.Literature Review

Several methods for identifying false records rely on the analysis of individual interpersonal organization profiles to identify characteristics or a combination of characteristics that aid in differentiating between real and fake records. Specifically, several characteristics are gathered from the profiles and posts before using machine learning methods to build a classifier capable of identifying fraudulent data.

For instance, identifying and defining phantom profiles in online social gaming apps is covered in Nazir et al. (2010) [1]. The article examines a Facebook application called "Fighters Club," an online game that is notorious for rewarding players who bring their friends into the game with bonuses and other advantages. The authors argue that by providing such incentives, the game encourages its users to create fictitious accounts. The user would boost the incentive for himself by introducing such false accounts into the game.

Adikari and Dutta (2014) [2] show unmistakable evidence of bogus LinkedIn profiles. The study shows that with limited profile information as input, phony profiles may be identified with 84% accuracy and 2.44% false negatives. Methods including principal component analysis, neural networks, and SVMs are used. Highlights like the number of languages a person speaks, training, skills, suggestions, hobbies, and accolades are used, among other things. As a starting point, the characteristics of profiles that are known to be false and are placed on unusual websites are used.

To distinguish between Twitter accounts run by humans, bots and cyborgs, Chu et al. (2010) [3] advocate this (i.e., bots and people working in concert). An Orthogonal Sparse Bigram (OSB) text classifier that employs pairs of words as features are used to identify spamming records as part of the formulation of the detection issue.

Twitter supporter markets are analyzed by Stringhini et al. (2013) [4]. They list the characteristics of Twitter aficionados who promote and classify the patrons of various company sectors. According to the authors, there are primarily two types of accounts that pursue the "client": hacked accounts and phony accounts (also known as "Sybils"), whose owners do not assume that the number of their followers is growing. Customers of follower markets may include politicians or celebrities who want to appear to have a larger fan following, or they may include crooks who want to make their profile appear more real so they can transmit malware and spam more quickly. Thomas et al(2013) .'s [5] investigated Twitter spam distribution by black market accounts.

By distributing honeypot pages, De Cristofaro et al. (2014) [6] examine Facebook-like cultivation.

According to Viswanath et al(2014) .'s [7] analysis of abnormalities in their like behavior, black market Facebook data can be found.

Two black hat online commercial hubs, SEO Clerks and My Cheap Jobs, are examined by Farooqi et al. (2015)[6]. Fayazi et al. (2015) consider online review manipulation.

Related Work

2.Material and Methods:

1.1 Dataset:

A collection of false and real profiles was necessary. The dataset contains several properties, such as the number of friends, followers, and status counts. Training and testing data are separated in the dataset. A training dataset is used to train classification algorithms, while a testing dataset is used to assess the algorithm's effectiveness. 20% of both profiles are utilized to generate a testing dataset, and 80% of both profiles are used to prepare a training dataset from the dataset.

1.2 Attributes Considered

The factors taken into account for fake profile identification are included in Table 2 along with a brief description of each attribute

Table 2: Attributes Considered for the fake profile Identification

S.No	Attribute	Description
1.	Profile ID	Profile ID of account
2.	Profile Name	Profile Name of account
3.	Status Count	Status Count of account
4.	Followers Count	Followers Count of account
5.	Friends Count	Friends Count of account

6.	Gender	Gender of account holder
7.	Language Code	Language Code of account

2. Algorithm

2.1 Convolutional Neural Network

A typical Deep Learning architecture for image classification and identification applications is the convolutional neural network (CNN). It has many layers, including fully connected, pooling, and convolutional layers.

To extract features from the input picture, the convolutional layer applies filters. To decrease computation, the pooling layer samples the image, and the fully connected layer then produces the final prediction. With the help of gradient descent and backpropagation, the network learns the best filters. The reader is presumed to be familiar with the idea of neural networks.

Artificial neural networks do incredibly well in machine learning. Artificial neural networks are used to classify text, sounds, and images among other things. Various forms of neural networks are employed for various tasks. For example, to predict the order of words, recurrent neural networks—more specifically, an LSTM—are used. Similarly, to classify images, convolution neural networks are employed. We're going to create the fundamental building element for CNN in this blog.

The input layer is where we provide input to our model. The entire number of characteristics in our data is equal to the number of neurons in this layer (number of pixels in the case of an image).

The hidden layer is then fed the information from the input layer. Depending on our model and the volume of the data, there may be numerous hidden levels. The number of neurons in each hidden layer might vary, although they are often more than the number of features. Each layer's output is calculated by multiplying the output of the layer below it by its learnable weights, adding learnable biases.

The output of each class is then converted into the probability score for each class using a logistic function, such as sigmoid or SoftMax, using the data from the hidden layer as input.

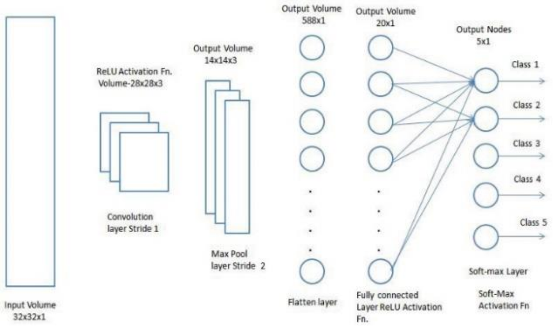

Figure 3: Working of CNN

4. Result:

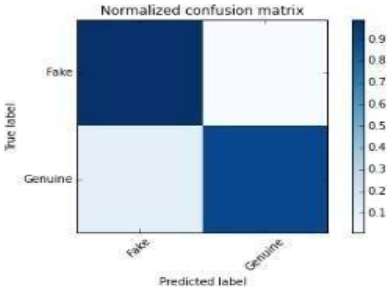

Figure 3: Normalized Confusion Matrix

ROC Curves:

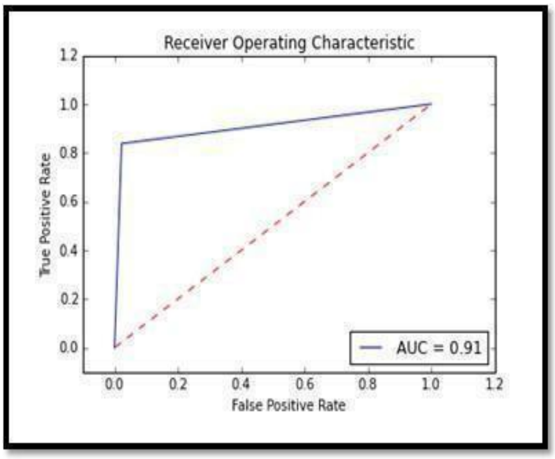

Figure 4: ROC Curve

The efficiency of neural networks in classifying data is 95%. We have taken 80% of the data for training neural networks and 20% for classification.

5. Conclusion:

For a variety of reasons, individuals or organizations establish fake profiles on social media platforms. The findings relate to determining if an account is real or false utilizing manufactured characteristics and trained machine learning models like CNN and Naive Bias. Predictions show a 95% accuracy rate for the neural network method. In the future, it is hoped that additional features will make the detection and identification process more precise when employing NLP approaches. As Facebook rolls out new features, it will be simple to spot phony accounts.

6. Future Work

The idea is of attaching an Aadhar card number when signingup for an account so that we can restrict the creation of a single account and there is no chance of fake profiles at any moment.

7. References:

[1] Nazir, Atif, Saqib Raza, Chen-Nee Chuah, Burkhard Schipper, and C. A. Davis. "Ghostbusting Facebook: Detecting and Characterizing Phantom Profiles in Online Social Gaming Applications." In *WOSN*. 2010.

[2] Adikari, Shalinda, and Kaushik Dutta. "Identifying Fake Profiles in LinkedIn." In *PACIS*, p. 278. 2014.

[3] Chu, Zi, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. "Who is tweeting on Twitter: human, bot, or cyborg?." In Proceedings of the 26th annual computer security applications conference, pp. 21-30. ACM, 2010.

[4] Stringhini, Gianluca, Gang Wang, Manuel Egele, Christopher Kruegel, Giovanni Vigna, Haitao Zheng, and Ben Y. Zhao. "Follow the green: growth and dynamics in Twitter follower market." In Proceedings of the 2013 conference on Internet measurement conference, pp. 163-176. ACM, 2013.

[5] Thomas, Kurt, Damon McCoy, Chris Grier, Alek Kolcz, and Vern Paxson. "Trafficking Fraudulent Accounts: The Role of the Underground Market in Twitter Spam and Abuse." *In* Presented as part of the 22nd {USENIX} Security Symposium ({USENIX} Security 13), pp. 195-210. 2013.

[6] Farooqi, Gohar Irfan, Emiliano De Cristofaro, Arik Friedman, Guillaume Jourjon, Mohamed Ali Kaafar, M. Zubair Shafiq, and Fareed Zaffar. "Characterizing Seller-Driven Black-Hat Marketplaces." arXiv preprint arXiv: 1505.01637 (2015).

[7] Viswanath, Bimal, M. Ahmad Bashir, Mark Crovella, Saikat Guha, Krishna P. Gummadi, Balachander Krishnamurthy, and Alan Mislove. "Towards detecting anomalous user behavior in

online social networks.” In 23rd {USENIX} Security Symposium({USENIX} Security 14), pp. 223-238. 2014