Identification of Fake vs. Real Identities on Social Media using Random Forest Algorithm

Priyanka Shahane
Department of Computer Engineering
P. E. S. Modern College of Engineering
priyankashahane04@gmail.com

Deipali Gore
Department of Computer Engineering
P. E. S. Modern College of Engineering
deipali.gore@moderncoe.edu.in

***Abstract*—Identity deception on various social media platforms has become growing problem with tremendous increase in number of accounts on these platforms. These fake identities are used by offenders for various malicious purposes, so it has become necessity of time to identify them. The fake identities can be categorized into two two main types i.e. fake accounts created by bots and fake accounts created by humans. This system removes accounts by bots from the corpus during preprocessing and performs classification of accounts by humans into two two categories i.e. Fake vs. Real using Random Forest algorithm based on different parameters as very little research has been done till date on fake identities by humans. Here, the dataset we test for is that of Twitter.**

***Index Terms*—Identity Deception, Social Media, Cyber Crimes, Machine learning, Random Forest.**

I. Introduction

Social media platforms such as Twitter are one of the most crucial means of communication and information dissemination over internet. Much can be learned about people's behavior by analyzing their profiles over social media.This helps offenders to create fake identities in order to commit various cyber crimes such as skewing perceptions, manipulation of credit worthiness of accounts, terrorist propaganda, cyber bullying, fraud, identity impersonation, dissemination of pornography, misdirecting people to some malicious website, spreading malwares and so on.

These fake identities can be created by bots or humans. The fake identities by bots generally target large group of people at a time. Whereas, fake identities by humans generally target specific individual or limited number of people. This system represents an approach to detect fake identities created by humans on Twitter.

II. Review of Literature

The problem of identification of fake identities can be solved by different classification techniques such as Support Vector Machine (SVM), Logistic Regression (LR), Multi Layer Perceptron (MLP), Random Forest (RF), K Nearest Neighbor (KNN), Adaboost, Gradient boosting, Neural Networks and so on.Here are some examples:

Estee et. al. [1] trained the classifier by applying previously used features for bot detection in order to identify fake accounts created by human on Twitter. The training is based on supervised learning. They have tested for three different classifiers i.e. Support Vector Machine (SVM) with linear kernel, Random Forest (RF) and Adaboost. For SVM, the svmLinear library in R is used. Here the boundary based on feature vectors is created for classification. For RF model, the RF library in R is used. RF model creates variations of trees and mode of class outcome is used to predict identity deception. For boosting model, the Adaboost function in R is used. Adaboost is used along with decision trees where each feature is assigned different weight to predict outcome. These weights are iteratively adjusted and output is evaluated for effectiveness of identity deception prediction at each iteration. This process is repeated until best result is obtained. Among these three classifiers RF gave the best result.

Sen et. al. [2] performed supervised learning based on features obtained from FakeLike data and RandLike data. They have experimented with different classification algorithms such as Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM) with RBF kernel, AdaBoost with Random Forest as base initiator, XGBoost and simple feed forward neural network i.e. Multi-Layer Perceptron (MLP) in order to detect the fake likes on instagram. For MLP they have used 2 hidden layers with 200 neurons each. Both layers use sigmoid activation function and output layer has a dropout of 0.2 in order to prevent over fitting. Here, MLP gave the best result compared to other methods.

Viswanath et. al. [3] uses unsupervised machine learning approach for training. The dataset used is from Facebook. They use K-Nearest Neighbors technique for this classification. In KNN data is classified based on majority vote of its neighbors, with test data being assigned to the class most common among its k nearest neighbors where k is a positive integer typically small in value. The classification is done into four classes i.e. Black market, Compromised, Colluding, Unclassified.

Sedhai et. al. [4] trained three different classifiers i.e. Naive Bayes (NB), Logistic Regression (LR) and Random Forest (RF) using semi supervised Learning. These three classifiers