Social Networks Fake Profiles Detection Using Machine Learning Algorithms

Chapter · February 2020

DOI: 10.1007/978-3-030-37629-1_3

3 authors:

Yasyn Elyusufi
Abdelmalek Essaâdi University
13 PUBLICATIONS **36** CITATIONS
SEE PROFILE

Zakaria Elyusufi
Abdelmalek Essaâdi University
4 PUBLICATIONS **4** CITATIONS
SEE PROFILE

Aït Kbir M'hamed
Abdelmalek Essaâdi University
57 PUBLICATIONS **112** CITATIONS
SEE PROFILE

Social Networks Fake Profiles Detection Using Machine Learning Algorithms

Yasyn Elyusufi(✉), Zakaria Elyusufi, and M'hamed Ait Kbir

LIST Laboratory, Faculty of Sciences and Technologies, Tangier, Morocco
elyusufiyasyn@gmail.com

Abstract. Fake profiles play an important role in advanced persisted threats and are also involved in other malicious activities. The present paper focuses on identifying fake profiles in social media. The approaches to identifying fake profiles in social media can be classified into the approaches aimed on analysing profiles data and individual accounts. Social networks fake profile creation is considered to cause more harm than any other form of cyber crime. This crime has to be detected even before the user is notified about the fake profile creation. Many algorithms and methods have been proposed for the detection of fake profiles in the literature. This paper sheds light on the role of fake identities in advanced persistent threats and covers the mentioned approaches of detecting fake social media profiles. In order to make a relevant prediction of fake or genuine profiles, we will assess the impact of three supervised machine learning algorithms: Random Forest (RF), Decision Tree (DT-J48), and Naïve Bayes (NB).

Keywords: User profiling · Fake profile detection · Machine learning

1 Introduction

Social media is growing incredibly fast these days. This is very important for marketing companies and celebrities who try to promote themselves by growing their base of followers and fans. The social networks are making our social lives better but there are a lot of issues which need to be addressed. The issues related to social networking like privacy, online bullying, misuse, and trolling etc. are most of the times used by fake profiles on social networking sites [1]. However, fake profiles, created seemingly on behalf of organizations or people, can damage their reputations and decrease their numbers of likes and followers. On the other hand fake profile creation is considered to cause more harm than any other form of cyber crime. This crime has to be detected even before the user is notified about the fake profile creation. In this very context figures this article, which is part of a series of research conducted by our team within the user profiling subject and profiles classification in social networks. Facebook is one of most famous online social networks. With Facebook, users can create user profile, add other users as friends, exchange messages, post status updates, photos, and share videos etc. Facebook website is becoming popular day by day and more and more people are creating user profiles on this site. Fake profiles are the profiles which are not genuine i.e. they are the profiles of persons with false credentials. The fake facebook

M. Ben Ahmed et al. (Eds.): SCA 2019, LNITI, pp. 30–40, 2020.
https://doi.org/10.1007/978-3-030-37629-1_3

profiles generally are indulged in malicious and undesirable activities, causing problems to the social network users. People create fake profiles for social engineering, online impersonation to defame a person, advertising and campaigning for an individual or a group of individuals [2]. Our research aims at detecting fake profiles at online social media websites using Machine learning algorithms. In order to address this issue, we have chosen to use a Facebook dataset with two thousand eight hundred and sixteen users (instances). The goal of the current research is to detect fake identities among a sample of Facebook users. This paper consists of three sections. In the first section, Machine learning algorithms that we chose to use to address our research issues are presented. Secondly, we present our architecture. In the third section, evaluation model guided by Machine learning algorithms will be advanced in order to identify fake profiles. The conclusion comes in the last section.

2 State of the Art

Fake profiles in social media are often used to establish trust and deliver malware or a link to it. Such fake profiles are also used in other types of malicious activities. To solve these problems, a significant body of research to date has focused on fake profiles detection in social media. Generally, following the taxonomy presented by Song et al. [3]. The approaches to identifying fake social media profiles can be classified into the approaches aimed analyzing individual accounts (profile-based techniques as well as graph-based methods), and the approaches capturing the coordinated activities spanning in a large sample of accounts. For instance, Nazir et al. describes in [4] detecting and characterizing phantom profiles in online social gaming applications. The work analyses a Facebook application, the online game "Fighters club", known to provide incentives and gaming advantage to those users who invite their peers into the game. The authors state that by providing such incentives the game motivates its players to create fake profiles, by introducing those fake profiles into game, the user would increase incentive value for himself. At first, the authors extract thirteen features for each user, and then perform classification using support vector machines (SVMs). This work concludes that these methods do not suggest any obvious discriminants between real and fake users. On the other hand Adikari, S. and Dutta, K., 2014 works on Known fake LinkedIn profiles, posted on special websites. The detection method uses the number of languages spoken, education, skills, recommendations, interests, awards, etc. These features are used to train neural networks, SVMs, and principal component analysis. The accuracy found in this work was 84% True Positive (TP), and 2.44% False Negative (FN). In the same context Stringhini et al. 2010 works on spam accounts registered by honeypots: 173 spam accounts in Facebook and 361 in Twitter [6]. In this research Random forest was constructed based on the following features: ratio of accepted friend requests, URL ratio, message similarity, regularity in the choice of friends, messages sent, and number of friends. The accuracy found in this work was 2% FP, 1% FN for (Facebook); and 2.5% FP, 3.0% FN for (Twitter). Also Yang et al. treat spam Twitter accounts defined as the accounts containing malicious URLs, about 2060 spam accounts are identified [7]. In This latter research authors uses graph based features (local clustering coefficient, betweenness centrality, and bi-directional links

ratio, neighbor-based features, also timing-based features were used to construct different classifiers. The accuracy found was 86% TP, 0,5% FP. The previous approaches assume that the machine learning techniques are too challenging because the attackers create patterns that cannot be trained by machines. But recent works have applied many standard machine learning algorithms, such as ensemble of classifiers, Random Forests, SVM, Decision Trees and Artificial Neural Networks. Several machine learning algorithms are used for the classification of profiles based on their features. The survey on efficient machine learning studies introduces several algorithms and discusses their processing ability with respect to prediction accuracy. Many machine learning algorithms are used in order to identify fake profiles in social networks. The different machine learning techniques used by various works are shown in Fig. 1.

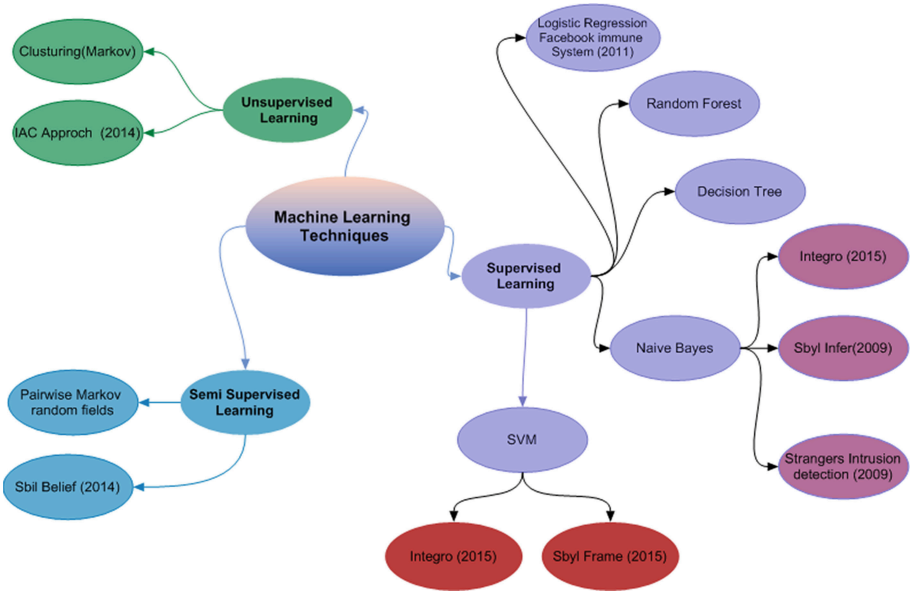

Fig. 1. Machine learning techniques used in recent works.

In several works, we have discussed profiling techniques and their applications [8–10]. In this work, we identify the minimal set of user profile data that are necessary for identifying fake profiles in facebook, about two thousand eight hundred and sixteen users including fake and genuine profiles are used. In order to identify the most efficient machine learning algorithms, we used three supervised machine learning Algorithms: Random Forest (RF), Decision Tree (J48) and Naïve Bayes (NB). The application section will be made with Jupter tool (Python 3) in order to choose the most efficient algorithms. Finally the comparison of the accuracy and confusion matrix of each model can better explain the interest of our work.

3 Using Naive Bayes Classifiers

Naive Bayes classifiers are a family of simple probabilistic classifiers used in machine learning. These classifiers are based on applying Bayes theorem with strong (naive) independence assumptions between the features. Naive Bayes is a simple method for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. It is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: all naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable [11]. Naive Bayes classifiers are a popular statistical technique of email filtering. They emerged in the middle of the 90s and were one of the first attempts to tackle spam filtering problem [11]. Naive Bayes typically use bag of words features to identify spam e-mail, an approach commonly used in text classification. Naïve Bayes classifiers work by correlating the use of tokens (typically words, or sometimes other constructions, syntactic or not), with spam and non-spam e-mails and then using Bayes theorem to calculate a probability that an email is or is not a spam message [11]. In our paper we will assess the impact of using Naïve Bayes classifiers in the prediction of fake or genuine profiles in social networks (Facebook Data set).

4 Using Decision Trees Classifiers

A decision tree is a popular classification method that generates tree structure where each node denotes a test on an attribute value and each branch represents an outcome of the test as shown in Fig. 2. The tree leaves represent the classes. This technique is fast unless the training data is very large. It does not make any assumptions about the probability distribution of the attributes value. The process of building the tree is called induction [12].

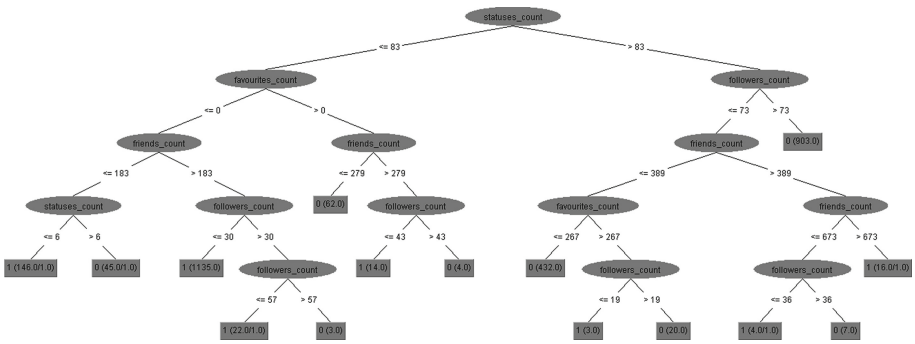

Fig. 2. Example of fake profiles identification using decision tree.

The decision tree algorithm is a top-down greedy algorithm which aims to build a tree that has leaves as homogenous as possible. The major step in the algorithm is to continue dividing leaves that are not homogeneous into leaves that are as homogeneous as possible until no further division is possible as shown in Fig. 2. In our approach we will assess the impact of using Decision Trees classifiers in the prediction of fake or genuine profiles in social networks (Facebook Data set).

5 Using Random Forest Classifiers

Random Forest is one of the most used machine learning algorithms, because its simplicity and the fact that it can be used for both classification and regression tasks. It's a supervised learning algorithm. As it's seen from its name, it's bagged decision tree models that split on a subset of features on each split, it creates a forest and makes it somehow random. The «forest» it builds, is an ensemble of Decision Trees, most of the time trained with the "bagging" method. The general idea of the bagging method is that a combination of learning models increases the overall result. To say it in simple words: Random forest builds multiple decision trees and merges them together to get a more accurate and stable prediction. For better understanding of the RF algorithm is necessary to explain what the main idea behind decision trees is. Depending on the features in each dataset, the decision tree model learns a series of questions to figure out the class labels of the instances. What makes this model successful is that it is non-parametric and it can handle heterogeneous data (ordered or categorical variables, or a mix of both). Furthermore decision trees fundamentally implement feature selection, making them at least to some extent robust to irrelevant or noisy variables and are robust to outliers or errors in labels [13].

To summarize, here is steps that Random Forest algorithm follows:

- Randomly choose n samples from the training set with replacement.
- Grow a decision tree from the n sample. At each node.
- Repeat the steps 1 to 2 k-times.
- Aggregate the prediction by each tree to assign the class label by majority vote.

6 Our Approach

In our research work, a novel approach has been presented for the identification of fake profiles on facebook using supervised machine learning algorithms. The proposed model has applied data preprocessing techniques on datasets before analyzing them. A technique has been applied to identify the non significant attributes in datasets and to do attribute reduction. The proposed model was trained using supervised machine algorithms individually for dataset including fake and genuine users. Ensemble classifier has been used to make the prediction more accurate. As shown in Fig. 3, the process of fake profile detection has three levels, in the first level profile features are extracted and then in the second level: Random Forest (RF), Naïve Bayes (NB) and Decision Tree (DT) are used to determine the fake and genuine profiles. The third level, we calculate and compare the accuracy rates across the results of both techniques.

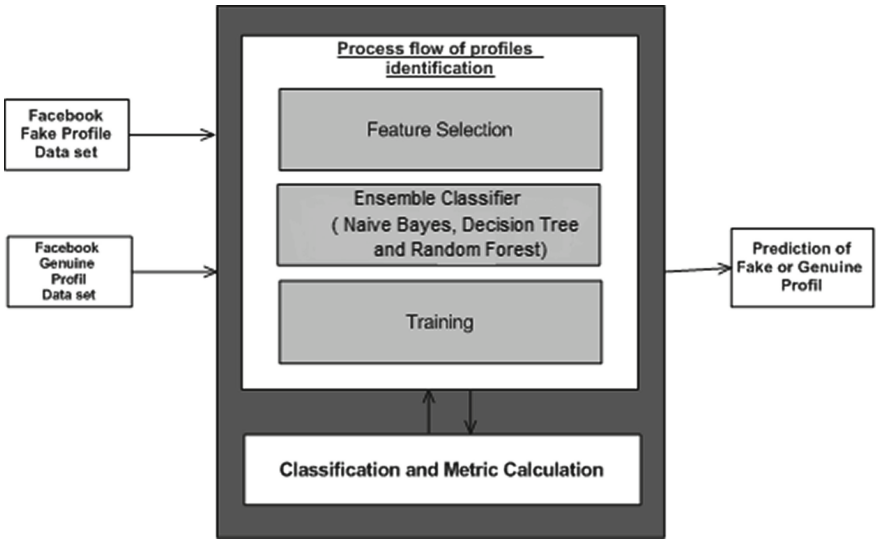

Fig. 3. The process flow in identification of the real or fake profile

6.1 Initial Features

After using the dataset, we proceed for feature selection phase. We noticed that there was much of unneeded features, either have no meaning for our subject or full of NaN values, so to make our models train well, we decided to drop them and to let only those of will affect directly on the results. As shown in Table 1, initially we have 33 profile features which will be used in the Facebook DataSet.

6.2 Features Selection

Feature selection is one of the basic concepts in machine learning which hugely impacts the performance of classification and prediction. In our work, and in order to make our models train well, we decided to use only features which will affect directly the results. The features on the final dataset were: statuses_count, followers_count, friends_count and favorites_count. Below is the meaning of each feature Table 2.

After all this steps, and in considering that the algorithm split the dataset into training and testing sets, and to make the row homogenous, we shuffled them. Finally we use the final dataset to train and evaluate the three machine learning algorithms: Random Forest, Decision Tree (J48), and Naïve Bayes.

Table 1. Table of initial features.

Features	Description
Id	Id of user
name	User name
screen_name	Screen Name
statuses_count	Statuses Count
followers_count	Followers Count
friends_count	Friends count
favourites_count	Favourites Count
listed_count	Listed Count
created_at	Date of account creation_at
url	Acount Url
lang	Language
time_zone	Time zone
location	Geographic Location
default_profile	Default profil
default_profile_image	Default Profil Image Status
geo_enabled	Géo localisation
profile_image_url	Profile Image URL
profile_banner_url	Profile Banner URL
profile_use_background_image	Profile Background Image
profile_background_image_url_https	Profile background Image Url
profile_text_color	Profile Text Color
profile_image_url_https	Profile Image Url Https
profile_sidebar_border_color	Profile Sidebar Border Color
profile_background_tile	Profile Background Title
profile_sidebar_fill_color	Profile Side Bar Fill Color
profile_background_image_url	Profile Background Image URL
profile_background_color	Profile Background Color
profile_link_color	Profile Link Color
utc_offset	Offset Status
Protected	Protection Status
verified	Verification Status
Description	Description of Account
updated	Update Date

Table 2. Table of selected features.

Features	Description
statuses_count	Statuses Count
followers_count	Followers Count
friends_count	Friends Count
favourites_count	Favourites Count

7 Results and Discussion

7.1 Splitting the Dataset

Before the training phase, we have to split the dataset. In this step we started with choosing the input and output values. The input contains the independent variables and the output contains the dependent variable (Fake or Genuine value) that takes the value of 0 and 1, then we split the dataset into the training set and test set. In our work the training test set is defined with 80%, while the test set is defined with 20%.

7.2 Evaluation Metrics

We present the metrics used to evaluate the results in order to select the best supervised machine learning algorithm. We first show the model accuracy and then use the confusion matrix shown in Formula 1. This matrix is used to visualize the performance of the different algorithms using the following metrics. This metric indicates the fraction of returned cases that are valid fake profiles.

The accuracy function (Formula 1)

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{1}$$

Table 3 represents the confusion matrix used to evaluate the efficiency of proposed methods.

Table 3. The confusion matrix

	Predicted class	
	Predicted	No predicted
Examples of this class	TP	FN
Example not belonging to this class	FP	TN

7.3 Experimental Results

In this research, we have compared the results of three machine learning algorithms (Random Forest, Decision Tree (J48) and Naïve Bayes) to determine the most appropriate approach to differentiate the legitimate profiles from fake profiles in Facebook dataset. Tables 4, 5 and 6 summarize the final confusion matrix values akin to each algorithm by calculating the correctly classified instances and incorrectly classified instances. In addition, it shows the accuracy calculation for each algorithm.

Based on the result tests of the tree algorithms, as shown in Fig. 4, it is obvious that Random Forest algorithm is better than Decision Tree (J48) and Naïve Bayes algorithms. It ranked first with an accuracy score of **99,64%**. Where Decision Tree (J48) algorithm give only 559 correctly classified instances with an accuracy of **99,28%**. Naïve Bayes is the last with an accuracy of **78,33%**.

Table 4. Classification of profiles on Facebook using (**Decision Tree**)

Decesion Tree (J48)	
Confusion Matrix	
260	**1**
3	**299**

Correctly Classified Instances **559** 99.2895%
Incorrectly Classified Instances **4** 0.7105%

Table 5. Classification of profiles on Facebook dataset using (**Naïve Bayes**)

Naïve Bayes	
Confusion Matrix	
260	**1**
121	**181**

Correctly Classified Instances **441** 78.3304%
Incorrectly Classified Instances **122** 21.6696%

Table 6. Classification of profiles on Facebook dataset using (**Random Forest**)

Random Forest	
Confusion Matrix	
260	**1**
1	**301**

Correctly Classified Instances **561** 99.6448%
Incorrectly Classified Instances **2** 0.3552%

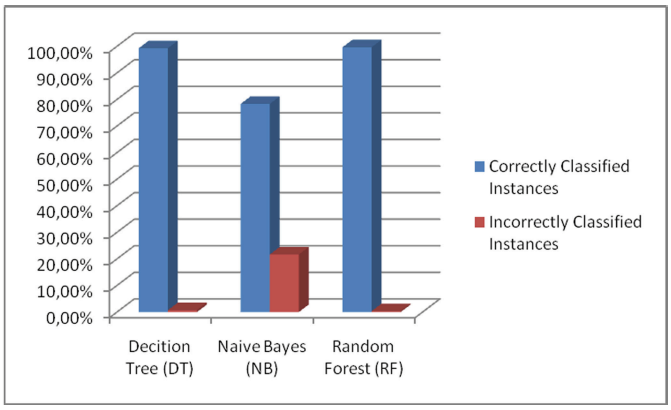

Fig. 4. Accuracy comparison between DT, NB and RF algorithms

8 Conclusion and Future Work

In this paper, we proposed an approach to identify the fake profile in social network using limited profile data, about 2816 users. As we concluded in our paper, we demonstrate that with limited profile data our approach can identify the fake profile with **99.64%** correctly classified instances and only **0.35%** incorrectly classified instances, which is comparable to the results obtained by other existing approaches based on the larger data set and more profile information. Our research can be a motivation to work on limited social network information and find solutions to make better decision through authentic data. Additionally, we can attempt similar approaches in other domains to find successful solutions to the problem where the least amount of information is available. In future work we expect to run our model using more sophisticated concepts such as ontology engineering, in order to semantically analyze user posts, and comportments. This later concept can improve the quality of prediction of fake or genuine profiles.

References

1. Boshmaf, Y., Muslukhov, I., Beznosov, K., Ripeanu, M.: The socialbot network: when bots socialize for fame and money. In: Proceedings of the 27th Annual Computer Security Applications Conference, pp. 93–102. ACM (2011)
2. Romanov, A., Semenov, A., Veijalainen, J.: Revealing fake profiles in social networks by longitudinal data analysis. In: 13th International Conference on Web Information Systems and Technologies, January 2017
3. Song, J., Lee, S., Kim, J.: CrowdTarget: target-based detection of crowdturfing in online social networks. In: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, CCS 2015, pp. 793–804. ACM, New York (2015)
4. Nazir, A., Raza, S., Chuah, C.-N., Schipper, B.: Ghostbusting Facebook: detecting and characterizing phantom profiles in online social gaming applications. In: Proceedings of the 3rd Conference on Online Social Networks, WOSN 2010. USENIX Association, Berkeley, CA, USA, p. 1 (2010)
5. Adikari, S., Dutta, K.: Identifying fake profiles in Linkedin. Presented at the Pacific Asia Conference on Information Systems PACIS 2014 Proceedings (2014)
6. Stringhini, G., Kruegel, C., Vigna, G.: Detecting spammers on social networks. In: Proceedings of the 26th Annual Computer Security Applications Conference, ACSAC 2010, pp. 1–9 (2010)
7. Yang, C., Harkreader, R.C., Gu, G.: Die free or live hard? Empirical evaluation and new design for fighting evolving Twitter spammers. In: Proceedings of the 14th International Conference on Recent Advances in Intrusion Detection, RAID 2011, pp. 318–337. Springer, Heidelberg (2011)
8. Elyusufi, Y., Seghiouer, H., Alimam, M.A.: Building profiles based on ontology for recommendation custom interfaces. In: International Conference on Multimedia Computing and Systems (ICMCS) Anonymous IEEE, pp. 558–562 (2014)
9. Elyusufi, Y., Alimam, M.A, Seghiouer, H.: Recommendation of personalized RSS feeds based on ontology approach and multi-agent system in web 2.0. J. Theor. Appl. Inf. Technol. **70**(2), 324–332 (2014)

10. Elyusufi, Z., Elyusufi, Y., Ait Kbir, M.: Customer profiling using CEP architecture in a Big Data context. In: SCA 2018 Proceedings of the 3rd International Conference on Smart City Applications Article No. 64, Tetouan, Morocco, 10–11 October 2018. ISBN: 978-1-4503-6562-8
11. Granik, M., Mesyura, V.: Fake news detection using naive Bayes classifier. In: Conference: IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), May 2017
12. Ameena, A., Reeba, R.: Survey on different classification techniques for detection of fake profiles in social networks. Int. J. Sci. Technol. Manage. **04**(01), (2015)
13. Beatriche, G.: Detection of fake profiles in Online Social Networks (OSNs), Master's degree in Applied Telecommunications and Engineering Management (MASTEAM), (2018)