https://doi.org/10.31449/inf.v43i1.2319 Informatica **43** (2019) 77–86 **77**

An Empirical Study for Detecting Fake Facebook Profiles Using Supervised Mining Techniques

Mohammed Basil Albayati and Ahmad Mousa Altamimi
Applied Science Private University, Amman, Jordan
E-mail: mohammed.sabri@asu.edu.jo, a_altamimi@asu.edu.jo

Keywords: data mining, online social networks, facebook, fake profiles, data science

Received: April 29, 2018

Our social life and the way of people communicate are greatly affected by the social media technologies. The variety of stand-alone and built-in social media services such as Facebook, Twitter, LinkedIn, and alike facilitate users to create highly interactive platforms. However, these overwhelming technologies made us sank in an enormous amount of information. Recently, Facebook exposed data on 50 million Facebook unaware users for analytical purposes. Fake profiles are also used by Scammers to infiltrate networks of friends to wreak all sorts of havoc as stealing valuable information, financial fraud, or entering other user's social graph. In this paper, we turn our focus to Facebook fake profiles, and proposed a smart system (FBChecker) that enables users to check if any Facebook profile is fake. To achieve that, FBChecker utilizes the data mining approach to analyze and classify a set of behavioral and informational attributes provided in the personal profiles. Specifically, we empirically examine these attributes using four supervised data mining algorithms (e.g., k-NN, decision tree, SVM, and naïve Bayes) to determine how successfully we can recognize the fake profiles. To demonstrate the validity of our conceptual work, the selected classifiers have been implemented using RapidMiner data science platform with a dataset of 200 profiles collected from the authors' profile and a honeypot page. Two experiments are developed; in the first one, the k-NN schema is applied as an estimator model for imputation the missing data with substituted values, whereas in the second experiment a filtering operator is applied to exclude the profiles with missing values. Results showed high accuracy rate with the all classifiers, however, the SVM outperforms other classifiers with an accuracy rate of 98.0% followed by Naïve Bayes.

Povzetek: Opisana je metoda iskanja lažnih profilov na Facebooku s pomočjo strojnega učenja.

1 Introduction

In recent years, social media technologies (e.g., Facebook, Twitter, LinkedIn, etc.) have become a vital part of our life [1]. They are designed and maintained by social media organizations presenting a portal for facilitating communication, interaction, sharing information, and entertainment via virtual communities and networks. Users typically utilize these services by creating their own profiles and then connecting them with others' profiles through various technologies that offer social media functionality [2]. By using such services, users can create digital contents, such as text posts, comments, digital photos, videos, or data generated through all online interactions [3].

Social media sites have presented a various service included with high levels of quality, consistency, and The Introduction should provide a clear background, a clear statement of the problem, the relevant literature on the subject, the proposed approach or solution, and the new value of research which it is innovation and availability. This results in huge registered users [4]. Some of the most popular social media websites are Facebook (and its associated Facebook Messenger), Gab, Google+, MySpace, Instagram, LinkedIn, and others. Statistics and surveys for example the one that conducted by the American Academy of Pediatrics exhibit that about 84% of adolescents in America registered on Facebook social online site [5], also showed that the average users spend more than two hours on social network and even more on social online sites such as Facebook, Twitter and else more than any other sites or platform [6]. The benefit of engaging and participating in social online sites have gone beyond simply social online activities, sharing information, or communication but to building careers, making business opportunity, financial income, etc. [7].

Historically, according to Mark Zuckerberg, a co-founder of Facebook which is the largest social network site, Facebook have more than 175 million active users registered in 2009 after just five years' time frame. Nowadays, Facebook has 1.94 billion users on the last official announcement on March 31st, 2017, according to Facebook newsroom administration [8], which exceeds the population of some big countries. With this rapid revolution in this technology, number of negative consequences and risks are raised such as security risks, privacy violation, cloning, hacking, spamming, and others [3]. For example, Spam on social media repeatedly posts the same thing over and over or causes a sudden spike in messaging activity [9]. Fake profiles on the other hand, allow scammers to infiltrate networks of friends to wreak all sorts of havoc such as: stealing valuable information, financial fraud, or entering other user's social graph [10].

It is important to mention here that according to the Facebook's Statement of Rights and Responsibilities; users should provide their real and legit information once they created their profiles. Facebook urges its users to be committed to these policies and terms in order to have an experience in an environment of safety, security, and privacy [11]. In this work, we focused on the problem of detecting fake profiles in Facebook and presenting a smart detection system (FBChecker) to handle this problem based on the prediction and classification techniques of data mining.

Our work is motivated by works presented in [12-15], where researchers employed data mining approach for extracting hidden knowledge within social media. For example, authors of [12] utilized data mining tools for accurately capturing the behavior of intrusions and normal activates in an anomaly detections approach. One can consider also the Web mining that applies data mining tools onto web resources to further developments in World Wide Web mining [15].

In our model, supervised mining techniques are applied to classify Facebook's profiles into fake and real profiles based on a set of behavioral and informational attributes. These attributes are provided in their personal profiles and used to identify the reality of user's identity such as: person's legal name, location, workplace, age, education, and others. The required data set for the training and testing purposes in our work has been collected from the authors' personal profiles considered as a source of real profiles, and from a created honeypot page, which is fake Facebook's pages used for the purposes of data harvesting [10] to attract and collect these profiles. As a collecting tool we wrote our own script to develop a special CRAWLER for gathering the required data set. To underscore the practical viability of our approach, the selected classifiers (e.g., SVM, Naïve Bayes, k-NN, and Decision Tree) have been implemented using RapidMiner data science platform for the mining tasks. These classifiers were evaluated using 10-folds cross validation method and conducted on the collected data set. It is important to mention here that 33 records have some missing values of their attributes.

To solve this problem, two empirical studies were developed, in the first one, the k-NN schema was used as an estimator model for imputation the missing data with substituted values. Results showed that the classifiers (SVM, Naïve Bayes, k-NN, and Decision Tree) achieved (0.9850, 0.9700, 0.8400, 0.9650), respectively. In the second experiment, a filtering operator is applied to exclude the profiles with missing values. Here, the classifiers showed (0.9880, 0.9641, 0.8443, 0.9461), which are relatively equal to the results of the first experiment. The numbers and the ROC graph (Receiver Operating Characteristics) which is a graphical plot utilized to assess the classifiers performance ability showed that, in the both experiments SVM classifier achieved the highest accuracy rates while, the k-NN performance showed the lowest accuracy detection rate among the classifiers. These experiments are discussed in more details in Section 5.

The remaining of this paper is structured as following: Section 2 reviews the related works to the proposed approach and to the fake profiles in online social networks, specifically the Facebook. Section 3 describes the background material of the research work along with brief description of the employed supervised algorithms. Section 4 explains the research methodology while the proposed system along with its main components presents in section 5. Section 6 discusses the implementation of the FBChecker system, the evaluation and experimental results are given in Section 7. Finally, section 8 offers the conclusion and the possible future work. introduction.

2 Related work

Many studies and works have been conducted focusing on the phenomena of fake profiles on online social networks, each researcher tried to came up with new way to detect and handle this problem. Studies in this field differ according to how they look at the problem from their own perspectives. Each of which is raised for solving a certain problem and faces certain challenges and difficulties. In this regard, many approaches presented in the literature for handling fake profiles.

One can consider for example, the work in [16]. Here, the authors present a machine learning pipeline framework consists of three components for detecting clusters of duplicate accounts (cluster level detection) rather than making a prediction for an individual account. Here, the pipeline uses simple information that is provided at the registration time, so the profile is detected before it is activated. Moreover, the classifier determines whether the clusters of accounts were created by the same actor, showing a strong evaluation on sample grouping based on the simple text information like name, email, company, etc. and the IP address. Practically. The system captures more than 250,000 fake accounts in practical use. In contrast, [17] proposed a behavioral approach for detecting fake accounts on Facebook. It is designed using information regarding user profile's activities and interactions with other users. Authors characterized these activities through an extensive set of 17 features like (likes, comments, shares, tag, and apps usage on Facebook). To ground their idea, these features are applied on a total of 12 supervised machine learning techniques. The system's performance showed an accuracy of 79%, which may not be impressive results, but the author considered it as a first step or baseline work for further improvements.

Detecting Spam profiles, which is one of the fake profiles types [10], has also considered in the literature. Authors of [18] proposed a statistical analyzing model with 14 generic features from Facebook and Twitter data set regarding 4 basic kinds of social interactions including (profile interaction features, posts/ tweets, URLs and tags & mentions). The model identifies spam profiles on Facebook and Twitter based on information collected manually through scanning these networks for both normal and spam profiles using three different supervised classification algorithms (naive Bayes, Jrip, and J48). Then two different experiments were performed: firstly, examining the role of the whole feature set and calculate

the accuracy of the proposed system. And secondly, removing each one of the features and analyzing the results of the system to discover the impact of each features and find out which one can play the key role in the classification model.

Detecting spam profiles is also presented in the literature as in [19], Presented Social Privacy Protector software (SSP) for detecting fake profiles on Facebook, the SSP consist of three protection layers: The software first identifies a user's friends who might pose a threat and then restricts this "friend's" exposure to the user's personal information (The Friends Analyzer Facebook Application). The second layer is an expansion of Facebook's basic privacy settings based on different types of social network usage profiles (The Social Privacy Protector Firefox Add-on). The third layer alerts users about the number of installed applications on their Facebook profile, which have access to their private information (The HTTP Server). The software present convenient method for restrict the users that may be suspected as fake profiles without removing it from the user's friends list.

The Friends Analyzer Application on the Facebook scans the user's friends list and returns a credibility score. Each friend analyzed by machine learning algorithms which takes into account the strength of the connection between the user and his friends. The strength of each connection is based on a set of fifteen connection features depends on three types of the collected dataset, such as the number of common friends between the user and his friend and the number of pictures and videos the user and his friend were tagged in together. Applying eight supervised algorithms such as (Naive-Bayes, Bagging, Random-Forest, J48, and others). The Social Privacy Protector add-on in the Firefox browser help improve the user privacy with simple steps. Finally, The HTTP server responsible for connecting the SPP Firefox Add-on to the SPP Facebook application. Authors of [20] proposed a framework for detecting spammers/ fake profiles on online social network using Facebook as test case in a machine learning approach by exploiting a behavioral and community-based features (attributes) that include the structure of the nodes and some topological features (attributes) in the network.

The framework implemented using WEKA tool as mining environment, using ten discriminative topological attributes (Total out-degree, Total in/out ratio, Total reciprocity, Core node, Community memberships, Foreign out-degree, Foreign in/out ratio, Foreign out-link probability, Foreign reciprocity, and Foreign out-link grouping) regarding the social interactive of the profiles like number of posts, number of sent/ received messages…etc. Four experiments are conducted using two datasets: Facebook dataset and Enron network (Email messages dataset). Four supervised classifiers are employed in this work (Naïve Bayes, J48, k-NN, and Decision Tree).

Ultimately, authors of [21] proposed a machine learning approach for detecting spam bots in Twitter online social network through exploiting two main spam features, which are: The graph-based features including the number of friends, number of followers and the follower's ratio (the ratio of the number of peoples following you to the number of peoples you follow). And the content-based approach which is the number of duplicated tweets, number of HTTP links, and the number of replays/mentions. Regarding the detection process, the approach applied different classification methods such as decision tree, neural network, support vector machines, and k-nearest neighbors to identify spam bots on Twitter. The evaluating results showed that the Bayesian classifier has a better overall performance.

3 Background material

Data Mining basically is the process of extracting Knowledge from a huge amount of data, by looking for a pattern, identified, validated, make a prediction and summarize it into useful information. Data mining process goes through a sequence of procedures, applying set techniques, combining several of discipline and fields like statistics, machine learning, database, algorithms visualization methods, pattern recognition and other disciplines [22].

3.1 Machine learning techniques in data mining

Machine learning is a branch of computer science, which deals with algorithms that have the ability to learn and adapt to make a decision [22]. One of the most common tasks that data mining offers is Classification & Predication in which they fall into the machine learning techniques.

In machine learning, there are two main techniques known as Supervised Learning, where the training dataset has a class label, and Unsupervised Learning, where the data are grouped together based on observable behavior or features. In other words, in supervised, a labeled set of training data is used to estimate or map the input data to the desired output. In contrast, under the unsupervised methods, no labeled examples are provided and there is no notion of the output during the process, instead the data with similar attributes or similar behavior are grouped together (clustered) [22, 23].

In this work only the supervised techniques have been employed as mentioned, particularly four supervised techniques that are: SVM, Decision Tree, k-NN, and Naïve Bayes. A brief description about these classifiers will be presented in the next subsection.

3.1.1 Supervised learning

Supervised machine learning is a heuristic process of mapping inputs to specific output, estimating unknowns based on labeling samples. The objective of supervised learning technique is to build a model with distinguished features and predefining labels with a known class, then using this model to classify or predict a new data with unknown class.

The process of classification and prediction in supervised machine learning involves two major steps:

- The learning step: the model constructed, analyzed and trained with known label dataset called "training set", then the classification and prediction rules are generated.
- The classification and prediction step: the model (classifier) used for classifying or predicting a given data based on the gained experience from the training set. The model's results are evaluated through testing and evaluating process to estimate its accuracy.

The test metrics use to assess how good or how accurate the classifier was. If the it reaches a level of accuracy that is acceptable based on specific standards, then the model can be deployed on new unknown labeled data, otherwise it will be modified [24].

In our work, we choose to employ the most common supervised machine learning algorithms that are:

1- Decision Tree is a predictive model takes a tree structure that generates the classification rule by breaking down the dataset into smaller and smaller subset until the decision node (class label) is met. Each node in the tree represents an attribute of the training set, however, leaf nodes hold the class label (final outcome), while the root node represents the attribute with highest information gain that determines the tree branches in which each branch represents one of the outcomes of the model.

2- k-NN is one of the simplest algorithms perform similarity functions, which store all cases with a known label and classifies new data based on the similarity measures or distance function. k-NN classify new data by using k value to find the nearest case in the data set, for example if (k =1) then simply assign the new case to the class of its first nearest neighbor, if the (k = 3) then k-NN calculate the distance of the nearest three cases and apply majority vote on the class of these cases to decide the class of the new data. The distance measures for finding the nearest neighbor for the numerical data is calculated by the Euclidian distance function and for the categorical data hamming distance measure.

3- Support vector machine algorithm is a classification technique designed to define a hyperplane that classify the training data vectors into classes, the goal or the best choice is to find a hyperplane with widest margin to separate the data classes. The support vector are the data points which are closest to the hyperplane.

4- Finally, Naïve Bayes or simple Bayesian classifier is considered also in the mining process as a supervised classification technique as it is simple and prove its effectiveness, Naïve Bayes is probabilistic algorithm depends on applying Bayesian theorem with naïve assumption that the occurrence of one of the attributes\ predictors are independent of the occurrence of other attribute and regardless of any correlation between these attributes in the classification process. Bayes rules adopted in this algorithm stated a conditional probability of certain event based on previous knowledge about that event [22, 23].

4 Research methodology

In this work, we developed a smart system (FBChecker) that enables users to detect the fake Facebook profiles by utilizing the supervised data mining techniques. To do so, the system firstly, collects the data of a set of behavioral and informational attributes derived from the user's friends' profiles (listed in table 1). To achieve this, a special purpose module (called CRAWLER) is developed to collect the required attributes from the user's friends list. CRAWLER is running at the user level for collecting this data. Secondly, the collected data is validated to increase the accuracy of the detection process. Specifically, the problem of missing values has been solved using two methods, the k-NN scheme and a special operator to exclude them. Ultimately, a set of supervised mining algorithms are implemented using the RapidMiner data science platform to detect the fake profiles. The main objective of using the supervised machine learning techniques is to build a model with distinguished features and predefining labels with a known class, then using this model to classify or predict a new data with unknown labels. This process involves two major steps. Firstly, the learning step that includes constructing, analyzing and training with known label data set (training set), then the classification and prediction rules are generated. Secondly, the classification and prediction step that the learner model (classifier) gives data based on the gained experience from the training set.

5 The FBChecker smart system

Figure 1 illustrates the main components of proposed FBChecker System. In this section, we discuss the steps that followed carefully to build up the system along with its main components.

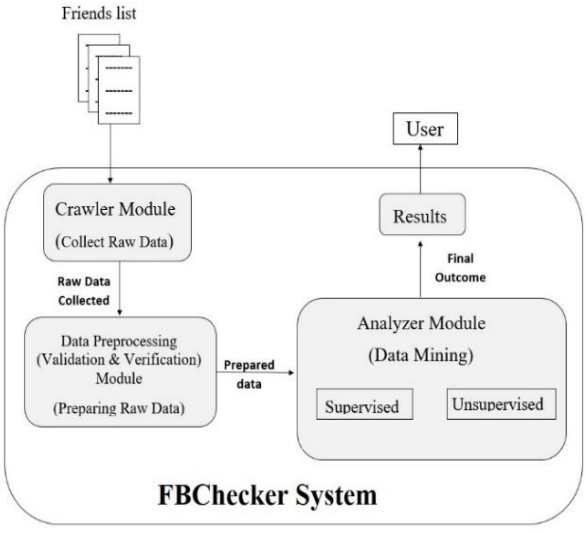

Fig. 1: FBChecker System Components.

1- Collecting the required data: first thing needed to be considered in building a machine learning system is collecting the required data for the training and testing purposes. In this regard, a special purpose module (CRAWLER) was developed and written in

Java Script for collecting the required attributes form the user's friends list. The considered attributes are listed in table 1 along with their description and their using justifications.

2- Preparing the data: the raw data need to be prepared and validated to increase the data quality and to be eligible for applying the mining techniques. Here, the preparation process is done as following:

- *Missing Values*: we note that some profiles have missing values due to privacy issues or the users did not fill these attributes with required information. To solve this problem, two methods have been applied, the k-NN schema is applied as an estimator model for imputation the missing data with substituted values, and a filtering operator is applied to exclude the profiles with missing values.
- *Profile Picture*: it is recognized by the user himself as a real picture or not.
- *Education*: it is validated according to a multi-lingual database of size ~10,000 records of colleges and universities existed around the world.
- *About "Bio." Section*: making a textual condition, if the number of words in this section greater or equal 5 return true/real otherwise false/fake value.
- *Other attributes*: such as *Relationship Status*, *Life Events*, *Living Place*, and *Check Ins* do not need to be validated as Facebook evaluates the attributes' values. So, the CRAWLER module retrieves them as is.

3- Training and Appling the Supervised data mining algorithms: after the data is prepared and ready for mining, a supervised data mining technique is applied (Analyzer module). The classifiers are trained with known class data that are (Fake, Real) profiles. At this step, the system gaines the experience and the ability to classify and detect the fake profiles. In addition, the classification rules are generated and prepared through applying the supervised algorithms. Finally, the selected supervised data mining algorithms are applied using the prepared collected data for detecting fake profiles.

6 The FBChecker implementation

6.1 Data set description

We note that there is no available standard data set with the required information. Thus, we choose to prepare our own one. The CRAWLER is employed on the author's profile for gathering real profiles and returns 151 profiles friends out of 151. However, 18 profiles were excluded as they were faked, underaged, or duplicated. This ends up with 133 real profiles. Regarding the fake profiles, a honeypot page is created and utilized as a source for collecting fake profiles. The inspecting of the fake profiles was finalized with selecting of 83 fake profiles as some of the collected profiles were not stable with their liking activity in which they drop their likes from our page after few days. As a result, 200 profiles were collected, 117 real and 83 fakes, as summarized in Figure 2.

Attribute	Description	Justification
Profile Picture	Visual identification of the user	Real users use their real pictures more often than fake users
Work place	Workplace or job title's information	Real users more often use their real workplace information than fake users
Education	Attended (school, college, university…etc.) information	Real users mentioned their education information in their Facebook profiles more often than fake users
Living Place	Living place address (city, town, state…etc.) information	Real users more often use their real living place information than fake users
Relationship Status	Social relation status (married, single, engaged, etc.) information	Real users share their real social relation status than fake users
Check In	Information for announcing user location	Real users check into places in their Facebook's profiles more often than fake users
Life Events	Information for the users to tell their stories	Real users share their life events more often than fake users.
Introduction "Bio."	Introduction information about Facebook's users	Real users are more often write something about themselves than fake users
No. of Mutual Friends	Number of the people who are Facebook friends with both users and the target profiles	Real users have more mutual friends with target profile than fake users, hence gives profile more incredibility
No. of Pages Liked	Number of pages liked	Real users usually liked more pages than fake users
No. of Groups Joined	Number of groups joined by the target profile.	Real users usually join groups more than fake users.

Table 1: Attributes used by FBChecker.

6.2 Building the FBChecker system

After collecting the 200 profiles data set, we are ready to generate the classification and prediction rules. In this regard, RapidMiner 8.0.1 platform was utilized as a mining tool, which offer the use of various machine learning algorithms easily and provides a flexible environment designed specifically for data science and

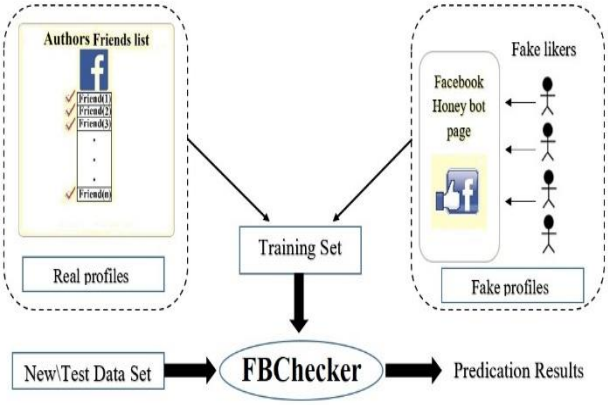

Figure 2: Collecting Training Data Set.

data mining purposes. For the training and testing processes, the (K fold) cross-validation with 10 folds was applied to evaluate the results accuracy as it is considered as one of the most effective methods for evaluating the predictive models with relatively small data set.

7 Evaluation process

To evaluate the FBChecker performance, the selected classifiers were tested with two experiments. In the first experiment, the k-NN schema was utilized to substitute the missing values, while in the second experiment, profiles with missing value were excluded. These experiments are discussed in detail in the following subsections. Finally, metrics for the validation process were calculated and proper justifications were provided.

7.1 Performance metrics

A group of common metrics are applied in the validation process, in this work the following metrics are used: Recall, Precision, Accuracy, F-measure, and specificity [25]. Next, we give a brief description for each one:

1) *Recall* true positive rate (total numbers of true positive divided by the total number of actual positives)
2) *Precision*: Measure the probability that the positive predications is correct (total numbers true positives divided of total number of predicted positives)
3) *Accuracy* Measure the performance of the classification model (total numbers of correct examples divided by total number of the example set)
4) *Specificity* true negative rates (total numbers of true negatives divided by the total number of actual negatives)
5) *F-measure* is an overall measure of a model's accuracy that combines precision and recall.

7.2 The experimental results

Four supervised algorithms were applied on the collected data set based on the following cases:

7.2.1 Estimating the missing values using k-NN schema

In this case, the k-NN schema is utilized for handling the missing values. After that the four supervised algorithms (e.g., Decision Tree, k-NN, SVM, and Naïve Bayes) are tested. In addition, the Cross-validation technique with 10 folds is used for performance assessments of these classifiers. The results showed that while the Decision Tree and Naïve Bayes exhibit close results with accuracy of 0.9650 and 0.9700 respectively, the SVM classification registered higher performance accuracy with 0.9850. On the other hand, k-NN algorithm with k=1 showed accuracy of 0.8400. Table 2 shows the complete results along with the validation metrics of these algorithms. Also, Figure 3 shows the accuracy of the classifiers and Figure 4 shows the ROC graph comparison of these classifiers Moreover, Figures 5, 6, 7, and 8 illustrate the ROC of each and every classifier's performance in this experiment. ROC graph is graphical plot that diagnosis the classifier performance by analysis the its work based on the rates of true positive predication against the true negatives predication [26].

Validation metrics	Decision Tree	k-NN	SVM	Naïve Bayes
Accuracy	0.9650	0.8400	0.9850	0.9750
Recall	0.9658	0.8291	1.0000	1.0000
Precision	0.9741	0.8899	0.9750	0.9590
F-measure	0.9700	0.8584	0.9873	0.9791
Specificity	0.9639	0.8554	0.9639	0.9398

Table 2. Supervised performance with k-NN estimator.

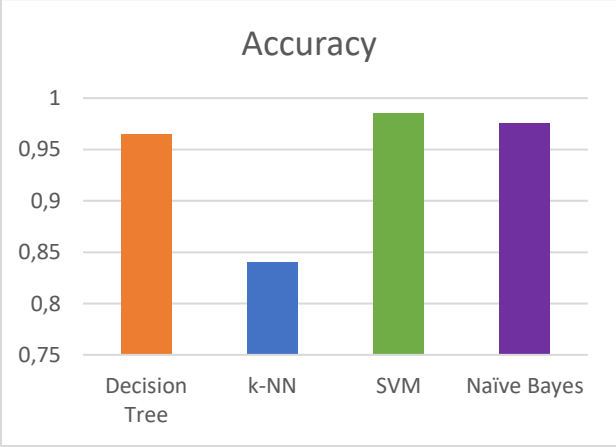

Figure 3: Supervised accuracy with k-NN estimator.

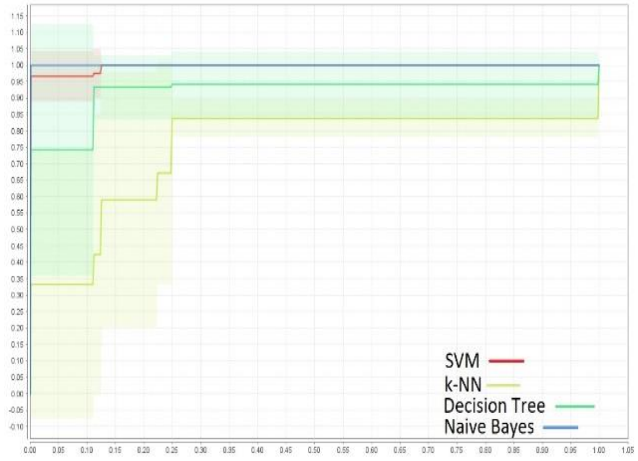

Figure 4: ROC graph comparison of the all classifiers with k-NN estimator.

Figure 5: ROC graph of the k-NN performance with k-NN estimator.

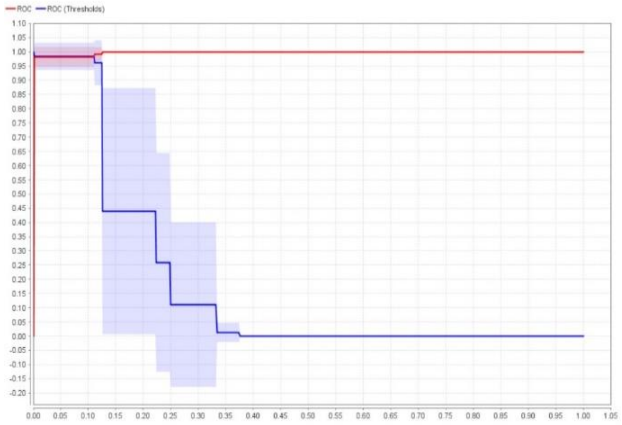

Figure 6: ROC graph of the Naïve Bayes performance with k-NN estimator.

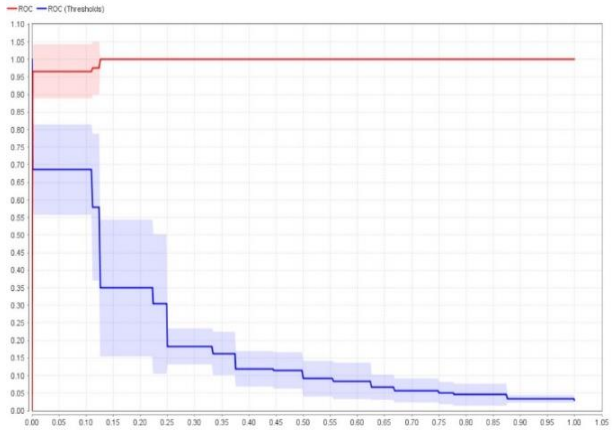

Figure 7: ROC graph of the SVM performance with k-NN estimator.

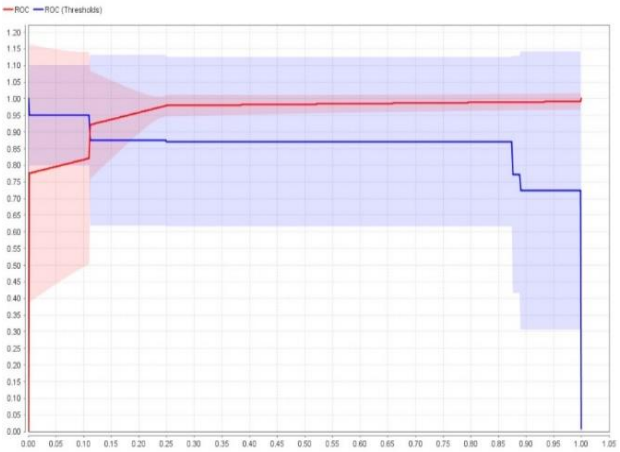

Figure 8: ROC graph of the Decision Tree performance with k-NN estimator.

Although all the classifiers achieve high accuracy rate, however, the SVM outperforms other classifiers as it employs "Nominal to Numerical" operator to map the different types of data to numerical type, so SVM can calculate the distance of these attributes to the hyperplane that separates the concept classes. Specifically, SVM proved its efficiency for application of two concepts classes due to find the optimal decision boundary (Hyperplane) that separate the two class in which are (Fake and real) and calculate the distance of each case (profile) to its nearest class label for the classification process.

7.2.2 Excluding the missing values using filtering operator

In the second case, the profiles with missing attributes are excluded by employing a special filtering operator provided by the RapidMiner, which filter the profiles based on specific conditions to keep/remove the profiles that met these conditions. Practically, the conditions of the Filter are set to remove any profile with missing values in anyone of their attributes. By applying this operator, a total of 33 profiles were removed from the collected data set leaving 167 profiles to be considered in this

experiment. The main purpose behind this experiment is to eliminate any factor that could affect the model's classification process or the accuracy because we estimated the missing values in the first experiment.

Validation metrics	Decision Tree	k-NN	SVM	Naïve Bayes
Accuracy	0.9461	0.8443	0.9880	0.9641
Recall	0.9406	0.8317	1.0000	1.0000
Precision	0.9694	0.9032	0.9806	0.9439
F-measure	0.9548	0.8660	0.9902	0.9712
Specificity	0.9545	0.8636	0.9697	0.9091

Table 3. Supervised performance with filtering operator.

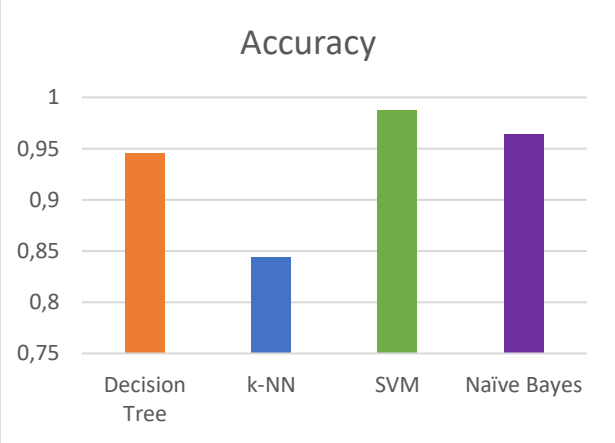

Figure 9. Supervised accuracy with filtering operator.

After that, the supervised algorithms are applied on the data, results showed the following accuracy rate (0.9461, 0.8443, 0.9880, and 0.9641) for Decision Tree, k-NN, SVM, and Naïve Bayes, respectively. Again, SVM exhibits the highest detection performance with accuracy of 0.9880, while K-NN the lowest with accuracy of 0.8433. Other performance indicators for these supervised algorithms are showed in table 3. And following the same vein of the previous experiment Figure 9 illustrates the accuracy results the employed classifiers and Figure 10 the ROC graph comparison of all classifiers employed in this experiment, Figures 11, 12, 13, and 14 shows the ROC graph for each one.

However, although our results are stable and good, one limitation that affects the validity of our study is that the used dataset is relatively small. Therefore, further validations over large datasets is required.

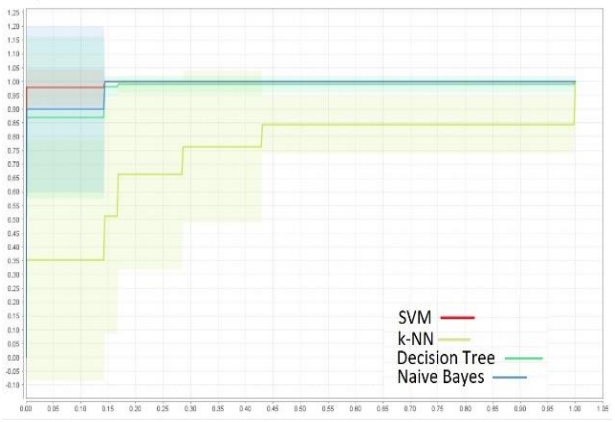

Figure 10. ROC curve of the supervised algorithms with filtering operator.

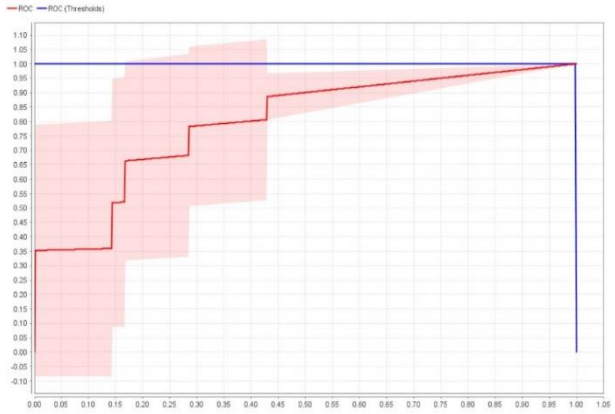

Figure 11: ROC graph of the k-NN performance with filtering operator.

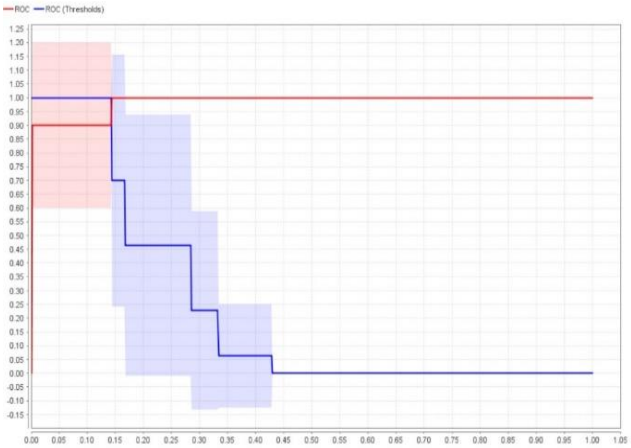

Figure 12: ROC graph of the Naïve Bayes performance with filtering operator.

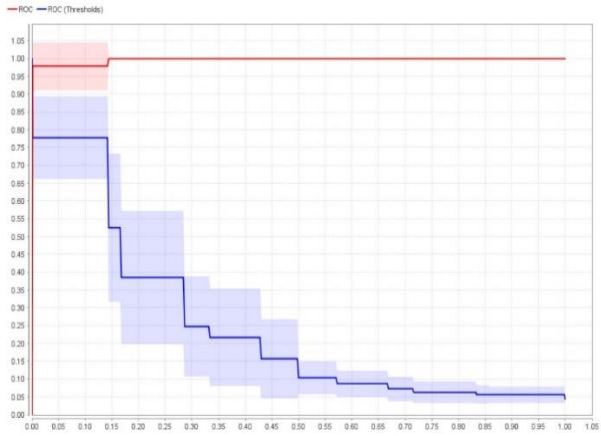

Figure 13: ROC graph of the SVM performance with filtering operator.

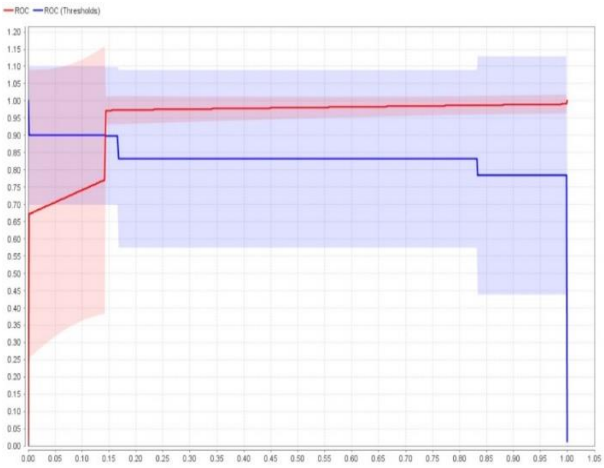

Figure 14: ROC graph of the Decision Tree performance with filtering operator.

8 Conclusion and future work

In this work, a smart system FBChecker is presented that have been designed specifically for detecting Facebook fake profiles. FBChecker consists of several components that collecting, preparing, validating, and mining the users' profiles using four supervised data mining techniques. These supervised techniques were implemented using the open source RapidMiner data science platform. The proposed system shows high efficiency performance for detecting fake profiles with accuracy rates reached %98, which represents a successful and promising result.

As a future work, we are aiming to use a large data set size and include more attributes that may employed in the detection model as discriminative features, and also apply more data mining techniques (unsupervised/Clustering algorithms) then evaluate which technique among them perform best.

ACKNOWLEDGMENT

The authors are grateful to the Applied Science Private University, Amman-Jordan, for the full financial support granted to cover the publication fee of this research article.

References

[1] Romero, Daniel M., Wojciech Galuba, Sitaram Asur, and Bernardo A. Huberman. "*Influence and passivity in social media*." In Proceedings of the 20th international conference companion on World Wide Web, 2011, ACM, pp. 113-114. https://doi.org/10.1145/1963192.1963250

[2] Ngai, E. W., Tao, S. S., & Moon, K. K. (2015). Social media research: Theories, constructs, and conceptual frameworks. International Journal of Information Management, 35(1), 33-44. https://doi.org/10.1016/j.ijinfomgt.2014.09.004

[3] Kaplan, Andreas M., and Michael Haenlein. "*Users of the world, unite! The challenges and opportunities of Social Media*." Business Horizons 53, no. 1: 59-68,.2010. https://doi.org/10.1016/j.bushor.2009.09.003

[4] Agichtein, Eugene, Carlos Castillo, Debora Donato, Aristides Gionis, and Gilad Mishne. *"Finding high-quality content in social media."* In Proceedings of the 2008 international conference on web search and data mining, 2008, ACM, pp. 183-194. https://doi.org/10.1145/1341531.1341557

[5] O'Keeffe, Gwenn Schurgin, and Kathleen Clarke-Pearson. *"The impact of social media on children, adolescents, and families."* Pediatrics 127, no. 4: 800-804, 2011. https://doi.org/10.1542/peds.2011-0054

[6] Hajirnis, Aditi. "*Social media networking: Parent guidance required*." The Brown University Child and Adolescent Behavior Letter 31, no. 12: 1-7, 2015. https://doi.org/10.1002/cbl.30086

[7] Tang, Qian, Bin Gu, and Andrew B. Whinston. "*Content contribution for revenue sharing and reputation in social media: A dynamic structural model.*" Journal of Management Information Systems 29, no. 2: 41-76, 2012. https://doi.org/10.2753/MIS0742-1222290203

[8] Facebook Newsroom. https://newsroom.fb.com/ company-info/ (24th July 2017)

[9] Aswani, R., Kar, A. K., & Ilavarasan, P. V. (2018). Detection of spammers in twitter marketing: a hybrid approach using social media analytics and bio inspired computing. *Information Systems Frontiers*, 1-16. https://doi.org/10.1007/s10796-017-9805-8

[10] Singh, N., Sharma, T., Thakral, A., & Choudhury, T. (2018, June). Detection of Fake Profile in Online Social Networks Using Machine Learning. In *2018 International Conference on Advances in Computing and Communication Engineering (ICACCE)* (pp. 231-234), IEEE. https://doi.org/10.1109/ICACCE.2018.8441713

[11] Facebook Terms of Service. https://www.facebook.com/legal/terms (18th august 2017).

[12] Moustafa, N., & Slay, J. (2016). The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the

comparison with the KDD99 data set. *Information Security Journal: A Global Perspective*, *25*(1-3),18-31. https://doi.org/10.1080/19393555.2015.1125974

[13] Fürnkranz, Johannes. "*Separate-and-conquer rule learning."* Artificial Intelligence Review 13, no. 1: 3-54,1999. https://doi.org/10.1023/A:1006524209794

[14] Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153-1176. https://doi.org/10.1109/COMST.2015.2494502

[15] Büchner, Alex G., and Maurice D. Mulvenna. "*Discovering internet marketing intelligence through online analytical web usage mining*." ACM Sigmod Record 27, no. 4: 54-61, 1998. https://doi.org/10.1145/306101.306124

[16] Xiao, Cao, David Mandell Freeman, and Theodore Hwa. "Detecting clusters of fake accounts in online social networks." In *Proceedings of the 8th ACM Workshop on Artificial Intelligence and Security*, pp. 91-101.ACM,2015. https://doi.org/10.1145/2808769.2808779

[17] Gupta, Aditi, and Rishabh Kaushal. "*Towards detecting fake user accounts in Facebook*." In Asia Security and Privacy (ISEASP), 2017 ISEA, pp. 1-6. IEEE,2017. https://doi.org/10.1109/ISEASP.2017.7976996

[18] Ahmed, Faraz, and Muhammad Abulaish. "*A generic statistical approach for spam detection in Online Social Networks*." Computer Communications 36, no.10:1120-1129,2013. https://doi.org/10.1016/j.comcom.2013.04.004

[19] Fire, Michael, Dima Kagan, Aviad Elyashar, and Yuval Elovici. "Friend or foe? Fake profile identification in online social networks." *Social Network Analysis and Mining* 4, no. 1 (2014): 194. https://doi.org/10.1007/s13278-014-0194-4

[20] Bhat, Sajid Yousuf, and Muhammad Abulaish. "Community-based features for identifying spammers in online social networks." In *Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference on*, pp. 100-107. IEEE, 2013. https://doi.org/10.1145/2492517.2492567

[21] Wang, Alex Hai. "Detecting Spam Bots in Online Social Networking Sites: A Machine Learning Approach." DBSec 10: 335-342, 2010. https://doi.org/10.1007/978-3-642-13739-6_25

[22] Han, Jiawei, Jian Pei, and Micheline Kamber. Data mining: concepts and techniques. Elsevier, 2011.

[23] Cook, Diane J., and Lawrence B. Holder, eds. Mining graph data. John Wiley & Sons, 2006. https://doi.org/10.1002/0470073047

[24] Kotsiantis, Sotiris B., I. Zaharakis, and P. Pintelas. "Supervised machine learning: A review of classification techniques." Emerging artificial intelligence applications in computer engineering 160: 3-24, 2007.

[25] Powers, David Martin. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." 2011.

[26] Hanley, James A., and Barbara J. McNeil. "The meaning and use of the area under a receiver operating characteristic (ROC) curve." Radiology 143, no. 1: 29-36: 198.