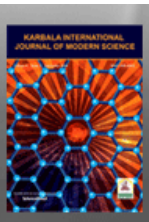

Karbala International Journal of Modern Science

Volume 6 | Issue 4 Article 8

Twitter Fake Account Detection and Classification using Ontological Engineering and Semantic Web Rule Language

MOHAMMED JABARDI
University of Kufa, mohammed.naji@uokufa.edu.iq

Asaad Sabah Hadi
University of Babylon

Twitter Fake Account Detection and Classification using Ontological Engineering and Semantic Web Rule Language

Abstract

Nowadays, Twitter has become one of the fastest-growing Online Social Networks (OSNs) for data sharing frameworks and microblogging. It attracts millions of users worldwide where subscribers communicate with each through posts and messages known as "tweets". The open structure and behaviour of Twitter cause it to be vulnerable to attacks from fake accounts and a large number of automated software, known as 'bots'. Bots are regarded to be malicious as they send spam to users of social networks over the internet. Data security and privacy are among the most critical issues of social network users, as the protection and fulfilment of these requirements strengthen the network's interest and, ultimately, its credibility. To overcome these issues, we need to build an efficient model to detect and classify fake twitter accounts. This paper presents a new approach with dual functions, namely to identify and classify the twitter bots based on ontological engineering and Semantic Web Rule Language (SWRL) rules. Web Ontology Language (OWL), Semantic Web Rule Language (SWRL) rules, and reasoners are deployed to inductively learn the rules that distinguish a fake account (bot) from a real one, as well as to classify fake accounts into fake followers or spam bot. Our approach could properly identify the false account with an accuracy of (97%) in the first stage, after which these fake accounts were classified into spam or fake follower bots with an accuracy rate of (94.9%). Furthermore, it has been found that he ontology classifier is a more interpretable model that offers straightforward and human-interpretable decision rules, as compared to other machine learning classifiers.

Keywords

Spambot, FakeFollower, Semantic Web Rule Language, Ontology, Reasoner

1. Introduction

OSNs play an essential role for users of internet in carrying out their daily activities such as sharing content, reading news, posting messages, reviewing items, and discussing events. The increase of web infrastructure and the development of Online Social Networks (OSN), such as Facebook, Twitter, and LinkedIn, have contributed to the emergence of unnecessary digital bots as human-robotic actors. These bots are in fact computer programs which quickly and efficiently execute various tasks that cannot be performed by humans easily. Statistics indicate that the average number of monthly active Twitter users was 336 million in total during the first quarter of 2018, with about 5 percent of fake or spam accounts [1]. Equally, 9%−15% of the tweets posted came from fake accounts [2]. In recent years, the Twitter platform has come under fire, particularly considering the cases of pressure and its alleged involvement in US politics. However, Twitter has decided to act in reaction, by cleaning up over 70 million fake accounts last year [3]. This massive population of OSN causes different kinds of problems in terms of data security and privacy. Bots designed for malicious activities have turned into a severe internet threat. Various OSN service providers used several ways to counter spambots. For example, Twitter and Facebook have enabled the option “report as spam” to help identify a spam bot. Bots make use of OSNs as an enticing tool for transmitting offensive information, biasing public opinion, manipulating consumer understanding, and performing fraudulent activities. They can generally be classified into Spam Bots, Social Bots, Like Bots, Influential Bots and Fake follower Bots [4].

Among the numerous tweets that are found in the Twitter stream, there are a number of posts that come from bots. A Twitter bot is a software that sends tweets consequently to users. Twitter bots can have a significant impact on public opinion on a particular topic, which can in turn be used to promote an ideology or business because of Twitter's power and reach, as well as very low-cost implementation. The most significant purposes of Twitter bots can be summarized as follows: 1) disseminating misinformation and fake news; 2) stigmatize someone's personality; 3) capturing subscribers' credentials and generating malicious communications; 4) misleading users towards counterfeit sites; 5) changing the intellect of an individual or gathering paradigm, impacting notoriety [5]. Data protection and privacy are among the critical concerns of social networks users, and preserving and fulfilling such criteria will enhance the trust of the network and consequently, its reputation. To overcome these issues, an efficient model needs to be built to detect and classify fake Twitter accounts.

In this paper, we presented a new ontology-based approach called Fake Account Detection and Classification Ontology (FADCO) on a Twitter online social network. First, the essential features are selected to fit our strategy. Secondly, the ontology is constructed so as to represent the relationships. Finally, the SWRL rules-based reasoner is used to determine whether the profile is a bot or real account under predefined constraints. Our study is based on the Fake Project dataset released by the Institute of Informatics and Telematics of the Italian National Research Council (IIT-CNR) Lab [6].

The rest of the paper is organized as follows: Section 2 is reserved for the literature review on the detection and classification techniques. Section 3 is concerned with the methodology used in this paper. Section 4 describes and evaluates the extent to which the model performance proved to be effective. Section 5 presents a summary of the obtained results, followed by a brief explanation of the expected outcomes. Finally, Section 6 offers future work conclusions and directions.

2. Related work

Fake account detection within social networks has received much attention in the last years, and different approaches have addressed this problem. The main focus currently lies in two approaches: a machine learning approach and semantic modelling approach.

2.1. Machine learning approach

Steganography methods are used to hide essential information of account into images. At all times, whenever someone else copies the picture in the future and tries to create a fake profile using stolen data, the program will automatically detect this deception and fraud [7]. A combination of time interval entropy and tweet similarity is used in the detection of bot spammers. Timestamp sets are used to measure the entropy of each user's time interval. The similitude based on

https://doi.org/10.33640/2405-609X.2285

uni-gram comparison can also be used to measure tweet similarity [8]. The URL blacklist matching system is presented with SVM (vector support devices). A system is split mainly into three components: mapping and assembly, pre-filtering, and classifying. Mapping and assembly are used for matching the URL blacklist. If these form a match indeed, then the specific URL will be identified as spam. In case they do not match, then the SVM will further analyze the URL [9].

The process of real-time cybersecurity account detection on Twitter is based on three different feature sets and three different methods of machine learning, which are decision trees, random forests, and SVM [10]. The DeepScan model splits the activity data of each user into many continuous-time intervals. DeepScan utilizes deep learning technology to use time-series features that are relatively more inclusive and descriptive than standard features [11]. A flexible metric is suggested for measuring any changes in user activity, as well as to design new features for measuring user evolution patterns. The unsupervised and supervised machine learning techniques are then used to distinguish spammers [12]. A novel framework proposed for the spammer detection system is based on Bagging Extreme Learning Machine ELM, which detects spammers within social networking systems. Several combined ELMs have been deployed in the bagging approach [13]. A semi-supervised technique developed for spam detection in Twitter makes use of an ensemble-based framework that is comprised of four classifiers [14]. Several different combinations of four machine learning techniques, namely "Bernoulli Naive Bayes", "Gaussian Naive Bayes," "Multinomial Naive Bayes," and "Decision Tree" have been presented. They used the voting classifier, which is a form of group learning, to measure the accuracy of the various classifier combinations [15]. Both fuzzy logic and evaluating multilayer perceptron via the neural network have been used to identify spam. It was found that the applied fuzzy logic efficiently managed the large data set and consumed relatively less time to detect spammers in seconds [16].

2.2. Semantic modelling approach

Social network analysis has become more popular nowadays. Nevertheless, quite few researchers have analyzed fake accounts detection as based on a semantic approach and ontology engineering. A novel ontology-based method was proposed to identify dubious content on Twitter during instances or events where tweets are linked to ontologies of specific themes, eventually in order to verify the similarities between tweet texts and ontologies dealing with relevant subjects [17]. A new spam detection algorithm has been presented, called Social Event Detection, and it is based on ontology Multiple steps are introduced, beginning with the creation of the ontology, attribute extraction, the correlation of words to the current class context and ending with the identification of whether it is spam or not [18]. There are two types of ontology spam filters that have been applied: global ontology filter and user-customized ontology filter. The user-customized ontology filter was created based on the background of the specific user and the filtering mechanism that was used in the creation of the global ontology filter [19].

The ontology-based phishing approach is proposed as a semantic system used for identification. It examines conceptualization activity for lexical characteristics, which are hypothesized to eliminate any confusion to superficial characteristic variance. The proposed solution adds semantics to the bag-of-words and part-of-speech strategies that are incredibly accurate [20]. A definition logic based on ontology proposes to describe potential phishing scenarios for the text. It starts by giving a generic taxonomy of processes of email phishing, followed by developing DL-based concept of the Tbox and Abox [21].

3. Methodology

Bots have developed exponentially over the past few years to the point that it has become difficult to distinguish them from real accounts. Supervised machine learning models are the most popular techniques used for the detection of bots. This section explains our new proposed approach, based on user attributes and ontology technologies, attempts to identify and recognize fake accounts on Twitter. The system is composed of three stages: data preprocessing and features extraction stage, ontology construction stage, and SWRL rules and reasoner as a classifier stage, as shown in Fig. 1.

3.1. Data preprocessing and features selection

The Fake Project dataset released by the Institute of Informatics and Telematics of the Italian National Research Council (IIT-CNR) used in this study. The dataset has three sets of Twitter accounts, social spam bots, fake follower bots, and real accounts. The dataset contains 11,737 Twitter accounts with 12, 030, 893 tweets, as shown in Table 1.

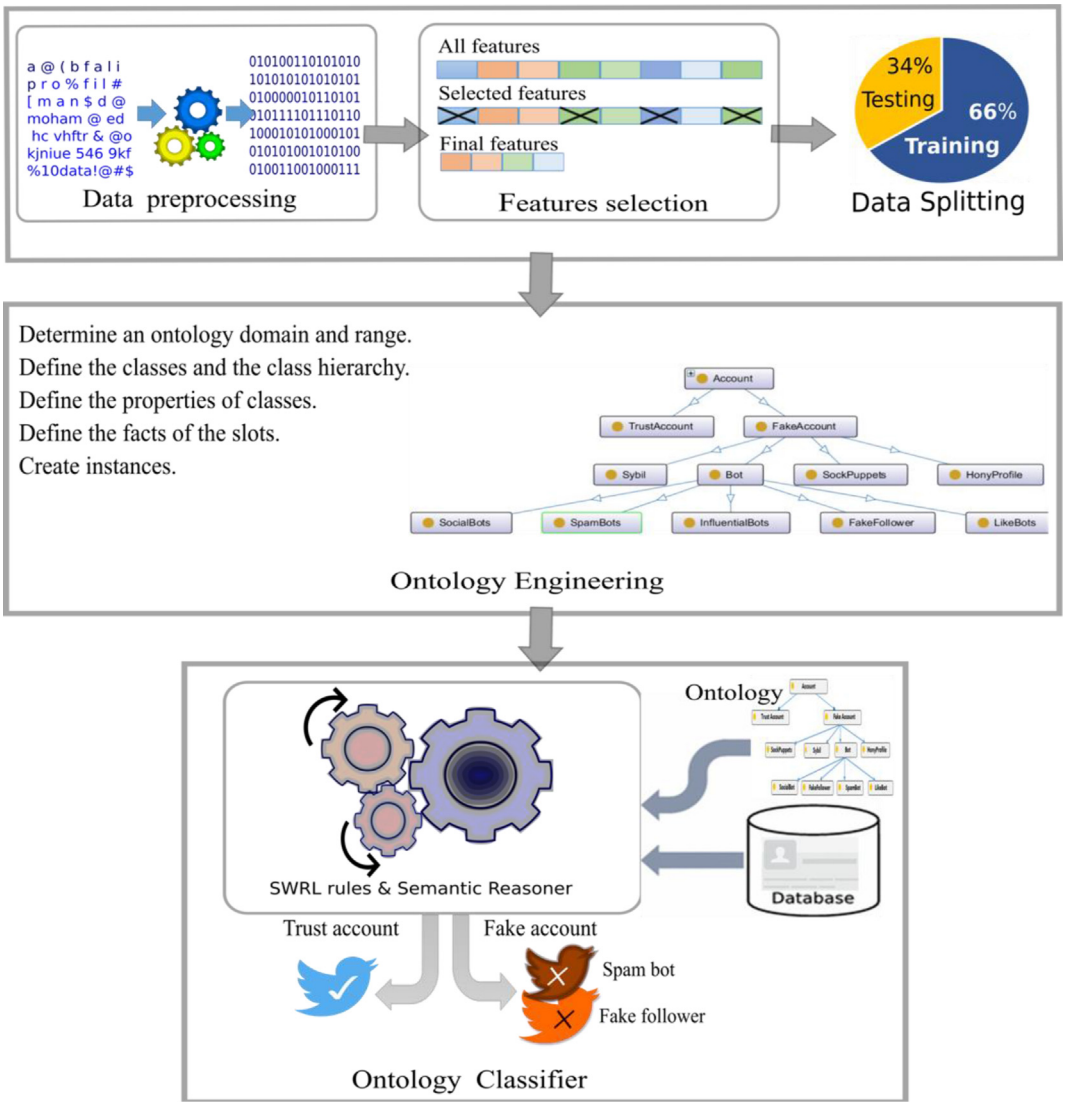

Fig. 1. Block diagram of the proposed approach.

Given the fact that data can be inaccurate, inconsistent, and incomplete; therefore, certain features are analyzed to select the best ones which may gain more favorable outcomes. The dataset preprocessing includes dropping unimportant features, type conversion, and too many missing values are dropped, as shown in Table 2.

Table 1
The dataset description. 42% Spambots, 28% Fake follower and 30% Trust account.

Grouping	Number of Accounts	Number of Tweets
Spambots	4912	3,457,344
Fake Followers	3351	196,027
Real Users	3474	8,377,522
Total	**11,737**	**12,030,893**

New driven features are extracted from the original features such as the friendship feature, which is represented by the ratio of friend's number to a number of followers. The cleaned data set with important related features and new driven features is shown in Table 3.

A crucial step towards practical algorithms in machine learning is the selection of instructive, discriminating and independent features. Feature selection functions help the machine learning algorithm to train more quickly. It reduces a model's complexity and makes understanding simpler. It enhances the accuracy of a model whenever the correct subset is selected [22].

The Area Under the ROC Curve (AUC) technique is a performance measurement for the binary classification problem at different thresholds [23]. AUC-ROC

Table 2
Preprocessing operation task of the dataset, including the removal of all features of string type, irrelevant features, and boolean to integer type conversion (0 or 1).

Feature	type	Pre-processing
Id	int	
name	String	Dropped
screen_name		Not useful
URL		
Lang		
time_zone		
Location		
default_profile		
profile_image_url		
profile_banner_url		
profile_background_		
image_url_https		
profile_text_color		
profile_image_url_https		
profile_sidebar_border_color		
profile_sidebar_fill_color		
profile_background_image_url		
profile_background_color		
profile_link_color		
utc_offset		
statuses_count	Int	No action
followers_count		
friends_count		
favourites_count		
listed_count		
Description	String	Converted to integer (0/1)
default_profile_image	Boolean	
geo_enabled		
profile_use_background_image_image		
profile_background_tile		
translator_type	String	Just too many missing
is_translator	Boolean	values dropped
follow_request_sent		
Protected		
Verified		
Notifications		
contributors_enabled		
Following		
has_extended_profile		
created_at	date	age of Account
crawled_at		

curve can be used for ranking the top essential features among many features according to a specific threshold (the cut-off point). Threshold classifications are monotonic; therefore, any feature that is ranked positive for a given threshold will also be ranked positive for all lower thresholds. A threshold value (or cut-off point) determines how expected posterior probabilities would be translated to class labels for binary scoring classification. Finally, the dataset is summarized, and only the essential features are selected from the original ones, as shown in Table 4.

3.2. Ontology construction

The ontology is developed from information gathered by domain experts and assigned to the ontology in the form of a set of concepts, relationships, and definitions. The proposed ontology, called Fake Account Detection and Classification Ontology (FADCO), was built using Protégé 5.0. Protégé is a free, open-source ontology editor and a framework for managing information, established by the Biomedical Informatics

Table 3
The extracted new hidden knowledge.

Driven feature	Description
friendShip	The ratio of friends counts to followers count
followerShip	The ratio of followers counts to friends count
Interestingness	The ratio of favourite counts to statuses count
Activeness	The ratio of statuses counts to AccountAge
friendRate	The ratio of Friends counts to AccountAge
followerRate	The ratio of followers counts to AccountAge
Reputation	The ratio of followers counts to the sum of friends count and followers count

Table 4
The ranking of the most important features.

Feature name	Ranking	weight
favouritesCount	1	0.931
interest	2	0.893
statusesCount	3	0.888
geoEnabled	4	0.877
followersCount	5	0.87
accountAge	6	0.862
friendRate	7	0.844
reputation	8	0.835
friendShip	9	0.814
listedCount	10	0.792

Research Center of Stanford [24]. The architecture of ontology consists of the following steps:

- Determining the domain and scope of the ontology.
- Defining the classes and class hierarchy.
- Defining the properties of classes.
- Defining the facts of the slots.
- Creating instances.
- Defining and writing SWRL rules of physiognomy.

The classification and validation apply ontology reasoners to check the ontology and extract knowledge, in order to make a knowledge base. We utilized the Stanford Protégé ontology editor to develop an OWL ontology to represent the classes and properties of the model. The main classes were Account and Person. A person is an object that can own an Account. Apparent Properties were primarily defined in the account class to set attributes such as the number of followers and the number of friends. Fig. 2 shows the classes' hierarchy of the FADCO ontology, as visualized in the WebVOWL ontology visualization tool [25]. WebVOWL is a plugin tab in protégé's editor for the interactive visualization of ontologies. It carries out the Visual Notation for OWL Ontologies (VOWL) by providing graphical representation for elements of the OWL. Different formats are provided to organize the ontology structure automatically. It supports different relationships: subclass, individual, properties of the domain/range object, and equivalence. We can filter relationships and types of nodes to create a specific view.

3.3. Classifier

3.3.1. The reasoner

OWL reasoners such as Pellet, FaCT++, and HerMiT [26] are required for executing SWRL rules and infer new ontology axioms. The Pellet reasoner, which has been applied in our approach, has a more direct functionality for working with OWL and SWRL rules, particularly because it allows defining custom SWRL built-ins. The process of detecting and classifying fake accounts depends on the hierarchical classification concept. The hierarchical classification has two main

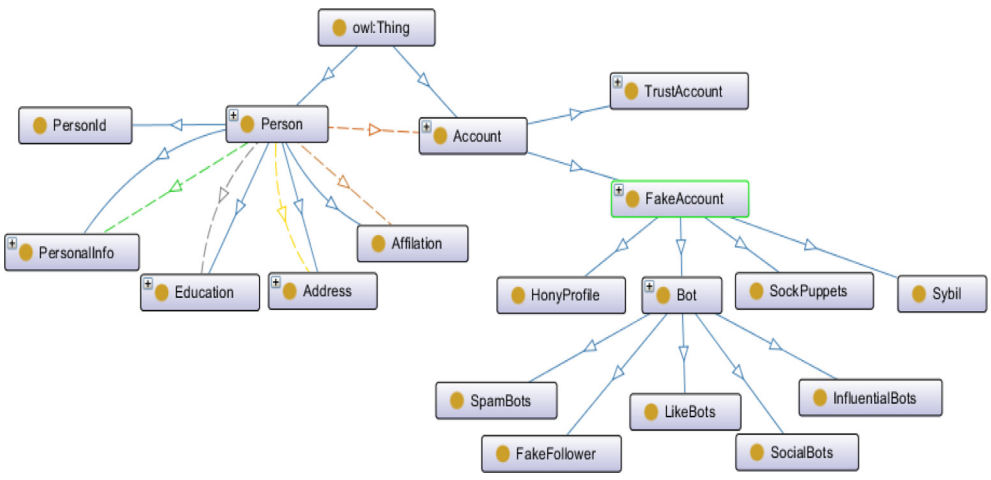

Fig. 2. The classes hierarchy of Fake Account Detection and Classification Ontology (FADCO).

advantages compared to flat classification. First, searching for some high-level classes and then for some relevant sub-level categories is much easier than performing a universal search for all current groupings. Second, it inverts the understanding of the connectedness of topics in the hierarchy. At each level of the category tree, one or more classifiers are constructed, and at that level, each classifier works as a flat classifier. The classifier would first classify a piece of root level knowledge into one or more categories below the level. The classifier would then organize it into the lower level category until it enters one or more final categories, which may be either a leaf or internal category.

3.3.2. Semantic rule

In this phase, SWRL rules will be written depending on the strong relationships in order to detect fake accounts from the given profile account features. The rules will be used to infer new knowledge from an existing ontology knowledge base. All rules are expressed in terms of ontology concepts (classes, properties, individuals). The SWRL rules will be stored as Web Ontology Language (OWL) syntax in the domain ontology. For the classification of multi-labels, we use an approach that focuses on estimates of a posteriori likelihood of an object belonging to a specific class. According to this approach, a set of rules (conditions) are used to check whether the instance is fit to these conditions. The rules are determined by some specific thresholds which are capable of predicting the probability of class membership. For example, SWRL rules for inferring that a particular account belongs to a reliable class can be built as follow: if an account ***hasStatusAccount*** value is higher than 144 and the ***hasFavritesCount*** is greater than three, as shown in Fig. 3.

4. Evaluation

The assessment methods play a critical role in the design of classification models. We assessed the performance of our model to obtain better results, and

Account (? a) ^
hasFavoritesCount (?a, ?v) ^
swrlb: greaterThan(?v, 3) ^
hasStatusesCount (?a, ?s) ^
swrlb:greaterThan(?s, 144) → Reliable(?a)

Fig. 3. SWRL rule for inferring new knowledge "Reliable" class membership.

here is where Confusion matrix comes to the spotlight. A confusion matrix is a method of summing up a classification algorithm's results [27]. The most fundamental terms used with a confusion matrix for a binary classifier are:

- True-positive (TP): the number of accounts correctly identified as Faked.
- False-positive (FP): the number of accounts incorrectly identified as Faked.
- True-negative (TN): the number of accounts correctly identified as Trusted.
- False-negative (FN): the number of accounts incorrectly identified as Trusted.

The assessment metrics are often computed from a confusion matrix:

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

$$F-Score = 2*\frac{Precesion*Rcall}{Precesion+Rcall} \tag{3}$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+TP} \tag{4}$$

5. Results and discussion

5.1. Results

This paper aimed to detect and classify fake Twitter accounts using ontology engineering and SWRL rules. We used the MIB Datasets of fake and legitimate Twitter accounts published by the Institute of Informatics and Telematics (IIT), Italian National Research Council's (CNR). The data set included 11,737 Twitter

Table 5
The ranking and accuracy of the most important features and their Cut-off points.

Feature name	Cutoff point	TP	FP	TN	FN	Accuracy
favouritesCount	3	2763	26	1114	88	**93.1%**
interest	0.0296	2600	174	966	251	**89.3%**
statusesCount	144	2454	47	1093	397	**88.8%**
geoEnabled	1	2789	428	712	62	**87.7%**
followersCount	26	2381	46	1094	470	**87.0%**
accountAge	77	2330	27	1113	521	**86.2%**
friendRate	0.45	2629	399	741	222	**84.4%**
reputation	0.1929	2256	60	1080	595	**83.5%**
friendShip	1	2644	533	607	207	**81.4%**
listedCount	0	2392	291	849	459	**79.2%**

Table 6
Confusion matrix of faked and trusted accounts. Correctly classified instances 3880 (97.5%) and incorrectly classified instances 111 (2.5%).

Faked and Trusted classification		Actual values	
		Faked	Trusted
Predictive value	**Faked**	TP = 2758	FP = 39
	Trusted	FN = 72	TN = 1122

accounts (see Table 1), divided into a training set 66% and testing set 34%.

The Area Under the ROC Curve (AUC) technique is a performance measurement for the binary classification problem at different thresholds. The AUC-ROC curve can be used for ranking the top essential features among many features according to a specific threshold (the cut-off point). Threshold classifications are monotonic; therefore, any feature that is ranked positive for a given threshold will also be ranked positive for all lower thresholds. A threshold value (or cut-off point) determines how expected posterior probabilities can be translated into class labels for binary scoring classification. The cut-off point represents the range of criterion values which determines a positive condition giving the highest accuracy, as shown in Table 5.

The proposed ontology approach for fake account detection and classification falls into two successive stages. In the first stage, a reasoner uses SWRL rules to infer fake and legitimate accounts, so as to classify them into Fake Account and Trusted Account classes. The results of the detection stage show that 2758 out of the 2797 accounts are correctly identified fake accounts with a 97.5% accuracy. The confusion matrix shows the findings of the first stage (see Tables 6 and 7).

In the second stage, the fake accounts class resulting from the previous stage are reclassified by the reasoner for inferring spam accounts and fake follower accounts. The second stage results indicated that 1672 out of 1776 fake accounts were classified as Spam bots, whereas the remaining 1016 accounts were classified as fake followers with an accuracy rate of 96.1%, as shown in Table 7.

Table 7
Confusion matrix: correctly classified instances 2688 (96.1%) and incorrectly classified instances 109 (3.9%).

Spam and Fake follower Classification		Actual values	
		Spam	FakeFollower
Predictive value	**Spam**	TP = 1672	FP = 104
	Fake follower	FN = 5	TN = 1016

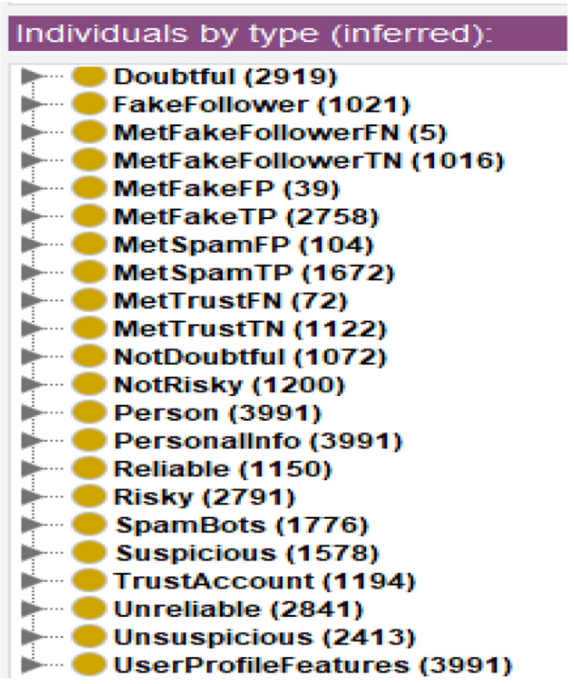

Fig. 4. The statistics of inferred concepts (classes) using existing facts and deduced axioms see Fig. 5.

If we discuss the results from the view of interpretable and clarity (How & Why), the ontology outcomes considered the best because the make of decision is understandable and easily described.

The ontology results of both detection and classification stages with all details are shown in Fig. 4. The results of the ontology clarify that all individuals (Twitter accounts) were inferred using SWRL rules for the top essential features.

5.2. Discussion

Previous research has focused on machine learning for detection and classification techniques, while this approach demonstrated the use of ontology engineering with semantic web rules. In the fake account detection phase, we compared the ontology results to different machine learning techniques using the

Table 8
Evaluation metrics show the accuracy of an ontology vs some machine learning techniques.

Technique	Recall	Precision	F-Measure	Accuracy
NAIVE-BAYS	0.990	0.930	0.959	0.945
Ontology	**0.975**	**0.986**	**0.980**	**0.972**
Logistic	0.981	0.987	0.984	0.978
SVM	0.984	0.987	0.985	0.979

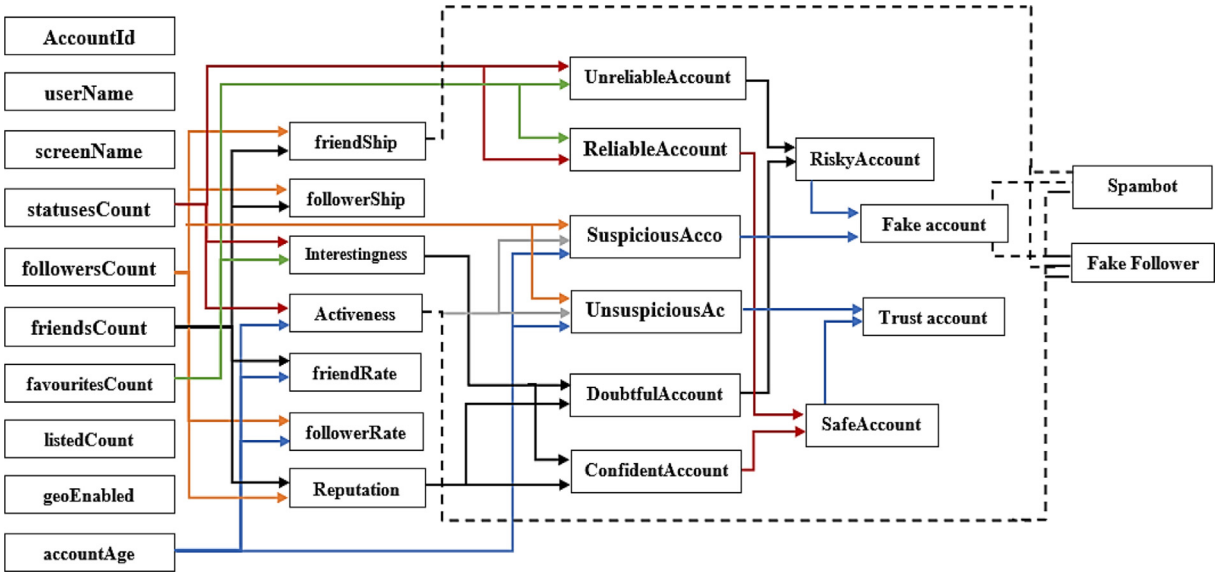

Fig. 5. The inferring ontology levels that are used to detect and classify a fake account and its type depending on the existing relationship (profile features).

Waikato Environment for Knowledge Analysis (WEKA) [28]. The machine learning techniques used are Naïve Bayes, SVM and Logistic. The results show that the ontology approaches achieved an accuracy of 97.5%, which is considered to be a much more favorable outcome as compared to alternative machine learning classifiers, as shown in Table 8.

In terms of comparison, despite the fact that Black-box classifiers can make excellent predictions, yet they do not offer human interpretable and straightforward decision rules and thus leave the reasoning behind such predictions rather incomprehensible [29]. The ontology classifier, on the other hand, is an interpretable model, which can thus provide insights into how the process takes the decision. The findings of the ontology approach are very similar and comparable to those of the machine learning techniques. The results are also human-interpretable, and rules for the decision can be quickly established.

The significant rules with the best outcomes are selected according to the important related features. All test rules are converted into SWRL rules, using a conversion technique known as rolification. A Pallet reasoner was applied to determine whether the account is fake or not. The power of ontology is to distinguish hidden knowledge from existing knowledge.

Different levels of knowledge are constructed hierarchically, starting from existing facts to construct the first level, and further progresses to infer the other levels. We deduced the final judgment by integrating the first-level hypothesis of evidence with the second-level inference and so forth, as described and summarized in Fig. 5.

We want to compare the current study findings with other works in the field of fake Twitter accounts detecting and classification. At this point, a meaningful comparison is not feasible, primarily for two reasons. Many types of research concentrate on either identifying or classifying bogus accounts, not both, so these topics are too different from our perspective to be comparable. The second reason that there is no standard public dataset for researchers, so each group use their own dataset. Therefore, the scale, date, and features of each dataset are different. The issue in our scenario is that the results from the detecting stage will be the input to the classification stage, so any errors in the detection stage will pass to the classification stage. For instance, those accounts that have been incorrectly identified as false accounts will pass to the next step. Therefore, they will be classified as either a fake follower or spam bot, meanwhile they are neither in reality.

6. Conclusion

In this paper, a new approach has been proposed to detect and classify fake accounts on Twitter social networks, using ontological engineering. We modeled an ontological approach of knowledge representation across the OWL language, SWRL rules, and reasoner.

We focused on the features of profiles that could be further translated into axioms and laws. The reasoner has been used for executing all OWL ontology queries so as to obtain the correct request answers. The Pellet reasoner was used as a classifier for Twitter account detection and classification.

The proposed approach was carried out based on the standard metrics, using 3991 Twitter profile accounts. The system correctly identified 2758 out of the 2797 accounts as fake accounts with an accuracy rate of 97.5%. While in the classification stage, results indicated that 1672 out of 1776 fake accounts were classified as Spam bots, whereas the remaining 1016 accounts were classified as fake followers with an accuracy rate of 96.1%.

Finally, the ontology classifier is an interpretable model that offers straightforward and human-interpretable decision rules. Ontology can work by clarifying the terminology for coherent and unified reasoning purposes. Furthermore, ontology allows the reuse and sharing of existing knowledge bases that describe specific circumstances.

References

[1] C. T, E. Dwoskin, Twitter Is Sweeping Out Fake Accounts like Never before, Putting User Growth at Risk, 2018. https://www.washingtonpost.com/technology/2018/07/06/twitter-is-sweeping-out-fake-accounts-like-never-before-putting-user-growth-risk/. (Accessed 13 October 2020).

[2] Z. Chong, Up to 48 Million Twitter Accounts Are Bots, Study Says, 2017. http://www.cnet.com/news/new-study-says-almost-15-percent-of-twitter-accounts-are-bots/. (Accessed 13 October 2020).

[3] Y. Lin, 10 Twitter Statistics Every Marketer Should Know in 2020, 2020. https://www.oberlo.com/blog/twitter-statistics. (Accessed 13 October 2020).

[4] A. Wani, S. Jabin, A. Nehaluddin, A Sneak into the Devil's Colony-fake Profiles in Online Social Networks, arXiv.org e-Print Archive vol. 5, 2018, pp. 26−39.

[5] E.V.D. Walt, J. Eloff, Using machine learning to detect fake identities: bots vs humans, IEEE Access 6 (2018) 6540−6549, https://doi.org/10.1109/ACCESS.2018.2796018.

[6] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, M. Tesconi, The paradigm-shift of social spambots: evidence, theories, and tools for the arms race, abs/1701, in: Proceedings of the 26th International Conference on World Wide Web Companion vol. 3017, 2017, pp. 963−972.

[7] E. Ahmadizadeh, E. Aghasian, H.P. Taheri, R.F. Nejad, An automated model to detect fake profiles and botnets in online social networks using steganography technique, IOSR J. Comput. Eng. 17 (2015) 65−71, https://doi.org/10.9790/0661-17146571.

[8] R.S. Perdana, T.H. Muliawati, R. Alexandro, Bot spammer detection in Twitter using tweet similarity and time interval entropy, Jurnal Ilmu Komputer dan Informasi 8 (2015) 19−25, https://doi.org/10.21609/jiki.v8i1.280.

[9] S.M. Asadullah, S. Viraktamath, Classification of twitter spam based on profile and message model using Svm, Int. J. Eng. Res. Technol. 4 (2017) 2862−2865.

[10] Ç.B. Aslan, R.B. Sağlam, S. Li, Automatic detection of cyber security related accounts on online social networks: twitter as an example, in: Proceedings of the 9th International Conference on Social Media and Society, 2018, pp. 236−240, https://doi.org/10.1145/3217804.3217919.

[11] G. Qingyuan, Y. Chen, X. He, Z. Zhuang, T. Wang, H. Huang, X. Wang, X. Fu, DeepScan: Exploiting deep learning for malicious account detection in location-based social networks, IEEE Commun. Mag. 56 (2018) 21−27, https://doi.org/10.1109/MCOM.2018.1700575.

[12] Q. Fu, B. Feng, D. Guo, Q. Li, Combating the evolving spammers in online social networks, Comput. Secur. 72 (2018) 60−73, https://doi.org/10.1016/j.cose.2017.08.014.

[13] S. Rathore, A.K. Sangaiah, J.H. Park, A novel framework for internet of knowledge protection in social networking services, J. Comput. Sci. 26 (2018) 55−65, https://doi.org/10.1016/j.jocs.2017.12.010.

[14] A. Singh, S. Batra, Ensemble based spam detection in social IoT using probabilistic data structures, Future Generat. Comput. Syst. 81 (2018) 359−371.

[15] V. Gupta, A. Mehta, A. Goel, U. Dixit, A.C. Pandey, Spam detection using ensemble learning, in: Harmony Search and Nature Inspired Optimization Algorithms, Springer ICHSA, 2019, pp. 661−668.

[16] P. Tehlan, R. Madaan, K.K. Bhatia, A spam detection mechamism in social media using Soft computing, in: 2019 6th International Conference on Computing for Sustainable Global Development, INDIACom), 2019, pp. 950−955.

[17] B. Halawi, A. Mourad, H. Otrok, E. Damiani, Few are as good as many: an Ontology-based tweet spam detection approach, IEEE Access 6 (2018) 63890−63904, https://doi.org/10.1109/ACCESS.2018.2877685.

[18] S. Selvam, R. Balakrishnan, B. Ramakrishnan, Social event detection-A systematic approach using ontology and linked open data with significance to semantic links, Int. Arab J. Inf. Technol. 15 (2018) 729−738.

[19] S. Youn, SPONGY (SPam ONtoloGY): email classification using two-level dynamic ontology, Sci. World J. (2014), https://doi.org/10.1155/2014/414583.

[20] G. Park, Towards Ontology-Based Phishing Detection, Doctoral Dissertation, Purdue University, 2018.

[21] F. Tchakounté, D. Molengar, J.M. Ngossaha, A description logic ontology for email phishing, Int. J. Inf. Secur. Sci. 9 (2020) 44−63.

[22] J. Brownlee, Feature Selection to Improve Accuracy and Decrease Training Time, *https://machinelearningmastery.com/feature-selection-to-improve-accuracy-and-decrease-training-time/*, 2014. (Accessed 12 October 2020).

[23] M. Sokolova, N. Japkowicz, S. Szpakowicz, Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation, in: Australasian Joint Conference on Artificial Intelligence, 2006, pp. 1015−1021.

[24] M.A. J.A.m. Musen, The Protégé Project: a Look Back and a Look Forward vol. 4, 2015, pp. 4−12.

[25] S. Lohmann, V. Link, E. Marbach, S. Negru, WebVOWL: web-based visualization of ontologies, Int. Conf. Knowl. Eng. Knowl. Manag. 8982 (2015) 154−158.

[26] B. Parsia, N. Matentzoglu, R.S. Gonçalves, B. Glimm, A. Steigmiller, The OWL reasoner evaluation (ORE) 2015 competition report, J Autom. Reas. 59 (2017) 455−482.

[27] A. Luque, A. Carrasco, A. Martín, A. de las Heras, The impact of class imbalance in classification performance metrics based on the binary confusion matrix, J. Pattern Recogn. 91 (2019) 216−231.

[28] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The WEKA Data Mining Software: an Update vol. 11, 2009, pp. 10−18, https://doi.org/10.1145/1656274.1656278.

[29] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A survey of methods for explaining black box models, ACM Comput. Surv. 51 (2018) 1−42.