ISSN 2278-3091
Volume 9, No.4, July – August 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse299942020.pdf
https://doi.org/10.30534/ijatcse/2020/299942020

Identification of Fake Identities on Social Media using various Machine Learning Algorithm

Bharat S. Borkar[1], Dr. Manish Sharma[2]
[1]Department of Computer Science & Engineering, Gyan Vihar School of Engineering & Technology, Suresh Gyan Vihar University, Jagatpura, Jaipur, India, borkarbharat.sgvu@gmail.com
[2]Department of Computer Science & Engineering, Gyan Vihar School of Engineering & Technology, Suresh Gyan Vihar University, Jagatpura, Jaipur, India, manish.sharma@mygyanvihar.com

ABSTRACT
In the current scenario, online social media platforms are the technology's most common and fastest tools for information exchange. The majority of people from all backgrounds expand their time on social networking platforms. An enormous amount of information is developed and shared worldwide through social networks. Such motives have contributed to unauthorized participants engaged in malicious acts against members of the social platform. False account formation is seen on social media as doing more damage than in any other form of Cybercrime. This offense must be identified well before the consumer is told about both the development of the fake identity. Numerous algorithms and approaches have been suggested for the identification of false identities, most of which use the vast amounts of raw data produced by social platforms. In this research, we proposed fake identity detection of social accounts on twitter dataset. Various machine learning algorithms have been used to evaluate the proposed results using NLP techniques. SVM, Fuzzy Random Forest, and Naïve Bayes have used for classification. The experimental analysis shows the effectiveness of the system and how it produces better accuracy than other machine learning algorithms as well as existing systems.

Key words : Machine Learning, Naïve Bayes, Fuzzy Logic, Random Forest, twitter dataset, fake identity, bots, Natural Language processing, classification.

1.INTRODUCTION

The social media platforms are now gradually the domain of our lives and include different facets of the day-to-day societal activity. The demands of people perform certain social positions in such these as media prefer what they do in actual life. Behaviors, perceptions, acts, and habits in these systems, individuals hold (Social culture in Social Media) When the value of such media rises, this concept is more important to research and to test Comprise. The paper pays intellectual attention to the issue by focusing on regional cultural aspects Google app discrepancies, among the most common social media. Social networking is a level higher quickly these days, which is vital for ad strategies and celebrities that try to boost it by increasing their number of followers and fans. Nevertheless, false accounts, produced apparently on behalf of organizations or individuals, can destroy their credibility and decreasing their numbers of friends and comments. They are still suffering from bogus notifications and excessive ambiguity with others. Fake profiles of all kinds generate negative effects that counteract the potential of social media in marketing and promotions for companies and lay the groundwork for online harassment. In an online world, consumers have specific questions concerning their privacy.

There are a few social media sites which include Twitter, Google+, Youtube, Instagram, Flickr, Facebook, and snap-chat. There were 823 million individuals who used social media on their smartphones every day that is an improvement over the previous fiscal quarter of 654 million these consumers. Social networking sites such as Facebook cannot yet offer real-time updates to false accounts, so for semi-technically advanced consumers it is impossible to differentiate between true so false accounts, In fact, other big data problems, namely data collection, how to manage data streams, and how to deliver instantaneous user replies, have to be addressed when running on vast quantities of data concurrently to obtain reliable profile recognition performance. Earlier work on fake accounts tackles experimentation that evaluates preventive measures against fake user behavior patterns. A Facebook social manipulation research project using the google maps API analysis quantifiable information about the number of men and women friends, data access documents, clustering algorithms of mutual acquaintances, details about education and work, location Facebook users knowledge, and mutual interests. The security measures to defend users from attackers include knowledge of data, privacy laws, techniques to enhance safety, and awareness-raising training.

Such work will also tackle the challenge of the rising amount of data of fast pace and range. The below figure 1 shows the activity of common procedures in the detection of real as well as fake profiles. The big information collected by social networks can appropriately be used To separate fraudulent from true profiles, from instead propose inquiries only from actual profiles. This is vital To non -specialist consumers, teenagers, and kids who do not recognize the protection settings. The research should be geared to helping User to differentiate between real and fake profiles in real-time, and guide the others about whether profiles should be reported And advise their partners to be related.

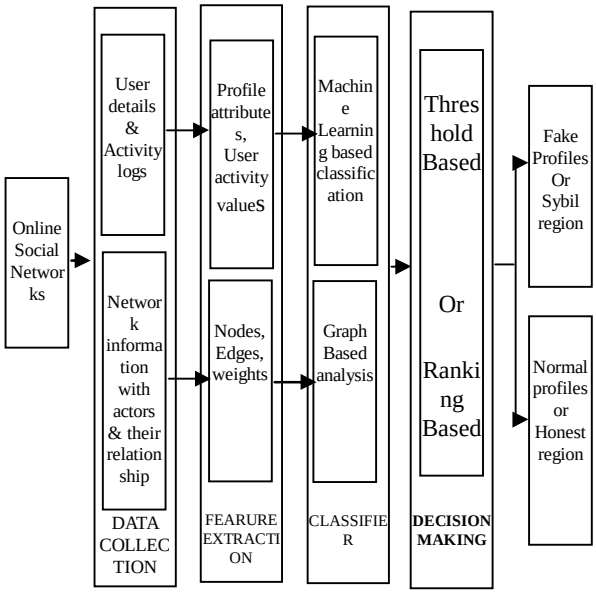

Figure 1 : General process flow in detection of the real or fake profiles on social media.

2.LITERATURE SURVEY

Drouin, Michelle, et al. [1] this system suggested initially focuses on each public web activity. Individual people experience a great amount of insincerity through a variety of various online platforms (i.e. social networking sites, online social websites, private chat rooms, and dating websites), claim people are less truthful than they should be at any of these online websites, and perceive specific forms of lies through multiple sites. In comparison, expectations of the dishonest actions of others were more critical than specific attributes in forecasting integrity through contexts, this system eliminates all those problems using machine learning algorithms and removes such identities.

L. M. Jupe et. al. [2] Potential explanations of identity manipulation, approach to verification, annexation manipulation. This method studied participants who were required in three situations to lie and say the truth regarding their invasion: choice lie, coerced lie, and fact. Recordings were evaluated using the Verifiability Technique, which defines the information that should be verified in the comments of the defendant. Even though it was speculated that liars would provide fewer provable details in their verbal statements than tellers of truth, the formal verification Approach could not differentiate between factual and manipulative declarations.

Y. Li, O. et. al. [3] Proposed system developments in two aspects: computationally, machine evolutionary techniques are implemented in a completely hierarchical way that is easy to compute and conveniently parallel; in fact, the method architecture is entirely generally applicable and can be applied to many behavioral cluster-based problems and implementations without too much modification.

T. Tuna et. al. [4] system analyses commonalities and dissimilarities among spammers and hackers, then discusses the faster access and/or context attention to detect hackers for better outcomes. Start investigating boundaries-spanning creeping (one person might be a long time reader in a friend's circle because of the variation of boundary weaknesses while becoming quite involved in a specific circle), build connection analysis and website crawling for twitter posts that will help simplify spam search and improve accuracy, and recognize the actual people who construct bot identities.

P. Galan-Garcia et. al. [5] carried out a current practice users identity anonymity issue creates scenarios, replicated regularly, in which individuals with false identities, or at least not relevant to their own identity, post articles, comments or multimedia content trying to ridicule or target certain persons who may not have been aware of the assault. Such acts may have a huge effect on the world of the perpetrators, producing circumstances where cyber assaults in real-life intensify into devastating consequences. In this present system for detecting and associating fake Twitter social media network profiles that are used for untruthful operations to real characteristics even in the same network by analyzing the text of the comment threads produced by both profiles. The methodology under guidance has been used to identify these profiles.

K. et. al. [6] Self-presentation Analyses of major five aspects of the element in psychopathology. The program describes the two methods below (1) misrepresenting yourself to mislead someone online and (2) looking for positive and enduring online relationships. While pursuing this aim, the program also attempted to build scales to quantify such structures, when currently no such measurements are usable. First, move in learning how to utilize human variations in temperament and

psychopathology to anticipate trickery and communication online. The performance of the suggested method would be roughly 85-88%, with minimal negative result ratio.

Thi Thu Thoung Le et. al. [7] the power disaggregation or Carried out to verify the system proposed is the perfect solution for reducing our power consumption. A number of algorithms are related to this area of machine learning. The classification results from all those algorithms are not as strong as planned, however. Throughout this article, we suggest a new approach to creating an energy segmentation classifier with a profound learning area. We use the Recurrent Neural Network (RNN) based Gated Recurrent Unit (GRU) to train the model using the UK DALE dataset on this area. Furthermore, we equate our method to the required energy propagation RNN.

Yong Zhang et. al. [8] proposed a design named Comprehensive Attention with Recurrent Neural Networks (CA-RNN) that can hold previous, successive, and local aspects of any sequential location is created. The Bidirectional Recurrent Neural Networks (BRNN) is used to view knowledge from the possible futures while a convolutional layer is being used to collect location features. The standard RNN is also composed of two new emerging variations of RNN, which include LSTM and gated recurrent unit, to significantly improve the new architectural design's efficacy. Another salient design of the new model is something without any human involvement, this can be prepared end-to-end. It is incredibly simple to incorporate.

Peddintiet. al. [9] designed a classifier that transforms the four-class classification process into two binary classifications: the one classifying each identity as anonymously or especially non-anonymous and the other classifying each account as recognizable or quasi-identifiable. To identify every account as 'anonymous,' 'identifiable', or 'unknown' for Twitter info, the values of two classifiers are merged. Almost all binary classifiers use Random Forest (RF) as a classification algorithm, with 100 tree branches. The option of the classifier and the number of trees is dependent on the results of the validation set and the bag failure. These classifiers are also cost-sensitive meta classifiers, where higher cost is imposed for misclassifying instances as anonymous or identifiable. The dataset used here was from Twitter.

Philogene Kyle Dimpas et. al. [10] Filipino as well as foreign languages click on bait identification using a recurrent long-term neural network with recurrent memory. In fact, this work has gathered Filipino and English Headlines and decides if it is a click trap. This used a neural network model based on a BiLSTM. The model uses Word2Vec to offer corpora word representation and embedding. The experimental findings revealed a precision of 91.5% using the standard.

Rajesh Purohit Bharat Sampatrao Borkar [11] promoted False Detection versus real Social Media identity utilizing Random Forest and Deep Convolutional Neural Network This method eliminates false identity by bots during classification and primarily focuses on the detection of fake identity by humans because a very little study has been performed so far on fake identity by humans. It searches with two separate algorithms for the description i.e. Random Forest (RF) and Neural Recurring Network.

B.Pandu Ranga Raju, B.Vijaya Lakshmi, C.V. Lakshmi Narayana [12] propose an efficient and versatile malicious URL detection system with a rich collection of features representing the diverse characteristics of phishing websites and their hosting platforms, including features that are difficult to forge. Using the Random Forests algorithm, the program benefits from both high detection capacity and low error levels. The results of the experiment show that the program can be used by the blacklist provider to create automated blacklists

Roobaea Alroobaea [13] To achieve the best possible classification, we adopt a variety of Machine Learning(ML) algorithms like 'Support Vector Machines(SVM)', 'Random Forest(RF)', 'Multilayer Perceptron's (MP)', 'Decision trees(DT)','Naïve Bayes(NB)', k-Nearest Neighbors(KNN) and Deep Learning technique' Long Short-Term Memory(LSTM)'. The results show that each data set (age or gender) have different results when machine learning techniques were applied. Deep learning techniques proved that they will be helpful when the data set is large.

3.RESEARCH METHODOLOGY

Social media websites have an influence on science, knowledge, managing the mass movement, job opportunities, business, etc. Researchers examined these social media to see what influence they are having on the people. Users can easily reach people by making it a friendly environment for students to study, teachers everywhere now get to know these sites, helping to bring online classroom pages, giving coursework, conducting conversations, etc. Most social businesses may utilize these social media networks to employ candidates who are skilled and interesting in the job, making it much easier to do long - term research. In this initiative, we aim to include a mechanism for the automatic classification of fake accounts in order to protect people's social identity and, leveraging this automatic recognition strategy, we can make it easier for the websites to handle the enormous amount of personalities that can be handled automatically. For

example, if we choose to evaluate false features based on their period, publishing date or tweets, language, and geo-position. And though these points rely on each and almost every variation or on the existence of the other factors, below we describe entire research follow to complete this research work including learning algorithms.

3.1 Tokenization

Tokenization is the mechanism of splitting dissemination of textual data into words, phrases, symbols, or numerous essential factors punctuation marks. The object of tokenization is to identify the phrases in an expression. The tokens sequence is data for any further analysis, which is analogous to sorting or sequential information mining. In both language studies and scientific research, tokenization is worthwhile as a part of text summarization. Linguistic knowledge at the start is a block of protagonists the simplest. All Know-how Extraction techniques include the terms of the data collection. For that function, a parser's prerequisite is record tokenization. This may be negligible speech since the text is already stored in understandable codecs for mobile devices. Nevertheless, several issues are still, such as adding exclamation points. Specific characters, such as brackets, hyphens and so on often need validation.

3.2 Stop Word Removal

Skip phrases are often widely employed than traditional expressions such as 'and,' 'are,' 'this,' etc. We may not seem to be important in document-classification. Nevertheless, they have to be removed. Furthermore, the construction of these recordings of stop phrases is complicated and contradictory among textual sources. This process also lowers the expertise of the text and enhances the quality of the approach. For these sentences, which are not critical for text analysis applications,-the textual content report provides.

3.3 Stemming and Lemmatization

The purpose of both stemmed and character segmentation is to scale down phrase forms of inflection and mostly derived variations to a crafted base form. Stemming usually refers to a simplistic heuristic process that cuts the ends of terms with the intention of reaching this aim more frequently than not and involves removing derived pronunciation quite often. Stemming is also corresponded to performing it efficiently through a language and anatomical study of words, in most circumstances interpretive at extracting intonation ends and contributing to the root or vocabulary form of a term that is also called technology

NLP Prepossessing Text prepossessing is an essential part of any NLP method and the significance of the NLP prepossessing follows above mention processes. Three different machine learning algorithms' have been used called Support Vector Machine (SVM), Fuzzy Random Forest (FRF), and Naïve Bayes (NB) algorithms that are used in this proposed framework.

3.4 Support Vector Machine (SVM)

An SVM classifies knowledge by discovering the extraordinary hyper-plane which distinguishes all aspects of 1 category knowledge from those of the other classification. In an SVM approach, the better hyper-plane is the one with the longest line between the two groups. An SVM classifies data by discovering the extraordinary hyper-plane which distinguishes all aspects of awareness in one group from those in the other. The support vectors are the types of data that are nearest to holding a hyper-plane separated.

3.5 Naive Bayes

Naive Bayes algorithm is the learning algorithm the probability of a particular crew/category contributing to an entity with designated features. In brief, the classifier is a deterministic one. The Naive Bayes method is applied "naive" since it allows the assumption that the frequency of a certain trait is irrespective of the incidence of certain circumstances. For example, if we choose to evaluate false features based on their period, publishing date or tweets, language, and Geo position. And though these points rely on each and almost every variation or on the existence of the other factors, in my opinion, all these characteristics add to the possibility that the incorrect profile will be eliminated automatically.

4.PROPOSED SYSTEM DESIGN

Various existing approaches have already defined to detect the malicious activity or bots defined in [5] and [6], but those systems still having various issues like false alarm rate and low classification accuracy. The below figure 2 shows the proposed system which works with various social network website dataset to detect the face account on the entire dataset.

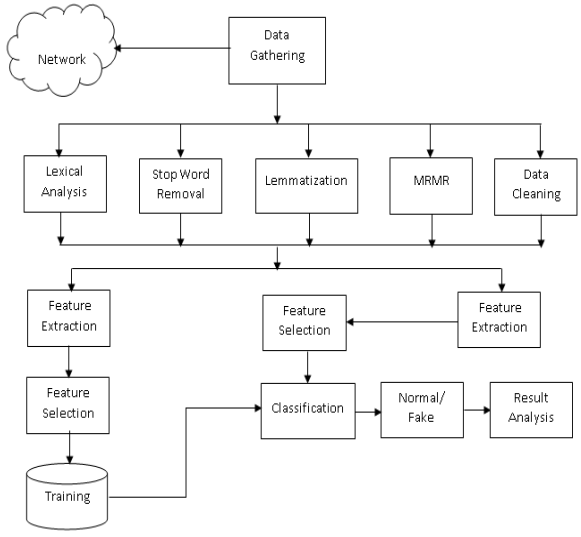

Figure 2: Proposed System Architecture Design

Initially, the system collects data from the Twitter account using Twitter API, which extracts the data from twit comments which is recently viewed by users. The main problem of social media applications is unable to detect bots or fake accounts. This research we carried out the combination of NLP and Machine learning algorithm to eliminate such issues in existing systems.

First, we gather the data from various social media sources once data has received it will be stored into the data repositories as well as data set files. The data has collected from various internet sources like Twitter, so it should be unstructured sometimes. It is mandatory to preprocess such a kind of data with a specific sampling technique as well as data filtration techniques. The systematic sampling technique has used to data separation and bloom filter has used to eliminate miss-classified instances. Electrical analysis should provide sentence detection as well as tokenization respectively, tokenize words has stored into to string array which provides hassle-free string checking. Additional NLP algorithms called stop word removal and lemmatization, this feature comes under natural language processing. After such a cleaning process system performs some feature extraction techniques TF-IDF, co-occurrence correlation is another technique used for this functionality to achieve. This election has been done using various quality thresholds and those features pass to the classification algorithm. A similar process has been followed for data training as well as testing respectively. We used three different machine learning algorithms like Fuzzy Random Forest, Naive Bayes, and support vector machine respectively.

5.ALGORITHM DESIGN

5.1 Training using updated NB

Input: Training dataset Train Data[], Various activation functions[], Threshold Th

Output: Extracted Features Feature_set[] for completed trained module.

Step 1: Set input block of data d[], activation function, epoch size,

Step 2 : Features.pkl ◄——ExtractFeatures(d[])

Step 3 : Feature_set[] ◄——optimized(Features.pkl)

Step 4 : Return Feature_set[]

5.2 Testing using updated NB

Input: Training dataset TestDBLits [], Train dataset TrainDBLits[] and Threshold Th.

Output: Resulset < class_name, Similarity_Weight > all set which weight is greater than Th.

Step 1: For each testing records as given below equation

$$testFeature(k)=\Sigma_{m=1}^{n}(featureSetA[i].....A[n]\leftarrow TestDBLits)$$

Step 2 : Create feature vector from testFeature(m) using below function.

$$ExtractedFeatureSet(x)[t....n]=\Sigma_{x=1}^{n}(t)\leftarrow(testFeature_{(k)})$$

ExtractedFeatureSetx[t] holds the extracted feature of each instance for testing dataset.

Step 3: For each train instances as using below function

$$trainFeature(l)=\Sigma_{m=1}^{n}(featureSetA[i].....A[n]\leftarrow TrainDBLits)$$

Step 4 : Generate new feature vector from trainFeature(m) using below function

$$ExtractedFeatureSet(y)[t....n]=\Sigma_{x=1}^{n}(t)\leftarrow(trainFeature_{(l)})$$

Extracted_FeatureSet_Y[t] holds the extracted feature of each instance for training dataset.

Step 5 : Now evaluate each test records with entire training dataset

$$weight=calcSim(FeatureSetx)\|\Sigma_{x=1}^{n}FeatureSety[y]$$

Step 6 : Return Weight

6.RESULTS AND DISCUSSION

The proposed implementation has done in Windows open-source environment, java Platform has used due to the availability of open source. The publicly available twitter API has used to extract the data from the Twitter web application. We create various data chunks to perform the system classification accuracy with three different machine learning algorithms. Three different kinds of cross-validation techniques have used for data splitting like 10 fold, 15 fold, and 20 fold separately.

6.1 Experiments using Naive Bayes

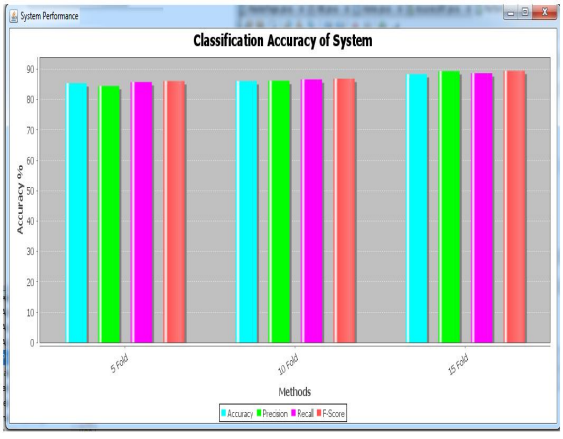

Figure 3 : Classification accuracy using Naive Bayes with different cross validation

The above figure 3 shows the classification accuracy of Naïve Bayes using the twitter dataset, the similar experiments have done with various cross-validation and results have illustrated in table 1. According to this analysis, we conclude 15 fold cross-validation provides highest 89.40% classification accuracy for Naïve Bayes.

Table 1 : Classification accuracy with confusion matrix for Naïve Bayes

(NB)	5-Fold	10-Fold	15-Fold
Accuracy	85.40	86.10	89.40
Precision	84.40	86.50	89.40
Recall	85.70	86.65	89.70
F-Score	86.50	86.90	89.50

6.2 Experiment using Fuzzy Random Forest

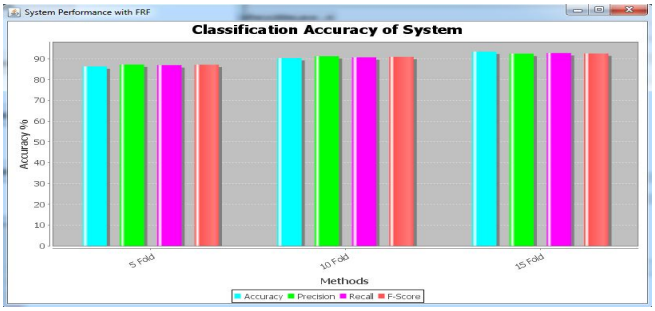

Figure 4: Classification accuracy using FRF with different cross validation

The above figure 4 shows the classification accuracy of FRF using the twitter dataset, the similar experiments have done with various cross-validations, and results have described in table 2. According to this analysis, we conclude 15 fold cross-validation provides highest 90.60% classification accuracy for FRF

Table 2 : Classification accuracy with confusion matrix for FRF

(FRF)	5-Fold	10-Fold	15-Fold
Accuracy	86.30	90.30	90.60
Precision	87.20	91.00	91.20
Recall	86.90	90.25	91.00
F-Score	87.10	90.80	91.60

6.3 Experiment using Support Vector Machine

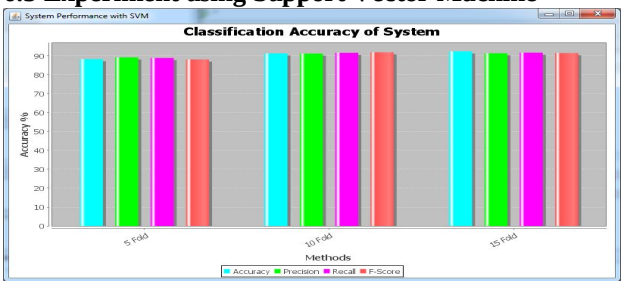

Figure 5 : Classification accuracy using SVM with different cross validation

The above figure 5 shows the classification accuracy of SVM using the twitter dataset, the similar experiments have done with various cross-validation, and results have illustrated in table 3. According to this analysis, we conclude 15 fold cross-validation provides highest 91.40% classification accuracy for SVM.

Table 3 : Classification accuracy with confusion matrix for SVM

(SVM)	5-Fold	10-Fold	15-Fold
Accuracy	88.80	91.10	92.40
Precision	89.00	91.15	91.40
Recall	89.30	91.20	91.70
F-Score	88.70	91.60	91.50

The all classification algorithm has used from the Weka tool, the 3.8 Weka environment has used for the utilization of machine learning algorithms. Eventually, collaborative effort and multiple platforms analysis of interconnected social media and the analysis of large amounts in the social sphere The present research reflects on the applications for the operation of various attacks and very few studies have concentrated on this. The science of data this is an evolving process, expected to be the next largest field of study during the next decade. The key social data is Big data channels over the Internet. Big data analytic and analysis helps diverse fields.

7.CONCLUSION

This research article describes the most prevalent common methods and reflects on the state of the works for identifying Sybil or false accounts and media networks. The different approaches are contrasted with their synthesized networking device and data set metrics. We have also concentrated on the strategies recently proposed and their positives and disadvantages. Comparison of those systems their analytical success is focused on. The accessible problems are recorded in the area of false profile identification in online social networks. We assume that, given several current systems, there is still no comprehensive solution for the identification of false profiles in social networks can ensure user information is recognized efficiently, efficiently, and simply. Using this experiment analysis we conclude SVM produces better results than NB as well as FRF in all cross-validations. In the future, machine efficiency can be enhanced by utilizing certain techniques such as deep learning with specific activation functions.

REFERENCES

1. Drouin, Michelle, Daniel Miller, Shaun MJ Wehle, and Elisa Hernandez. **Why do people lie online?"Because everyone lies on the internet,** Computers in Human Behavior 64 (2016): 134-142.

2. Jupe, Louise Marie, Aldert Vrij, Galit Nahari, Sharon Leal, and Samantha Ann Mann. **The lies we live: Using the verifiability approach to detect lying about occupation.,** Journal of Articles in Support of the Null Hypothesis 13, no. 1 (2016): 1-13.
3. Li, Yixuan, Oscar Martinez, Xing Chen, Yi Li, and John E. Hopcroft. **In a world that counts: Clustering and detecting fake social engagement at scale.,** In Proceedings of the 25th International Conference on World Wide Web, pp. 111-120. International World Wide Web Conferences Steering Committee, 2016.
4. Tuna, Tayfun, Esra Akbas, Ahmet Aksoy, Muhammed Abdullah Canbaz, Umit Karabiyik, Bilal Gonen, and Ramazan Aygun.,**User characterization for online social networks. ,**Social Network Analysis and Mining 6, no. 1 (2016): 104. https://doi.org/10.1007/s13278-016-0412-3
5. Galan-Garcia, Patxi, Jose Gaviria de la Puerta, Carlos Laorden Gomez, Igor Santos, and Pablo García Bringas. **Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying.** Logic Journal of the IGPL 24, no. 1 (2016): 42-53.
6. Stanton, Kasey, Stephanie Ellickson-Larew, and David Watson. **Development and validation of a measure of online deception and intimacy.,** Personality and Individual Differences 88 (2016): 187-196. https://doi.org/10.1016/j.paid.2015.09.015
7. Kim, Jihyun, and Howon Kim., **Classification performance using gated recurrent unit recurrent neural network on energy disaggregation.,** In 2016 International Conference on Machine Learning and Cybernetics (ICMLC), vol. 1, pp. 105-110. IEEE, 2016.
8. Zhang, Yong, Meng Joo Er, Rajasekar Venkatesan, Ning Wang, and Mahardhika Pratama. **Sentiment classification using comprehensive attention recurrent models.,** In 2016 International joint conference on neural networks (IJCNN), pp. 1562-1569. IEEE, 2016.
9. Peddinti, Sai Teja, Keith W. Ross, and Justin Cappos. **Mining Anonymity: Identifying Sensitive Accounts on Twitter.,** arXiv preprint arXiv:1702.00164 (2017).
10. Dimpas, Philogene Kyle, Royce Vincent Po, and Mary Jane Sabellano. **Filipino and english clickbait detection using a long short-term memory recurrent neural network.,** In 2017 International Conference on Asian Language Processing (IALP), pp. 276-280. IEEE, 2017. https://doi.org/10.1109/IALP.2017.8300597
11. Rajesh Purohit Bharat Sampatrao Borkar , **Identification of Fake vs. Real Identities on Social Media using Random Forest and Deep Convolutional Neural Network,** in International Journal of Engineering and Advanced Technology, Issue-1 7347-7351 IJEAT 2019 https://doi.org/10.35940/ijeat.A9739.109119
12. B.Pandu Ranga Raju, B.Vijaya Lakshmi, C.V.Lakshmi Narayana, **Detection of Multi-Class Website URLs Using Machine Learning Algorithms,** *International Journal of Advanced Trends in Computer Science and Engineering*, pp. 1704-1712 ,Volume 9, No.2, 2020. https://doi.org/10.30534/ijatcse/2020/122922020
13. Roobaea Alroobaea, **An Empirical combination of Machine Learning models to Enhance author profiling performance,** *International Journal of Advanced Trends in Computer Science and Engineering*, pp. 2130-2137,Volume 9, No.2, 2020. https://doi.org/10.30534/ijatcse/2020/187922020