A Generalized Proposed Approach to Identify the social media Fake Accounts using SVM-NN: Machine Learning Perspective

Anu Sharma[1, *], Dr. M.K Sharma[2], Dr. Rakesh Kr. Dwivedi[3]
[1]Research Scholar, Uttarakhand Technical University, Dehradun, Uttarakhand, India
[2]Professor, Amrapali Institute, Haldwani, India
[3]Professor, CCSIT, Teerthanker Mahaveer University, Moradabad (UP), India
*Corresponding Author: er.anusharma18@gmail.com

Abstract: The social network platforms Google+, Facebook, Twitter are surprisingly consideration in the most recent decade pertained to the combination of social behavior and computational systems. Leading organizations, industries, and promoters use fake social media accounts to promote services and social activities like marketing, political movements, and sale promotions. Recently, many fake accounts created on social media platforms. This is the main securities challenge for the users that attracted cybercriminals and attackers to carry malicious activities, steal personal data and information, share false news, and request money. Several techniques and algorithms were proposed to detect fake accounts and check users' behavioral, non-behavioral activities and characteristics on social media networks. The main contribution is machine learning, support vector machine (SVM), K- nearest neighbor (k-NN), decision tree, Neural Network, Artificial Intelligence, Convolutional neural network, Naive Bayes to detect fake accounts. This paper proposed a generalized approach to detect fake accounts using a SVM and NN. Our proposed system uses fewer features and steps to identify the account's identity (real or fake) compared to other methods with higher accuracy. To evaluate the results using our proposed approach, we used the Twitter dataset available on (S. Cresci et al., 2015) and found our approach outperformed.

Keywords: Data mining, social media data, online social networks, Machine learning, Neural Networks, fake profile detection, Fake account detection, Support vector machine.

Tob Regul Sci. ™ 2021;7(6-2): 63-77
DOI: doi.org/10.18001/TRS.7.6.2.8

INTRODUCTION

Nowadays, maximum activities are performed online through internet resources and social media platforms. It is one media that connects the globe with friends, groups, professionals, teaching, and e-learning used for sales and marketing, business, events, invitations. The main social media platforms are FACEBOOK, YOUTUBE, WHATSAPP, INSTAGRAM, LINKEDIN (Dean, B,

2020 & Simon Kemp, 2020 & Dave Chaffey, 2020). By the region social media growths in Asia +16.9%, Afrika +13.9%, America +6.9%, Australia +4.8% and Europe +4.3% in year 2020. 50.6% of the 7.7 billion people use social media age 13+ (Eligible Criteria), and more than 60% are active users. Out of 3.89 billion users, 99% access the Apps and Websites using mobile devices and others via desktops or laptops. Globally around 2 Hrs 24-minute time is spent on social media by a one-person and 5.7 years of their whole life. Around 2.7 billion monthly users are active on Facebook, 2 billion on YouTube, 1.3 billion on FB Messenger, and 1.2 billion on WeChat. Facebook noticed abusive words, false news, spreading fake information, adult nudity, terrorist propaganda, malicious activities, and fake accounts (CBC news, 2012). According to some reports, more than 15% of users are active monthly with fake accounts. In 2019, Facebook deleted more than 3 billion accounts, and Twitter has suspended over 7.5 billion fake accounts weekly. Google blocked and removed more than 2.7 billion bad advertisements. Table 1 shows social media's positive and negative impacts. Researchers focus on social media networks to identify fake accounts through various activities and features.

Table 1 Impact of Social Media Networks

Negative Impact	Positive Impacts
Social media's growing impact on lives	Education Help
Impact on Mental Health	Promotion and Marketing
Making more people insecure	NGO
Cyberbullying & Hacking	Govt help to fight with crimes
Addiction	Business model and reputation
Fraud and Scams	Building social relations
Reputation	Open, honest customer feedback
Cheating and Relationship Issues	
Glamorizes Drugs and Alcohol	

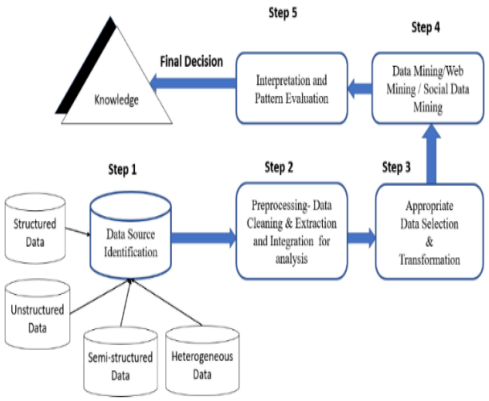

Figure 1 Data Mining Steps and Tasks

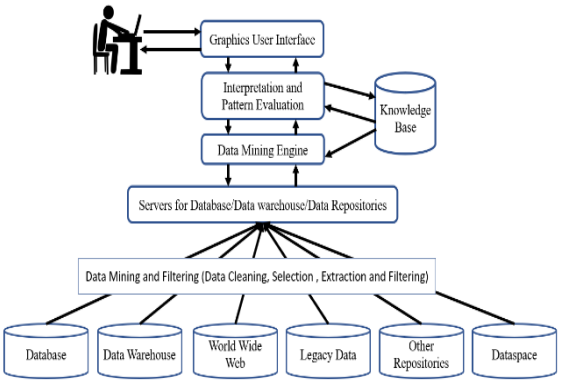

Figure 2 Data Mining System Architecture

Latest technologies like data mining and Machine learning (El-Tazi, Neamat et al., 2018), graph level structure (El-Tazi, Neamat et al., 2018), the Neural Network (NN) (Adikari, S. and Dutta, K. 2014), Support vector machine (SVM) (S. Cresci et al., 2015) are introduced to detect fake accounts on social media networks. The basic steps of data mining tasks performed are depicted in Figure 1 with architecture (see Figure 2) (Zhao, Yongqiang et al., 2020) and Data mining techniques (See figure 3) (Zhao, Yongqiang et al., 2020).

The rest of the paper is organized as follows. Section 2 covered the literature review, background, findings of various the articles and global challenges in the social media network and platforms. Section 3 presented the recent trends of social median network and Machine learning. Section 4 introduced the proposed algorithm and framework to detect fake accounts using the SVM-NN machine learning algorithms with result, and finally, Section 5 concludes our paper.

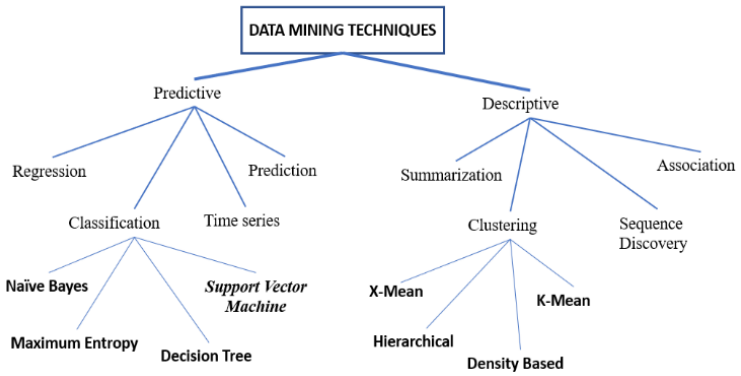

Figure 3 Modeling of Data Mining Techniques

LITERATURE REVIEW AND FINDINGS

Social media companies like Facebook, YouTube, Instagram, Twitter, etc. are trying to stop users to create fake accounts and false information, Government to deliver government services to citizens (Telangana Today, 2021 & Singh V., et al., 2021). We have studied many reports and research papers, which motivated us to highlight the issues of social media platforms and

integrating of latest technology trends using machine learning approaches. Research work done in social media networks using machine learning, and heterogeneous data mining and various approaches to detect fake account on social media platforms, algorithm used with finding from year 2014 to 2020 are summarized in Table 2 and Table 3.

Table 2 Existing research work, Technology used and findings

#	References	Description of the paper	Research points Details	Parameters / Features used	Algorithm / Technology	Findings
	(De S.S. and Dehuri S., 2014)	discussion on social media networks	Research on Users and Groups links	Analysis through graph theory on data	Graph Theory	Better results Graph
	(Anu Sharma et al., 2014)	Facebook	Sentiment analysis	Users' sentiments	Sentiment analysis Classifiers	NA
	(Ximing Wang et al., 2014)	Data Mining approaches	data points are uncertain	Description of Supervised algorithm	SVM Algo	NA
	(Mohammad Noor et al., 2016)	Unstructured data and systematic pattern mining	Data mining of social media network	Hidden patterns	data mining techniques	Data mining approach
	(B.Umadevi and P.Surya, 2017)	DM technique for Social Media	Large, noisy, and dynamic problem	Social media Characteristics	controlling the social media effect	Review Paper
	(S. Qamar et al., 2018)	Clustering, Mining	heterogeneous data mining	Data format and sources	K-mean and K-medoids	K-mean and K-medoids
	(N. Lal and N. Kaur, 2018)	Social Networking, and Big data	clustering for social media networking	Big data	Clustering using SparkR	SparkR for Big data analysis
	N. Lal et al., 2020)	Unstructured data	Dataspace	Clustering based of data	Vector space model (VSM)	New Framework using VSM

Table 3 Comparative study of previous research and findings for Fake Account Detection and Identification

#	Author's Name & References	Social media Platform	Data collection Sources	Parameters / Features used	Algorithm / Technology Used	Results Accuracy
	(B. Y et al., 2019)	Facebook	Not available	user and graph	Random walk approach	96%
	(B. Anand et al., 2019)	Twitter	scraping framework using python	Tweets avg. length, URL's retweets	EGSLA	90%
	(S.R., Gupta et al., 2019)	Twitter	Twitter API and Crawler	ID, creation date, tweets, followers,	Random forest and regression	97.1%
	(Aktaş et al., 2017)	Twitter	API	16 features collected own and using API	Naïve Bayes algorithm	91%
	(Kaushal et al., 2017)	Facebook	FB API	16 features	SVM Algorithm	79%
	(Kaushal, R et al., 2015)	Twitter	Existing data set collected	URLs,Followers, followers count, and Hashtags	Naïve Bayes, and Decision Tree Algo.	87.9%
	(Iqbal, M.A et al., 2017)	Twitter	Twitter's API used	Based on the user, graph, and contents	J48 and Naïve Bayes	98%
	(El-Tazi, N et al., 2018)	Twitter	Existing data set collected	16 numerical features	SVM and NN	NA
	(Wilson, C et al., 2014)	Renren	Not available	Normal and Sybils users	SVM, MLE	96%
.	(Dutta, K et al., 2014)	LinkedIn	Web and internet sources	11 plus features	SVM-NN	84%

In this paper, we have studied more than 80 papers, and many researchers have worked in data mining, the social media network, and the detection of fake accounts. Many challenges overcome by many researchers using mining and machine learning techniques by (Injadat, Mohammadnoor et al., 2016) but existing algorithms have not trained the machine properly. No generalized approach is proposed for social media networks that are somewhat challenging. Some global challenges are identified and specified in this paper.

Challenge 1-Single user handles multiple social accounts for malicious activities
Challenge 2-Making robust system to protect from attackers.
Challenge 3-Difficulties in features selection for the classification of accounts.
Challenge 4-Heterogeneity of sources, features, and data format.
Challenge 5- Preventing approach selection for fake account detection.
Challenge 6-Features integration problem with a social media platform.
Challenge 7-General approach that can be used with any social media platform.
Challenge 8-Real time determination of fake account request
Challenge 8- Technology selection for fake account detection

MACHINE LEARNING PERSPECTIVE TO DETECT FAKE ACCOUNTS ON SOCIAL MEDIA NETWORKS

In today's era, various trends and technologies are integrated to solve problems. Multiple approaches like data mining, machine learning approaches (K-NN,K-means, and SVM) and statistical methods extract hidden patterns and discover patterns (Mohammadnoor et al., 2016). The sociogram representation is shown in Figure 4 (De S.S., Dehuri S., 2014). Where points are people and lines are used to establish the relationship among the points. It has various features like colors, symbols, size, dotted lines: multi-lines, bold lines, and shades to specify different network properties (De S.S., Dehuri S., 2014). This section covered an overview of terminologies related to social media platform. ML algorithms and techniques

Machine Learning: Machine learning algorithms for social media fake accounts are used for prediction like Artificial Neural Networks (ANN), Convolution neural network (CNN) covered in (Singh V., Shanmugam R., Awasthi S. (2021). Naive Bayes Classifiers Algorithm (Granik, M., and Mesyura, V., 2017). Principal Component Analysis (PCA) (Joshi S. et al., 2020)., Neural network (NN) and Support vector machine (SVM) (El-Tazi, Neamat and Mokhtar, Hoda, 2018).

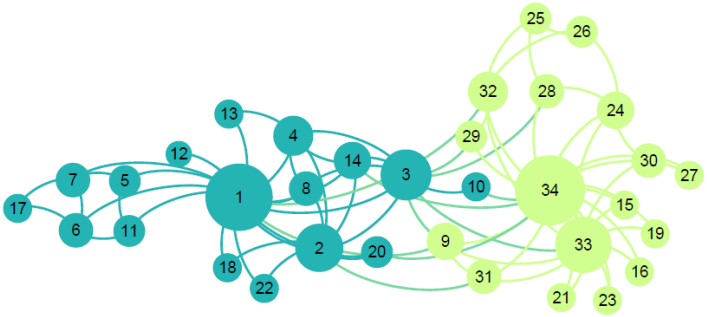

Figure 4 Sociogram representation of social networks

Graph representation and analysis: This method helps in various applications like to create circuit diagrams, connectivity, image processing and solution for social network analysis (Sebei, H., et al., 2018). A social network graph represented as G = (E, V), where V- Social network users, and E-edge (E⊂ N × N shows the relationships). Some graphs are 1) Sparse adjacency matrix, 2) Friendship graph (FG), 3) Common friends, (4) Total no of friends (TF), 5) Jaccard similarity (JS) given in equation 3, 6) Cosine similarity given in equation 4, and 7) L1 norm similarity (JS) given in equations.

$$\text{Jacc-coef (v.u)} = \frac{|FG(u).N \cap FG(v).N|}{|FG(u).N \cup FG(v).N|}$$

$$\text{(1) CosineSim(v.u)} = \frac{|FG(u).N \cap FG(v).N|}{\sqrt{|FG(u).N \cup FG(v).N|}} \quad (2)$$

$$\text{L1Norm(v.u)} = \frac{|FG(u).N \cap FG(v).N|}{|FG(u).N|. \ |FG(v).N|} \quad (3)$$

Feature-Based detection and Reduction: This approach is based on machine learning for the performance of prediction and classification. After feature extraction, classifiers are used to train data using clustering and machine learning technique and unwanted features may be decreased to reduce the redundant parts that create a better machine learning model.

Fake profiles and Fake account creation are more harmful than other cybercrimes on social media platforms. This crime needs to be detected before the users' notification and malicious activities, which is very significant. It is necessary to discover and stop these fake accounts creation, and even if an account is detected as a fake account, it would not be created. Detecting and identifying fake accounts is necessary to avoid malicious activities to protect personal data and secure user accounts (Hall, J. A. et al., 2021). Therefore, this paper focuses on detecting and identifying fake accounts on online social media platforms.

Proposed Methodology

Social media networks attract various malicious and illegal activities, and in the last few years, multiple solutions also have been proposed to detect fake accounts. To detect fake accounts on social media networks, we have proposed a general approach depicted in Figure 5. The main steps of our proposed process are summarized below:

1) Data Collection- Real time data collected from social media platforms needed. *2) Features selection and Extraction-* Social media account features are required like Account profile, date of creation, photo, age, followers count, etc. Initially, we chosen 23 elements of Twitter MIB dataset [84] mentioned in Table 4. We removed and reduced the unneeded features for better prediction mentioned in Table 5. *3) Technology selection (Machine learning)-* Machine learning SVM and NN approaches are selected, and finally *4)* Account decision (Fake or Real).

Table 4 Initially selected features

Features	Features	Features
User's_profile id	Geographical_location	Name_of_the_user
Default_Profile_status	Name_displayed_on_the_screen	Profile_image
Counts_of_status	Geo_location_status	Total_number_of_followers
Url_of_the_profile's_image	Total_Number_of_friends	URL_of_Profile_banner
Favourites_count	Background_image	Account_creation_date
URL_of_profile's_bg_image	Age_of_Account	Color_used_for_text
Account_URL	Profile_updated_status	Language
Color_used_in_profiles	Time_zone	

Table 5 Selected features after reduction

Features	Features
Counts_of_statuses	Total__followers
Total__friends	Favourites_count

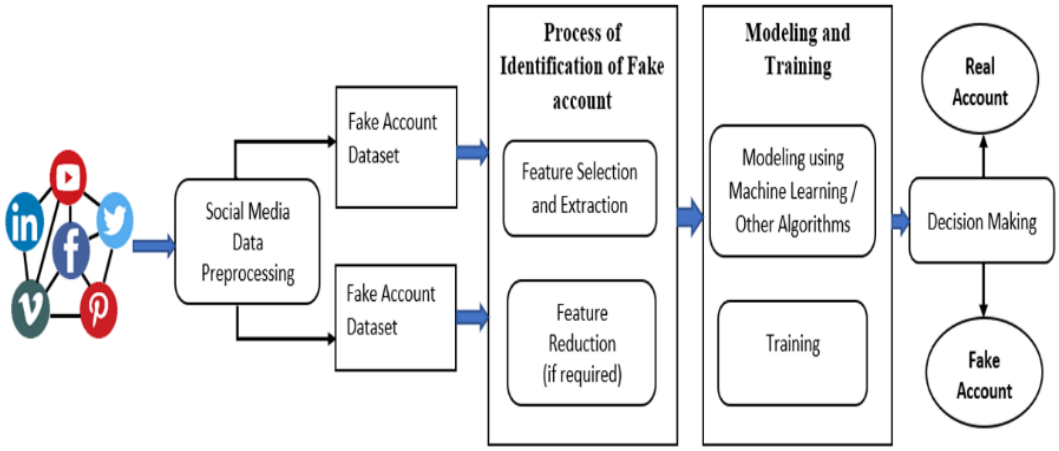

Figure 5 Proposed steps to detect fake accounts

Proposed Algorithm and Framework

In this paper, we presented a machine learning (SVM-NN, Naïve Bayes algorithm, PCA) & natural language processing system to detect fake accounts on social media networks. A proposed algorithm for accurate calculation and counterfeit account detection using SVM-NN is given in Algorithm 1, and the generalized proposed framework is shown in Figure 6.

Algorithm 1: Fake Account Detection using SVM-NN

1. Select the social media platform;
2. Data collection step;

3. Select the initial features;
4. Identified the select final selected features using PCA, Regression, and SVM techniques;
5. Create features subset *s*;
6. Split the collected data into a Training set *(fLable),* and Testing set *(vLable)* using 8 folds
7. **for each subset (*s*), do**
8. Train model using SVM algorithm and identify the *fLable*.
9. Store the previous step output into *decisions*
10. Now train the Neural Network model using *decision* and identify *fLables*
11. *Find the testing set output with the help of SVM and store it to variable testDecisionL*
12. Test Neural network using *testDecisionL* & NN and store output into *nnPredicted*
13. Find the neural prediction for each *s* using vLable and *nnPredicted*
14. Finally, calculate the average (avg.) accuracy foreach fold.
15. **end**

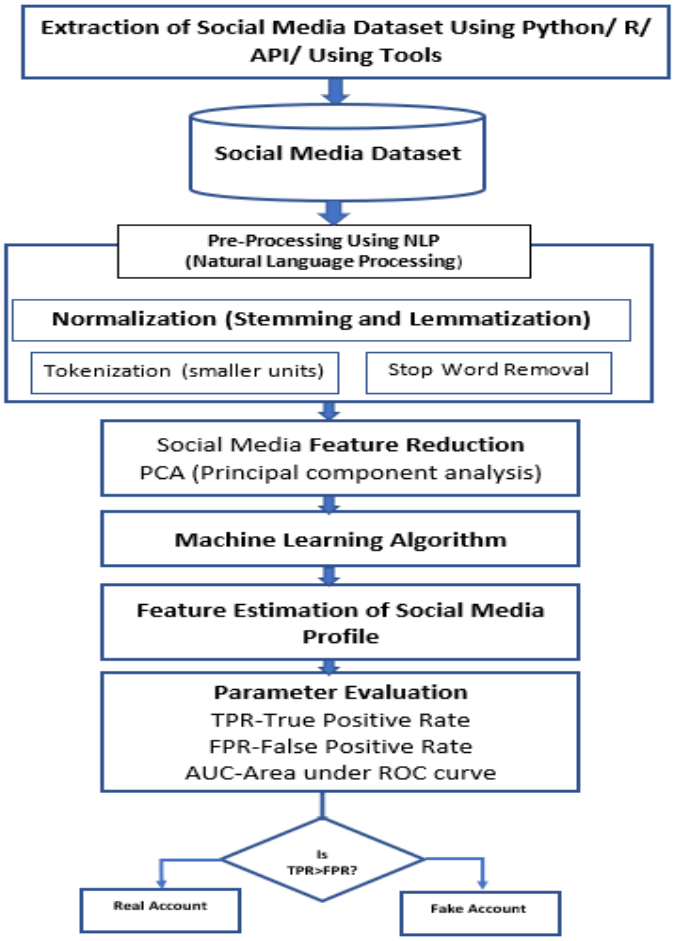

Figure 6 Proposed framework to detect fake account

Result analysis and discussion

Evaluation Metrics. We calculated the confusion matrix and other matrices that are associated with True/False rate (Positive/Negative) using 1) True Positive (TP)- Identified Actual node (Real Account) as Fake Nodes (Fake Accounts), 2) False Positive (FP)- Identified Normal node (Normal

accounts) as Fake Nodes, 3) True Negative (TN)- Identified normal node, Normal nodes and 4) False Positive (FN)- Identified Fake nodes as Normal nodes with true or false positive or negative rates. Below, matrices are used to calculate the ROC and performance;

$$TNR = TN/ (TN+FP) \quad (4)$$

$$FPR = FP/ (FP+TN) \quad (5)$$

$$TPR = TP/ (TP+FN) \quad (6)$$

$$FNR = FN/ (FN+TP) \quad (7)$$

We have selected the Twitter dataset MIB [84]. The description of the collected dataset is given in Table 6. As explained in the paper, the collected dataset is split into multiple datasets [84].

Table 6 Dataset description

Dataset	Detail	Accounts Count
TFP- The Fake Project dataset	100% Humans	469 users
E13- 2013 Election Dataset	100% Humans	1481 users
INT-Inter Twitter	100% Fake Followers	1337 users
FSF-Fast Followers	100% Fake Followers	1169 users
TWT- Titter Technology	100% Fake Followers	845 users

Initially, we divided the dataset 70% training and 30% testing using SVM-NN on python shown in Figure 7. Performance accuracy based on Feature subsets is shown in Figure 8, which shows our approach gives more than 96% accurate results to detect the fake account.

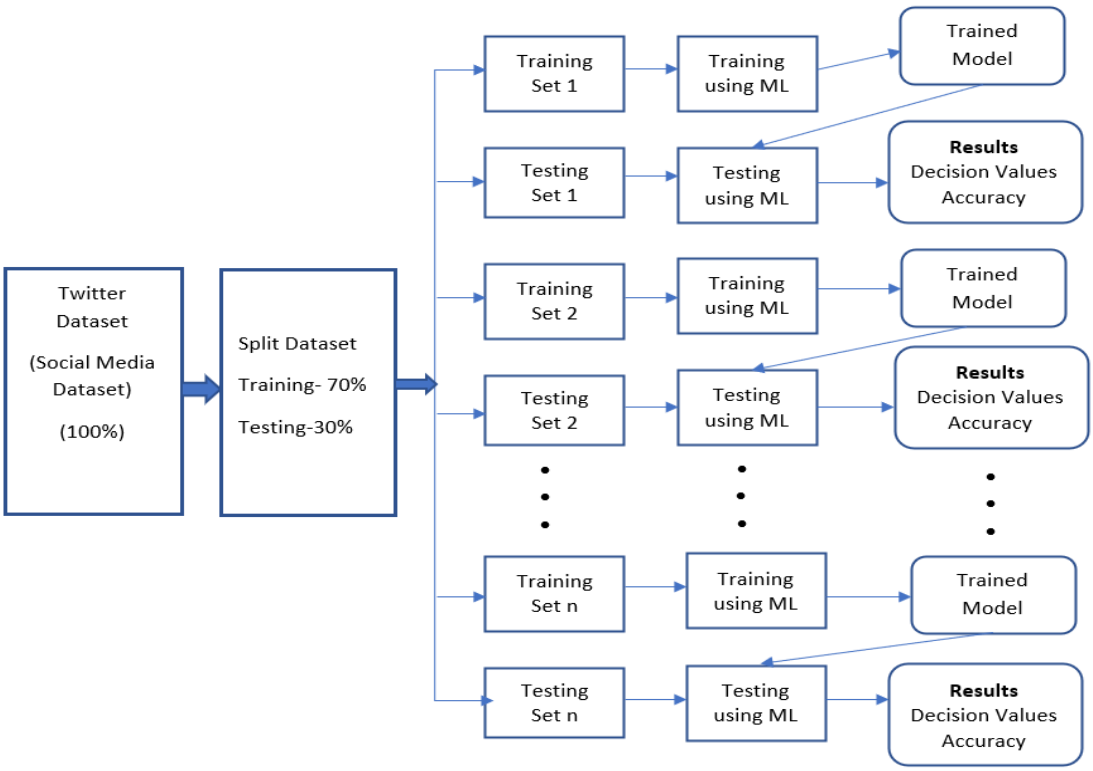

Figure 7 Training and testing of dataset

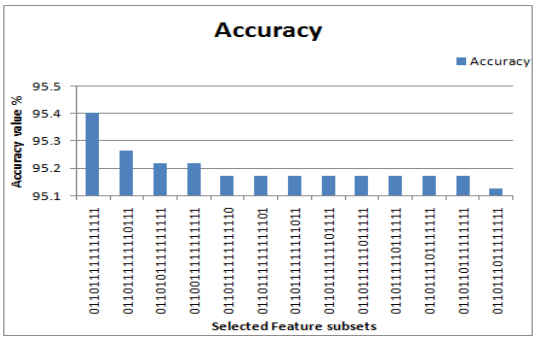

Figure 8 Performance accuracy

The below accuracy formula given in Equation 8 and 9 is used to find the accuracy of correctly or not correctly accounts and confusion matrix in Table 7.

$$\%\text{Accuracy} = \frac{Correctly\ detected\ and\ identify\ social\ media\ account}{Total\ number\ of\ social\ media\ accounts} * 100 \quad (8)$$

$$\text{Accuracy} = \frac{TruePostive\ (TP) + TrueNegative\ (TN)}{TP+TN+FN+FP} * 100 \quad (9)$$

Table 7 Confusion Matrix description

	Predicted Class	
	Not predicted	Predict
Belongs_to_this_Class	TP	FN
Not belongs This Class	FP	TN

All steps to calculate accuracy using our proposed algorithm (SVM-NN) are depicted in Figure 9.

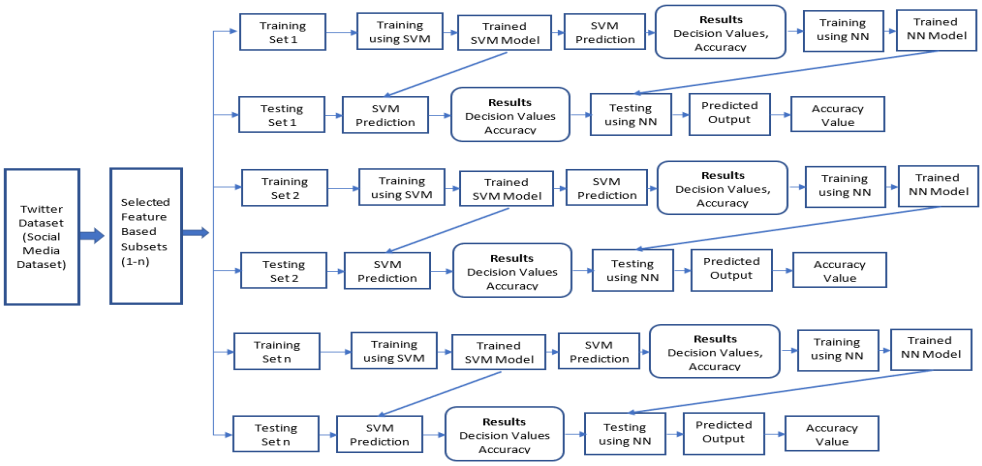

Figure 9 Accuracy Calculation using Proposed SVM-NN algorithm

Comparative results of SVM, NN, and SVM-NN with the Existing approach (Gianluca Stringhini et al., 2010 & Yang, C., et al., 2013). are given in Table 8, Figures 10, and accuracy results are illustrated in Figure 11. Our proposed SVM-NN method outperforms existing approaches covered by (Kruegel et al. **2010,** and (Gu et al., 2013)**,** depicted in Figure 10 and Figure 11.

Table 8 Result comparison

Selected Feature Set	SVM-Results			NN Results			Proposed approach SVNM-NN results		
	FP	FN	Accuracy	FP	FN	Accuracy	FP	FN	Accuracy
(G. Gu et al., 2013)	0.111	0.001	0.887	0.059	0.204	0.738	0.086	0.002	0.913
(Kruegel et al., 2010)	0.100	0.002	0.890	0.049	0.200	0.778	0.087	0.001	0.915
Correlation	0.037	0.047	0.924	0.076	0.097	0.823	0.014	0.003	0.982
PCA	0.038	0.046	0.915	0.652	0.277	0.923	0.032	0.044	0.924
Regression	0.016	0.034	0.946	0.040	0.070	0.889	0.028	0.010	0.960
SVM-Wrapper	0.038	0.004	0.956	0.052	0.114	0.833	0.028	0.006	0.966

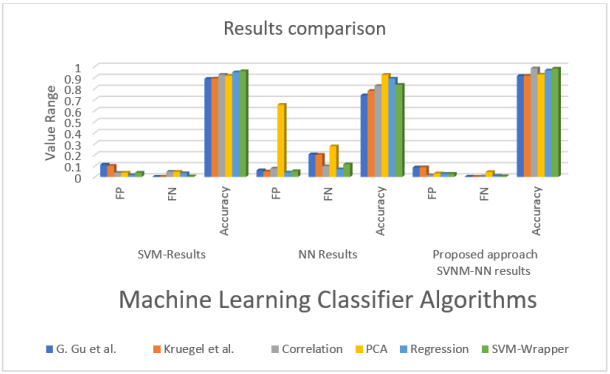

Figure 10 Machine learning Classifier algorithms

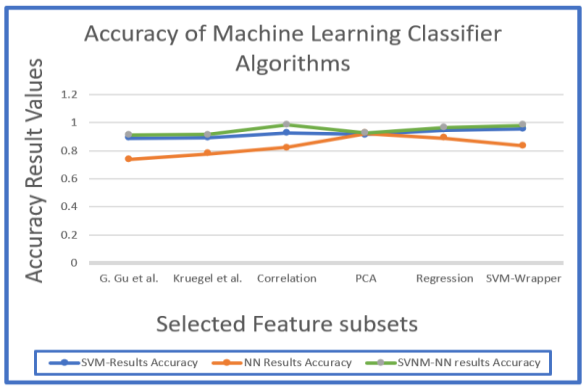

Figure 11 Accuracy of Classifier Algorithms

CONCLUSION

In this paper, we have discussed the data mining techniques, social media network statistics, and impact of these social media on human life, and insight into existing machine learning algorithms for detecting fake accounts. This paper proposed new generalized algorithms and approaches to detect the account's identity (real or fake) based on some features. Our proposed method is compared with existing Naïve Bayes, Neural network, PCA, regression algorithm to detect the fake accounts. We have used the "MIB" dataset (Cresci, S., et al., 2015). Our proposed (SVM-NN) approach outperforms existing machine learning algorithms, depicted in Figure 10 and Figure 11, which gives more than 96% accuracy to detect fake accounts of social media platform (Facebook, Twitter, and YouTube). Our proposed model also will help new researchers, companies, industries, business communities, practitioners, new integrated application designers, and the global community solve the new research problem and reduce the design failure rate of 80% by large through social media mining and networks.

REFERENCES

1. **Hall, J. A., Mihailova, T., & Morella, 2021.** "Relationships and Technology Lab, College of Liberal Arts & Sciences:" https://randtlab.ku.edu/jeffrey-hall. [accessed Jan 18, 2021].
2. **Summer Allen, 2019.** "Social media's growing impact on our lives" https://www.apa.org/members/content/social-media-research-series. [accessed Jan 18, 2021].
3. **Zhao, Yongqiang et al., 2020.** "IEEE Access Special Section Editorial: Advanced Data Mining Methods for Social Computing" IEEE Access. 8. 228598-228604. 10.1109/ACCESS.2020.3043060.
4. **Dean, B , 2020.** Social Network Usage & Growth Statistics: How Many People Use social media in 2020? [Online] Available: https://backlinko.com/social-media-users [Jan 23, 2021]
5. **Simon Kemp, 2020.** GLOBAL SOCIAL MEDIA OVERVIEW [OCT 2020] [Online] Available: https://datareportal.com/social-media-users [Jan 23, 2021].
6. **Dave Chaffey, 2020.** Global social media research summary August 2020 [Online] Available: https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/ [Jan 23, 2021]
7. **Statista , 2018.** Political advertising spending on facebook between 2014 and 2018. Internet draft. [Online]. Available: https://www.statista.com/statistics/891327/
8. **CBC news, 2012.** "Facebook shares drop on news of fake accounts. Available: http://www.cbc.ca/news/technology/facebook-shares-drop-onnews-of-fake-accounts-1.1177067
9. **Investor 2018.** Quarterly earning reports. Internet draft. Available: https://investor.fb.com/home/default.aspx
10. **Statista, 2018.** Twitter: number of monthly active users 2010-18. Available: https://www.statista.com/statistics/282087/number-of-monthlyactive-twitter-users/

11. **El-Tazi, Neamat & Mokhtar, Hoda., 2018.** Detecting Fake Accounts on Social Media. 3672-3681. 10.1109/BigData.2018.8621913.
12. **S. Cresci et. al., 2015.** "Fame for sale: efficient detection of fake twitter followers," Decision Support Systems, vol. 80, pp. 56–71.
13. **Injadat, Mohammadnoor & Salo, Fadi & Nassif, Ali., 2016.** Data Mining Techniques in Social Media: A Survey. Neurocomputing. 214. 10.1016/j.neucom.2016.06.045.
14. **Lal, N., Singh, M., Pandey, S., & Solanki, A.K., 2020.** A Proposed Ranked Clustering Approach for Unstructured Data from Dataspace using VSM. 2020 20th International Conference on Computational Science and Its Applications (ICCSA), 80-86.
15. **Kaur, N., & Lal, N., 2018.** Clustering of Social Networking Data Using SparkR in Big Data.
16. **Sebei, H., Taieb, M.A., & Benaouicha, M., 2018.** Review of social media analytics process and Big Data pipeline. *Social Network Analysis and Mining, 8*, 1-28.
17. **Telangana Today, 2021**. "Detect fake profiles on social media networks" Available: https://telanganatoday.com/detect-fake-profiles-on-social-media-networks. [Accessed 14 Feb 2021].
18. **Singh V., Shanmugam R., Awasthi S., 2021.** Preventing Fake Accounts on social media Using Face Recognition Based on Convolutional Neural Network. Karuppusamy P., Perikos I., Shi F., Nguyen T.N. (eds) Sustainable Communication Networks and Application. Lecture Notes on Data Engineering and Communications Technologies, vol 55. Springer, Singapore. https://doi.org/10.1007/978-981-15-8677-4_19
19. **Granik, M., & Mesyura, V., 2017.** Fake news detection using naive Bayes classifier. *2017 IEEE 1st Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, 900-903.
20. **De S.S., Dehuri S., 2014.** Machine Learning for Auspicious Social Network Mining. In: Panda Social Networking. Intelligent Systems Reference Library, vol 65. Springer, Cham. https://doi.org/10.1007/978-3-319-05164-2_3.
21. **Anu Sharma Dr M.K Sharma, 2014.** "CRITICAL REVIEW OF SENTIMENT ANALYSIS TECHNIQUES", Proceeding of the ICAICS.
22. **Wang, Ximing & Pardalos, Panos, 2014.** A Survey of Support Vector Machines with Uncertainties. Annals of Data Science. 1. 293-309. 10.1007/s40745-014-0022-8.
23. **Gianluca Stringhini, et al., 2010**. Detecting spammers on social networks. In Proceedings of the 26th Annual Computer Security Applications Conference (ACSAC '10). Association for Computing Machinery, NY, USA, 1–9. DOI: https://doi.org/10.1145/1920261.1920263.
24. **Dr B.Umadevi & P.Surya, 2017.** "A Review on Various Data Mining Techniques in Social Media", in International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 4, PP: 8082-8086.
25. **Muthu, Bala Anand et al., 2019.** An enhanced graph-based semi-supervised learning algorithm to detect fake users on Twitter. The Journal of Supercomputing. 75. 10.1007/s11227-019-02948-w.

26. **Erşahin, B., Aktaş, Ö., Kilinç, D., & Akyol, C., 2017.** Twitter fake account detection. *2017 International Conference on Computer Science and Engineering (UBMK)*, 388-392.
27. **Gupta, A., & Kaushal, R., 2017.** Towards detecting fake user accounts in facebook. *2017 ISEA Asia Security and Privacy (ISEASP)*, 1-6.
28. **Mateen, M., Iqbal, M.A., Aleem, M., & Islam, M.A., 2017.** A hybrid approach for spam detection for Twitter. *2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, 466-471.
29. **Gupta, A., & Kaushal, R., 2015.** Improving spam detection in Online Social Networks. *2015 International Conference on Cognitive Computing and Information Processing(CCIP)*, 1-6.
30. **Sahoo, S.R., & Gupta, B.B., 2019.** Hybrid approach for detection of malicious profiles in twitter. *Comput. Electr. Eng., 76*, 65-81.
31. **Yang, C., Harkreader, R.C., & Gu, G., 2013.** Empirical Evaluation and New Design for Fighting Evolving Twitter Spammers. *IEEE Transactions on Information Forensics and Security, 8*, 1280-1293.
32. **Adikari, S., Dutta, K, 2014.** Identifying fake profiles in Linkedin. In: PACIS, p. 278.
33. **Joshi S., Nagariya H.G., Dhanotiya N., Jain S., 2020.** Identifying Fake Profile in Online Social Network: An Overview and Survey. In: Bhattacharjee A., Borgohain S., Soni B., Verma G., Gao XZ. (eds) Machine Learning, Image Processing, Network Security and Data Sciences. MIND 2020. Communications in Computer and Information Science, vol 1240. Springer, Singapore. https://doi.org/10.1007/978-981-15-6315-7_2
34. **Sehgal D., Agarwal A.K., 2018.** Real-time Sentiment Analysis of Big Data Applications Using Twitter Data with Hadoop Framework. In: Pant M., Ray K., Sharma T., Rawat S., Bandyopadhyay A. (eds) Soft Computing: Theories and Applications. Advances in Intelligent Systems and Computing, vol 584. Springer, Singapore. https://doi.org/10.1007/978-981-10-5699-4_72
35. **Sehgal, Divya & Agarwal, Ambuj, 2016.** Sentiment analysis of big data applications using Twitter Data with the help of the HADOOP framework. 251-255. 10.1109/SYSMART.2016.7894530.