ISSN 2229-5518

Twitter Spam Detection Using Machine Learning

Asfa Falak, Dr. Hamid Ghous, Dr. Mubashir Malik

Abstract— Now a days, social media platforms have become an important part of our existence. The social media networks like Facebook, Instagram, Twitter, SanpChat and YouTube are used for communication among people and source of promoting businesses. Twitter is an excessive communication and sharing platform, where people can share their emotions and promote their businesses by using 140 character messages. More than 42 millions Twitter accounts are created every month. Twitter's receptiveness to spamming has prompted the prominence of activities on Twitter. Twitter spam is a very a sophisticated issue however it's difficult to unravel. So far, previous research has suggested a number of detection and defense methods that secure the Twitter users from spammers. So, we are going to work on spam detection techniques of Twitter. This study consists of 3 sections: 1- Background about spam detection on Twitter. 2- A literature review comparative analysis of machine learning, deep learning and hybrid algorithms. 3- Discussion on limitation of previous studies and future directions.

Index Terms— Social networks, Twitter Spam detection, Feature extraction, security, Data mining algorithms, Machine Learning, Deep learning, Hybrid models.

1 Introduction

In recent years, millions of Internet users have been able to communicate and collaborate on social media online networks (OSN) [4]. Today, we have entered the age of online social networks OSN [36]. Interest in this issue has been growing misinformation spread online on social media. Facebook, Twitter, and LinkedIn are the most prominent social media platforms that enable users to communicate with each other, use information and communicate in a meaningful way. Twitter is a great communication platform and sharing, it attracts profiles when provided services for spreading 140-character messages. Every month, the number of new accounts increasing more then 42 millions on Twitter [35]. Companies and individuals impressed the supremacy of quickly sharing information, it also performance as a smart power for the sender unsolicited and uncontroversial messages over the Twitter. This kind of data or messages are understood as spam data or messages. Though, due to the immense fame of Twitter, it also attracts the attention of cybercriminals (such as spammers).

Manual filtering of messages or data from Twitter is the starting point of spam detection, then there are some popular features that can detect a spam message with the help of modest filtering guidelines. Tranditional machine learning methods used for spam detection models and automated spam detection methods are also started with the utilization. For the spam filtering simple blacklisting, content-based and conversational spam detection techniques of data mining methods used. These type of methods done fairly on large data or emails messages, but, identifying the spam is being a big challenge from small and noisy spam detection platforms day by day. In short, from the domain of Twitter and SMS, it is more difficult to identify the cause of the spam, noise and small length of messages and emails. In this research, use different machine learning, deep learning techinques and compare their performance on larger datasets. Also, squeeze and compare performance as well as the number of features extracted.

- *Asfa Falak is presently tracking masters degree program in computer science in Institute Southern Punjab, Multan, Pakistan, PH- +92 306 7753720. E-mail: asfarana8@gmail.com*
- *Dr. Hamid Ghoush is presently employed as assistant professor at Institute of Southern Punjab, Multan, Pakistan, PH- +92 315 6098599. E-mail: hamidghous@isp.edu.pk (He did his PHD from University of Technology Sydney.He got more than ten years of research experience from oversease and Pakistan.*
- *Dr. Mubasher H. Malik is presently employed as assistant professor in the department of Computer Science at Institute of Southern Punjab, Multan, Pakistan. PH- +92 301 8630005. E-mail: mubasher@isp.edu.pk*

2 Background

Social Media platforms are digital-base innovation that encourgh the sharing of thoughts, considerations and data through the structure of virtual organizations and networks. By plan, Social media platforms are we based and provides customers brisk electronic correspondence of substance. Social media paltforms are an aggregate term for sites and presentation which center on correspondence, local area based information, communication, content-sharing and coopration. Social media platforms without a doubt has become a necessary piece of our every day lives. The workplace of interchanges and showcasing manages the fundamental like Instagram, Youtube, Facebook, Twitter and SnapChat accounts.

Span can be characterized as superfluous or spontaneous messages sent over the internet. Thses are normally sent to an enormous number of customers for an assortment of utilization cases, for example promoting, phishing, spreading malware and so forth.

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

Spammers the entire heart has gone to this stage and versatile organizations, brining about a multiplier increase in the measure of spam. From fruad accounts deluding posts, social media spam makes superfluous clamor that overwhelms real content and commitments.

Twitter is an American microblogging and person to person communication administration on which clients post and associate with masseges know as "tweets". Over the previous years, Twitter has pulled in an ever increasing number of cilents to post messages, turning into another style of web administrations for online correspondence and spread data. Starting at 2018, Twiter had more then 321 million month to month dynamic users. Twitter is very microblogging organization, given that most Twitter's post are composed by a minority of users. The fame of Twitter engage the spammers which have prompted to the increasing of spam. There is a number of misrepresentation or utilization of fraud accounts by spammers and advertisers. While I have describe social media spamming is undeniably more viable then traditional techniques like email spamming. Currently, fruad reviews and spam has expending and are turing into a major issue. As per to research, 15% of the Twitter users are robot and normally one out of 20 tweets is spam. Spam on Twitter effect both online social experience and cyberspace. In September 2014, the New Zealand Internet liquefied down because of the spread of malware download spam. Such spam tempted users to tab on URLs that professed to contain Holywood star photograph, yet actullay users were told to download malware to dispatch DDos assaults.The bellow tweet is an illustration of spam: "RT@Stormzy1: The clean hearted alwaysssss win in d end. U badmind lil weirdos wid u r bad energies are gonna destroy urselves trust", additional illustration, "Aft I finish my lunch then igo str down lor. Ard 3 smth lor. U finish ur lun`ch already?"

Several techniques have been suggested to combat spam. To automatically detect spam, researchers have implemented data mining algorithms to make spam detection a classification issue. Data mining has many types but in this research, spamming extraction from tweets by using deep learning and machine learning. Machine learning uses two main techniques: Supervised learning allows you to collect data or produce a data output from a previous ML deployment. Unsupervised learning find the hidden pattern or data grouping without the requirement for human intercession. Machine learning have further types used in spam detection research like CatBoost, GBM, KNN, K-Means, LightGBM,SVM, Random Forests, dimensionality reduction algorithem, Naïve Bayes, Descision Tree, XGBoost, linear regression, gradient boosting aldorithem, logistic regression. The function of artificial intelligence is deep learning that impersonate the function of human cerberm in handling information and making designs for use in making decision. It also known as deep neural network or deep neural leraning that utilized various layers of nonlinear handling units to remove features from information. Deep learning have further types used in spam detection research like Reccurrent Neural Networks, Generative Adversarial Networks, Deep Belief Networks, Multilayer Perceptron< Self-Organizing maps, Convolutional Neural Networks, Long Short Term Memory Networks and others.

In social media paltforms much research has been done demonstrating spam yet some work has never really on spam tweets. The framework of most techquies are equivelant. Mostly infrastructure use step by step procedure for spam detection on Twitter. First, Preprocessing techniques are applied on collected text or data. Second, after applying preprocessing techniques, extract several features and based on extracted features. At the final stage, for dectecting messages or post are spam or not applied machine learning and deep learning algorithems.

3 Literature Review

In previous research, data mining algorithms are applied on tweets dataset. We have examined previous work on the bases of machine learning, deep learning and hybrid algorithms. These algorithms comparison structure is below in Fig 3.1:

3.1 Structure

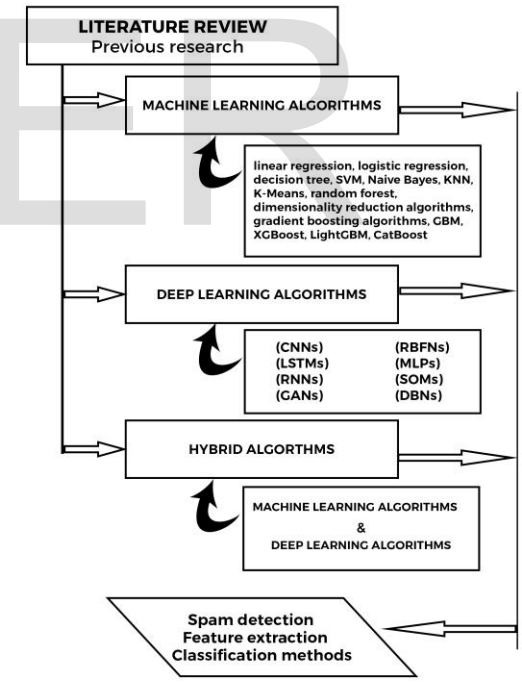

FIGURE 3.1

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

3.1 Machine Learning Algorithms

[1]Proposed included extraction steps and preprocessing methods for distinguished wheter tweets were spam or not spam. The feature extraction was ordered into different five distinct classes of account information based features, user profile based feature, user interaction based feature, and user activity based feature, tweet content based features and 28 different featuers included. Learning process through two polynomial kernels and gaussians a support vector used. At the final stage research method compare with Naïve Bayess, Random Forests, K-Nearest Neighbors and MultiLayer Perceptron methods. The acquire result shows the excellence of the research method by using polynomial kernels and SVM algorithems with .96 accuracy, .93 efficiency, .988 precision and F- .969.

[2] Suggested a better way to abolish misused technologies and search new ways to give results in progress. They proposed four modules: Data Evaluation that analyzes data, Pre-handling that handles the missing data in datasets, feature engineering that discounted the selection feature to machine learning algorithm and prediction module just tested the all processing step that applied on datasets not used for training. The given architecture just tells the way of detecting spam data. They did not implement any method on the explained module, they just suggested how to detect spam data.

[3] Presented the whole process as dependent on Learning and Classifying. It categorized the Twitter spam detection approaches and afterward sorted spam tweets as URL based spam detection and fake content based detection. Fake user based detection is also compared with methods depedent on a few features such as time features, content features, structure features and user features. The datasets about breast cancers cells that were collected from Twitter. Two classified modules applied on datasets that were SVM (Support Vector Machine) and Naive Bayes. Both comparison performance results were SVM Accuracy 83% and Naive Bayes Accuracy 92%. Hence, Naive Bayes Accuracy was higher than SVM.

[4] Introduce a new camping detection model that depends on vector-based qualities for sentence installing. The whole research depends on 3 basic steps: Firstly, to analyze the similarity of Twitter accounts in which posts or tweets are on the same topic. This similarity helps to build a graph. Second step, to classify campaigns, the graph was built on the basis of similar accounts. Third step, classifying the detecting tweets as spam campaigns. Ground-Truth Twitter dataset from Twitter obtained by using a real-life 3-day. Two-step semantic similarity function applied on datasets. The Sent2vec model is used for found similarity and manhattan lstm model is used for recalculating the similarity. These models provided the result of 58 candidate campaigns: A Precision was 0.945, A Recall was 0.93 and AF was 0.946. These models were compared with the U & T Based Model that provided the precision was 0.909, A Recall was 0.873 and AF was 0.89.

[5] Slove the issue of categorized news articles identified with disinformation and standard news by surveying dissemination contrivances on Twitter. Italian dataset was collected from US mainstream articles and disinformation articles, IT mainstream articles and disinformation articles. Multi-layer diffusion network global networks may be viably misused to recognize online disinformation. off-the-shelf classifiers for example, logistic regression on dataset relating to two diverse media scenes (US & Itely) produce exceptionally exact arrangement results (AUROC up to 94%) which are much better than our baseline with upgrades up to 20%.

[6] Applied 5 different feature extraction on 2 different datasets, the first dataset is collected from SHP and the other one is custom collected. Feature extractions are: account based features are used to collect outer information about accounts. Stylistic features are used to identify the symmetric variations of NL. Hashtag based features allow the user to apply tagging facilitates. Word embedding based features where words have the same meaning and representation. Topic word based feature used as important keywords. The proposed model has a total 4 steps: tweets extracted from different Twitter accounts, preprocessing techniques stop words and tokenization applied on extracted tweets, Feature extraction using LSA, LDA and glove applied on collected datasets and in the last step datasets is ready for train test splitting. For best results applied evaluation metrics and MLP recorded the highest accuracy 93%, 98% accuracy was observed for the SPD dataset and Classification probability 97 accounts correctly classified and only 3 misclassified.

[7] Proposed an Ontology-Based framework for criminal intention classification (OFCIC) framework for detection of spam and suspicious posts or tweets from Twitter. Ontology of Criminal Expressions (OntoCexp) presented for execution of above framework. This research had two parts: function and content. Function part is used in OFCIC for characterized the intention of the speaker and specify the illocution. OntoCexp used a content part which presented the meaning of the post. ML techniques are used to automatically illocution class to tweets posts. The best ML configurations presented F1-score around 0.5 and the result obtained 0.72 of general F1-Score by combining gloVe and ANN techniques.

[8] Used a Denstream known as density-based grouping technique for sorting floods of tests. Summarized the whole model into five main steps: 1) by arriving the primary window of data, two or three bunches are made by Euclidian distance, since no genuine microcluster has been made now. 2) By arriving the second part of data, if the general population of all the made microcluster outperformace "MinC" an INB classifier will be given out to it 3) for every microcluster whose populace surpasses "MinC," a full INB is prepared. 4) To try not to occupy the memory, the examples are killed aside from its markers like population, timestamps, and mean. 5) Updated to force an extremely low-computational cpmlexity to the proposed system in connection with standard DenStream. Randomly four types of employee datasets collected from Twitter. SimThersold esteems bigger than 0.8and lower than 0.5. The parameters are set within the range of [0.6–0.8]. The given methods gives the equivalent or more noteworthy outcomes then the DenStream.

[9] Proposed a hybrid approach for identifying the spam based profiles on the bases of similarity. Cluster approaches are used for selecting the initial spam accounts for classification purposes. Three classifiers were used in the proposed model: multilayer perceptron (MLP) used to solve the linear and nonlinear classification problems, support vector machine (SVM) analyzed the data and detect the pattern, Random Forest is the branch of decision tree and it works on tree structure. The datasets are collected from public figures and ground truth data. 100% of F-measure did'nt get from proposed model but it improved performance of the classifier with reduction in error rate.

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

[10] Various machine learning models and gathered datasets sampled by recommended features set. Its execution is reliable through the different models and there is gigantic enhancement above the pattern. The combination of high quality features and features learnt in an unsupervised way using word implanting appears to basically improve benchmark execution and to perform comparatively to the best performing feature set using more humble number of features.

[11] Used seven classification models for spam detection: Naive Bayesian (NB) required less computational time for training data and performance is good. K-Nearest Neighbor (k-NN) was used to classify new sample based similarity measures and store all available sample based. Decision Tree (DT) using less memory space than other classifiers for good results. Random Forest (RF) use because it is very simple and for regression and classification tasks it can be used. Logistic Regression is used for calculating the probabilities of events that are used. Support Vector Machine (SVM) performs grouping by conclusion the hyperplane that supports the lead concerning two classes. EXtreme Gradient Boosting (XGBoost) corrects the previous model and adds prediction. Twitter Social Honeypot Dataset is used because it is before categorized as spammers and authentic clients dependent on tweet content, user behavioral and topological featuers. Datasets are labelled as Y and Z.top 10 datasets are labeled as Y and top 7 datasets labeled as Z. For Y, XGBoost is 91% and RF is 92% that is the highest of F-score. For Z, RF highest of F-score is 94% and XGBoost lowest of F-score is 74%.

[12] Used Lfun (Learning from Unlabeled) for spam detection that has two components: Learning from Unlabeled Tweets (LDT) classified the tweets as spam or non-spam. Learning from Human Labelling (LHL) is used to label the data from human annotators and the number of unlabeled data. K-means cluster is used to delete old samples from training sets and Naive Bayes is used to improve the Lfun approaches. In terms of detection rate Naive Bayes give better results and F- measure is 2% and it reduces the processing time also.

[13]Proposed a framework that can detect spam and advanncment campings.This framework comprises of three fundamental advances: Firstly, detect URLs of accounts that have similar posts. Secondly, detect user camping that may be for spam or promotions. Utilized graph based methods for competitor campings detection. Finally, posting accounts links to similarity measurement. It can separate among advancements and spam campaings from ordinary ones dependent on SVM. They confirmed the feasibility of their framework on a geather datsets and anticipated results showed that this prposed technique removed the campaings at that point classified them into typical, advancement and spam classifications with high exactness. Sample feature strategy precision is 0.98. Its performance is not good rather than all features strategies. SVM classifiers got a higher precision and recall rates.

[14] Conducted a study to arrange messages with an end goal to recognize among ham and spam email by building a productive and delicate classification model with high accuracy and low false positive rate. Text preprocessing, tokenization and filtering of stop words that build the feature dictionary and documents of feature vectors. For better results, stop words removed and two end results predicted: false negative or false positive. False positive is the worst case because ham SMS goes into spam. SVC additionally delivered no bogus positives with less false positives contrasted with Naïve Bayes. At last, the utilization of a number of consolidated SVC classifiers in an Adaboost model produces a more adjusted result.

[15] Proposed a way to detect the spam on social media applications in real time. It devises the detection on tweet level by using a framework two types of module: real-time modworking with spam detection module and batch mod working with model update module. Four lightweight detectors classified by spam detection. Data is updated in batch and that's how it learns from patterns and detects the spam on tweet levels. The results achieved from experiments shows that confidently labeled clusters and provide good accuracy.

[16] Proposed a method to analyze the data for Twitter spamming in real time. They identify the spammers inside the Twitter traffic by using examinig gray box machine learning system and Random forests algorithems. To experiment their detection technique, they used two different sheets. One is from other researchers and the second one is built by them. They assigned the different benchmarks to assess the spamming. A non-uniform feature sampling method gets better effective predator rather than other conventional approaches.

[17] Proposed a scalable framework for both promotion campaigns and spam detection. The three steps of the process include linking those accounts who post URLs for the same purposes, extracting those campaigns of the candidates which might be used for spam, and differentiating their intent. Related large datasets from Twitter have been used for this purpose.

[18] Design application structure that permits designers to construct Twitter spam implement their own spam identification and classification application library articles. This proposed method execute two classification strategies: Naïve Bayes and K-Nearest Neighbor. The trending feature of Twitter was detect spamming. The Naïve Bayes and K-Nearest Neighbor classification startegies can detect spam and ham content with 82% and 71% precision.

[19] Suggested a way to deal with discover fake Twitter accounts. The intention to show optional consequences for the proposed approaches Naïve Bayes calculate position via social media measurement. Export the data by using entropy minization discretion with minimum description length for stop the quality. Some experiments have been done and has expanded the accuracy to Naïve Bayes from 85.55% to 90.41% using just their dataset discretization prodecure in chose features. Tha data damage from discrimination is so insignificant the increase in accuracy is great. It's a huge Increase using only numbers and a very promising result Social media data.

[20] For spamming activity detection, there are various techniques related to spam. Spam detection techniques used some models such as Binary detection, Lfun and ASLLfun among them the strategy is superior to other detection methods the rate was 4% and F-measure was reliable. Lfun Twitter spam wipes out the problem of drifting and the accuracy of this models is less then Binary detection model.

[21]Twitter is different from other social media platforms because it has some unique features that's why traditional methods of spam detection is not suitable for Twitter. Therefore, a spam detector framework designed specifically for Twitter this research suggests a TwitterSpamDetector. A framework of spam detection from Twitter a

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

designed known as TwitterSpamDetection and for rank spammer's legitimate users use Naïve Bayes classifiers that depend on Twitter's special features. On the report of diagnostic results, Twitter-SpamDetector's accuracy is 0.943 and sensitivity is 0.913.

[22] Develop spam profiles detection models. 82 Twitter profile dataset are collected for analyzing. Two methods Relief and Information Gain are used for feature selection. Classification algorithe,m were applied after feature extraction related to Multilayer Perceptron, Naïve Bayes, Decision Tree and K-Nearest Neighbor and compare there results. Naive Bayes achieve high evaluation rates compared to other classification algorithms. The results were provided by usining Naïve Bayes algorithem an accuracy, a precision and a recall rates were 95.7%, 94% and 96% respectively.

[23]A novel Asymmetric Self-Learning approach was proposed for spam detection of Twitter. In proposed ASL, the secheme was reinstructed by the additional modified "changed spam"tweets, this way to decrease the impact of "spam development" fundamentally. After the proposed approaches of ASL Apapplied and showed the experimental results that detection rate and F-measurement performance improved. For Bayes networks 10% improved F-measurement and for all tested three ML algorithems 20% improved detection rate.

[24] Evaluated the Twitter spam detection algorithm. They have done it by collecting over 600 million public tweets. For spam detection, they have removed 12 lightweight featuers and named around 6.5 million spam tweets. They used six machine learning algorithems that were used for expriments such as K-Nearest Neighbor, Decision Tree, Naïve Bayes, Support Vector Machine, Random Forests and Naïve Bayes Networks. Under different conditions to more comprehend their viability and soft spot for ideal Twitter spam detection.

[25] Presented the machine learning algorithm that has been developed to detect fake followers on Twitter. Firstly, collected a huge sample consisting the 13000 number of fake followers and 5386 number of real followers gathered and authorized all collected data manually. Secondly, for differentiating between real followers and fake followers identified several features. Thirdly, used identified features as an attribute of machine learning methods to categories as real or fake followers. Finally, using machine learning methods get high detection accuracy rate approximating SVM is 60.48%, Simple Logistic is 90.02 % and k-nearest neighbor is 98.74% and using others achieved low accuracy.

In Table 1 illuminate the before researches algorithms using deep learning algorithms.

TABLE 1
RESEARCH PAPERS USING MACHINE LEARNING METHODS

Author & Year	Preprocessing	Methodology	Dataset	Results	Future Work
Saleh Beyt Sheikh Ahmed, Mahnaz Rafie, Seyed Mojtaba Ghorabie (2021)	-User Profile Features -Account Information Features -User Activity Based Feature -User Interaction Based Feature -Tweet Content Base Feature	-Gussian -polynomial kernels -Multi-Layer precpetron -Naïve Bayes -Random Forest -K-Nearest Neighbors Methods	Real-world datasets	Precision = 0.988 Efficiency = 0.953 Accuracy = 0.96 F-measure = 0.969 ROC = 0.985	
Rumi Juwairiyyah, Nanditha Sriram, Jyotshna Bhushan Sharma, Babeetha (2020)	feature functions -Scikit learn preprocessing for missing data -EDA	-Review-behavioral model -User-behavioral model -Review-linguistic model	-Spam reviews	Betterment of mankind to integrate Various dimensions of prob-	

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

	-LDA -ANOVA	-User-linguistic model		lem.	
Deepali Prakash Sonawane Dr. Baisa L. Gunjal (2020)	-URL based spam -Labelling of spam -Feature extraction Shortlisting	-SVM -Naive Bayes algorithm	Tweets about breast cancer cells	Accuracy is Support Vector Machine = 83% Naïve Bayes = 92%	
M Mostafa, A Abdelwahab, H M Sayed (2020)	-Cluster Method - Ignore all special characters -Softmax output function by negative sampling -Word N-gram	-Semantic Similarity -Unsupervised model -LSTM -NB -Siamese Recurrent Network -Non-linear SVM algorithm with (RBF) kernel	Ground-Truth Twitter dataset	-58 candidate campaigns provides A Precision = 0.945 -A Recall =0.93 -AF = 0.946	add the Tweet timestamp Improve text similarity method to extract new strange words More dataset that is less bias ensemble learning to solve spam drift problem
Francesco Pierri, Carlo Piccardi, Stefano Ceri (2020)	-Multi-Layer diffusion network -Global network properties -Tuple of features topological features	-Off-The-Self Classifier - Logistic Regression -Disinformation -Mainstream networks	US mainstream articles & disinformation articles IT / Italian dataset	Get high accuracy up to 94%	Investigate three directions employ temporal networks extensive comparison of the diffusion of disinformation.
Ratul Chowdhury, ,Kumar Gourav Das Banani Saha, Samir Kumar Bandyopadhyay (2020)	-account-based -content-based -URL-based -graph-based feature -Feature extraction -lightweight features	-Novel hybrid learning framework -Machine learning -Deep learning algorithms -Analytical model based -Logistic regression	-social honey pod(SHP) -manually created dataset via tweedy API	MLP records the highest accuracy 93%. 98% accuracy SPD dataset Classification.	
Ricardo Resende de Mendonca, Daniel Felix de Barito,	-Semantic Web -Ontologies -Resource Descrip-	-MLA -SVM -RNN	employ ciphered posts 8,835,290	-ANN provided 0.54 accuracy, -SVM provided	

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

Ferrucio de Franco Rosa, Julio Cesar dos Reis, Rodrigo Bonacin (2020)	tion Framework -Ontology Web Language	-DTDNN -OFCIC	tweets 1000 Criminal Slang Expressions	0.56 Accuracy - Random Forest provided 0.52 accuracy.	
Hadi Tajalizadeh and Reza Boostani (2019)	-former methods -Naïve Bayes (INB) classifier -symmetric distribution	-CluStream -DenStream -State-of-the –art methods -StreamKM++	Randomly collected employed datasets	SimThreshold values < 0.8 and > 0.5. The parameters within the range of [0.6–0.8].	Current stream clustering methods will be used to enhance the performance of ral-time distance learning.
Isa Inuwa-Dutse, Mark Liptrott, Ioannis Korkontzelos (2018)	-feature extraction -(UPF) -(AIF) -(EwF)- (EbF)	-lexical richness -(TTR) -Mean Word Frequency	-Spam-posts detection -Honeypot dataset	significantly improve baseline performance	
Zulfikar Alom, Barbara Carminati, Elena Ferrari (2018)	-Graph- based features -Content-based features	- NB - KNN - DT - RF - LR - SVM - XGBoost	Twitter Social Honeypot dataset	-For Y, the highest F1 Score = 92% XGBoost =91% -For Z, the 91% = Highest F1. 74% = Lowest F1	- Will deal with altering the machine learning calculations -Apply technique to various Social media platforms.
Rutuja Katpatal Aparna Junnarkar (2018)	-Lfun scheme - (LDT) - (LHL) -Naive Bayes classifier -k-means clustering	machine learning based classifier	training dataset	In terms of detection rate Naive Bayes give better results and F-measure is 2%	This method will be applied on other social media sites.
Xianchao Zhang, Zhaoxing Li, Shaoping Zhu, Wenxin Liang	-URL-Based Method -Campaign Extraction -graph-based approach	-two-level classifier -Baseline Algorithms -U&T-based method -SVM	Benchmark Tweets2011 dataset	Sample feature strategy precision = 0.98. SVM classifi-	Will extend this framework for other social media networks.

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

(2016)				ers got a higher precision and recall rates.	
Olubodunde Stephen Agboola (2020)	-Naïve Bayes -SVM -spam Classification -AdaBoost	-machine learning (ML) -machine learning classifier	-SMS Spam Collection	Adaboost models produce a more balanced result.	deep learning model that can offer a better classification rate
Surendra Sedhai and Aixin Sun (2017)	-lightweight detectors -user-level spam detection -Tweet-level spam detection	-semi-supervised approach -fine-grained approach -unigram and bigram	HSpam14	spam detection good accuracy	
Claudia Meda, Edoardo Ragusa, Christian Gianoglio, Rodolfo Zunino, Augusto Ottaviano (2016)	-No uniform feature sampling method -uninformative content -Random Forests Algorithm	-gray box Machine Learning System	-Tweets Dataset of others -Customized Dataset	-a non-uniform feature sampling method gets better effective predator rather than other conventional approaches.	-Further experiments, more ineffective subsets are chosen, and will compare to the random forest algorithm.
Xianchao Zhang, Shaoping Zhu, Wenxin Liang (2012)	-linking accounts -extracting candidate -distinguishing	-URL-driven estimation method -Graph-based Approach -SVM Algorithems	Huge Number of datasets		-Will to improve similarity estimation method -Will add more highlights for the arrangement.
Aryo Pinandito, Rizal Setya Perdana, Mochamad Chandra Saputra, Hanifah Muslimah Az-zahra (2017)	-K-Nearest Neighbor -Naïve Bayes	-Android application framework -TFIDF weighting method	Training dataset	spam and ham contents 82% and 71% accuracy	
Buket Ersahin, Ozlem Aktas, Deniz Kilinc, Ceyhun Akyol	-numerical features -Naïve Bayes algorithm	-supervised discretization technique - (EMD) - (MDL)	Social media datasets	Naïve Bayes accuracy = 85.55% to 90.41%	Can be extended to detect a fake account in other social media platforms.

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

(2017)					
Rutuja Katpatal, Aparna Junnarkar (2018)	Twitter Spam Drift ML classifier	-Lfun Techniques -Asymmetric Self-Learning Techniques -Binary Detection Model	Tweets	Accuracy Lfun=83%. ASL =81%. Binary detection=92%	
Abdullah Talha Kabakus, Resul Kara (2019)	-Naïve Bayes classifier -Twitter4J16 -MongoDB18 -NoSQL	-TwitterSentiDetector -Cosine Similarity -clustering	77,033 tweets dataset	Twitter-SpamDetector's accuracy = 0.943 and sensitivity = 0.913	The system can be extended by consolidating delicate figuring strategies.
Ala'M. Al-zoubi, Ja'far Alqatawna, Hossam Fairs (2017)	-Feature engineering -Binary and Simple Features feature selection Methods -Relief -Information Gain	Classification algorithms: -Decision Tree -Multilayer Perceptron -K-Nearest Neighors -Naïve Bayes	82 Twitter's profiles	Naive Naive Bayes gets -accuracy rate = 95.7% -Precision = 94% -recall = 96%.	It aims to collect large amounts of data in different languages by using the same methodology that is used.
Chao Chen, Jun Zhangt, Yang Xiang, Wanlei Zhou (2015)	-statistical features based mechanisms -Changed Spam -Spam Drift	-ML based Classifiers -asymmetric self-learning (ASL) -Bayes Network	large dataset of Twitter	improves F measure 10% for Bayes Network improves Detection Rate more than 20% for all three tested ML algorithms	Will develop practical system for ASL approach
Chao Chen, Jun Zhang Xiao Chen, Yang Xiang, Wan lei Zhou (2015)	-12 Lightweight Features -CDF figures -Trend Mirco's Web Reputation System	machine learning algorithms: -RF - DT -BN -NB -K-NN -SVM	600 million public tweets		

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

Ashraf Khalil, Hassan Hajjdiab, Nabeel Al-Qirim (2017)		Machine learning algorithms: -SVM -Simple Logistic -Instance-based classifier using 1 nearest neighbor	13000 fake follow And 5386 genuine followers	High accuracy rate SVM = 60.48% Simple Logistics = 90.02% nearest neighbor = 98.74%	

3.2 DEEP LEARNING ALGORITHMS

[26] Proposed a new framework to solve this issue based on Convolutional Neural Networks (CNN) and CNN consists of three layers: Firstly, CONV layer takes input and calculates dot products, Secondly, POOL layer reduces the input size and Lastly, FC layer activation function was used to generate the output. CNN was exploited by two classifiers. Text-based classifiers embedding the text before sending the CNN, CNN made neurons with learnable weights and biases and softmax function used for class score prediction in classification. Combined classifiers use meta-data as input, normalized input data in 0, and 1 form, combine the metadata classifier and text-tweets and send then as input for classification. Social honeypot dataset and 1 KS to 10 KN dataset were used in this study. Combined classifier provided the high accuracy rate was 99.68% and 93.12% for dataset I, II.

[27] Proposed a new architecture model with help of other three different architecture models. Firstly, Convolutional Neural Networks with semantic layers and known as Semantic Convolutional Neural Networks. By using ConceptNet knowledge-based and WordNet the initial text was enhanced and is signified with word2vec based. Secondly, LSTM neural networks with semantic layer known this framework Semantic Long Short Term Memory. It enhances the semantic representation of the words. Finally, present the combination of above two models that model is named Sequential Stacked CNN-LSTM Model. Above models used for spam detection from social media and it take the advantages from above both models. Twitter dataset and SMS datasets were collected for implementation of hybrid models and this model compared with traditional models and get good results. SMS dataset accuracy rate was 1.16% and Twitter dataset accuracy rate was 2.05% that was increased rate.

[28] For detecting spams from Twitter proposed a Neural Network-based technique with traditional features-based method and deep learning methods. CNNwas used for experiments with multi word embedding. Machine learning algorithms are frequently one-sided toward majority class. 1KS10KN and HSpam data sets are used in this research paper.In CNN, 1KS10KN recall was low and HSpam recall was high. The feature-based methods perform ineffectively when analyzed 14 datasets of HSpam against the deep learning methods. F-mesure was 0.984 get from results.

[29] To find the varieties of spam activities proposed a novel technique based on deep learning tenhniques. Word vector traing mod was learned the structure of each tweet. After this on representing a dataset binary classifier based builted. In investigations, 10-day real tweet datasets are collected and implemented to evaluate proposed methods. To compare proposed methods to other existing text-based methods and proposed methods better performed to existing methods

[30] For spam Twitter detection suggest a new automated feature engineering machnism. A Deep Neural Network was trained by using classifying the dataset. Remove from hidden features layers it to represent tweets. To guess deep-learnt features performance, in this research also compare its set feature with statistical features and word2vec features. Deep Neural Networks with the help of word2vec featuers and statistics features detect the spam from twittier extra accurately.

[31] For detecting spam from social media netwoks have provided a Semantic Convolutional Neural Network as it has become increasingly important. For processing Convolutional Neural Networks was used and it is also used in NLP tasks. Word vector gets enriched semantically as it was finised with assistance of WordNet, Word2Vec andConpectNet. Datasets of Twitter providing 94.40% accuracy and SMS Spam dataset (UCI repository) providing 98.65% accuracy perform the evaluation of the architecture.

[32]Spam activity on Twitter detetion proposed a Deep Learning technique.Word Vector will enable the learning of the syntax of each tweet and deep learning will be used for its training purpose. For methodology and evalution a10-day real tweets datsets has been used. Observed examination with sequence of comparisons with different techniques including non-text-based techniques of detection, performance of different classifiers, and other text-based methods, have been used. It was found out that the features used here were unique among all detection methods.

[33] Long Short Term Neural Networks and Convolutional Neural Networks based with recommended a Novel Deep Learning architecture.This model is unheld by presentation of semantic data in the sample representation with the assistance of knowledge-based like ConceptNet and WordNet. Procedure of these bases improve execution by addressing better semiconductor vector; this is not the reason for testing words that were previously of random value Training. Two datasets are collected in this research one is SMS spam dataset

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

and second is Twitter dataset. proposed experimental results of training datasets and shows SSCL model has increased 1.16% in SMS spam dataset SSCL model accuracy has increased 2.05% in Twitter dataset.

In Table 2 clarify the previously researches algorithms using deep learning algorithems.

TABLE 2
Research Papers using Deep Learning Methods

Author & Year	Preprocessing	Methodology	Dataset	Results	Future Work
Gauri jain, Manisha Sharma, Basant Agarwal (2019)	-Semantic word vectors -Knowledge-bases -WordNet -ConceptNet	-Novel Deep Learning Model -Long Short Term Neural Networks -Convolutional Neural Networks	-SMS spam -Twitter dataset	SSCL model 1.16% in SMS -2.05% in Twitter accuracy has been increased.	
Sreekanth Madisetty and Maunendra Sankar Desarkar (2018)	-CNN Models -Feature-Based Models -User-Based Features -Content-based Features -N-gram Features	-Neural network-based	HSpam data set 1KS10KN	F-score = 0.894	By adding extra information about Twitter improve performance.
Tingmin Wu, Shigang Liu, Jun Zhang, Yang Xiang (2020)	-WordVector -Training Mode -binary classifier based -machine learning based classifiers	-Novel technique based -Machine-learning based methods -Text-based methods -Blacklist Techniques	-Ground-truth datasets	Outperformance of this model is largely. -Experiment results show more accurate results than others.	-Will look at more classifiers and different techniques later on.
Xinbo Ban, Chao Chen, Shigang Liu, Yu Wang, Jun Zhang (2018)	-URL-Based -Meta-Data-Based -Social Relation-Based -Word2vec -Statiscal	-Machine learning network -Deep Neural Networks -Security -Bi-LSTM	labelled dataset	Recall = 90% F1-score = 93%	
Gauri Jain, Manisha Sharma,Basant Agarwal, (2018)	-random word vectors -Word2vec -ConceptNet	- (CNN) - (SCNN)	-SMS Spam dataset -Twitter dataset	-on SMS spam Dataset 98.65% accuracy -on Twitter dataset 94.40% accuracy	The big data experiments of SCCN will be perfom in future that could not be done in recent work.

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

TingminWu, ShengWen, Shigang Liu, Jun Zhang, Majed Alrubaian (2017)	-text-based methods -non text-based techniques -WordVector -binary classifier	-novel technique based -deep-learning technique	-10-day real tweet dataset	-Existing Features were stronger than other comparison features.	-will compare more classifiers. -will collect more real data from social media
Gauri Jain, Manisha Sharma, Basnt Agrwal (2019)	-Semantic word vectors -Knowledge-bases -WordNet -ConceptNet	-Novel Deep Learning Model -Convolutional Neural Networks -Long Short Term Neural Networks	-SMS spam -Twitter dataset	-SSCL model has increased 1.16% in SMS spam dataset. -SSCL model accuracy has increased 2.05% in Twitter dataset.	

3.3. HYBRID ALGORTHMS

[34] By using community-based feature for detection of automated spammers proposed a hybrid approach besides further features like content-, metadata-, and interaction-based features. Followings, followers and other activities of the user provide the information. The research revolves around such characterization of the spammer that is based upon its neighboring nodes and their respective interactions. For spam detection analized to be the most effective features were Community-based features and metadata-based but metadata is the least effective for spam detection.

[35] They hybrid systems based on social honeypots used to detect the spam tweets, content filtering to detect similar tweets and classify the results that were provided by above two layers. The API streaming dataset of 100000 Twitter profiles that had malicious and legitimate profiles was trained by the preprocessing technique of spam filtering, text-based spam filtering, content filtering, extract characteristics and word N-gram. The model was tested by four algorithms that were random forest, bayes naive, treesJ48, classification via the regression and CNN-LSTM. Accuracy of classification regression was 99%, a positive rate & negative rate was (100%), recall and f-measure was 99%, precision was 99% and error rate was 1.7965%.

[36] Giving a spam detection system which detects spam tweets in near real time by using raw data capture. To design a training model on a large number of detecting spam tweets data for experiments. After preprocessing, real-time pulling data is used to collect 200 tweets at a time and it also helps the user to detect whether the tweet is spam or not. Before applying the above techniques, lightweight feature extraction extract 13 features on collected dataset of ground truth data. Nine machine learning algorithems used for spam or non-spam tweets and for training used ground truth data. Supervised machine learning algorithms classifications are: K-Nearest Neighbor-based algorithem, boosting algorithem Naïve Bayes, Neural Network, Deep learning, Gradient Boosting machine, Boosted Logistic Regression, Random Forests and Decision Tree-based algorithem. The probability of spam tweets combined with nine algorithms results that showed accuracy was 80% and F-meaure and TPR values were above 80%.

[37] Said some researchers and industries use different approaches that base on only tweets-based features and some base onuser-based features. In this research proposed a new framework that contains tweet-based features and user-based features alongside text-based features for classification of tweets. Text-based features can detect the spam tweets even spammers use new accounts with the help of tweet-base features and user-based features. HSpam 14 datset are based on tweets that was colletecd from Twitter. Train models of machine learning are: Gradient Support, Support Vector Machine, Randome Forests and Neural Natworks. These applied on dataset processing and got the predicted results with neural network accuracy was 91.65% and surpassed was near about 18%.

[38] Mentioned the usage of Tweeter that it has millions of users and it makes it easy for spammers to get into it. This paper suggests Genetic algorithem, a decision tree and particle swarm optimization make a combination that is used in Twitter spam in real time. Real-time information was extracted from API of Twitter for research purposes. They created six hundred million tweets using URL security tools. For comparison of their result, they compared their data with other hybrid techniques too. As they used a large data set to perform their technique which characterize the empirical cumulative distribution. PSG-DT stands out in all classifiers and less accuracy is shown by GA-DT.

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

[39] Decide the most useful features for the detection of spam as spammers began abusing Twitter by spreading infeactions, undesirable advertisement and phishing assaults. These features ensure the development of strong and accurate models of spammer detection. Popular classifiers including Decision Tree, K-Nearest Neighbor, Super Vector Machine, Naive Bayes, Multilayer Perceptron Neural Networks, Random Forests, are used for the testing of data sets. Further three methods including Relief-F, CoM and IG, are used for the analysis purpose. The features of the data set are related to Content-based features, user behavior-based features and graph-based features. Test results shown the significant part of feature like the status of accounts, usual mention per tweet and length of the tweet, average time betwwen post and age of the account, identifyning the spammers on Twitter get the highest ratio of graph-based.

[40] Proposed the method for detecting the spam on Twitter using intelligent Twitter spam technique. It will help the users to secure their personal information. It gives the data of the spammer's profile which comprises not a single classifier by using a hybrid technique. To make it secure, they used Google security APIs before getting the tweets as samples. They classified the information extracted from tweet into 2 categories i.e., content-based features and user-based features. They utilized the hashtags too to get the tweets and analyzed them. With the multilayer approaches 87.30% high system accuracy get.

[41]For spam profiles detection a hybrid technique with user-graph, conten-base features and graph-based feature used. Used dataset of Twitter with 11000 uses and more than 400000 tweets for experiments and Results have shown that high classification accuracy with low false positive.

[42]Firstly spot the “Spam Drift” issues of Twitter spam detection in statistical feature based. Proposed Lfun approaches to solve spam detection issues. In term of rate and F-measurement detection assess the performance of Lfun approaches. In experiments when applying proposed Lfun approaches that shows the improvement in detection rate and F-measurement. In this research, proposed Lfun approaches compare and other four traditional machine learning methods and locates high beats each of the four methods as to generally precision, F-meaurment and detection rate.

[43] Recommended some user-friendly and content-based features that can do this used to separate between a spammer, a popular and legitimate user on Twitter. They audit the value of these features spammer detection using traditional machine learning methods like K-NN, Support Vector Machine, Random Forests, Naïve Bayes, structures use the Twitter dataset they have plural. Using this rating, their recommended features precision was 95.7% and achievable f_measurment was 95.7%.

[44] The area had for implementation and evoulation identified as deep learning methods. A special framework of Recursive Neural Networks and Long Short Term Memory for spam detection was used. Before using LSTM for rating work text has been changed to meaning word semantic word vectors forConceptNet, Word2Vec and WoerdNet. The positioning outcomes are constrasted and the benchmark evaluation such asANN, Naïve Bayes, K-NN, Random Forests and SVM. For result evoulation SMS spam collection and Twitter datasets ere used. Assesment of the outcomes shows that LSTM can improve the traditional machine learning algorithems to distinguish spam with sufficient edge.
In Table 3 explain the earlier researches algorithms using hybrid algorithems.

TABLE 3
RESEARCH PAPERS USING HYBRID METHODS

Author & Year	Preprocessing	Methodology	Dataset	Results	Future Work
Zulfikar Alom, Barbara Carminati, Elena Ferrari (2002)	-Novel Deep Learning Framework - DT - RF - LR - SVM	-Machine Learning approach -Deep Learning approach -Convolutional Neural Networks	-social honeypot dataset -1 KS to 10 KN dataset	-Accuracy = 99.68% for Dataset I -Accuracy = 93.12% for Dataset II	Order different sorts of spammer's various kinds of Social media Platform.
Keyode Sakariyah Adewole, Tao Han, Houbing Song, Arun Kumar,	-hybrid approach Random Forest model -Classification based -Clustering approach	Machine learning classifiers	Public dataset -ground-truth data	Did not get the 100% F-measure but it improved the performance of the classifier with reduction in	In future, the research will be helpful in addressing model scalability without comparison accuracy performance.

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

Sangaiah (2018)				error rate.	
Mohd Fazil and Muhammad Abulaish (2018)	-Metadata-Based Feature -Content-Based Features -Interaction-Based Features	-hybrid approach -Random Forests -Decision Tree - Bayesian Networks	-Twitter Datasets such as .username .location use rid	-Community-based features -metadata-based analysis effective for spam detection -metadata least effective.	Spammers characterization at different levels of granularity utilized on the behalf some interesting patterns relieved by spammers.
Fatna Elmendili, Younès El Bouzekri El Idrissi (2020)	-Spam Filtering -Text-based spam Filtering -Content Filtering -Extract Characteristics -Word N-gram	-Random Forest -classification via the regression -Bayes Naive -TreesJ48 -CNN-LSTM	API streaming	-Accuracy classification Regression= 99% -a positive rate, negative rate = (100%), recall -F-Measure = 99% -Precision = 99% -error rate=1.796%.	-Conduct more theoretical studies on the out performance of methods -better understand the social honeypots based on malicious user's detection framework.
Nan Sun, Guanjun Lin, Junyang Qiu, Paul Rimba (2020)	-Light-weight feature extraction -URL base Extraction -Account-based features content-based features	-MLA, -RF -DT , - (KNN) -Boosting algorithms - (NB), - (NN) and (DL)	real-time Twitter data	-Accuracy is 80% -TPR & F-Measure values above then 80%.	-Increasing stability of Deep Learning and Random Forest. -Near real-time Twitter spam detection could also be improved by using this system to gather more tweets data.
Himank Gupta, Mohd. Saalim Jamal, Sreekanth Madisetty (2018)	-Support Vector Machine -Extracting Lightweight Features9	-Neural Networks -Random Forests -Gradient Bossting	HSpam14	Neural Networks gets: Accuracy = 91.65% Surpassed = 18%	-Will update the spam tweets bag of words. -Will perform URL crawl mechanisms.
N. Senthil Murugan G. Usha Devi	-user-based features -tweet-based features	-Hybrid Algorithms -PSO, GA and DT	Twitter dataset	GA-DT showed = less accuracy	Will extent this work in future

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

(2018)	-Features Statistics			PSG-DT = best classifier	
Wafa Herzallah, Hossam Faris, Omar Adwan (2017)	-Naive Bayes -SVM -Decision Tree -Random Forests -KNN	-Change of Mean Square Error -Information Gain -Relief-F method -MLP neural network	comprehensive dataset	highest ratio of Graph-Based Feature = 47%, -Conten-Based Feature = 40% -User Behaviour Feature = 13%	-Address the issue with a lot bigger informational indexes. - To examine the impact of the awkwardness information appropriation.
Varad Vishwarupe, Mangesh Bedekar, Milind Pande, Anil Hiwale (2018)	-User-based features -Content-based -Single-tier -Single-classifier approaches	Hybrid approach Machine learn	local dataset Google Safe Browsing Toolkit	Accuracy = 87.30%	More classifiers can be added that can make twitter spam detection more valuable for users.
Malik Mateen, Muhammad Aleem, Muhammad Azhar Iqbal (2017)	-User-Based Features -Content-Based Features -Graph-Based Features	-a hybrid technique	-Guofei Gu's data -real Twitter dataset	-high classification accuracy -with low false positive	-these techniques will be used for other social networking site.
Chao Chen, Yu Wang, Jun Zhang, Yang Xiang, Wanlei Zhou (2015)	statistical features based	Lfun approach	10-day ground truth	Lfun scheme improve accuracy in real-world scenarios	Will Incorporate incremental adjustment to adjust the training data
M. McCord and M. Chuah (2011)	-User-Based Features -Content-Based Features -Detection Scheme Based	-Random Forest classifier -Naïve Bayesian -SVM -K-NN	active Twitter users most recent 100 tweets	using the Random Forest classifier -F-measure = 95.7% -Precision = 95.7%	detection scheme using larger Twitter dataset
Gauri Jain, Manisha Sharma, Basant Agarwal (2018)	-Semantic Word Vectors -Word2vec -WordNet -ConceptNet	-Long Short Term Memory (LSTM) -Recursive Neurel Network (RNN	-SMS Spam Collection -Twitter dataset	LSTM can beat customary AI techniques	

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

	-Naïve Bayes -ANN -K-NN -RF - SVM				

4 Discussion

In previous work, more variables needed to add in the framework to enhance the accuracy of the model and classification rate. Need to improve text similarity for extracted new strange words from tweets. In previous researches, data mining algorithms were applied on small amounts of collected dataset and limited tweets. So, large amounts of data set need to be tested for the accuracy of previous algorithms.

In Future, we can collect the dataset of tweets in different languages. We can apply data mining algorithms on other social media platforms like Facebook, Instagram, LinkIn, YouTube and WhatsApp. More classifiers can be added that can make Twitter spam detection more valuable for users. Research will help to solve model scalability without performing comparative accuracy. Can use the characteristics of spammers at different levels of granularity have been used by some interesting patterns released by spammers.

5 Significant

Implementing spam detection is essential for any social media platform especially Twitter. Spam detection not only helps keep detect spams from tweets, but also helps improve the quality of life of social media accounts because they run smoothly and are only used for their intended purpose. Therefore, we are going to implement data mining algorithms for detecting spam tweets, messages and URLs from Twitter. In the future, these algorithms can help for spam detection on other social media platforms.

6 Conclusion

Today is the time of social media and Twitter is the most well known social media network where anyone can post their thoughts, send their messages and promote their business. Followers have been increased on Twitter to capture attention of the spammers. In previous research, there are many algorithms of data mining that are used for spam detection on Twitter's collected datasets. In literature review, we have compared the different data mining algorithms in the category of machine learning, deep learning and hybrid algorithems. All of these algorithms researchers use for different types of spam detection. But the previous algorithms are not enough to extract and detect the spam on Twitter accurately. So, we need to expand the research for the high classification rate of spam detection. In future, we will apply previous methods on further social media stages like Instagram, SnapChat, Facebook and YouTube.

Reference

[1] Saleh Beyt Sheikh Ahmad & Mahnaz Rafie & Seyed Mojtaba Ghorabie, "*Spam detection on Twitter using a support vector machine and users' features by identifying their interactions*" 06 January 2021.

[2] Rumi Juwairiyyah, Nanditha Sriram, Jyotshna Bhushan Sharma, Babeetha, ***"Social media spam detection using deep learning"* June 2020.**

[3] Deepali Prakash Sonawane Dr. Baisa L. Gunjal, "***New Approach for Detecting Spammers on Twitter using Machine Learning Framework"*** May 31st 2020.

[4] M Mostafa, A Abdelwahab, H M Sayed, "***Detecting spam campaign in twitter with semantic similarity***" 2020.

[5] Francesco Pierri, Carlo Piccardi, Stefano Ceri, "***A multi-layer approach to disinformation detection on Twitter"***2020.

[6] Ratul Chowdhury, Kumar Gourav Das ,Banani Saha, Samir Kumar Bandyopadhyay, "***A Method Based on NLP for Twitter Spam detection***"2020.

[7] Ricardo Resende de Mendonça, Daniel Felix de Brito, and Ferrucio de Franco Rosa, Júlio César dos Reis, Rodrigo Bonacin, " ***A Framework for Detecting Intentions of Criminal Acts in Social Media: A Case Study on Twitter*** "2020.

[8] Hadi Tajalizadeh and Reza Boostani," ***A Novel Stream Clustering Framework for Spam Detection in Twitter***" 2019.

[9] Isa Inuwa-Dutse, Mark Liptrott, Ioannis Korkontzelos, "Detection ***of spam-posting accounts on Twitter*"** 2018.

[10] Zulfikar Alom, Barbara Carminati, Elena Ferrari, "Detecting *spam accounts on Twitter* "2018.

[11] Rutuja Katpatal, Aparna Junnarkar, "*An Efficient Approach of Spam Detection in Twitter*" 2018.

[12] XIANCHAO ZHANG, ZHAOXING LI, SHAOPING ZHU, WENXIN LIANG,"Detecting *Spam and Promoting Campaigns in Twitter* "2016.

[13] Olubodunde Stephen Agboola, "*Spam Detection Using Machine Learning*" 2020.

[14] Surendra Sedhai and Aixin Sun, "*Semi-Supervised Spam Detection in Twitter Stream*" 2017.

[15] Claudia Meda, Edoardo Ragusa, Christian Gianoglio, Rodolfo Zunino, Augusto Ottaviano, "*Spam Detection of Twitter Traffic: A Framework based on Random Forests and non-uniform feature sampling*" = 2016.

[16] Xianchao Zhang, Shaoping Zhu, Wenxin Liang, "*Detecting Spam and Promoting Campaigns in the Twitter Social Network*" 2012.

[17] Aryo Pinandito, Rizal Setya Perdana, Mochamad Chandra Saputra, Hanifah Muslimah Az-zahra, *"Spam Detection Framework for Android Twitter Application Using Naïve Bayes and K-Nearest Neighbor Classifiers"* 2017.

[18] Buket Ersahin, Özlem Aktas, Deniz Kılınç, Ceyhun Akyol "*Twitter Fake Account Detection*" 2017.

[19] Rutuja Katpatal, Aparna Junnarkar, "*Spam Detection Techniques for Twitter*" 2018.

[20] Abdullah Talha Kabakus, Resul Kara ""*TwitterSpamDetector" A Spam Detection Framework for Twitter"* 2019.

International Journal of Scientific & Engineering Research, Volume 12, Issue 2, February-2021

ISSN 2229-5518

[21] Ala' M. Al-Zoubi, Ja'far Alqatawna, Hossam Faris, "*Spam Profile Detection in Social Networks Based on Public Features*" 2017.

[22] Chao Chen, Jun Zhangy, Yang Xiang and Wanlei Zhou, "*Asymmetric Self-Learning for Tackling Twitter Spam Drift*" 2015.

[23] Chao Chen, Jun Zhang Xiao Chen, Yang Xiang and Wan lei Zhou, "*6 Million Spam Tweets: A Large Ground Truth for Timely Twitter Spam Detection*" 2015.

[24] Ashraf Khalil, Hassan Hajjdiab, and Nabeel Al-Qirim, "*Detecting Fake Followers in Twitter: A Machine Learning Approach*" 2017.

[25] Zulfikar Alom a, Barbara Carminati b, Elena Ferrari, "*A deep learning model for Twitter spam detection*" 2002.

[26] Gauri Jain, Manisha Sharma, Basant Agarwal, "*Spam detection in social media using convolutional and long short term memory neural network*"2019.

[27] Sreekanth Madisetty and Maunendra Sankar Desarkar, *"A Neural Network-Based Ensemble Approach for Spam Detection in Twitter"* 2018.

[28] Tingmin Wu, Shigang Liu, Jun Zhang and Yang Xiang, *"Twitter Spam Detection based on Deep Learning"* 2020.

[29] Xinbo Ban, Chao Chen, Shigang Liu1, Yu Wang, and Jun Zhang (2018) "*Deep-learnt features for Twitter spam detection"*

[30] Gauri Jain, Manisha Sharma, Basant Agarwal, *"Spam Detection on Social Media Using Semantic Convolutional Neural Network"* 2018.

[31] TingminWu, ShengWen, Shigang Liu, Jun Zhang, Majed Alrubaian, *"Detecting spamming activities in twitter based on deep-learning technique"*2017.

[32] Gauri Jain, Manisha Sharma,Basant *Agarwal, "Spam detection in social media using convolutional and long short term memory neural network"* 2019.

[33] Kayode Sakariyah Adewole, Tao Han, Houbing Song, Arun Kumar, Sangaiah, "*Twitter spam account detection based on clustering and classification methods"* 2018.

[34] Mohd Fazil and Muhammad Abulaish*"A Hybrid Approach for Detecting Automated Spammers in Twitter*" (2018

[35] Fatna Elmendili, Younès El Bouzekri El Idrissi *"A Framework for Spam Detection in Twitter Based on Recommendation System"*2020.

[36] Nan Sun, Guanjun Lin, Junyang Qiu & Paul Rimba, *"Near real-time twitter spam detection with Machine learning techniques"* 2020.

[37] Himank Gupta, Mohd. Saalim Jamal, Sreekanth Madisetty, *"A Framework for Real-Time Spam Detection in Twitter"* 2018

[38] Nan Sun, Guanjun Lin, Junyang Qiu & Paul Rimba*"Near real-time twitter spam detection with Machine learning techniques" 2020.*

[39] Himank Gupta, Mohd. Saalim Jamal, Sreekanth Madisetty, *"A Framework for Real-Time Spam Detection in Twitter"* 2018.

[40] N. Senthil Murugan, G. Usha Devi, *"Detecting Streaming of Twitter Spam Using Hybrid Method"* 2018.

[41] Wafa Herzallah, Hossam Faris, Omar Adwan, *"Feature engineering for detecting spammers on Twitter: Modelling and analysis"* 2017

[42] Varad Vishwarupe, Mangesh Bedekar, Milind Pande and Anil Hiwale, *"Intelligent Twitter Spam Detection: A Hybrid Approach"* 2018

[43] Malik Mateen, Muhammad Aleem, Muhammad Azhar Iqbal, *"A Hybrid Approach for Spam Detection for Twitter"* 2017.

[44] Chao Chen, Yu Wang, Jun Zhang, Yang Xiang, Wanlei Zhou, *"Statistical Features Based Real-time Detection of Drifted Twitter Spam"* 2015.

[45] M. McCord and M. Chuah "*Spam Detection on Twitter Using Traditional Classifiers*" 2011

[46] Gauri Jain, Manisha Sharma, Basant Agarwal, *"Optimizing semantic LSTM for spam detection"* 2018.