An Ensemble Approach to Detect Fake News Spreaders on Twitter

Muhammad Nabeel Sarwar[1], Riaz UlAmin[1] and Sidra Jabeen[1]

University of Okara, Punjab, Pakistan
Corresponding Author: Riaz UlAmin riazulamin@gmail.com

ABSTRACT
Detection of fake news is a complex and a challenging task. Generation of fake news is very hard to stop, only steps to control its circulation may help in minimizing its impacts. Humans tend to believe in misleading false information. Researcher started with social media sites to categorize in terms of real or fake news. False information misleads any individual or an organization that may cause of big failure and any financial loss. Automatic system for detection of false information circulating on social media is an emerging area of research. It is gaining attention of both industry and academia since US presidential elections 2016. Fake news has negative and severe effects on individuals and organizations elongating its hostile effects on the society. Prediction of fake news in timely manner is important. This research focuses on detection of fake news spreaders. In this context, overall, 6 models are developed during this research, trained and tested with dataset of PAN 2020. Four approaches N-gram based; user statistics-based models are trained with different values of hyper parameters. Extensive grid search with cross validation is applied in each machine learning model. In N-gram based models, out of numerous machine learning models this research focused on better results yielding algorithms, assessed by deep reading of state-of-the-art related work in the field. For better accuracy, author aimed at developing models using Random Forest, Logistic Regression, SVM, and XGBoost. All four machine learning algorithms were trained with cross validated grid search hyper parameters. Advantages of this research over previous work is user statistics-based model and then ensemble learning model. Which were designed in a way to help classifying Twitter users as fake news spreader or not with highest reliability. User statistical model used 17 features, on the basis of which it categorized a Twitter user as malicious. New dataset based on predictions of machine learning models was constructed. And then Three techniques of simple mean, logistic regression and random forest in combination with ensemble model is applied. Logistic regression combined in ensemble model gave best training and testing results, achieving an accuracy of 72%.
Keywords: *Fake news detection, Fake news Spreaders, Ensemble learning, Statistical Model, Feature Analysis of fake news spreaders, identification of twitter users.*

1. Introduction

Social media platforms often act as news sources. Research shows that most of the social media posts are related to daily news. Its role is increasing as a news provider and news amplifier. Basically, on social Media Platforms, anyone, anywhere, can produce and help circulate content for other people to read. Barriers of traditional publishing as in print media and electronic media has vanished. This leads to an extinction of quality control processes. Individuals and organizations publish content on social media platforms like Facebook, Twitter etc. without following basic journalism principles such as source verification, fact checking and accountability. These principles are easily bypassed or ignored by them on social networks. In present era social media is a prime source of spreading information and making public announcements. According to research, number of discussions on twitter are directly linked to headline news, results showed 85% conversations are about recent headlines [1]. Analysis of social media especially twitter may lead to movie success prediction [2], analysis of stock market over a time period[3], understanding outbreaks of certain diseases[4], tracking crisis & pandemic situations[5]. Owing to easy accessibility of social platforms information and news are at distance of click these days. Easy creation and proliferation of the content causes increased number of malicious users on social networks, these users infect the network by propagating rumors and misinformation. Falsity of information leads to many financial, political and social issues [6-8]. This unreliability of the content is now referred as “Fake News”.

Fake news has been a topic covered significantly by media ever since USA presidential elections of 2016. “Fake news” circulates largely due to social networks almost 30-40% [9]. Facebook defines fake news as:

“Articles that purport to be factual, but contain intentional mis-statements of fact with the intention to arouse passions, attract viewership, or deceive” [10].

New Media Companies such as Twitter, Facebook, Instagram enable news, stories, and rumors to be propagated fast globally without proper verification procedures. A report by the Jump shot Tech Blog exposed that 50% of the total d of the Facebook stories referrals are to fake news websites and 20% to the reputable sites. Almost 62% of the U.S. adults consume social media as news source [11], Thus, capability to identify fake content among available online sources is a tenacious need. Researchers proved existence of a strong correlation between twitter activity and real-world events.
Verified impacts of fake news on society urged many researchers, developers and organizations to work in this

Manuscript received May 5, 2022
Manuscript revised May 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.5.41

domain. Google, Twitter, Facebook and PAN currently consume a lot of money in research and development to overcome this current era problem. There are multiple approaches and methods explored by researchers in the domain of content analysis, source credibility and fake news. As there is variety of contexts in which research is being carried out by the researchers it is needed to summarize them. This paper presents an overview of strategies used to identify fake news, methodologies along with their performance results. Moreover, a comparison of state-of-the-arts techniques to guide upcoming researchers (newbies) which methodology performed better in the past years. This paper focuses on different most promising branches of the problem and their initial proposed solutions. This paper also aims at presenting a future perspective of the problem with a focus on next steps of research in this field.

2. Literature Review

First hand reports of well-known on-going events such as natural disasters, debates, public gatherings, terrorist attacks, and active shooting etc. are posted on social media. However, these reports are not always correct and factual. Twitter has added a feature of being "verified" account to be believed as "credible" information being posted by authentic account. Nonetheless, twitter accounts of most genuine users are not verified even though there many posts may be factual and correct. An important and on-going challenge is to identify non-credible users who spread fake news and negativity on the network.
Fake news has gained immense importance, and consequently getting exponential attention, following US Presidential elections of 2016 [12]. Detection of fake news is tough, almost impossible for human beings. Now a days, citizens consume social media for information gaining and learning about current events around the world. There is a rapid increase in use of social media. An advent of executives, officials, celebrities, political and public figures on social media has become prevalent. Public and private organizations show their information to the public. Due to this widespread applications and powerful use of social media, it is obligatory to pay attention to the source i.e., where the news is coming from. There is a dire need of checking credible news sources in terms of providing correct information on remarkably sensitive topics
This review covers distinct sub-areas of non-credible information on social network. Only most relevant computer science papers are selected for review excluding already published review papers on the topic while prioritizing most recent publication.

I. Definitions of Fake News:

One of the most agreed definition of fake news is such news items which are deliberately written or created and can be verified as false; for the purpose of mislead consumers. This definition leads to two main features of the definition: "authenticity and intent".
Fake news contains such an information which can be authenticated as false. Secondly, its creation is based on a purpose of misleading the readers. Recent researchers assume this definition of fake news in their studies.
"Fake news is such an article that exhibits or provide misleading information for the human being" . "Fake news is misinformation or manipulated news that is spread across the social media with an intention to damage a person, agency and organization" [13].
Fake news is where individuals or organizations intentionally publish hoaxes, propaganda and other misinformation and present it as factual. This can include blog and social media posts and fake online media releases.
Fake news is classified into three general categories.
1. News which is completely fake and generated by writer.
2. Satire news, created with an intention of humor for the audiences
3. Ill- articles, written poorly and contains some real news but not accurate. Designed intentionally to promote some agenda or influenced opinion [14].
Rubin and team deliberate three kinds of fake news, representing erroneous or deceiving reportage[15] .

II. Phase 1 -Stance Detection

Evaluating the authenticity of a news is a multifaceted, cumbersome and complex task for experts and even for trained humans. Auspiciously, fake news detection process can be fragmented stages or set of steps to achieve overall goal. Proposed first step identifying fake news is to comprehend opinion of other news sources or organizations about specific news. This automatic verification, named Stance Detection, provide base for further news checking process considering it as a building block and assist in completing AI based fact checking pipeline.
PAN held a competition of FAKE NEWS DETECTION shortly as FNC. Thus, first stage of the Fake News detection Challenge (FNC-1) is Stance Detection, announced as task for teams and individuals to participate.
First Fake News Challenge (FNC-1) was organized by Pomerleau and Rao (2017) to promote AI based system development for automation detection of fake news. This announced challenge got special attention of researchers in the field of natural language processing. Challenge got participation from both academic world and industry. Almost 50 teams participated [16]. As per the description of FNC organizers, participants were not supposed to be very much subjective and categorize news/news headline as fake or real directly as it is highly subjective and

complex question. Instead, this challenge is organized for "stance detection," involving a comparison of headline and article's body to determine either there exists a relationship (if any) between the two.
FNC organizers recognized two research papers which may provide foundations for stance detection. Ferreira and Vlachos [17] used the "Emergent" dataset laying useful work bases for comparisons of real news articles and false rumors. These news articles were first identified by critics with an estimate of their veracity. They claim the stance of the article towards the rumor as either ""for," "against," or "observing,"", achieved an accuracy of 73%. Similarly, Augenstein et al.[18] did almost identical work of stance detection but with another dataset having shorter text strings. They predicted tweets as "positive", "negative" or "neutral" with their short topic string considered as the "target."
[19] devised methods to separate related from unrelated headlines and further classification of the related headlines. They performed computations on a publicly available data set which contained the headlines and corresponding articles. Stance detection of headlines with article bodies achieved a (weighted) accuracy score of 89.59%.

III. Phase 2

S. Karodia focused on developing a system that can algorithmically assess the credibility of tweets on Twitter and present the assessment results to the user. A classifier was trained using an annotated dataset, generated through a crowdsourcing mobile application, and results were displayed in the Twitter interface via a Web browser extension [20].
Rubin et al. deployed a prototype which identify specific classes of news articles such as satire and humor. They scrutinized 360 satirical news articles from four realms of life namely science, civics, business, and entertainment/gossip articles. They proposed an SVM classification model and achieved highest precision of 90% [21].
H. Ahmed et al. presented an n-gram features-based approach for detection of fake news. A text analysis of N-gram features and implementation of machine learning algorithms to accomplish the purpose. Six supervised machine learning techniques are applied, i.e., K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Logistic Regression (LR), Linear Support Vector Machine (LSVM), Decision tree (DT) and Stochastic Gradient Descent (SGD). This teams implemented these algorithms and a performance comparison is presented. Evaluation of these algorithms is examined with a compiled data set from different news websites. These outperforms much better and yields inspiring results, achieving an accuracy of 92% [22].
The main purposes of [23] was the implementation of machine learning algorithms for fake news detection. Popular algorithms (Naïve)Bayes, Neural Network and Support Vector Machine (SVM)) are implemented. Before applying algorithms, main advantage of this system was data normalization achieved at pre-processing step by cleaning the data. The results showed that Naïve bayes achieved lower accuracy among all other, an accuracy of 96.08% while other two achieved 99.90 % prediction accuracy.
PAN tasks predict user traits on the basis of their social media posts focusing on twitter. The rapid expansion of blogging and electronic data in Web 2.0 is extensive and therefore, it is important to identify the profile of the author. The problems of automatically recognizing the gender and age of the author are based on a more interesting model of language and style in recent years. Such methodologies are also useful for some other applications like criminal detection, security and author detection, etc. Lexical, synthetic and structural characteristics provide research bases to identify personal, sociolect and demographic aspects of the author.
In 2014 shared task of PAN at CLEF online reputation of the user lead to authors profiling. dataset contained twitter accounts with their almost 600 allied tweets. Developed task solution must be able to classify user's as one the given set of categories such as politician, celebrity, company, organization, client or authority etc. Knowing category of tweeting user helps in determining opinion of user and importance of his/her point of view on certain topics as an added advantage [24].
Iftene and his team [25] aims at analyzing twitter using neural network to provide key notes in assessing the credibility of news as well as users. Technically, researcher applied two methods for said purpose i.e., by sentiment analysis and by neural network. Authors provide basic details about building dataset, potential feature extraction which might affect the output of the model. Moreover, paper provides a mathematical formula to compute the credibility of the tweets. Besides identifying fake news and fake users, this model presents statistics about the evolution of the fake news around the globe. Researcher built a real-time heat-map on the basis of information collected from users of twitter to display and differentiate between real and fake tweets[25].
[26] This paper refers to an approach involving techniques of text mining and human intelligence. It uses Python's Natural Language Toolkit (NLTK) for better application of natural language processing algorithms. This piece of research has an added advantage of focusing on "WHO" is fake or not-credible instead of classifying "What" is credible. Credible information is delivered, on the bases of news articles, official news releasing authorities, domestic records, criminal profiles records, Twitter accounts that are Verified. Official records of the organizations are considered as a Golden Standards for comparison and declare a news source as credible. Moreover, results of

automated model are compared with real-time data using crowd-sourcing.

Mendoza measured trust while spreading information using Twitter [27]. They focused on special disaster event of Chile earthquakes in 2010. Aggregate analysis of data of tweets performed to detect tweets and rumors. Later they devised a method of prediction tweet's credibility using improved set of features. Model attained 60%-70% precision in predicting the credibility[28].

Kang presented two different models to determine Twitter credibility, i.e. social features-based model and content based model [29]. First mentioned model uses weighted computation of positive credibility depending on the underlying features of network such as number of followers, re-tweets etc. Other model used a probabilistic language-based approach to identify credibility using sentiments, intense words, URLs, images, punctuation marks, or emoticons. Experiments showed a substantial difference between performance of these two models. Social model predicted results with an accuracy of approximately 88.17% whereas content model outputs 62%. This significant difference of performance leads to a conclusion that re-tweet widely social factors are of prime importance in predicting the credibility of tweets[29].

After above mentioned research, Kang researched on another feature of social domain factors. The considered number of followers and friends and calculated a correlation of this factor with credibility of the tweet. They plotted the results showing validity or suspicion tweets. Low credible users have large number of followers and follow many people except verified accounts of celebrities etc. as they are followed by many and they follow less number of people [29].

Some researches investigate credibility of a tweet and the user of social media based on both content of the tweet and data of tweeting user. In [30] analyzing text of the tweet and source of the text a statistical set is obtained about authenticity of the tweet. User-score and tweet scores leads to decision about a tweet which are computed using mathematical formulas. Investigations on the data is verified by applying on different use-cases.

[31] devised a reliable model to categorize the news story as fake or real. Machine learning algorithms such as logistic regression, SVM, Naïve Bayes Classification with Lidstone smoothing were used as classifiers trained and tested over Kaggle dataset. Computations and calculation based this model was tested with different variations. Naïve bayes with lidstone smoothing and without any hyperparameters achieved an accuracy of approximately 83%. Whereas without lidstone smoothing it achieved an accuracy of 74%. Other algorithms have a scope of almost 60-75%.

[32] Worked on the detection of news veracity of images in the news stories, by exploring and authenticating over the web. Proposed model extracted text from the images and check top 15 google search results for authentication. This images dataset based novel algorithm computed reality parameter for a news iteratively, if Rp crossed certain threshold value, news is classified as real otherwise false. They observed accuracy of the algorithm with different values of reality parameter (Rp), achieving best accuracy 85% with Rp value of 40.

There is a possibility of such users on social media who uses right information and mixes it with some false information for their own purpose and post it. This misleading news or information will also spread on social media just like other real news. This may lead to confusion and misperception among readers. It's better to control the spread of such confusing, negative and ambiguous news. PAN at CLEF 2020 competition focuses on controlling the propagation of fake news by identifying users who disseminates false news. According to [33] main aim of the shared task of PAN is to develop a system which automatically identify fake news spreaders on twitter and to check how difficult it is to limit them.

[34] built a model as a participant of CLEF 2020. To achieve the target, they used statistical analysis of the language used by the user of twitter who spreads fake news. Central tendency measure is used for statistical analysis of the language. Sentiment and polar classification technique alongside the vectors obtained after processed words are used to make a set of features which lead to find profiles spreading negative opinions. Participants pre-processed and cleaned the data before feeding to machine learning and deep-learning models. Feature's analysis yields a set of features such as URL's, emojis, hash-tags, polarity of each message as positive comments, negative comments or average. These features processed as text-vectors to input model consisting of a set of classifiers such as Logistic Regression, K-Neighbors, Random Forest, Decision Tree, Naïve bayes, LDA, and SVM etc. Results of these experiments on the given data is presented in terms of accuracy, a comparison of resulting accuracies is also presented by three authors. According to them, Random Forest Classifier performs best and classifies profiles with an accuracy of 76% for Spanish and 71.7 % for English language dataset.

Inna Voghal and Meghna They applied diverse feature extraction procedures and experimented with learning techniques with both languages i.e., Spanish and English. They used N-grams features with linear Support Vector Machine algorithm and Logistic Regression (LR). They achieved an accuracy of 79% on Spanish and 73% on English language. Their model secured 3rd position among 72 competitors of the competition[35].

Duan and his team proposed two-step solution to the problem of identifying fake news spreaders. the first step is at tweet level, sentiment analysis and political influences are analyzed. Second step generates features at profile level for binary classifier. This model achieved an accuracy of 70% with 10-fold cross validation of SVM classifier [36].

3. Methodology

Fake news spread on social media is vital and wide research area, this piece of research concentrates on Twitter. To assess reliability of a user, deep learning algorithms are implemented. Developed models are trained on a given corpus of tweets. There are several steps to determine if the tweet is true. Our focus is on user-based features. Determining credibility of the user is directly related to reliability of the users.

I. DATASET AND PRE-PROCESSING

This piece of research used input dataset of PAN Challenge 2020 [33]. Folder contains 300 XML files for each twitter user with 100 tweets from their feed/timeline. The XML file is named for corresponding author id which is unique for each twitter user and a truth.txt file with ground truths and authors list. To avoid over-fitting cross-validation techniques are used. Moreover, due to smaller size of the dataset whole corpus (without splitting as training set and development set) is used for development and training of the model.

Dataset was already pre-cleaned, URL's, hashtag and all user mentions are transforming into standard tokens.

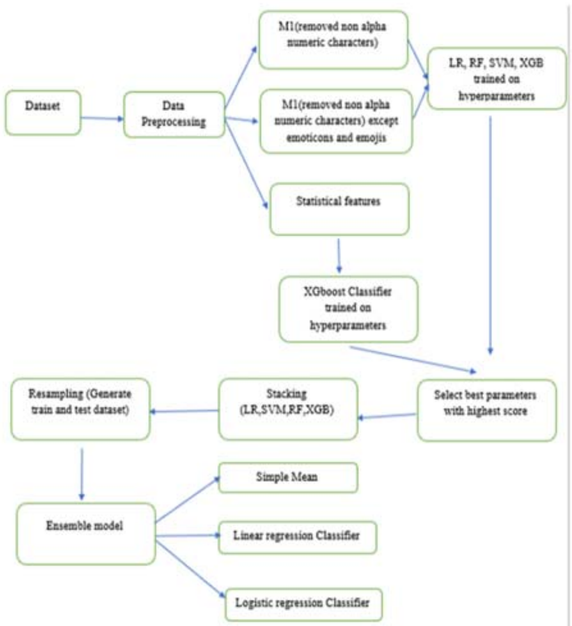

Figure 1: Work flow diagram

II. ARCHITECTURE

To train the models three types of algorithms are used N-gram based models, User statistical model and Ensembled learning model. Multiple models are trained and experienced their behavior on their n-gram models. Logistic regression, Random Forest, XBboost and Linear support vector machine are trained on n-gram models. Deep XGboost model is trained on statistical features. To develop all of these models', grid-search with 5-fold cross validation is applied to find optimal parameters, text preparation and vectorization for building a model.

III. HYPERPARAMETERS OPTIMIZATION

N-grams based machine learning algorithms are dependent on hyperparameters to make prediction. Above explained algorithms are implemented with different values of corresponding hyperparameters. A pipeline is made up for the implementation of architectural framework. by using the extensive grid search to find the best hyper parameters for the algorithms. Grid search with cross validation and with different values of document frequency and n-gram ranges are implemented. Machine learning models are experimented on 56 different n-grams variants {(1,1),(1,2),(2,2)} and observe their behaviour on different minimum document frequency.

Their hyperparameter details are shown in the below tables.

Table 1: HYPERPARAMETER VALUES FOR LOGISTIC REGRESSION

HyperParameters Name	Value
Regularization of coefficient (C)	{0.1,1,10,100,10000}

Table 2 HYPERPARAMETER VALUES FOR SUPPORT VECTOR MACHINE

HyperParameters Name	Value
Regularization of coefficient (C)	{1,10,100,10000}

Table 3: HYPERPARAMETER VALUES FOR XGBOOST

HyperParameters Name	Value
Learning rate [37]	{0.01,0.1,0.3}
Number of estimators {n_estimators}	{200,300}
Maximum depth of a tree (max_depth)	{3,4,5,6}

Subsample ratio {subsample)	{0.6,0.7,0.8}
subsample ration of columns(colsample_bytree)	{0.5,0.6,0.7}

Table 4 HYPERPARAMETER VALUES FOR RANDOM FOREST

HyperParameters Name	Value
Boosting rounds numbers (b)	{100,300,400}
Minimum number of cases on each leaf (min_samples_leaf)	{5,6,7,8,9,10}

Initially, testing with the combination of these parameters, model with highest yielding accuracy was applied for whole training dataset. However, later on a decline in the performance (approximately 5% to 7%) of these models wasobserved on testing dataset. Which leads to a need of some better technique to be considered.

1. USER STATISTICAL MODEL

A model consisting of statistical variable explaining all given tweets of a user is also developed. It gives another prediction for each author on the basis of 100 tweets provided in the corpus. Calculation based on 17 variables other than on n-gram variables are:

1. Both words and characters mean **length**
2. **Minimum length** of words and characters
3. Words and characters' **Maximum length**
4. The **standard deviations** from maximum and minimum lengths of both words and characters
5. **Range of tweets** in both words and characters
6. **Retweets frequency** by author in dataset
7. **URLs or website** links
8. **hashtags**
9. **Emoji's**
10. **Mentions**
11. **Number of ellipses** at the end of the tweet text
12. **Lexical diversity** measured by type-token ratio, a stylistic feature

Different values of hyperparameters used for the statistical model and the best parameters found are:

Table 5: HYPERPARAMETER VALUES AND BEST PARAMETERS FOR STATISTICAL MODEL

Parameters Name	***Best Parameters***	***HyperParameters***
Column sample by node	1	{0.8,0.9,1}
Column sample node	0.9	{0.8,0.9,1}
Gama	2	{0,1,2,4}
Learning rate	0.2	{0.3,0.2,0.1}
Max depth	4	{2,3,4}
Min child weight	4	{2,3,4,5}
Number of estimators	200	{100,150,200}
Alpha	0.1	{0.1,0.3,0.7}
Subsample	0.8	{0.6,0.8,1}

V. ENSEMBLE LEARNING

A machine learning technique that trains various models (also called "weak Learners"), to solve a problem jointly for improved results. Mainly it assumes that by proper joining of weak learners, we can acquire more precise and/or strong models. There are mainly three types of ensemble learning i.e. Bagging, Boosting and stacking. It can be described as the core focus of bagging is to get an ensemble model with less variance than its component weak learning models whereas the focus of stacking and boosting is to generate strong models with low biasedness than their component models along with a reduction in variance. In this research, we are going to use the stacking ensemble method tocombine multiple models in a single one to get precise results.

we used several models such as LR, RF, SVM, and XGB and identify their hyperparameters with cross-validation, so that it is required to find a reliable ensemble model to combine them a produced a single precise model. For this purpose, the stacking ensemble model technique has been used. To make the model free from overfitting the proposed ensemble model is not trained on the prediction of previously trained models instead a new dataset is created that produces a prediction of the proposed ensemble model. For this purpose, all the sub-models are refitted with the cross-validated hyperparameters five times on altered chunks of the inventive training data (each comprising of tweets from 240 users). The predictions produced by these models to the 60 leftover users were then added to the training data of the ensemble model, therefore, this training set comprised of predictions assumed to all 300 users in the training data, however, these predictions were assumed by all these different models in case of each model type. After that these built training and test, sets are used to find the best ensemble from the following three methods: majority voting, linear regression of predicted probabilities (this includes the simple mean), and a logistic regression model. The best and most accurate results are produced by the logistic model; so, this model is used as the final ensemble method. The below table shows the performance co-efficient of the final ensemble model (logistic regression).

Continuous experimentation with models, we selected the parameter that had accomplished high accuracy with cross validations. And models are fit on these parameters for further training.

Table 6: PERFORMANCE CO-EFFICIENT OF THE FINAL ENSEMBLE MODEL

Model	***Coefficient Values***
LR	0.8
SVM	0.48
XGB	1.07
STYLISTIC XGB	0.2
Random Forest	0

4. Results and Discussion

The core aim of our research was to establish a model that efficiently predict either an author is contributor in the spreading of fake news or the author is a credible source of information, from the available dataset. Authors implemented the software in two phases. In first phase, n-gram models (LR, SVM, RF and XGB) with cross validation grid search were implemented.

XGB classifier is used for the statistical model that is trained different hyperparameters with grid search cross validation to achieve best score on M3 dataset. Results of developed statistical model with 6 k-fold that partitioning test dataset and training dataset is 89 %.
Logistic Regression model is trained on multiple hyperparameters and on different variants of word N-Grams and minimum document frequencies. This model achieves 76% accuracy on the best parameters.
Random Forest model is trained on multiple hyperparameters and on different variants of word N-Grams and minimum document frequencies. Random forest attains highest accuracy 74.333% on M2 dataset.
Support Vector Machine trained on multiple hyperparameters cross validation to attain best results. SVM achieve highest 75.66666666666666% accuracy on M1 dataset.
XGB trained on multiple hyperparameters to attain best results. XGB achieve 73.66666666666666% accuracy on M1 dataset.
To achieve better performance and an enhancement of all existing systems is done by implementing ensemble learning model. All the results during optimization of the parameters with cross-validated extensive grid-search leads to highest accuracy. The constructed training and testing dataset M4 to select best ensemble model from these two-classifier linear regression and logistic regression and with simple mean. The results of simple mean, linear regression and logistic regression are computed as 71%, 72% and 72.9999999% respectively.

4. Conclusion

In proposed framework we apply multiple machine learning algorithms to classify an author as either a contributor in spreading the fake news. Architecture is implemented in two phases: first is n-gram based model and second is feature based model. Feature based model helps in achieving best scores.

Cleaning of dataset helps in development of stable models. After cleaning M1, M2 and M3 datasets are generated from the original dataset. N-gram based models (LR, RF, SVM, XGB) apply on M1 and M2 dataset and check on multiple hyperparameters with grid search cross validation. LR train on both dataset (M1 and M2) and achieve 76% accuracy on M1 dataset, RF achieve 74.333% on M2 dataset, XGB achieve 73.666% accuracy on M1 dataset and SVM achieve 75.66666666666666% accuracy on M1 dataset. XGB classifier is applied on statistical dataset i.e., M3. Hyperparameters grid search with cross validation showed a variety of results that leads us in selection of best performing parameter. XGB achieve 76% accuracy on 50 user test dataset and attain 89% accuracy on best parameters. Ensemble learning technique helps to overcome the weaknesses of other algorithms. To Avoid overfitting, instead of predictions of n-gram and statistical 74 models, predictions of five sub models with 5-fold stratified is used. This regenerated a dataset for training and testing of final ensembled model. This dataset consists of probabilistic results generated by five sub-models. This dataset uses as feature for ensemble model. Simple mean, linear regression and logistic regression are use as ensemble. Simple mean achieves 71% accuracy on training set and 70% on test set. Linear regression achieves 73% accuracy on training set and 71% on test set. And Logistic regression achieved best accuracy 72% on both training and testing sets. That why we use it as final model.

References

1. Kwak, H., et al. *What is Twitter, a social network or a news media?* in *Proceedings of the 19th international conference on World wide web*. 2010.
2. Asur, S. and B.A. Huberman. *Predicting the future with social media*. in *2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology*. 2010. IEEE.
3. Zhang, X., H. Fuehres, and P.A. Gloor, *Predicting stock market indicators through twitter "I hope it is not as bad as I fear".* Procedia-Social and Behavioral Sciences, 2011. **26**: p. 55-62.
4. Signorini, A., A.M. Segre, and P.M. Polgreen, *The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic.* Plos one, 2011. **6**(5): p. e19467.
5. Bodnar, T., et al. *On the ground validation of online diagnosis with Twitter and medical records*. in *Proceedings of the 23rd International Conference on World Wide Web*. 2014.
6. Garrett, R.K. and B.E. Weeks. *The promise and peril of real-time corrections to political misperceptions*. in *Proceedings of the 2013 conference on Computer supported cooperative work*. 2013.
7. Chatfield, A.T., C.G. Reddick, and K. Choi. *Online media use of false news to frame the 2016 Trump Presidential Campaign*. in *Proceedings of the 18th Annual International Conference on Digital Government Research*. 2017.
8. Cerf, V.G., *Information and misinformation on the internet*, 2016, ACM New York, NY, USA.
9. Turk, Ž., *Technology as enabler of fake news and a potential tool to combat it*2018: European Parliament.
10. Weedon, J., W. Nuland, and A. Stamos, *Information operations and Facebook.* Retrieved from Facebook: https://fbnewsroomus. files. wordpress. com/2017/04/facebook-and-information-operations-v1. pdf, 2017.
11. Gottfried, J. and E. Shearer, *News use across social media platforms 2016.* 2016.
12. Allcott, H. and M. Gentzkow, *Social media and fake news in the 2016 election.* Journal of economic perspectives, 2017. **31**(2): p. 211-36.
13. Ananth, S., et al., *Fake news detection using convolution neural network in deep learning.* International Journal of Innovative Research in Computer and Communication Engineering, 2019. **7**(1): p. 49-63.
14. Schow, A., *The 4 Types of 'Fake News'*. online] Observer. Available at: http://observer. com/2017/01/fake-news-russia-hacking-clinton-loss, 2017.
15. Rubin, V.L., Y. Chen, and N.K. Conroy, *Deception detection for news: three types of fakes.* Proceedings of the Association for Information Science and Technology, 2015. **52**(1): p. 1-4.
16. Hanselowski, A., et al., *A retrospective analysis of the fake news challenge stance detection task.* arXiv preprint arXiv:1806.05180, 2018.
17. Ferreira, W. and A. Vlachos. *Emergent: a novel data-set for stance classification.* in *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies*. 2016.
18. Augenstein, I., et al., *Stance detection with bidirectional conditional encoding.* arXiv preprint arXiv:1606.05464, 2016.
19. Bourgonje, P., J.M. Schneider, and G. Rehm. *From clickbait to fake news detection: an approach based on detecting the stance of headlines to articles.* in *Proceedings of the 2017 EMNLP workshop: natural language processing meets journalism*. 2017.
20. Karodia, S., M. Lu, and K. Kinmont, *Fake News Detection on Twitter.*
21. Rubin, V.L., et al. *Fake news or truth? using satirical cues to detect potentially misleading news.* in *Proceedings of the second workshop on computational approaches to deception detection*. 2016.
22. Ahmed, H., I. Traore, and S. Saad. *Detection of online fake news using n-gram analysis and machine learning techniques*. in *International conference on intelligent, secure, and dependable systems in distributed and cloud environments*. 2017. Springer.
23. Aphiwongsophon, S. and P. Chongstitvatana. *Detecting fake news with machine learning method.* in *2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. 2018. IEEE.

24. Rangel, F., et al. *Overview of the 2nd author profiling task at pan 2014*. in *CEUR Workshop Proceedings*. 2014. CEUR Workshop Proceedings.
25. Iftene, A., et al. *A real-time system for credibility on Twitter*. in *Proceedings of the 12th Language Resources and Evaluation Conference*. 2020.
26. Thompson, K.M., *Assessment of information credibility in Twitter*, 2014, University of Georgia.
27. Mendoza, M., B. Poblete, and C. Castillo. *Twitter under crisis: Can we trust what we RT?* in *Proceedings of the first workshop on social media analytics*. 2010.
28. Castillo, C., M. Mendoza, and B. Poblete. *Information credibility on twitter*. in *Proceedings of the 20th international conference on World wide web*. 2011.
29. ODonovan, J., et al. *Credibility in context: An analysis of feature distributions in twitter*. in *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*. 2012. IEEE.
30. Atodiresei, C.-S., A. Tănăselea, and A. Iftene, *Identifying fake news and fake users on Twitter*. Procedia Computer Science, 2018. **126**: p. 451-461.
31. Soni, V.D., *Prediction of Geniunity of News using advanced Machine Learning and Natural Language processing Algorithms.* International Journal of Innovative Research in Science Engineering and Technology
2018. **7**(5): p. 6349-6354.
32. Vishwakarma, D.K., D. Varshney, and A. Yadav, *Detection and veracity analysis of fake news via scrapping and authenticating the web search.* Cognitive Systems Research, 2019. **58**: p. 217-229.
33. Rangel, F., et al. *Overview of the 8th author profiling task at PAN 2020: profiling fake news spreaders on Twitter*. in *CEUR Workshop Proceedings*. 2020. Sun SITE Central Europe.
34. Moreno-Sandoval, L.G., et al. *Assembly of Polarity, Emotion and User Statistics for Detection of Fake Profiles*. in *CLEF (Working Notes)*. 2020.
35. Vogel, I. and M. Meghana. *Fake News Spreader Detection on Twitter using Character N-Grams*. in *CLEF (Working Notes)*. 2020.
36. Duan, X., et al. *RMIT at PAN-CLEF 2020: Profiling Fake News Spreaders on Twitter*. in *CLEF (Working Notes)*. 2020.
37. Gaspay, A., L. Legorreta, and S. Dardan, *Collaboration Systems and Technologies Track Co-Chairs: Jay F. Nunamaker Jr. and Robert O. Briggs.*

Muhamad Nabeel Sarwar received his B.S degree in Computer Science from The Bahauddin Zakariya University Multan (BZU), Pakistan, in 2019. His Final Year Project "Fake news detection using machine learning". He earned his M.S degree in Computer Science from University of Okara, Punjab, Pakistan. His research interest includes machine learning, deep learning, natural language processing and computer vision. He has been part of research projects such as lungs cancer detection, early diagnosis of covid-19 and sentimental analysis. During his MS degree He was research assistant at university of Okara. Currently he is teaching as visiting faculty at university of okara Pakistan alongside his business setup as entrepreneur.

Riaz UlAmin (Associate Professor, Faculty of Computing, University of Okara) has been serving in various departments for over 17 years. He has over 30 publications and completed several funded research projects. His core research expertise is in the field of distributed systems, System Security in general and Digital forensics in particular. He has worked for industry providing viable solution using different process mining techniques and tools such as DISCO and PROM. While working with different forensic tools and approaches he has an extensive exposure to different ISO standards such as ISO 17799 and COBIT. Currently, he is leading research group who is investigating digital forensics to support local context with problem in several dimensions such as complexity, diversity, consistency and correlation; volume and universal time lining.

Sidra Jabeen earned her bachelor's degree in computer science with distinction from COMSATS institute of information technology, Lahore Pakistan in 2013. Her final year project" Enhancing IEEE 802.15.7 using OFDM and WDM" won grass root research initiative award in 2013. She is currently a student of master in computer science at University of Okara. Her areas of research interest are computer communication and networking, machine learning, Artificial intelligence and practical application of computer science in education. She is working as Secondary school educator in Punjab school education department since 2015 in okara. Her pervious publications are in local conference as well as in international journal of science and technology.