International Journal of Advanced Science and Technology
Vol. 29, No. 8, (2020), pp.4550-4558

An Ensemble Approach For Fake News Detection

Anil Singh Parihar[1], Swapnil Biswas[2], Vaishali Yadav[3]
Machine Learning Research Laboratory, Deptt. of Computer Science and Engineering
Delhi Technological University
New Delhi, India
Email: [1]*parihar.anil@gmail.com*, [2]*swapnil_bt2k16@dtu.ac.in*, [3]*vy074670@gmail.com*

Abstract

"Fake News" is a term used to represent false news or propaganda-driven news which leads to the spread of misinformation via traditional media like print or non-traditional media like social media. This has become a major challenge in recent years due to the lack of computational tools. In this paper, we have proposed an Ensemble Approach for detecting fake news. Different classifiers are combined in order to filter the noise and to overcome the drawbacks of a single classifier. Extensive experiments are conducted on the dataset and different deep learning approaches such as Bag of Words, Sentimental Analysis, and Long-Short Memory Networks (LSTM) have been incorporated for the implicit and explicit features of text and ensembled using Max Voting by the use of weights. We have compared our results with existing models and our model gives accuracy of 92% on "Getting Real about Fake News" Dataset.

Keywords— *Fake News, Deep Learning, Bag of Words, Sentiment Analysis, Recurrent Neural Network, LSTM*

I. INTRODUCTION

Over recent years, social media platforms emerge as hot spots for news websites and articles. Due to the lack of verification, there has also been a huge increase in cases of fake news across these platforms. The rise of fake news over different media is a matter of serious concern today and various techniques and methodologies are being worked on in the world to curb fake news in the world. Unaware people can be easily misled by false news and information that they consume online. The German government in 2017 said, *"We are dealing with a phenomenon of a dimension that we have not seen before."* Each and every country in the world is facing the challenge of countering the rise of fake news. It has become a popular tendency to create and spread false information and manipulate the people's perception and thinking by giving an illusion that what they are consuming is actually true

We can have a look at the statistics provided by Twitter regarding fake news. Abstudybconducted by Massachusetts Institutebof Technologyb onb thebspreadb ofb fakeb newsb onb Twitterb showedbthe pointsbthat:

- Fake news articlesbor postsbare 70%bmore likely tobbebretweeted.
- Time taken by actual stories is about six times more than fake stories to reach 1,500 people.
- Posts with false informationbare retweeted more widely thanbb trueb onesb statementsb at every depthbof a "cascade," whichbisbhow it refers to unbrokenb tweetb chains.b Theseb chains travelb10btob20 times more quicklybthanbfacts.

It has been noted in the study that false or fake information is more enticing. "False news is more novel, and people are more likely to share novel information.", says Sinan Aral, professor at MIT Sloan School of Management. Twitter also reports creation of unverified and "spammy" accounts on its platform. Recommender Systems, which are "information filtering software tools responsible for delivering information to end users in a safe environment" [2], are also proving to be inadequate in combating the rise of false information across platforms.

Fig 1. Statistics on how Twitter detects and deals with the spam accounts

We wish to work on distinguishing between real and fake news using the linguistic features and the sentiment extracted from the text content of the news. The sentiment may be Document Level, Sentence Level or Aspect Level [1]. Just like every other field, Convolutional Neural Networks (CNN) also find application in fake news detection and it's based on "Artificial Neural Networks (ANN), with neurons having weights and biases, which can be learned" [3].

II. Related Works

Criticalbamountbofbworkbhasbbeenbperformedbin the fieldbofbfakebnewsbdetection. This problem relies more on finding the various features of the news so that appropriate models can be applied to detect the news. As individuals have used their own set of extracted features, the algorithms and models are applied in the same manner. As stated in, Parihar et al have applied various machine learning approaches, Ayat Abedalla, Aisha Al-Sadi & Malak Abdullah] stated some ensemble methods which used Convolutional Neural Network (CNN) and Dense Neural Networks (LSTM) for classification and feature extraction of the text and in order to solve the overfitting problem, Bi-LSTM and hybrid approach of CNN were combined [1] [6] [7] [8].

Few papers have used Data mining approaches while many have used different Machine Learning algorithms. There exists some papers which have focussed more on user behaviour i.e. the users promoting the fake news and the source from where it is generated. One such model is the "CSI model" [13]].

In general, deep learning based methods give higher accuracy compared to other hand crafted feature -based approaches. In [15]], the authors proposed a framework which could learn and comment emotion for both publishers as well as users. It could also exploit content and social emotions simultaneously for fake news detection. In [14], the author presented a web application that detected fake tweets and users who were continuously promoting fake information on twitter. In [11], the authors proposed a Recurrent Neural Network with attention mechanism and they also incorporated image features into text and social context for fake news detection. Outputs of LSTM were utilized by using it with visual features. While some have also used an ensemble method of Machine learning algorithms as in [12]. We have proposed an ensemble approach of classifiers which outperform the state-of-art. Images associated with news posts can also be first enhanced [4] [5] and then deep learning methods may be applied on them to classify the news posts as fake/real.

III. Proposed Methodology

In this section, we introduce and compare the components of our methodology. We will be using the **Boosting technique** with dynamic weights instead of hard-coded weights as an improvement of previous papers. The models that we have used for training the explicit and implicit features of text are mentioned in this section.

A. *Model Overview*

The objective is to extend the field of work in detecting the fake news using linguistic features and sentiment extracted from the text content of the news. For this purpose, we have ensembled three classifiers i.e. Sentiment analysis, Bag-of-words, and Long-short-memory networks (LSTM) which will be discussed below.

B. Explicit Features of Text: Sentimental Analysis

Understanding the belief, opinion, and emotions of people is necessary for detecting whether the news is fake or not. This is analysed by Sentiment Analysis where it tells us about the degree of positiveness and negativeness in the news. Thebbasisbofbsentimentbanalysisbisbto determine theb polarityb ofb newsb inb order bto know bhow positivebor negativebthebnewsbis whether itbis ab wholeb document,b ab paragraphb or ab single clause.bItbis the most common classification tool which analyzes the text and tells the sentiment whether it is positive or negative. It provides a more comprehensive view as it can create a multitude of feelings as a feature vector.

Sentiment analysis works by first breaking the text into its component parts i.e. sentences, tokens, phrases followed by assigning a sentiment score each phrase and document from -1 to +1. The system works in the same way as our mind works. A sentiment library is there which contains a large number of phrases (good work, nice team etc.) and adjectives (wonderful, awesome etc.) that have been hand-scored by coders. This sentiment library is used by the system to identify the polarity of the text. The guidelines are decided by the coders in some systems. The sentiment library is also updated from time to time by adding new phrases and adjectives and the scores are combined for a multi-layered system. This way, the polarity of the text is determined by using sentiment analysis. The output of the sentimental analysis goes well with human evaluation.

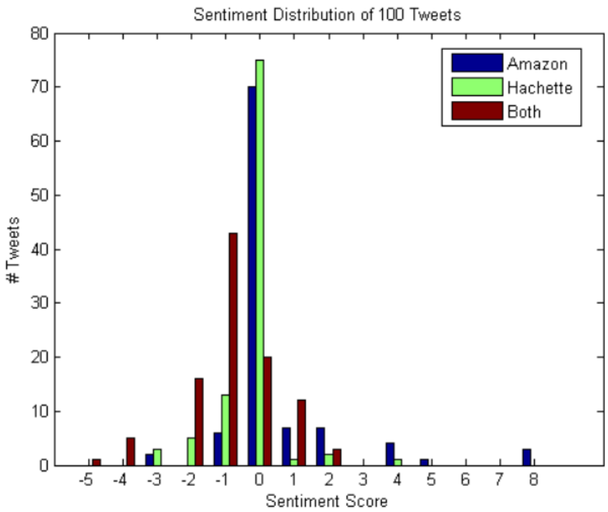

Fig 2. Sentiment Analysis of News Graph

C. Bag of Words for Linguistic Features of Text

As the text cannot be fed directly into the algorithm along with the Machine Learning algorithms do not work on texts. So Bag Of Words (BoW) is used for preprocessing the text. Bag of words model used for retrieving the information from the text by making the bag (multiset) of words of the text where ordering and grammar are disregarded but the multicity of words are kept. For each text in a news item we find the tokenized word and frequency of each token and create a matrix. The frequency of each word in the text is used as a feature in the classifier. It is also the main feature used by Fontanarava et al [10] in their model for credibility assessment.

Let d1, d2, and d3 are the documents containing some text. The frequency of each word is stored in a matrix for each document.

d1- it was the best time
d2- it was the worst time
d3- the age of goodness

The text is tokenized and the unique words are created as: “it”, “was”,” the” “best”, “time”, “worst”, “age”, “of”, “goodness”. And a matrix is created:

	it	was	the	best	time	worst	age	of	goodness
d1	1	1	1	1	1	0	0	0	0
d2	1	1	1	0	1	1	0	0	0
d3	0	0	1	0	0	0	1	1	1

Fig 4. Bag of Words Model

Bag of Words is simple to use and it provides a lot of flexibility for specific text data. It is known to show better results for text classification and so as in predicting the news whether it is fake or real.

D. Implicit Features of Text

Other than the polarity of the text, there are implicit features of the text which helps in detecting the news. So, We have used LTSM to take care of the implicit features of the text with the higher accuracy displayed by Yang et al [16] (as shown in the table below), this higher accuracy can daily be attributed to the long term dependencies supported by LSTM while forgotten by the Vanilla RNNs.

Long Short Memory Networks are mainly used in deep learning where there exist interdependencies in the data. It has feedback connections unlike feedforward neural networks which consists of a cell and gates. LSTM introduces gates which leak as to how much is to be learned and how much needs to be forgotten thus reducing this problem. The gates regulate the flow of the information in the unit and the cell remembers the values over the arbitrary interval of time.
Input Gate: This gate caters to the problem of whether the input should be updated and regulate the flow of information inside the cell.
*Forget Gat*e: This gate decides should memory be set to 0 or not.
Output Gate: It handles if the current info will be displayed to the next cell or not.

Here, the sigmoid function is used as the activation function of the LSTM unit. We chose LTSM instead of RNN because of higher accuracy displayed by Yang et al. This higher accuracy can daily be attributed to the long term dependencies supported by LTSM while forgotten by the Vanilla RNNs.

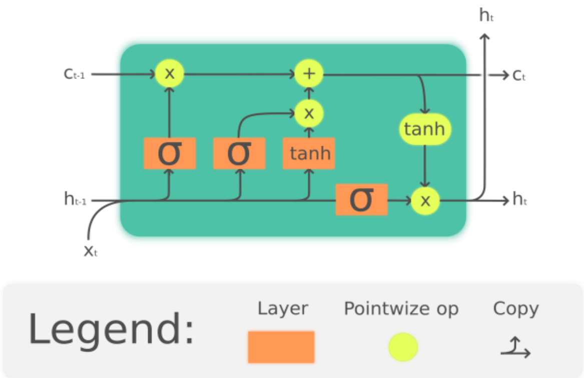

Fig 5. LSTM Architecture

The figure shows the LSTM working where:
C_{t-1} : memory from the last LSTM unit.
H_{t-1} : the output of the last LSTM unit
X: current input
σ : sigmoid layer
tanh: Tanh layer
C_t : updated memory
h_t : current output

Also, the interdependent tuning of hyper-parameters i.e. Learning rate and Network size make it more appropriate to use. The implicit text features cover a large range of features and act efficiently to provide holistic text attributes compared to just the sentiment analysis used as the explicit feature.

E. Design Architecture

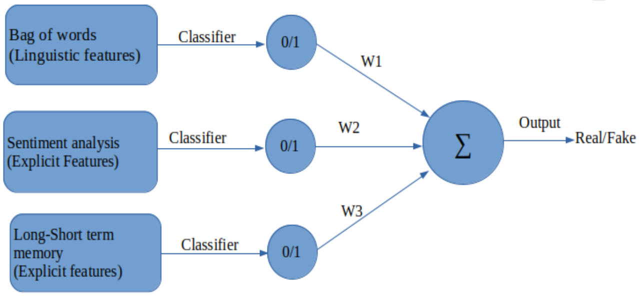

Fig 6. Design Architecture of Proposed Methodology

Different machine learning algorithms are applied for implicit and explicit features of the text of the news. The concept of weights is incorporated after this to ensemble the models using Max Voting. If any of the two models say the news is false then it is false. This way final output is generated to detect the news whether it is real or fake.

IV. IMPLEMENTATION

As discussed in Sec. III “Proposed Methodology”, we have attempted to build an ensemble approach to detecting fake news. Before explaining the results carried out, let us first describe the dataset.

A. Dataset and Tools

We have used TensorFlow which is a python library [17] Google CoLab was also used for the purpose of implementation.

The dataset we have used is called “Alldata”. Theb datasetb containsbmetadata band btextb from 244 websites and representsb12,999bpostsbin total for ab30bdays’ time period.bThebdatabwasbextracted usingbthe *webhose.io* API; sincebit's coming from their crawler,bnot allbwebsites identified by the BSbDetectorbarebpresentbin this dataset. Each websitebwasblabeled accordingbto thebBS Detectoras documented here.bData sourcesbthat were missingbablabel werebsimply assignedba labelbof "bs".bAlso some stories covering fake news in the news have been added to the dataset.

Dataset has 20 columns which includes image urls, likes, comments, shares, type etc. This whole dataset is neatly arranged in a CSV file. We are using only id, text, image and label information in our project.

B. Sentimental Analysis for Explicit Features of Text

As mentioned by Parihar et al [1], sentimental analysis can be performed at various levels. We have used Textblob API followed by the AdaBoost classifier tobgetbsentiment frombthe textbextractionbmodule.

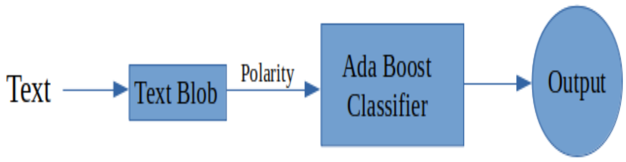

Fig 7. Sentiment Analysis Flowchart

C. Bag of Words Linguistic Features

We have used a CountVectorizer function of the sklearn library for this. We get the count of the frequency of each word and the length of the whole vocabulary as its output. Gradient Boost Classifier gives the best results for this.

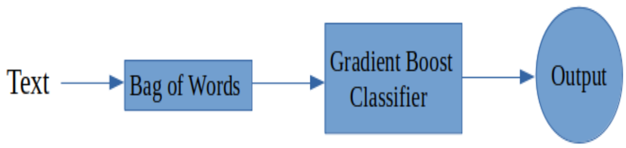

Fig 8. Bag of words flowchart

D. LSTM for Implicit Features of Text

Text is first tokenized followed by an embedding layer so as to obtain a vector to get the idea of its meaning and to get the recurrent structure of LSTM, we applied the dropout layer. Then a dense layer is applied using Softmax as activation in order to generate the output in the terms of fake or real.

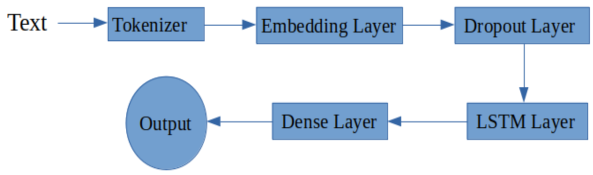

Fig 9. LSTM Flowchart

V. RESULTS AND EXPERIMENTS

A. Fake News Detection Metrics

Through our implementation of the above techniques, we obtained encouraging results in detection of fake news. The metrics of measuring correctness in detecting the presence of fake or false news and information that we are using are given below. The metrics have been used as per their feasibility in different cases:

- *Confusion Matrix*: It measures the performance of any classification algorithm or model on a test data whenever actual values are known. In a way, it provides the summary of the results that are predicted by an algorithm as compared to the actual values, essentially giving us an insight into the performance of the model/algorithm.

	Classb1 Predictedb	Classb2 Predicted
Classb1 Actual	*TruebPositive (TP)*	*FalsebNegative (FN)*
Classb2 Actual	*FalsebPositiveb(FP)*	*TruebNegative (TN)*

All other metrics of performance that we have used to find their basis in the Confusion Matrix itself.

- *Accuracy*: It is a measure of how correctly the classes are being predicted by the model. Itbisbgiven bybthe followingbformula:

$$\textbf{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- *Recall*: Recallbcanbbebdefinedbasbthebratiob ofb the total bnumberb of bcorrectly classified positive examplesb dividedb to thebtotal number of positivebexamples. It is given by the following formula:

$$\textbf{\textit{Recall}} = \frac{TP}{TP + FN}$$

- *Precision*: It is given bybthebratiobof correctly classifiedb positiveb examples tobthe total numberbofbpredicted positivebexamples. It is given by the formula:

$$\textbf{Precision} = \frac{TP}{TP + FP}$$

- *F1 Score/F-Measure*: It is the Harmonic Mean of Precision and Recall scores and is given by the following formula:

$$\textbf{F - measure} = \frac{2*\text{Recall}*\text{Precision}}{\text{Recall} + \text{Precision}}$$

B. Results

We now present our results and a comparative view of our results is provided with previously existing works at the end. Table I shows the output results of the matrices:

TABLE I. RESULTS

Method	Accuracy	Precision	Recall	F1- score
Bag of Words	92%	93%	92%	90%
Sentimental Analysis	88.65%	79%	89%	83%
LSTM	95.5%	95%	95%	95%

Finally, the 3 methods that we have implemented were ensembled by Boosting using Max Voting. It is one of the primary methods of combining predictions from multiple machine learning algorithms and is majorly used in classification problems. In max voting, each model makes a prediction and votes for a sample. The sample with the maximum number of votes is then chosen in the final prediction. Hence, the final results after ensembling are given below:

Accuracy: 91.8%
Precision: 83.89%
Recall: 99%
F1-score: 90.8%

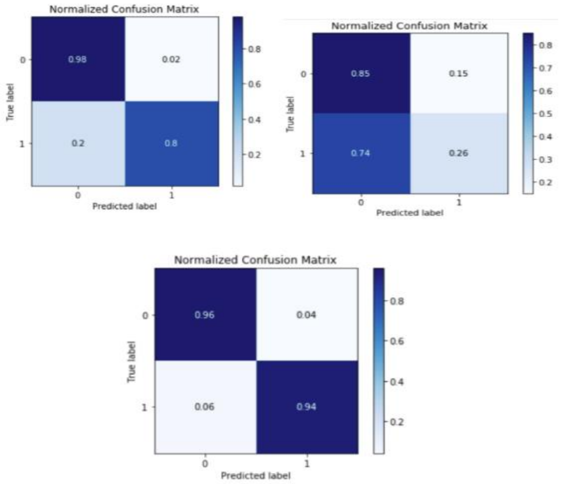

Fig 10. Confusion matrices of (a) Bag-of -words, (b) Sentiment analysis (c) LSTM

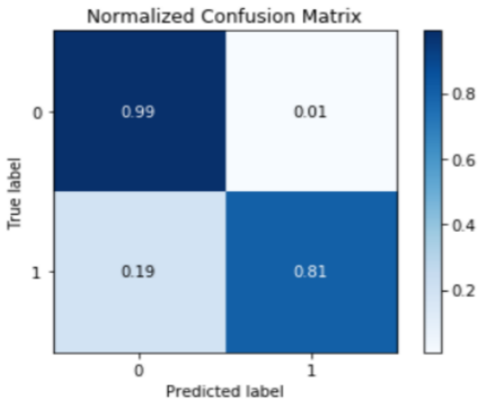

Fig 11. Confusion matrix for Ensemble method of three classifiers

VI. Conclusion and Future Work

In this paper, we have discussed in detail about the spread of fake and false news across the internet. We saw the negative implications and consequences attached to fake news spreading to unaware people. We, in this project, have proposed a novel ensemble approach to identify false news pieces and have come up with promising results. We worked on ensembling the techniques we worked on and then Boosting (using max-voting), which serves as our novelty feature in the model architecture that we propose. Our proposed methodology performs satisfactorily well with an accuracy measure value of 91.82%.

Now, we all know that news isn't just composed of text, but is also accompanied by images. Whileb allb theb featuresb explainedb upb untilbnow haveb shownb tob haveb ab largeb effectb on determiningb fakeb news,b ab lotb ofb featuresb are implicitbinbnaturebtobimagesband can'tbbe found explicitly.bHence, analyzing the features of images (both explicit and implicit) and using it in conjunction with our proposed model can help in developing a high-performance system for detecting fake news, Studying the features of images can be done by using Convolutional NeuralbNetworksb(CNN)band that is something we wish to work on in the future.

References

[1] A. S. Parihar and Bhagyanidhi, "A Study on Sentiment Analysis of Product Reviews," 2018 International Conference on Soft-computing and Network Security (ICSNS), Coimbatore, 2018, pp. 1-5, doi: 10.1109/ICSNS.2018.8573681.
[2] R. Katarya and A. S. Parihar, "Study on privacy preserving recommender systems datasets," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, 2017, pp. 335-338, doi: 10.1109/ICICI.2017.8365367.
[3] A. S. Parihar, "A study on brain tumor segmentation using convolution neural network," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, 2017, pp. 198-201, doi: 10.1109/ICICI.2017.8365336.
[4] A. S. Parihar and K. Singh, "A study on Retinex based method for image enhancement," 2018 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, 2018, pp. 619-624, doi: 10.1109/ICISC.2018.8398874.
[5] Parihar, Anil & Singh, Kavinder. (2020). Illumination Estimation for Nature Preserving low-light image enhancement. 10.36227/techrxiv.12236780.v1.
[6] Abdullah-All-Tanvir, et al. "Detecting Fake News Using Machine Learning and Deep Learning Algorithms." *2019 7th International Conference on Smart Computing & Communications (ICSCC)*, 2019, doi:10.1109/icscc.2019.8843612.
[7] Abedalla, Ayat, et al. "A Closer Look at Fake News Detection." *Proceedings of the 2019 3rd International Conference on Advances in Artificial Intelligence*, 2019, doi:10.1145/3369114.3369149.
[8] Ajao, Oluwaseun, et al. "Fake News Identification on Twitter with Hybrid CNN and RNN Models."

Proceedings of the 9th International Conference on Social Media and Society, 2018, doi:10.1145/3217804.3217917.

[9] Dizikes, Peter. “Study: On Twitter, False News Travels Faster than True Stories.” *MIT News*, 8 Mar. 2018, news.mit.edu/2018/study-twitter-false-news-travels-faster-true-stories-0308.

[10] Fontanarava, Julien, et al. “An Ensemble Method for the Credibility Assessment of User-Generated Content.” *Proceedings of the International Conference on Web Intelligence*, 2017, doi:10.1145/3106426.3106464.

[11] Jin, Zhiwei, et al. “Multimodal Fusion with Recurrent Neural Networks for Rumor Detection on Microblogs.” *Proceedings of the 2017 ACM on Multimedia Conference - MM '17*, 2017, doi:10.1145/3123266.3123454.

[12] Kaliyar, Rohit Kumar, et al. “Multiclass Fake News Detection Using Ensemble Machine Learning.” *2019 IEEE 9th International Conference on Advanced Computing (IACC)*, 2019, doi:10.1109/iacc48062.2019.8971579.

[13] Ruchansky, Natali, et al. “Csi.” *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2017, doi:10.1145/3132847.3132877.

[14] Saez-Trumper, Diego. “Fake Tweet Buster.” *Proceedings of the 25th ACM Conference on Hypertext and Social Media - HT '14*, 2014, doi:10.1145/2631775.2631786.

[15] Guo, Chuan & Cao, Juan & Zhang, Xueyao & Shu, Kai & Yu, Miao. (2019). “Exploiting Emotions for Fake News Detection on Social Media”.

[16] Lei Zheng, Jiawei Zhang, Qingcai Cui, Zhoujun Li, Philip S. Yu Yang Yang, "TI-CNN: Convolutional Neural Networks for Fake News Detection," arXiv preprint arXiv:1806.00749, June 2018

[17] Abadi, Martín, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin et al. "TensorFlow: A System for Large-Scale Machine Learning." In OSDI, vol. 16, pp. 265-283. 2016.