

Fake News and Rumour Detection on Social Media

M.Uma Devi, Harsh Tyagi, Alok Jindal

Abstract: *PC Mediated Communication (CMC) advances like, for example, online journals, Twitter, Reddit, Facebook and other web based life presently have such a large number of dynamic clients that they have turned into an ideal stage for news conveyance on a mass scale. Such a mass scale news conveyance framework accompanies a proviso of faulty veracity. Building up the unwavering quality of data online is a strenuous and an overwhelming test yet it is basically essential particularly amid the time-touchy circumstances, for example, genuine crises which can have destructive impact on people and society. 2016 US Presidential race is an encapsulation of the previously mentioned crisis. In a study, it was concluded that the public's engagement with phoney news through Facebook was higher than through standard sources. So as to battle the spread of malevolent and unplanned falsehood in online networking we built up a model to recognise counterfeit news. Counterfeit news recognition is a procedure of classifying news and estimating it on the continuum of veracity. Detection is done by classifying and clustering assertions made about the event followed by veracity assessment methods emerging from linguistic cue, characteristics of the people involved and network propagation dynamics..*

Keywords: *Fake news,rumour detection,machine learning,twitter*

I. INTRODUCTION

With the advent of social media and the freedom to write anything, comes the issue of credibility. Whether the news is trustworthy, whether the news is actionable or is it to provoke certain unnecessary actions. Initially when the print media was the main source of news, vetting process was very rigorous. The credibility of the source was the most important factor as it could make or break the future of the newspaper publisher. But, in case of social media, where there are millions and millions of user and thus millions of sources of news, it is impossible to verify the veracity of the news and segregating them from fake news. There are various sources of news such as television, radios, news channels, social media websites like, facebook, twitter or whatsapp, blogging sites etc. There are numerous sources of news.

According to online sources, there are around 500 million tweets a day. This is just on one normal day. Now, if

incase of any natural calamity or any unfortunate event, the count of daily increases many folds. And these enormous amount could contain various malicious news. And the motive behind those tweets or blogs is to spread panic, or hate during the time of national distress.

For example, a company called "Infibeam" in India, lost their stock value by 70% just because of a fake news which went viral. The second example, during the Boston marathon bombing in USA, fake news was spread and caused a lot of panic.

To prevent all this from happening, this survey paper catalogues various techniques employed by previous research paper and shows the algorithm uses, dataset used and accuracy of those techniques. At last, this paper contains a section called proposed model, which contains techniques which could be used in future applications.

II. RELATED WORKS

This section contains the survey of papers we have done. In further sub-sections we will discuss about the most common machine learning algorithms used, frequently used datasets and the results of those papers.

A. Automatically identifying fake news in popular twitter threads

The first paper we studied was "automatically identifying fake news in popular twitter threads" was written by Cody Buntain, Jennifer Golbeck published in 2017, IEEE. Dataset "pheme" was used using python language. Algorithm used was classification model with Accuracy rate of 70.28% in identifying fake news. Advantages of the paper, leveraging non-expert crowd sourced workers rather than journalists.

B. Fake news identification on twitter with hybrid CNN and RNN models

The second paper we studied was "Fake news identification on twitter with hybrid CNN and RNN models" was written by Dluwaseun Ajao, Deepayan Bhowmik, Shahrzad Zatrgeri published in 2018. Dataset employed by them was, Charlie Hebdo case, Ottawa shooting case and many others as shown in table 1. They were used using python language and algorithm used by them was CNN, RNN with accuracy rate of 82% in classification task. Advantages of this paper is that the plain LSTM model achieved the best performance in terms of accuracy f1 score and recall score.

Revised Manuscript Received on September 22, 2019.

Mrs.M.Uma Devi, Department of CSE, SRMIST, Chennai, India, umadevi.mu@ktr.srmuniv.ac.in

Harsh Tyagi, Department of CSE, SRMIST, Chennai, India, HarshTyagi_sa@srmuniv.edu.in

Alok Jindal, Department of CSE, SRMIST, Chennai, India, aj.alokit27@gmail.com

C. Evaluating machine learning algorithms for fake news detection.

The third paper we studied was "Evaluating machine learning algorithms for fake news detection", was written by Shlok glida published in 2017, IEEE. The Dataset used by them was signal media and used language python. Algorithms used were SVM, stochastic gradient descent and decision tree with accuracy rate of 77.2% in identifying fake news. Advantage, PCGs are good for fake ones classification.

D. Study of hoax news detection using naive bayes classifier

The fourth paper we studied was "Study of hoax news detection using naive bayes classifier" and was written by Iggrid yanuar risca pratiwi in 2018 IEEE. Dataset used was of its own of 250 pages of hoax articles was used using python language. Algorithm used was Naive bayes with weighted approach with accuracy rate of 82.6%.

E. Improving spam detection in Online social networks

The fifth paper we studied was "Improving spam detection in Online social networks" was written by Arushi gupta, rishabh kaushal published in 2015 IEEE. Twitter dataset was used using python language and Java. Algorithm which was used, Naive bayes with accuracy rate of 83.6%.

F. Automatic Online Fake News Detection

Combining Content and Social Signals The sixth paper we studied was "Automatic Online Fake News Detection Combining Content and Social Signals" written by Macro.L and published in 2018 IEEE. Dataset used was, fakenews.net, politifact and buzzfeed news using python language. Algorithm used, Logistic Regression with accuracy rate of 81.7%.

G. Fake news detection with deep diffusive network model.

The sixth paper we studied was "Fake news detection with deep diffusive network model" written by Jiawei Zhang et.al. 2018 and was published in ACE conference. Dataset used was politifact news and used python language. Algorithm used, back propagation RNN and gave 63% accuracy.

E. Common algorithms

• Naïve Bayes:-

Naive Bayes, these classifiers are a group of fundamental "probabilistic classifiers" in perspective on applying Bayes' hypothesis with strong (opportunity assumptions between the features. Naive Bayes classifiers are exceedingly adaptable,

requiring different parameters straight in the amount of elements in a learning issue. Most noteworthy likelihood getting ready should be conceivable by surveying a closed casing explanation, which takes straight time, instead of by expensive iterative gauge as used for some various types of classifiers.

▪ Recurrent Neural network:-

Recurrent Neural Network (RNN) is one of the class of Artificial Neural Network (ANN) in which the connection between the nodes of the neural network forms a directed graph along a sequence. RNN cannot be used to process the entire article because it is computationally infeasible therefore certain methods like cross-validation were used to determine a hyper- parameter that could limit its time steps. $ht = (xtWx + ht-1Wh + b)$

▪ Convolutional Neural Network:-

Convolutional Neural Network is one of the types of deep feed forward neural network. They uses multilayer perceptron but a special one in order to minimize preprocessing. There are 3 kinds of layers in the neural network, input layer that takes the input weights and holds them, hidden layers, these performs operations on the input weights and transfer it to other hidden layers, output layer which gives the classification. With max pooling a layer is added that selects the largest from the scaled inputs. The results are added and transformed via a ReLU unit, whose output is finally sent to a sigmoid nonlinearity to predict the label.

▪ Back Propagation neural network:-

Back Propagation is another class of neural network in which the weights are adjusted by propagating the errors backwards. The difference between the observed and desired output is calculated which is nothing but the deviation or error. This error is then used to adjust the weights. This kind of algorithm is commonly used to train deep neural networks. This neural network also uses gradient descent optimisation algorithm to adjust the weights of the neurons.

▪ Random Forests:-

Random forest is a learning strategy for arrangement, relapse and different assignments, that work by developing a large number of decision trees at preparing time and yielding the class that is the method of the classes (classification) or mean prediction (regression) of the individual trees. Random forest corrects for decision trees' propensity for overfitting to their preparation set.

▪ Decision Trees:-

Decision Tree is an algorithm which uses a structure similar to tree structure, in which each root node represents a class or a final decision and the intermediaries represents the test on an attribute. The path from root to leaf indicates the classification rules. These are one of most important tool which is used in machine learning. They are also used in operational research, to identify the strategy to reach the desired goal.

▪ Logistic Regression:-

One of the simplest model is this logit model. It is a linear model and was implemented so that it could serve as the baseline for other models which are far more sophisticated. Logistic (sigmoid) cross-entropy loss is used which is defined as:

TABLE 1

PAPER	AUTHOR	ALGORITHM USED	YEAR OF PUBLICATION	DATASET	LANGUAGE USED	FINDINGS	ADVANTAGES
Automatically Identifying Fake News in Popular Twitter Threads	Cody Buntain, Jennifer Golbeck	Classification model	2017, IEEE	Pheme	Python	70.28% Accuracy in identifying fake news	Leveraging non-expert crowdsourced workers rather than journalists.
Fake News Identification on Twitter with Hybrid CNN and RNN Models	Dluwaseun Ajao, Deepayan Bhowmik, Shahrzad Zargari	CNN, RNN	2018	CharlieHebdo, SydneySiege, OttawaSooting, Germanwings Crash, Ferguson Shooting	Python	82% accuracy in classification task.	The plain LSTM model achieved the best performance in terms of accuracy, f1 score and recall score.
Evaluating machine learning algorithms for fake news detection	ShlokGilda	SVM, Stochastic gradient descent and decision tree.	2017, IEEE	dataset published by Signal media	Python	Identifies fake news with an accuracy of 77.2%	PCGs are good for fake ones classification.
Study of hoax news detection using naive Bayes classifier	Ingrid Yanuar Risca Pratiwi	Naive bayes with weighted approach.	2018 IEEE	Own dataset of 250 pages of hoax	Python	82.6% Accuracy	
Improving spam detection in Online Social Networks	Arushi Gupta, Rishabh Kaushal	Naive bayes	2015, IEEE	Twitter dataset	Python, Java	83.6% Accuracy	Combined the accuracy of major machine learning algorithms.
Automatic Online Fake News Detection Combining Content and Social Signals	Marco. L	Logistic regression	2018, IEEE	Fakenews Net, Politifact and buzzfeed news	Python	81.7% accuracy rate	Work can be extended to chatbots and clickbait titles.
Fake News Detection with Deep Diffusive Network Model	Jiawei Zhang, Limeng Cui, Yanjie Fu, Fisher B. Gouza	Back Propagation RNN	2018 ACM Conference	Politifact Satsset	Python	63% on news articles	Shows that model can be created for articles as well.

$$J = y \log y^{\wedge} (1 - y) \log(1 - y^{\wedge})$$

B.Frequent dataset

• PHEME dataset

The dataset incorporates fake tweets, gathered and annotated for the use in the training model. These bits of fake tweets are related with 9 distinctive breaking news. It contains Twitter discussions which are started by a rumourous tweet; the discussions incorporate tweets reacting to those rumourous tweets. These tweets have been clarified for help, assurance, and evidentiality. The dataset contains 330 conversational strings (297 in English, and 33 in German).

• CredBank dataset

The CREDDBANK data was gathered between October 2014 and 2015. It is a gathering of gushing tweets followed over this period, themes in this tweet stream, points named occasions or non occasions, occasions clarified with believability evaluations. The information is spread crosswise over four documents.

• BuzzFeed dataset

Buzzfeed contains the data from facebook. This data includes political fake news. Buzzfeed reviews news from the various facebook pages.

• Politifact dataset

This is a website which tests the accuracy of a political fact. Many papers have used their API to fetch the news and used them to train their model.

• Twitter dataset

Twitter provides API to fetch live tweets as well as past tweets. Research paper by Arushi Gupta et. al used twitter to create their own dataset. Soroush Vosoughi fetched tweets related to Boston Bombing and ebola epidemic to create a self-annotated dataset.

III. FINDINGS

The accuracy of the models studied for this paper ranges from 63% to 85%. The paper by Jiawei Zhang et. al. used back propagation neural net and politifact data, gives and accuracy of 63%. Whereas, the paper by Arushi Gupta et. al. boasts the accuracy of about 83.6% with naïve bayes as the classifier and self-annotated twitter dataset.

IV. PROPOSED MODEL

The model proposed in this paper uses, buzzed dataset. This dataset contains various buzzed articles. The articles are divided into two folders, 'fake' and 'real'. The data is present in json format. When json file is parsed, the news is present in the 'text' key. We focused mainly on the bellow mentioned algorithm and then after implementing them, their accuracies are compared. The models are compared based on their 'f1-score, accuracy, and precision score'. All these metrics are defined in the coming sections.

A. Dataset.

The dataset used here is buzzed fake news dataset. From this dataset, we used the textual content and them stored them in the data-frame. The final data- frame has a text column and fake/real column. The later is the dependent variable and uses 0 to show real news and 1 for fake news. The text contains noise, which is removed during the preprocessing part.

B. Algorithms.

We analysed various different machine learning algorithms. The models used are, "Naive-bayes, decision tree, random forest, stochastic gradient descent and SVM". These models are tuned for the best results. The data fed into these models is converted to bag of words model and TF-IDF model. Hence, each machine learning model is applied on both, 'bag of words' as well as on 'TF-IDF model'. The results are them compared and shown in table 2.

C. Scoring.

- The scores are used to compare various models used in the implementation. The metrics on which the models are compared are:-

- F1 score: In the binary classification, the f1 score the measure of test's accuracy. It is the harmonic mean of 'recall' score and 'precision' score. It is calculated as :-
- $f1 = ((\text{recall} + \text{Precision}) / 2) - 1$
- Precision score: The precision score the number of true positives found among the total positives (true positive and false positives).
- Precision = true Positive / True Positive + false Positive.
- Accuracy score: It is the ratio of correct predictions to the total observations

V. IMPLEMENTATION

In this model, the used dataset is available online. It contains blogs or posts and are segregated into fake news and Real news folder. Each post or blog is present in son format. So, first the json files are parsed. The fake news posts are made into one data-frame and the real news blogs or posts are made into second data-frame.

In the following step, the data was cleaned. The data was made into lower case, all the numbers were removed and only textual content is used. At the last step, the data is lemmatised. The "stemming" was not chosen because "lemmatisation" preserves the root word.

To verify whether there is some kind of similarity between, fake news and real news, cosine similarity was calculated. The "cosine similarity", is the measure that calculates the cosine between two given vectors. This method does not measures the magnitude but the orientation. When two vectors have the same orientation, there cosine similarity is 1 and when the are orientated at 90 degrees, there cosine similarity is 0. Here the vectors means, the count vector tokenises the words and shows the word count in each senescence or row of a dataframe. It also encodes the data.

In our case the cosine similarity is quite high, which is around, 0.81. Hence, we can say that, the content and way of writing a fake news blog and real news blog is quite similar. In the next step we merged the fake news data frame and real news data frame. This is done so as to create a machine learning model. Then, the final data frame was converted into bag-of-words model and TF-IDF model. Here TF-IDF stands for "frequency-inverse document frequency". It is a weight is is often used for text-mining. The weight denotes the importance of the particular word in the whole cluster of words. The town different models are created just to compare the performance of each model.

The bag-of-words model and TF-IDF model is then fed into many machine learning model and the results are then compared. All the above process was done using the "scikit-learn" library of python. The results of each model's performance can be seen in the next section.

VI. RESULT

The machine learning algorithms used are, "Decision-tree classifier", "Random forest classifier", "K neighbour classifier", "Support vector machine classifier", "Stochastic gradient descent" and at last, the "naive Bayes classifier". Th results are shown in the table below.

Table II

	DT-		RF-		KN		SV		SG		NB	
	C		C		N		M		D			
	G W	F- IDF	G W	F- IDF	G W	F- IDF	G W	F- IDF	G W	F- IDF	G W	F- IDF
1	3.67	1.02	8.36	8.36	8.29	8.36	3.18	1.48	3.18	8.78	5.67	8.98
re cisi on	3.67	0.99	8.5	8.5	5.38	8.50	5.73	2.97	5.73	1.66	6.17	9.84
c cur acy	3.78	1.08	8.37	8.37	1.35	8.37	3.78	5.95	3.78	9.18	5.67	9.18

VII. CONCLUSIONS

Hence, we can conclude by saying that the best model for fake news classification is "naive-bayes" when used with "TF-IDF" instead of "Bag-of-word" model. The "stochastic gradient descent" model also performed extremely well, but the f1 score us slightly less than the f1 score of "naive-bayes" model.

REFERENCES

- 1) Oluwaseun Ajao, Deepayan Bhowmik, Shahrzad Zargari, Fake News Identification on Twitter with Hybrid CNN and RNN Models, UK, SMSociety, 2018.
- 2) Shlok Gilda, Evaluating machine learning algorithms for fake news detection, IEEE, 2017.
- 3) Arushi Gupta, Rishabh Kaushal, Improving spam detection in Online Social Networks, IEEE, 2015.
- 4) Ingrid Yanuar Risca Pratiwi, Study of hoax news detection using naïve bayes classifier, IEEE, 2018.
- 5) Marco L., Automatic Online Fake News Detection Combining Content and Social Signals, IEEE, 2018.
- 6) Jiawei Zhang, Limeng Cui, Yanjie Fu, Fisher B. Gouza, Fake News Detection with Deep Diffusive Network Model, ACM Conference, 2018.
- 7) Cody Buntain, Jennifer Golbeck, Automatically identifying fake news in popular twitter threads, IEEE, 2017..