MCA
Explore–Journal of Research
Peer Reviewed Journal
ISSN 2278 – 0297 **(Print)**
ISSN 2278 – 6414 **(Online)**

https://www.patnawomenscollege.in/journal

Fake News Detection using Machine Learning

• **Arshiya Kumari** • **Zainab Raza** • **Shruti Sinha**

Received : January 2021
Accepted : February 2021
Corresponding Author : Tapan Kant

Abstract: *Social media is allowing users to share content without checking the reliability of information over the internet. It has revolutionized the spread of news or other information without fact checking. People are unaware that they are consuming and sharing more fake news than real news. The rapid spread of any news (may be fake or real) can mislead the reader. Therefore, fake news detection on social media has become trendy in research community. It presents unique characteristic (i.e., linguistic and semantic) to distinguish fake or real. To probe the reliability of news or any information shared by users. We present a classification model and Natural Language Processing approach to detect fake news. We found that more than 57.7% of news are fake on social platform and only 42.3% are real news. The proposed technique has 98.3% of accuracy for training dataset. We also found that the posts which have more likes are shared frequently.*

Keywords : *hoaxes, ML, NLP, social-media, twitter, fake-news.*

Arshiya Kumari
MCA- IV Semester, Session: 2019-2022,
Patna Women's College, Patna University, Patna- 800001, Bihar, India

Zainab Raza
MCA- IV Semester, Session: 2019-2022,
Patna Women's College, Patna University, Patna- 800001, Bihar, India

Shruti Sinha
MCA- IV Semester, Session: 2019-2022,
Patna Women's College, Patna University, Patna- 800001, Bihar, India

Tapan Kant
Assistant Professor, Department of MCA,
Patna Women's College, Bailey Road,
Patna - 800001, Bihar, India
E-mail : tapan.mca@patnawomenscollege.in

Introduction:

Fake News represents fabricated news or propaganda, in which misinformation is communicated through traditional media channels such as print, television and also through non-traditional media channels, i.e., social media. The motive of spreading such news is basically misleading the readers, damaging the reputation of any entity, or to gain from sensationalism. It has the potential to become one of the greatest threats to democracy, free debate and the western order (K Langin, 2018).

Nowadays, fake news is being flooded through social media platforms such as Twitter, WhatsApp and Facebooks (Hunt Allcott et. al., 2017). People are sharing their sentiments and thoughts using these social media without checking the facts. There is much news shared across social media as compared to traditional news media. In a study, it is found that tweets that contain false entropy reach much faster (K Langin, 2018).

Natural Language Processing (NLP) could be used potentially to construct a method to automatically perceive fake news or hoax. However, the task to detect false news is challenging and requires models and comparison from real news or original news. It is also a challenging task to characterize the language of quotes and news written with different intention and degree (Hal et. al., 2014 and Jesse et. al., 2016).

In this paper, we are using NLP and Machine Learning Approach to hoaxes detection. The main idea behind our work, which constitutes its novelty, is based on the likes and dislikes of tweets and the use of NLP for the detection of hoax. We have used traditional dataset available in literature to train our model. We have also used traditional media sources (The Hindu, NewsLaundary, TheQuint, USNews, etc.), for verification and found that our technique has 98.3% of accuracy for training dataset. We have used more than 1 million tweets and more than 5000 likes or dislikes per tweets.

Background and Related Work:

There can be many forms of fake news, such as accidental errors made by news aggregators, utter false stories, or misleading stories developed to mislead or influence opinions. The different forms of fake news can affect people, government body, organizations, etc. which is different from facts (Kai et. al., 2018).

The main challenge behind the fake news detection is to define fake news. There are following types of fake news:

- A made-up story with an intention to deceive (Sabrina et. al. 2016).
- News articles that are intentionally and verifiably false, and could mislead readers (Hunt Allcott et. al., 2017).
- Fake news is a type of yellow journalism or propaganda that consists of deliberate misinformation or hoaxes spread via traditional print and broadcast news media or online social media (Hal et. al., 2014 and Jesse et. al., 2016)

In literature, there are many approaches have been proposed for detection of fake news or hoaxes as shown in fig. 1. These technique are broadly categorized as data-oriented, feature-oriented, model-oriented, and application oriented. Further, these techniques are categorized on the basis of methodology used for detecting fake news.

Initially, the hoax detection was performed in e-mail messages and webpages that usages keyword-based methods (Alan, 2004 and Pale et. al., 2005). A distance based methods was proposed by Ishak et al., 2012. A neural network and advanced text processing method was applied by Marin et. al., 2009. An anti-spam filters using evolutionary algorithms was used by Iryna et al., 2013. A logistic regression to automatically detect scam on web pages was proposed by Sharifi et al., 2011.

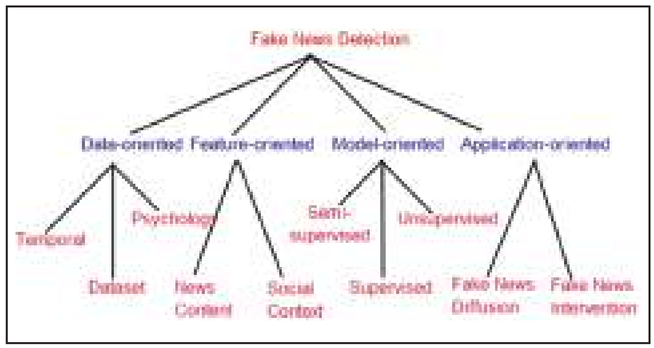

Fig. 1. Techniques of fake news detection on social media (Kai et. al., 2017)

A certain speech pattern of a speaker could be a sign to purposefully obscure the truth was proposed by Buller and Burgoon, 1994. Vague and Hedge words are qualifiers of statements of obscure meaning Choi et al., 2017 and Recasens et al., 2017. Unreliable news sources created with contrary intention and the level of accuracy of linguistic aspects was proposed by Nadia et al., 2015.

On the basis of the above study, we have set our objective as follows:

1. Detection hoax on user likes and dislikes.
2. Detection of hoax on linguistic aspects.

Our Approach :

In this section, we discuss the fake news detection process mathematically. We consider a tweet T as a news element tweeted by a user U at time t. Therefore, a tweet T can be represented as T = {p, u,rt, l, d, t}, where p is the content of a tweet, u is the id of twitter handle, rt is the number of times the tweet has been re-tweeted, l is the number of likes, d is number of dislikes and the t is the time of tweet.

In order to predict whether a tweet T is a fake or not, i.e., F(T): $\in$->{0,1} such that,

$$F\ (T)=\begin{cases}1, \textit{T is a fake news}\\ 0, \textit{Otherwise}\end{cases}$$

where, F is prediction function.

Therefore, the process of detecting fake news can be considered as a binary classification problem (Shapiro et. al., 2016; Kai et. al., 2018 and Kai et. al., 2017). This approach is accomplished into two phases: (i) feature extraction and (ii) construction of model. The process flow of our approach is shown in fig. 2. Our approach takes tweets as input. Further, the tweet is divided into two parts as **#tags** used in the tweet and body text of the tweet. Afterwards, weight score is assigned to each input. If the score is more than 25% and related to our research scope then the tweets are considered for processing else marked as unrelated. The weight score is calculated on basis of words and tags used in the tweets. In the final step, the relationship between tags and body text is established and again weight score is assigned to it. If the weight score is more than 75% then tweets are considered for classification otherwise discarded.

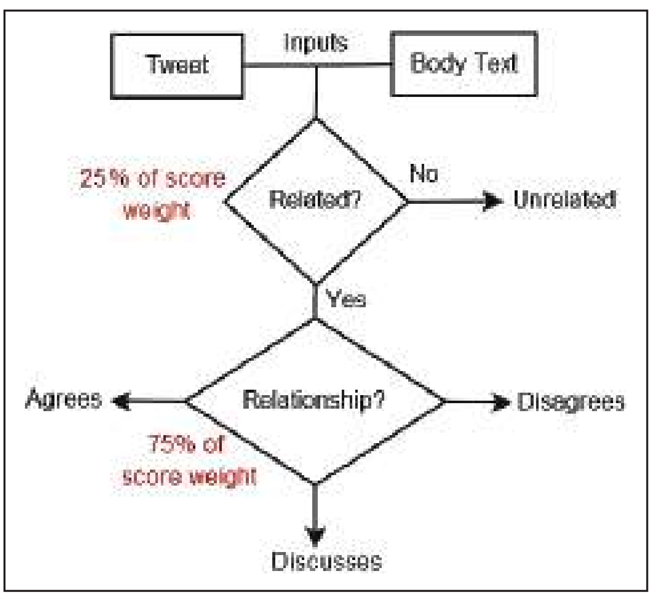

Fig. 2. Process of Fake New detection

Feature Extraction :

In traditional news media, fake news detection process generally relies on news contents. On the other hand, in case of social media, news is extracted in social context, auxiliary data and some additional data also helps in to detect fake news (Kai et. al., 2017). The features of news content in traditional media are:

Sources: Publisher and the Author of the news article (Kai et. al., 2017 and Hannah et. al., 2017).

Headline: It describes the content in crisp form (Kai et. al., 2017 and Hannah et. al., 2017).

Body Text: The detailed contents of the news are described (Kai et. al., 2017 and Hannah et. al., 2017).

Images: These are the visual content of the news.

In case of social media, we can derive many features of news such as number of likes or dislikes, number of times shared, date & time, the social engagements of the user, veracity, trends, etc (Kai et. al., 2017).

Model Construction:

In this subsection, we discuss the process of model construction of fake news detection. Our model is based on knowledge-based. The knowledge-based model mainly relies on comparison the content and features of news with reliable sources of news. It require external sources to check the facts that claims is true (Andreaset, 2014).

Classification Model :

The process of detecting fake news can be considered as a supervised binary classification problem (Shapiro et. al., 2016 and Kai et. al., 2017). A tweet T by user U, is considered as news with associated set of features {x_{tu} | u$\in$U}, where x_{tu} = 1 if u liked, re-tweeted post, and x_{tu}= 0 otherwise. The tweets are classified on the basis of their features. We have used Naïve Bayes classification algorithm.

$$P(h|D) = \frac{P(D|h)\ P(h))}{P(D)}$$

where,

P(h) = prior probability of hypothesish
P(D)= prior probability of training data D
P(h|D)= probability of h given D
P(D|h)= probability of D given h.

Generally, the most probable hypothesis given the training data *Maximum a posteriori* hypothesis h_{MAP}:

$$h_{map} = \arg \max_{h \in H} P(h|D)$$

$$= \arg \frac{\max_{h \in H} P(D|h)P(h)}{P(D)}$$

$$= \arg \max_{h \in H} P(D|h)P(h)$$

Result and Discussion :

In this section, we describe the data collection process and the relationship between data. Further, we discuss the results and our observation. Our experiment begins through data collection process. The collected data was noisy therefore, we have performed data cleaning process. Most of the tweets contain special characters, special symbols (i.e.,&, #,!, etc.), and numbers.

In order to perform classification on textual data, the lexical and syntactic features should be extracted. Therefore, we have performed word-level features such as frequency of words, uniqueness of words and phrases (ie. n-grams, and bag-of-words) approaches are used (Jakub et. al., 2020). In the second phase of our experiment, we train our model for the classification process of dataset. We have used training dataset available in (Kai et. al., 2018) to train our model and obtain results.

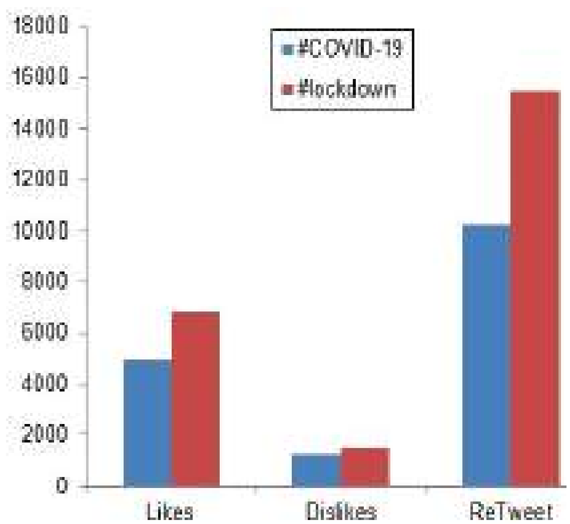

Fig. 3. The Average Number of Likes, Dislikes and Retweet on Search Text #COVID-19 & #lockdown

DataSet Collection :

Our dataset consists of public tweets' likes and dislikes collected from Twitter social media since March 2020 to December 2020. We collected data by using Twitter API tweepy in python2. The dataset collected with the search content **#COVID-19** and **#lockdown.** The resulting dataset, is composed of approx 1 million tweets from more than 5000 user and above 5000 likes per tweet and retweet. Among the tweets collected as dataset 42.3% of tweet were found true news and 57.7% were detected either as fake, false information or hoaxes as shown in fig. 4. At a first observation, we can observe the number of likes, dislikes, and retweets in fig. 3.

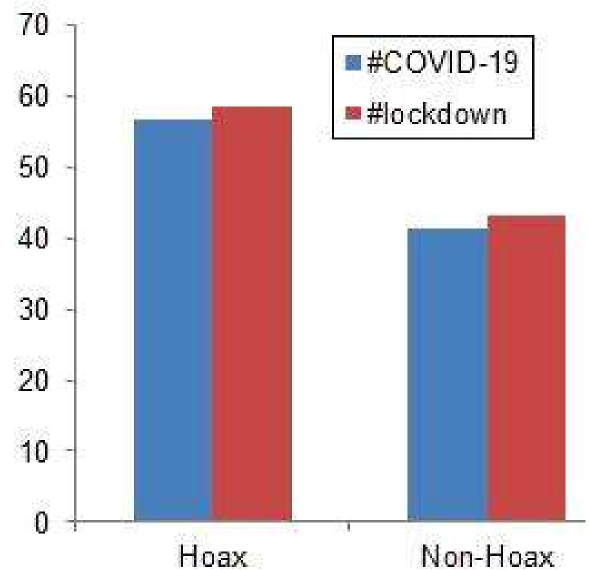

Fig. 4. The percentage of hoaxes and Non-Hoax

In fig. 5 the relation between number of likes, dislikes, and retweet on search test #COVID-19 & #lockdown per tweet can be observed. The percentage of likes, dislikes and retweets per user and per tweet is shown in fig. 6.

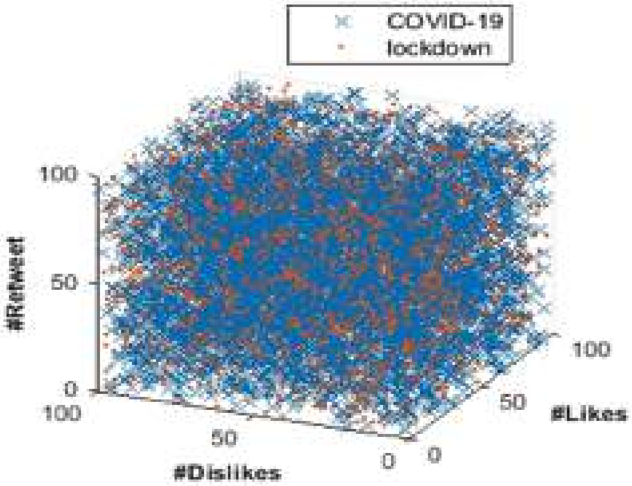

Fig. 5. The Number of Likes, Dislikes and Retweet on Search Text #COVID19 & #lockdown per tweet

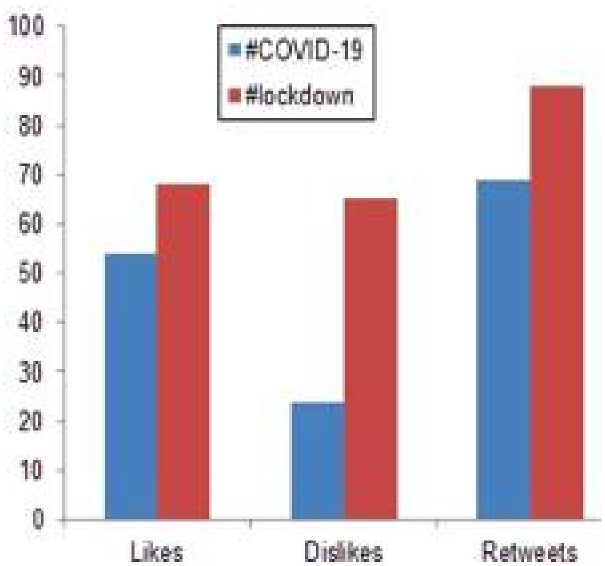

Fig. 6. The percentage of likes, dislikes and retweets per user, per tweet.

In order to analyse the accuracy of our classification model we performed standard cross-validation analysis. In this process, the dataset is divided into 80% training and 20% testing, and performed 5-fold cross-validation analysis. The performance of our approach is well with accuracy level of 98.3%.

Conclusion:

The popularity of social networking sites and the social medial over the internet has shown that people in our country consume most of the news from internet via social media. However, most of the fake news or hoaxes are being spread on social media and people are unaware that they are consuming and sharing fake news than real news.

In this work, we observed that fake news mainly depends upon the people choices i.e., their likes and dislikes. We also observed that fake news can also be shared or spread unintentionally. The fact checking should be considered before sharing. In order to improve the accuracy of fake news detection the use of NLP and deep neural network is required.

Future Work

In our future work, we will try to improve the accuracy of the results through the application of natural language processing and neural network. We promise to expand the application area of our research by including reliable media sources for fact checking so that our model can be applied in real life application. We would like to build up an application software for naive user who can handle more easily to detect fake news and its origin.

References:

Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. (2017). Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter, 19(1):22-36.

K Langin.(2018). Fake news spreads faster than true news on twitter thanks to people, not bots. https://news.mit.edu/2018/ study-twitter-false-news-travelsfaster-true-stories-0308.

Hunt Allcott and Matthew Gentzkow. (2017). Social media and fake news in the 2016 election. Journal of Economic Perspectives, 31(2):211-36.

Hal Daume III Mevan Babakar, Nada Bakos. (2014). Fake news challenges. http://www.fakenewschallenge.org/.

Jesse Holcomb Amy Mitchell, Michale Barthel and Rachel Weisel. (2016). Many americans believe fake news is sowing confusion. Pew Research Center.

Kai Shu, Deepak Mahudeswaran, Suhang Wang, Dongwon Lee, and Huan Liu. (2018). Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media. arXiv preprint arXiv:1809.01286.

Sabrina Tavernise. (2016). Spreads lies, more readers shurg at truth. https://www.audible.in/pd/As-Fake-News-Spreads-Lies-More-ReadersShrug-at-Truth-Audiobook/B07K3VWRZ1.

Alan Schwartz. (2004). Spam Assassin. O'Reilly Media, Inc.

P. Pale T. Petkovíc, Z. Kostanjcar. (2005). E-mail system for automatic hoax recognition.

A. Ishak, Y. Y. Chen, and Suet-Peng Yong. (2012). Distance-based hoax detection system. In 2012 International Conference on Computer Information Science (ICCIS), volume 1, pp.215-220.

Marin Vukovi¢, Kre imir Pripu i¢, and Hrvoje Belani. (2009). An intelligent automatic hoax detection system. In Juan D. Velásquez, Sebastián A. Ríos, Robert J. Howlett, and Lakhmi C. Jain, editors, Knowledge-Based and Intelligent Information and Engineering Systems, pp. 318-325, Berlin, Heidelberg. Springer Berlin Heidelberg.

Iryna Yevseyeva, Vitor Basto-Fernandes, David Ruano-Ordás, and José R. Méndez. (2013). Optimising anti-spam filters with evolutionary algorithms. Expert Systems with Applications, 40(10):4010-4021.

M. Sharifi, E. Fink, and J. G. Carbonell. (2011). Detection of internet scam using logistic regression. In 2011 IEEE International Conference on Systems, Man, and Cybernetics, pp. 2168-2172.

Judee K. Burgoon, David B. Buller, Amy S. Ebesu, and Patricia Rockwell. (1994). Interpersonal deception: V. accuracy in deception detection. Communication Monographs, 61(4):303-325,.

Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova, and Yejin Choi. (2017). Truth of varying shades: Analyzing language in fake news and political fact-checking. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pp.2931-2937, Copenhagen, Denmark, Association for Computational Linguistics.

Hannah J. Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova, and Yejin Choi. (2017). Truth of varying shades: Analyzing language in fake news and political fact-checking. https://www.osti.gov/biblio/1440673.

Nadia K Conroy, Victoria L Rubin, and Yimin Chen. (2015). Automatic deception detection: Methods for finding fake news. Proceedings of the Association for Information Science and Technology, 52(1):1-4.

Shapiro Matthew, Gentzkow Jesse M and Stone Daniel F. (2016). Media bias in the marketplace: Theory.

Kai Shu, Suhang Wang, and Huan Liu. (2017). Exploiting tri-relationship for fake news detection. arXiv preprint arXiv:1712.07709.

Andreas Vlachos and Sebastian Riedel. (2014). Fact checking: Task definition and dataset construction. In Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science, pp.18-22, Baltimore, MD, USA, Association for Computational Linguistics.

Jakub Kruczek, Paulina Kruczek, and Marcin Kuta. (2020). Are n-gram categories helpful in text classification? In Valeria V. Krzhizhanovskaya, GáborZávodszky, Michael H. Lees, Jack J. Dongarra, Peter M. A. Sloot, Sérgio Brissos, and João Teixeira, editors, Computational Science - ICCS 2020, pp. 524-537, Cham, Springer International Publishing.