

Fake News Detector in Online Social Media

R.Kanniga Devi

Abstract: Spreading of fake news in online social media is a major nuisance to the public and there is no state of art tool to detect whether a news is a fake or an original one in an automated manner. Hence, this paper analyses the online social media and the news feeds for detection of fake news. The work proposes solution using Natural Language Processing and Deep Learning techniques for detecting the fake news in online social media.

Keywords : Convolutional Neural Network, Deep learning, Fake news, Online Social media, Natural Language Processing..

I. INTRODUCTION

In this era of technology, every one's day begins and ends with the smart phone. Almost all the services are delivered to the people at the touch of the smart phone. With the advent of social media websites and applications in smart phones namely, Facebook, Twitter and WhatsApp, people share their content online, communicate with people in any part of the world and connect with like-minded people, usually for free [1]. The social media increases the awareness of social and political issues happening around the people and in turn the information are getting updated rapidly and becomes viral across the cyber and physical world. By taking this speed of communication among the people as an advantageous feature, the traditional business services, which heavily relied on the infrastructure setups are being transformed to online business services. Without Brick and Mortar infrastructure setup, social media helps the organizations to develop their businesses by promoting the brand awareness and reputation in a cost-effective way [2],[3]. The "Stay Connected" feature of social media helps the political parties to conduct online campaigns to know about their popularity among the society [4],[5]. Also, people organized the protests using social media, namely Turkey's political protest in 2013 [6] and India's Jallikattu protest in 2017 [7] and witnessed the country's history. Hence, social media is considered to be a boon for this technology era.

On the downside, the social media decreases the real-world personal connections by addicting the people in the cyber

world and lessens the human values [8]. It increases the possibilities of cyber bullying, stalking and hacking. The spread of fake news is the most dangerous of all the disadvantages [9]. The negative feedback damages the reputation of products and people [10]. The rumours spread on Facebook and WhatsApp have led to the lynching deaths

in India. Recently, the viral rumours about child kidnappers cost the life of about 20 people across India [11]. Despite the awareness campaigns and public alerts, the Indian Government had very limited success in deterring the spread of misinformation. The Indian authorities and officials blamed the social media giant WhatsApp for the irresponsible and explosive messages being shared by its 200 million Indian users [12]. The Indian Government has insisted the WhatsApp to act urgently to halt the spread of irresponsible and explosive messages on its platform. To take action over this series lynching matter, the WhatsApp made an announcement that the company offers \$50,000 for ideas to stop fake news spread [13]. Hence, the problem is still alive with no existing solutions for finding fake news in online social media.

II. PROPOSED WORK

The spread of fake news in online social media is a major nuisance to the public and there is no state-of-the-art tool to detect whether a news is a fake or an original one in an automated manner. This motivated us to propose a solution to detect the fake news in online social media. Generally, the false information in social media comes in either of the two formats namely, plain text message and an image format

A. Information source

To identify the fake news, user needs an information source, which provides the genuine news. This source of information should be in a computer readable format, which generally comes in the following two ways:

1. RSS (Rich Site Summary) feed of famous and authentic newspapers like Times of India, The Hindu etc.,
2. Twitter accounts of famous persons/ celebrities like the President of the United States and the Prime Minister of India etc.,

Data from the above two sources are downloaded on the daily basis using an automated script. Generally, a news data, whether it is true or false will represent an actor(s) or thing(s).

Sample news from Times of India on August 6, 2018: "During a debate on no-confidence motion against Narendra Modi government, Congress president Rahul Gandhi walked up to the PM and hugged him!"

The above news contains some actors and are mentioned in Table I.

TABLE- I: ACTORS

Information	Actor
During a debate on no-confidence motion against Narendra Modi government, Congress president Rahul Gandhi walked up to the PM and hugged him!	Narendra, Modi, Rahul, Gandhi

B. Feature Extraction

B.1 Language Detection



Revised Manuscript Received on December 16, 2019.

* Correspondence Author

R.Kanniga Devi*, Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, India. Email: rkannigadevi@gmail.com

One cannot expect to receive a fake news only in English language, but it is being spread in many other languages in social media. Hence, as a first step, an automated solution is needed to detect the language of the news. There is also a downside in detecting the language of the news, since one may send news in combination of various local languages along with the global languages. To deal with this issue, the information will be tokenized using a tokenizer provided by the Natural Language Processing (NLP) toolkit and the language of the individual words is checked. Subsequently, the most frequent language is selected to get the most out of the information. The algorithm 1 represents this process.

Algorithm 1:

```
function_detect_language(message)
{
    language_dict = {}
    words = tokenize(message)
    for word in words
    {
        language = find_language(word)
        if language in language_dict:
            weight = language_dict[language]
            weight += 1
            language_dict[language] = weight
        else:
            language_dict[language] = 1
    }
    return max(language_dict)
```

The languages, which are different from English will be converted using publicly available translation Application Program Interface (API).

B.2. Extraction of actors from information

An actor is a person or a place, which plays a major role in the information, irrespective of whether an information is false or genuine. The proposed work makes use of the Natural Language Processing toolkit provided by Python to extract the names of the actors, which play a major role in the information. A list of actor names is collected and stored in a serialized format for individual languages for easier separation.

B.3 Removal of stop words

An actor is a person or a place, which plays a major role in the information, irrespective of whether an information is false or genuine. The proposed work makes use of the Natural Language Processing toolkit provided by Python to extract the names of the actors, which play a major role in the information. A list of actor names is collected and stored in a serialized format for individual languages for easier separation. The words like “is, was, there, etc.,” are called stop words. The only role of stop words is to provide a stabilized grammar to the language, but it a useless thing in a feature extraction process. The stop words for various languages are also stored in a serialized format for easy removal.

After all the above steps are done, user will get the pure information. The resulting words are stemmed and stored in the database. The same preprocessing is done for the unverified messages and will be checked against the verified message. If anything seems to be matched, then it will be marked as verified, else it will be marked as unverified.

B.4. Picture Verification

By using the Convolutional Neural Networks (CNN), huge set of memes and non-offensive messages will be trained and whenever a new image is floating in the Internet it can be used to detect whether it is a meme or not. Fig.1 illustrates the flow diagram of the proposed methodology.

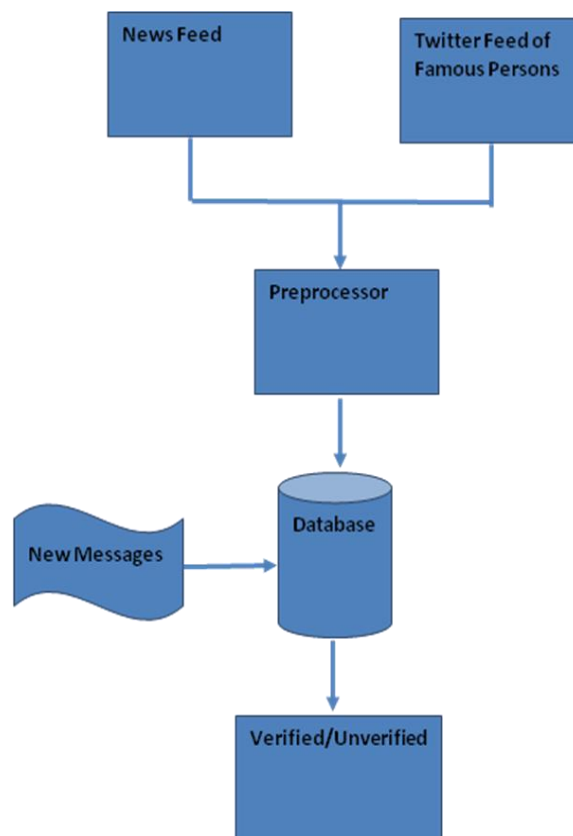


Fig. 1. Flow Diagram of the Methodology

III. EXPERIMENTATION AND RESULTS

A. Experimental setup

The choice of programming is Python. If a news is given to an ordinary person to identify whether it is fake or not, what he or she will do is refer a news paper or other genuine sources like famous Twitter accounts etc. So, the work automates the whole process of identifying the genuinity of the news. The server will run periodically for downloading the news and news feed data along with other authenticated sources like Twitter accounts of famous personalities. The downloaded data is stored in a database. Once the client requests the server to verify the genuinity of the news, the server will preprocess the news sent by the client and verify it with the server's database. If



there is any matching entry, the server will response from where the data came from like an URL of an news feed. If the news is not matched, the server will reponse unverified news message.

A.1 How originality of text is detected

The preprocessed original tweet data along with RSS feed is stored in the servers database. Once a message is passed to the server it will preprocess the message sent by the client and check it in the database, if the news is present in the database it will return the response as “Original News” else it will return “Non-Verified News”.

A.2 How originality of the image is detected

Huge collection of images of the famous personalities along with malicious memes which causes trouble to the public is downloaded. All the images are resized to 128 x 128 height and width and a CNN model is trained with the images. The client will send the images to the server and the server will predict the whether the image is malicious or not. The trained model is stored in a H5 file model.

A.3 H5 file format

The H5py package is a Pythonic interface to the HDF5 binary data format. It lets you store huge amounts of numerical data, and easily manipulate that data from NumPy. For example, you can slice into multi-terabyte datasets stored on disk, as if they were real NumPy arrays. Thousands of datasets can be stored in a single file, categorized and tagged however you want.

H5py uses straightforward NumPy and Python metaphors, like dictionary and NumPy array syntax. For example, you can iterate over datasets in a file, or check out the .shape or .dtype attributes of datasets. You don't need to know anything special about HDF5 to get started.

In addition to the easy-to-use high level interface, h5py rests on a object-oriented Cython wrapping of the HDF5 C API. Almost anything you can do from C in HDF5, you can do from h5py.

Best of all, the files you create are in a widely-used standard binary format, which you can exchange with other people, including those who use programs like IDL and MATLAB.

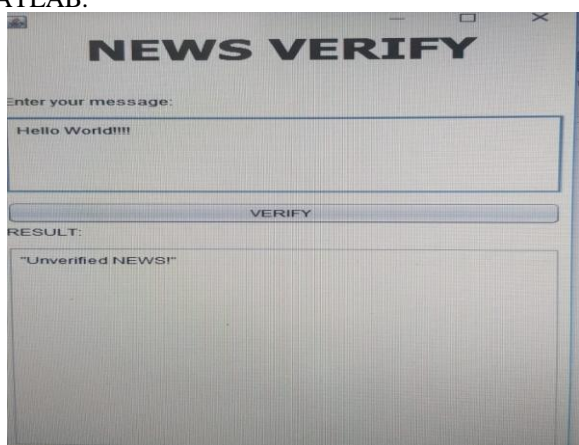


Fig. 2. Client side results

IV. CONCLUSION

Spreading of fake news in online social media is a major nuisance to the public; there is no state of art tool to detect

whether a news is fake or a original one in an automated manner. Hence, this paper proposes a solution to detect fake news in online media. This work uses Natural Language Processing techniques and Deep Learning for detecting the fake news in social media. The society could use our tool to prevent the spreading of fake news in online social media.

REFERENCES

- [1] Xiping Hu, Terry H. S. Chu, Victor C. M. Leung, Edith C.-H. Ngai, Philippe Kruchten, Henry C. B. Chan, A Survey on Mobile Social Networks: Applications, Platforms, System Architectures and Future Research Directions, IEEE Communications Surveys & Tutorials, Volume 17, Issue 3, pp: 1-25, 2015
- [2] Ali Abdallah Alalwan, Nripendra P. Rana, Yogesh K.Dwivedi, Raed Algharabat, Social media in marketing: A review and analysis of the existing literature, Telematics and Informatics, Elsevier, Volume 34, Issue 7, pp: 1177-1190, 2017,
- [3] Eun-Ju Seo, Jin-Woo Park, A study on the effects of social media marketing activities on brand equity and customer response in the airline industry, Journal of Air Transport Management, Elsevier, Volume 66, pp: 36-41, 2018.
- [4] Sebastian Stier, Arnim Bleier, Haiko Lietz and Markus Strohmaier, Election Campaigning on Social Media: Politicians, Audiences and the Mediation of Political Communication on Facebook and Twitter, Political Communication Journal, Taylor & Francis Online, pp: 50-74, 2018.
- [5] Karen Ross, Tobias Bürger, Face to face(book) Social media, political campaigning and the unbearable lightness of being there, Political Science, SAGE journals, Vol.66(1),pp:46-62, 2014.
- [6] John T. Jost, Pablo Barbera, Richard Bonneau, Melanie Langer, Megan Metzger, Jonathan Nagler, Joanna Sterling, Joshua A. Tucker, How Social Media Facilitates Political Protest: Information, Motivation and Social Networks, Advances in Political Psychology, Wiley, Vol. 39, Suppl. 1, pp:85-118, 2018
- [7] Sathish. R, Social Media and Peoples Movement: A view on Role of Social Media in Jallikattu Protest, Research Journal of Humanities and Social Sciences, 9(1), pp:131-134, 2018.
- [8] Shabnoor Siddiqui, Tajinder Singh, Social Media its Impact with Positive and Negative Aspects, International Journal of Computer Applications Technology and Research, Volume 5, Issue 2, pp: 71 - 75, 2016.
- [9] Soroush Vosoughi, Deb Roy, Sinan Aral, The spread of true and false news online, Social Science, Vol. 359, Issue 6380, pp: 1146-1151, 2018.
- [10] Alina Campan, Alfredo Cuzzocrea, Traian Marius Truta, Fighting fake news spread in online social networks: Actual trends and future research directions, IEEE International Conference on Big Data (Big Data), pp: 4453- 4457, 2017.
- [11] <https://www.straitstimes.com/asia/south-asia/mob-kills-5-people-in-india-on-suspicion-of-child-kidnapping> [online: July 2018]
- [12] <https://www.bbc.com/news/world-asia-india-44709103> [online: July 2018]
- [13] <https://www.telegraph.co.uk/news/2018/07/04/india-calls-whatsapp-hel-p-end-spate-lynchings-sparked-rumours/> [online: July 2018]

AUTHOR PROFILE



R.Kanniga Devi received her Bachelor of Engineering in Computer Science and Engineering from Madurai Kamaraj University, Madurai in the year 2001 and Master of Engineering from Anna University, Chennai in the year 2008. She pursued her doctoral program at Kalasalingam Academy of Research and Education, Tamilnadu, India in the year 2018. She is a member of CiE, CSI, IAEng, ISTE and IEEE.

She has 16 years of teaching experience. She has published 12 research papers in International Journals and Conferences. She got Science Academy Summer Research Fellowship and Tamilnadu Young Scientist Fellowship. Her research interest include Cloud computing, Graph theory applications in Cloud Computing, and Machine Learning. She is currently working as an Associate Professor in the Department of Computer Science and Engineering at Kalasalingam Academy of Research and Education, Tamil Nadu, India.