ORIGINAL ARTICLE



Spatio-temporal approach for classification of COVID-19 pandemic fake news

I. Y. Agarwal¹ · D. P. Rana¹ · M. Shaikh¹ · S. Poudel¹

Received: 1 February 2022 / Revised: 3 May 2022 / Accepted: 15 May 2022 / Published online: 27 June 2022 © The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract

The spread of Fake News during this global pandemic COVID-19 has dangerous consequences on economy and health of public. From origin of virus, spread, self-medication to hoaxes on vaccination, it created more panic than the fatality of the virus. For better infodemic preparedness and control, it is necessary to mitigate fear among people, manage rumours, and dispel misinformation. A survey on Fake News during COVID-19 was made by Poynter Fact Check institute. It stated that major chunk of the fake news on COVID-19 originated majorly in Brazil, India, Spain, and the United States. Fake news menace is severe in countries where the trust on online media is high such as Brazil, Kenya and South Africa. Based on these observations, this study provides preliminary insight on the co-relation of the spatial and temporal meta-information of the news like the news source country, the name of the countries specified in the news, and date of publish of news to the credibility of news. The main contribution of this study is to analyse the impact of spatial and temporal information features for classification of fake news, which to the best of our knowledge has not been explored yet. Also, these features are directly not available in any news article available online. Hence, these features are handcrafted. Meta-data of the news article such as origin of news is considered. Additional spatial information is extracted from the news article using NER tagging. Temporal information such as date of origin of news is given as an input to the LSTM model. These features are given as an input to Long Short-Term Memory (LSTM) model along with GloVe vectors and word length vector. A comparative analysis for accuracy is tested of the models with and without spatial and temporal information. The model with spatial and temporal information has achieved noteworthy results in fake news detection. To ensure the quality of prediction, various model parameters have been tuned and recorded for the best results possible. In addition to accuracy, the spatial and temporal information for fake news detection offers several other important implications for government and policy makers that will be instrumental in simulating future research on this subject.

Keywords Spatio-temporal information · COVID-19 · Infodemic · Fake news detection · Deep learning

I. Y. Agarwal d17co004@coed.svnit.ac.in

> D. P. Rana dpr@coed.svnit.ac.in

M. Shaikh maruf3141@outlook.com

S. Poudel saugatpoudel2054@gmail.com

¹ Sardar Vallabhbhai National Institute of Technology, Surat 395007, Gujarat, India

1 Introduction

The COVID-19 epidemic is a one-of-a-kind occurrence in human history. It is critical considering its two main reasons. Firstly, the COVID-19 epidemic has affected practically every country on the planet. Secondly, the crisis has brought in an indefinite time of ambiguity and anxiety for people all around the world (Kharasch and Jiang 2020). When people are in a state of ambiguity and fear, they are more likely to share and any information they get about health, financial system, community, and other aspects of the environment without testing it on the anvil of veracity. This encourages them to share and consume misinformation about health, economy, society, and surroundings (United Nations 2020a). The WHO had coined a new term for the spread of fake news during the COVID-19 pandemic-"Infodemic" stating that this is more deadly than the virus itself (Zarocostas 2020). It has affected people physically as well as mentally. Not only anxiety and fear levels have increased, the toll of death count has also amplified. During the first three months only, 800 people lost their life and more than 5800 hospitalized infected by "Infodemic" (BBC 2020). 60 cases of people loosing their eyesight were reported by drinking methanol, which was a misinformation circulated as cure of COVID-19 on online media (Aljazeera 2020). For better pandemic preparedness and control, it is necessary to mitigate fear among people, manage rumours, and dispel misinformation (Sakurai and Chughtai 2020; Islam et al. 2020). Also, for defining rules and regulations for governing this situation, it is of utmost importance to analyze the mechanisms driving the generation and consumption of fake news.

A study reveals that the tendency for creating and consuming COVID-19 related fake news varies greatly around the globe (Brennen et al. 2020). A survey carried out by Poynter Institute stated that a major chunk of fake news on COVID-19 initiated from United States, India, Spain, and Brazil (Poynter 2020a, b). The spread and consumption of fake news is high in countries such as Brazil, Kenya and South Africa where people have tendency to trust online media blindly and rely less on authentic sources (Newman et al. 2020). Countries with highest internet infiltration during "Infodemic" are Argentina and Chile. The rate of internet usage is 92.0-92.4%, respectively. These countries also topped the chart for relying on social media as their primary source of news 28.0% and 32.0%, respectively. Brazil and Colombia had a moderate figure for both metrics. Among all the countries, Mexico ranked first in use of social media, while Peru and Colombia had the highest indices of incapacity to identify fake news (Nieves-Cuervo GM et al. 2021). In countries with high literacy rates such as Germany, Denmark and Netherlands, the propensity of fake news is less (Newman et al. 2020).

In case of Infodemic, an observation can be drawn that credibility of information has a relevance with the location of news and location in the content. The main finding of this study is that the influence of spatial and temporal information features on the classification of fake news has not been investigated before, to the best of our knowledge. Furthermore, these features are not directly available in any online news article. As a result, these features are handcrafted. To prove this hypothesis, first research question raised and solved in this paper is:

#RQ1 – Is the spatial information related to news (location of origin of news or location information inscribed in news) linked to the credibility of the news?

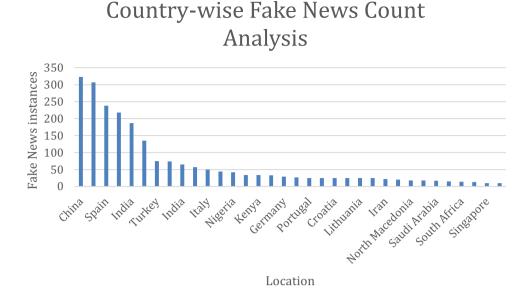
The second key observation drawn from the data was that there was a relevance of date with the credibility. However, such instances of fake news were not limited to the time when scientists were struggling to find ways to counter the coronavirus but also when the news around COVID-19 vaccines started hitting the news headlines. For example, DW Fact Check recently debunked many fake scientists' twitter accounts spreading misinformation about the coronavirus vaccine altering one's DNA or the news about Peru forcing its people for compulsory COVID-19 vaccinations (Mudge and Weber 2020). For example, the spread of fake news has resulted in instigating tensions and violence against minority communities and healthcare workers in certain parts of the world, while 5 G conspiracy theories of coronavirus propagation have resulted in significant property losses to national telecommunication infrastructures in the UK (Islam et al. 2020; Newman et al. 2020). It has also been reported that during the COVID-19 pandemic, fake news phenomenon has been opportunistically capitalised by individuals to further their personal, political and business agendas (UN News, 2020). Hence, the second question addressed in this work is:

#RQ2 – Is the temporal information such as publish date of news linked to its credibility?

To examine our research questions, a thorough survey of contributions made in the field of information mining using spatial and temporal information was carried out.

The objective of this work is to propose use of spatial and temporal information for classification of Fake news using LSTM. For this analysis, the dataset COVID19FN was compiled and published online for research community consisting of 2800 labelled news articles on COVID-19. For spatial information, the country information was extracted from the text. Temporal information was obtained from the date of publishing of the news article. These features along with global vectors (GloVe) for word embeddings and word length vectors formed the input feature set. A comparative analysis of the models-with and without these features to analyze the influence of spatial and temporal features is carried out. The proposed framework with spatial and temporal features has obtained promising results. Various model parameters were calculated and reported to achieve optimal results in order to ensure prediction accuracy. Source website used for data compilation of COVID19FN are used by other research of the fake similar research area like fake news detection. Also our dataset is having locations like China, Brazil, US, Australia, France, Philippines, Sri Lanka, Indonesia, Canada, Germany, Columbia, Japan, Thailand and Singapore. It can be observed that these are countries most popular among the tourists as shown in Fig. 1. This information can prove to be critical to help further dissipation of fake news and provide location based aid for the aid of affected people.

The rest of paper is organized as follows. The following section discusses survey of work done in Information Mining using spatial and temporal information. Formulation of model for addressing the research questions is explained in Fig. 1 Histogram of fake news depicting country wise instance of Fake News



the following section. Experimentation carried out on the dataset and its analysis is in the following section. Conclusion and directions for future research work are drawn at last.

2 Literature review

To examine these research questions formulated in the previous section, firstly a thorough survey of contributions made using the spatial and temporal information for Information Mining is carried out to get in-depth understanding of the use and implications of these features for credibility assessment.

Choi et al. discovered how repeated online news reporting has affected perceptions of the integrity of the press and news aggregators. The time and space feature were used to determine the repeated publication of identical news content by the same news companies (Choi 2017). Lie et al. developed a coding scheme to turn instances of representation into spatial codes to show their latent space positions. Given its spatial codes, a multimodal auto-encoder is constructed to generate the description of a representation instance. The auto-encoder incorporation makes the framework able to deal with different data types. Experiments are performed under various design conditions and different representational learning models to show the versatility and efficacy of their system.

Moreover, an ensemble approach is proposed that effectively incorporates and validates results from decomposition of various tensor in clean and coherent groups of fake news. This work studies the impact of spatial feature within the news as a classifier. The name of country is extracted from the text using NER. For the articles with multiple extracted values by NER, manual annotation was done. As a temporal cue, the publish date was passed as an input to the model. Next section throws light on the dataset followed by the methodology of experimentation. Comito analyzed geo-tagged movements to discover people and community behavior. This work proposed a novel methodology to mine popular travel routes from geo-tagged posts. The approach infers interesting locations and frequent travel sequences among these locations in a given geo-spatial region, as shown from the detailed analysis of the collected geo-tagged data (Comito et al. 2021). Falcone et al. used human activity dynamics to correlate categories with GPS positions of posts on OSN platforms. The Twitter API was used to collect geo-tagged tweets. A supervised learning framework analyses the spatial-temporal features of tweets to determine individual behaviors, which are then utilised to infer the location category: eating, entertainment, or work place (Leetaru et al. 2013). Another significant contribution using the spatial and temporal information was made prove correlation between tweets and real COVID-19 data, proving that Twitter can be considered a reliable indicator of the epidemic spreading and that data generated by user activity on social media is becoming an invaluable source for capturing and understanding epidemics outbreaks (Comito 2021). Another interesting work proposed an online algorithm that incrementally groups tweet streams into clusters. The approach summarizes the examined tweets into the cluster centroids. The assignment of a tweet to a centroid uses a similarity measure that takes into account both the cluster age and the terms occurring in the tweet. Experiments on messages posted by users in the Manhattan area show that the method is able to extract events effectively taking place in the examined period.

Thus, the spatial and temporal information are potential indicators to gain important insights from the text data and a lot of interesting pattern analysis and behavior mining can be carried out using these information.

From the thorough literature survey, following significant findings were made:

- i. The spatial and temporal information are significant features to gain insights from the textual data, which can be further used for pattern mining or behavior analysis.
- ii. The spatial information is not directly and distinctly available in the news data.
- iii. The spatial information can be co-related to the credibility of the news and can be used for automatic verification of news.
- iv. The spike and rise of fake news can also be linked to significant real-time occuring events like pandemic spread, war, election, etc.

These findings motivated for the use of these signals for credibility assessment of Fake News. To the best of our knowledge, no prior contribution use of these indicators for classification of fake news.



Fig. 2 Choropleth plot of fake news spread in various countries

The dataset used in this work is a Fake News on COVID19 [(Agarwal et al. 2020)]. COVID19FN consists 1740 Fake news articles and 1160 actual news articles on the novel coronavirus pandemic. The fake news articles are from fact-check engine Poynter while the real ones have been obtained from well-known sources such as FactCheck, Observador, Snopes and the like. For each article, the dataset contains several features like titles, text, date published, country and source URL. This work focuses on providing spatial and temporal features for classification. The exploratory data analysis was carried out on the dataset to analyses the statistical significance of the proposed features. Figure 2 describes the distribution of fake news articles according to countries. By observing this graph from the data collected, it can be noted that few of the regions contribute more in proliferation of fake news. Thus, adding country information to our fake news detection model could yield positive results. Figure 3 is the date-wise histogram of fake news about COVID-19 from January to April 2020. It can be observed here that there are spikes at particular dates on which the fake news was cooked up. Hence, this could also act as a potential classifier.

The process of cleaning the data was trivial as the COVID-19FN dataset is well curated. The next section briefs about Deep Learning Models and gives.

a brief intuition of the foundations of each one.

3 Proposed approach

The motivation for the dataset is to facilitate the development of machine learning models for automatic fake news detection during the time of crises. In this task, we frame this as a binary classification problem.

The architecture diagram and algorithms below explain how the features are extracted and the flow of work. Since LSTM model have obtained the state-of-the art results on many text classification datasets, the proposed approach is built on LSTM neural networks. The algorithm 1 explains

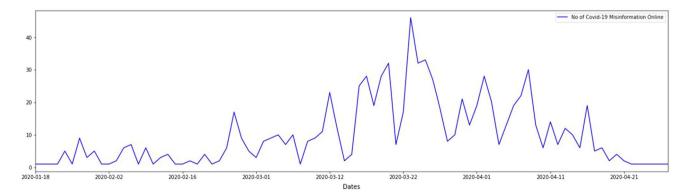


Fig. 3 Number of COVID-19 misinformation online

1: procedure GET_COUNTRY_INFO 2: for sentences in news_db do 3: words = tokenize(sentences[i]) 4: for word in words do 5: tagged[i] = pos_tag(words[i]) 6: chunks = ne_chunk(tagged) 7: ner = EntityRecognizer(nlp.vocab, entity types = ['GPE'])
 3: words = tokenize(sentences[i]) 4: for word in words do 5: tagged[i] = pos_tag(words[i]) 6: chunks = ne_chunk(tagged) 7: ner =
 4: for word in words do 5: tagged[i] = pos_tag(words[i]) 6: chunks = ne_chunk(tagged) 7: ner =
5: tagged[i] = pos_tag(words[i]) 6: chunks = ne_chunk(tagged) 7: ner =
6: chunks = ne_chunk(tagged) 7: ner =
7: <i>ner</i> =
EntityRecognizer(nln vocab_entity_types = $[GPE']$)
Emilyneeognizer (mp.voeuo, emily_types [01 E])
//GPE tags all the countries, cities etc.
8: <i>if (ner in pycountry.countries)</i>
9: location[i] = pycountry.countries
10: encode (location)
11: <i>embedding</i> \leftarrow GloVe (articles, location)

In the above algorithm, Step 1 through 3 tokenizes all the sentences in the news articles. The word vectors are tagged according to the respective categories. The NER tagger identifies named entities (people, places, organizations etc.) from a chunk of text, and classifies them into a predefined set of categories. The categories regarding the geographic locations can be chunked together and parsed with the help of entity recognizer as shown in line 6 and 7. The list of countries is present in the pycountry.countries. This list is then taken as filter to get the valid country information and a spatial feature vector is generated as shown in line 8 and 9. This vector is added as a classifier to the model. The entire work flow is explained in the next algorithm.

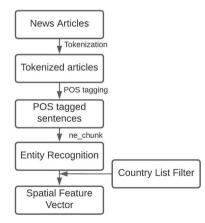


Fig. 4 Spatial Feature Extraction

Algorithm 2	2 Fake	News	Detection	using	Deep	Learnin	g and
GloVe Embe	edding						

1: procedure FAKE_NEWS_DETECTION				
2: for label in labels do				
3: if <i>label in list</i> _{<i>false</i>} then				
4: $label \leftarrow 'False'$				
5: if <i>label in list</i> _{true} then				
6: $label \leftarrow 'True'$				
7: else				
8: drop(label)				
9: for a in articles do				
10: W .append($w(h)$)				
11: $L.append(l(h))$				
12: normalize (<i>articles</i> , <i>date</i> , <i>country</i>)				
13: encode (<i>date, country</i>)				
14: <i>embedding</i> \leftarrow GloVe (articles, country)				
15: <i>Model</i> \leftarrow <i>date, country, W, L, embedding</i>				

In the above algorithm, Step 1 through 8 indicate the conversion of multi-class fake news detection to a binary classification by conversion of labels of different truth measures (partially true, barely true) to true or false. The reason this step was required was because data are collected from different websites and have different measures of truth. Step 9 is to append the word length and the number of words to feature vector in article for all the news data points in the dataset. Step 12 is to normalize the values of articles date and country value and append it to the feature vector. Step 13 indicated one-hot encoding for the categorical data of country and date values. Step 14 generates the GloVe vectors taking input the article text and the country values. Step 15 trains the model taking input the encoded date, country, embedding, and number of words and length of article features. Diagrammatic representation of the algorithm is presented in Fig. 4. Figure 5 shows the model integrating the spatial and temporal feature along with the embedding.

As shown, there are two inputs to the model, one which includes an LSTM layer and another that includes structural information about the model that is fed into two dense layers. Vectors generated by GloVe embedding are taken as an input. It is allocated a maximum size of 1500. The other two inputs are average word length and number of words in an article. A concatenation of the two inputs is then done, which is again fed to a dense layer. Dropout layer is used for reducing any over-fitting that might occur during training. Figure 6 describes the model architecture of the deep neural network with spatial and temporal information.

As shown in the architecture, there are four inputs to the model. Two of them are textual information that are to be fed into an LSTM, namely, the article body and spatial information that consists of the country name. One of the two inputs

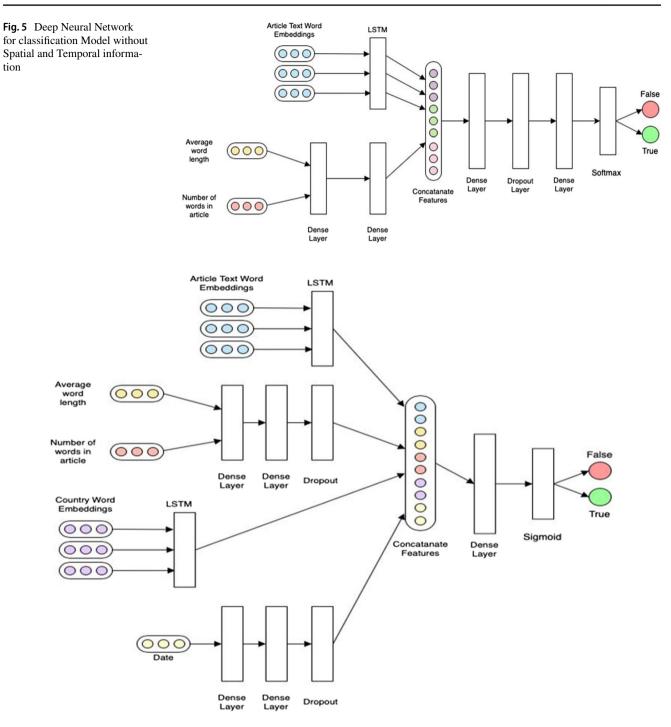


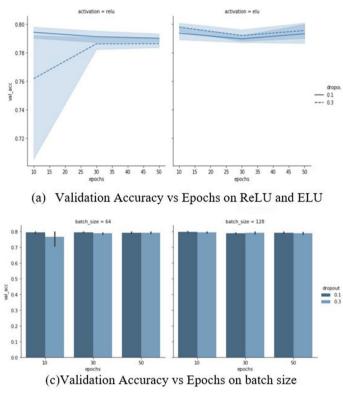
Fig. 6 Deep Neural Network for classification Model with Spatial and Temporal information

contain structural information about the text, which includes average word length and, number of words in an article and the other includes temporal information, that is, day of the week. Dropout layers are provided to the model to avoid any over fitting conditions. All the inputs are then concatenated and given to the final output dense layer.

Labels is a list containing the feature label, $list_{false}$ is a list varying degrees of falsehood like pants on fire, conspiracy

Deringer

theory, partly false, false etc. Similarly, list_{true} is a list of varying degrees of truth. Since the model is a binary classification model, we change any degree of falsehood or truth to a singular False or True label. Moreover, the data set also had other labels which were dropped, specified by the drop function in the algorithm. W and L are lists of average word length of a word in an article and number of words in an article, respectively. Normalize is a function that removes



0.82 0.81 0.80 0.79 0.78 0.77 0.76 0.0 0.1 0.2 0.3 0.4

(b) Validation Accuracy vs Dropout

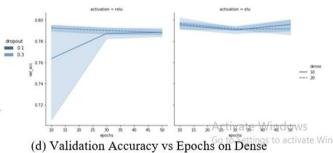


Fig. 7 Comparison of hyperparameters

any non-alphabets, symbols, and converts the text into lower case and Tokenize is a function which is used for converting words into a vectors. GloVe is function that is used for generating glove embedding (COVID-19 Archives 2020) given length. Finally, different features are fed into the model, the structure of which is shown in Fig. 7.

The following section presents the results and analysis of the proposed approach.

4 Result analysis

The proposed work is to analyse the impact of spatial and temporal features for the classification of fake news. Experiments were carried on the same setup with and without the features to testify the impact. The result analysis is divided into two subsections for better clarification (1) Classification without spatio-temporal features and (2) Classification with spatio-temporal features.

4.1 Classification without spatio-temporal features

In the initial run of the experiment, test accuracy of 79.64% was achieved by using activation function ReLU in the dense layers and activation function Softmax in the last output

Table 1 Hyperparameters

LSTM units	64, 128
Dense layer input	10, 20
Activation Function	ReLU, ELU
Dropout	0.1, 0.3
Batch Size	64, 128
Epochs	10, 30, 50

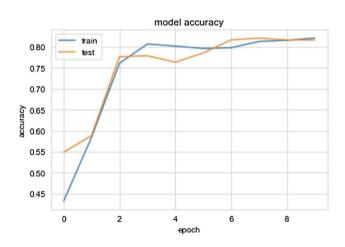


Fig. 8 Model accuracy vs. epochs

Table 2Hyperparameters formodel with Spatio-Temporalfeatures

LSTM units	32, 64, 128
Dense	10
Activation	Sigmoid, Softmax
Dropout	0, 0.1
Batch size	64
Epochs	10, 30, 50
Optimizers	Adam, Nadam

 Table 3
 Comparison of accuracy of models with and without spatiotemporal features

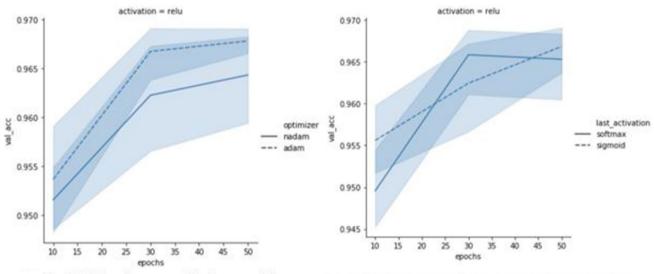
Spatio- temporal features	F1-score	Val F1-score	Accuracy	Val accu- racy	Val loss
Yes	0.974	0.968	0.974	0.967	0.104
No	0.826	0.804	0.826	0.803	0.464

layer. The model accuracies on train and test dataset is as shown in Fig. 8.

To achieve the best possible performance metrics, a number of hyperparameters were chosen and the model was

trained on them. The list of all the hyperparameters are given in Table 1.

Out of the different hyperparameters, the hyperparameters that work best for proposed model are dropout of 0.1, ELU activation function, batch size 64, epochs 50 and dense layer



(a) Validation Accuracy vs Epochs on optimizers

(b) Validation Accuracy vs Epochs on Activation function

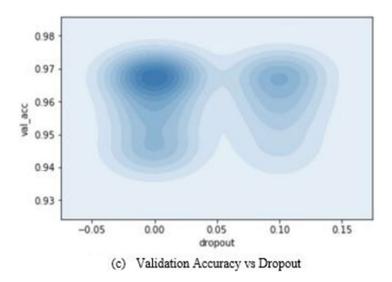


Fig. 9 Comparision of different hyperparameters

input of 20 to achieve a validation accuracy of 80.3% and f1score of 0.826 (Fig. 7).

4.2 Classification with spatio-temporal features

Working with Spatio-Temporal features, a validation accuracy of 95.24% was achieved by using activation function ReLU in the dense layers and activation function Sigmoid in output layer. To achieve the best possible performance metrics, a number of hyperparameters were chosen and the model was trained on them. The list of all the hyperparameters are given in Table 2.

The hyperparameters that work best for proposed model are Dropout of 0.1, Sigmoid activation function, batch size 64, epochs 50 and Dense layer input of 10 and Adam optimizer to achieve a validation accuracy of 96.78% and f1-score of 0.974 (Fig. 9). The comparative analysis of the proposed model with and without spatial and temporal features is shown in Table 3.

5 Conclusion and future scope

This work proposes a novel approach of use of Spatio-Temporal features for classification of Fake News. After the incorporation of spatial and temporal data, proposed model has shown promising results to predict domain specific fake news prediction on COVID-19 data. On addition of spatial and temporal features to the proposed model, we were able to achieve much higher accuracy and F1-Score measures of 97.4% and 0.974, respectively, concluding that spatial and temporal features do affect the detection of fake news.

These spatial and temporal information signals derived from the true or factual news can be further used to get early warnings of other outbreaks such as new variant Omicron. A large-scale pandemic assessment can also be performed to detect predicted events and enhance preventative efforts. Enabling cooperation across time and geography can lead to an increase in coordinated clinical governance.

References

- Agarwal, Isha; Rana, Dipti (2020). "COVID19FN", Mendeley Data, v3. https://doi.org/10.17632/b96v5hmfv6.3
- Animal Político | Periodismo libre para el ciudadano. (2020). Retrieved from https://www.animalpolitico.com/
- Bhatt, Gaurav, et al (2018). "Combining neural, statistical and external features for fake news stance identification". In: Companion Proceedings of the Web Conference 2018
- Brennen, J. Scott, et al (2020). "Types, sources, and claims of Covid-19 misinformation." Reuters Institute
- Chen, Emily, Kristina Lerman, and Emilio Ferrara (2020). "Covid-19: The first public coronavirus twitter dataset." arXiv preprint arXiv: 2003.07372

- Choi S, Kim J (2017) Online news flow: temporal/spatial exploitation and credibility. Journalism 18(9):1184–1205
- Comito C (2021) How COVID-19 information spread in US The Role of Twitter as Early Indicator of Epidemics. IEEE Trans Serv Comput. https://doi.org/10.1109/TSC.2021.3091281
- Comito C, Falcone D, Talia D (2016b) Mining human mobility patterns from social geo-tagged data. Pervasive Mob Comput 33:91–107
- Comito C, Pizzuti C, Procopio N (2016a). Online Clustering for Topic Detection in Social Data Streams. In: IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI) pp. 362–369, doi: https://doi.org/10.1109/ICTAI.2016.0062
- COVID-19 Archives Snopes.com. Snopes.com. (2020). Retrieved 13 June 2020, from https://www.snopes.com/tag/covid-19/
- Falcone, Deborah, Cecilia Mascolo, Carmela Comito, Domenico Talia, and Jon Crowcroft (2014) "What is this place? Inferring place categories through user patterns identification in geo-tagged tweets." In 6th International Conference on Mobile Computing, Applications and Services, pp. 10–19. IEEE
- Hosseinimotlagh, Seyedmehdi, Evangelos E. Papalexakis (2018). "Unsupervised content-based identification of fake news articles with tensor decomposition ensembles." In: Proceedings of the Workshop on Misinformation and Misbehavior Mining on the Web (MIS2). 2018.
- IFCN Covid-19 Misinformation Poynter. (2020a). https://www.poynt er.org/ ifcn-covid-19-misinformation/
- IFCN Covid-19 Misinformation Poynter. (2020b). https://www.poynt er.org/ifcn-covid-19-misinformation/
- Kaliyar, Rohit Kumar (2018). "Fake news detection using a deep neural network." In: 2018 4th International Conference on Computing Communication and Automation (ICCCA). IEEE
- Kim, Yoon (2014). "Convolutional neural networks for sentence classification." arXiv preprint arXiv:1408.5882
- Leetaru, Kalev, and Philip A. Schrodt (2013). "GDELT: Global data on events, location, and tone." ISA Annual Convention
- Mikolov, Tomas, et al. (2013) "Efficient estimation of word representations in vector space." arXiv preprint arXiv:1301.3781
- News 24. (2020). Retrieved 13 June 2020, from https://hindi.news2 4online.com/ tag/coronavirus/
- Roy, Arjun, et al (2018). "A deep ensemble framework for fake news detection and classification." arXiv preprint arXiv:1811.04670
- Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams (1986)."Learning representations by back-propagating errors." nature 323.6088: 533–536
- Shahsavari, Shadi, et al (2020). "Conspiracy in the time of corona: Automatic detection of covid-19 conspiracy theories in social media and the news." arXiv preprint arXiv:2004.13783
- Volkova, Svitlana, et al (2017). "Separating facts from fiction: Linguistic models to classify suspicious and trusted news posts on twitter." In: Proceedings of the 55th annual meeting of the association for computational Linguistics (Volume 2: Short Papers)
- William Yang Wang. (2017) "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Association for Computational Linguistics, 422–426. https://doi.org/10.18653/v1/P17-2067
- Zarocostas J (2020) How to fight an infodemic. The Lancet 395(10225):676
- Zhang, Jiawei, Bowen Dong, and S. Yu Philip (2020). "Fakedetector: Effective fake news detection with deep diffusive neural network." 2020 IEEE 36th International Conference on Data Engineering (ICDE). IEEE

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.