

DETECTION OF TWITTER-BASED FAKE NEWS WRITTEN IN HINDI AND MARATHI LANGUAGE USING DEEP LEARNING TECHNIQUES

Gouri Amol Vaidya

Department of Science, ICCS, Pune, India.

Email: vaidya.gouri@iccs.ac.in

Abstract

In our paper I have proposed the methodology of creating a dataset for Fake news written in low resource languages like Marathi and Hindi. The dataset is used to verify the tweets. We are also proposing a model based on Deep Learning Techniques: Transformer approach, CNN and LSTM and using programming Language: Python libraries for social networking services (SNS) and Evaluation Metric: Accuracy, F1-Score, Precision, Recall. We are working on Fake information on Tweeter during last 5 years for the Tweets written in Devnagri Script (Hindi and Marathi). Also we are verifying the probable fake tweets by authenticating the truthfulness of the tweet with dataset of online edition of popular Print media from Maharashtra like Maharashtra Times, Lokmat and Sakal.

Keywords: fake news, low resource languages, twitter

I. INTRODUCTION

Any kind of public communication medium like news articles, news channels, magazines, etc are the major sources to spread the information about the events happening around. In older days the print media have been widely used to publicize the information. The authenticity of such mediums was used to get validated by means of thorough study of authentic historical as well as live facts and figures. There were also the chances of some wrong information being spread either intentionally for personal or financial purpose or unintentionally passing on the misinformation [1]. The manipulated information can create biased views among the readers and hence affect the social health in negative manner. In 2016, the UK EU referendum and the US Presidential election were both marked by social media misinformation campaigns, which have subsequently reduced trust in democratic processes [2]. Social media has become a prime source of propagation of news since people makes use of it to connect to the friends, family as well as strangers through groups and communities [3]. Social media platforms like Twitter, Facebook, WhatsApp, Telegrams can play as a source of mis-information and fake news now a days that are being spread among user communities. It is essential to identify fake news at the early stage so that the spread of news should not affect the social health of the nation.

II. REVIEW OF LITERATURE

The work done by researchers in the area of Fake News Identification is binary of multi-class classification. Also there is work done on misinformation as well as fake news; whereas there is a thin line between the two. Misinformation is incomplete information passed on with or without any ill intention whereas fake news is deliberately crafted to mislead people.

We have divided our review in two categories based on Fake news detection techniques applications for-

- Data in the form of Image/Audio/Video
- The text in English language
- The text written in Non-English language platform

A. Data in the form of Image/Audio/Video

Classification was done on Twitter images [4] related to natural disaster and it classified the images as fake and true based on Naive Bayes Decision tree to give 97% accuracy. [5] worked on picture, video, text and title using classic ML techniques like Naïve Bayes and Random forest. They compared their results by developing model based on deep learning techniques like LSTM DROP and CNN.

B. The text in English language

[6] worked in a direction of identification of fake account and fake news over the platform of social media with reliable deep learning techniques. Hybrid method was carried out with different machine learning (ML) algorithms to categorize the text as true or fake. [7] proposed a Framework for Robust Rumor Detection where the authors target the issue of difference in data distribution of deployment to that of the training dataset. The authors also take into consideration the varying broadcast of messages by ambiguous remarks using multi-task and reinforcement learning. The feature extraction process checks for the feature of noisy comments and captures subjective information of comments.

Automated classification of news articles as fake or legitimate was performed by ensemble methods by [8]. The datasets used were DS1-ISOT Fake news Dataset (political), DS2-Kaggle (all domains) and DS3-Kaggle (sports, entertainment, and politics). One set of ensemble classifier combined logistic regression, random forest, KNN; while another set of classifier included combination of logistic regression, classification and regression trees and linear SVM.

English tweets [9] about Breaking news based on algorithmic techniques of sequential classifiers-hawkes processes, LSTM, linear and tree CRF. [10] worked on PHEME dataset of English breaking news tweets using NLP and methods of data mining. The objective of the work is to derive the impact of a rumour on readers as to make each tweet either of supporting, denying, querying or commenting on a rumour's truth value. LIAR dataset with 12,836 records was explored by [11] using SVC, Decision Tree, MLP Classifier and Gradient boosting to increase the previous performance by 16%. [12] worked on 11000 English Fake news detection using Naive Bayes, Support Vector Machine and Neural Network. [13] worked on dataset of BuzzFeed (news articles in English) for the input of News using KNN, Naïve Bayes, Random Forest, Support Vector Machine and XGBoost to give accuracy of 86%. [14], et al worked on fake news detection of social media posts in English related to COVID-19 and proved the accuracy of 97% using Deep learning algorithms and corresponding inbuilt library XGBOOST for working with multilingual platforms.

[15] worked on extracting WhatsApp data containing Misinformation about COVID-19. The messages were considered only from Brazilian Portuguese language. The algorithmic techniques used were RNN, BiLSTM neural networks, pooling operations and attention mechanism to give accuracy of 83%. Identification of Political fake news [16] from WhatsApp Public groups was carried out with features of group name, display picture, description. They contributed with manual WAM group filtering strategy with noisy user-generated data on WAM public groups. They also presented a fine-grained annotated dataset with multidimensional labels like language, malicious activity, political orientation, and political inclination.

An attempt was made to detect fake messages [17] circulated on WhatsApp using Supervised ML techniques considering the behavioural and contextual patterns. [18] identified if an article is authentic or fake using Supervised ML techniques considering the features like words, phrases, sources, sentences, and titles. [19] developed an automating fake news detection system to propose multi-level voting ensemble model to identify the fake text from websites pertaining to social media. Input features selected to decide as a candidate of fake news were categorical attributes, omitted data, headlines, the article content and the publisher details. Three feature extraction techniques were used viz. Supervised ML algorithms were implemented, for binary classification of text as fake or real, to give accuracy of 80%.

Combination of Deep learning and Machine learning algorithms were applied [20] on 7796 news articles based on dataset of fake and real news available at Github, to give 92% accuracy.

[21] tried to build a prediction model for the type of users who are more likely get fooled by the fake news. Authors have identified features such as user, network, and content; and grouped such users by multinomial classification achieving accuracy of 82%.

C. The text written in Non-English language platform

A web based system was developed [22] for a Fake News Detection for Hindi new articles given by media organizations. The proposed framework includes classic ML algorithms like Logistic Regression, Random Forest, Cosine Similarity, Gradient Boost, Decision Tree and Recurrent Neural Network. The news text features selected were author, summary, title and keywords. Another fake news detection in Hindi language was explored by [23] on Hindi News Dataset of Facebook using Naïve Bayes, logistic regression and Long Short-Term Memory (LSTM) to give accuracy 92.36%.

Fake news detection was carried out in the early stage of propagation for the tweets written in Arabic language were explored by [24] . The features selected were user-profile, content-based, and words-features. They used the dataset of 1862 previously annotated Arabic tweets for training the model. The authors followed NLP techniques, ML models, and a feature selection approach, Harris Hawks Optimizer (HHO). TF, BTF, and TF-IDF were used as text vectorization models and KNN, RF, SVM, NB, LR, LDA, DT, and XGboost were the machine learning algorithms implemented in the proposed approach. The authors suggested necessary future implementation for automatic processing of Arabic tweets, capturing the text’s semantic meaning using word embedding and comparing it with BoW models. The author also informed about future proposed method of using ensemble learning and deep learning behaviour to improve performance of prediction of fake news written in Arabic language.

[25] worked on Fake news detection for the text written in Urdu language. The approaches used included ensemble methods, CNN, and non-Urdu specialized Transformers (BERT, RoBERTa) as well as Urdu-specialized (MuRIL, RoBERTa-urdu-small) . For the dataset curation, they collected real news by crawling through authentic sources and for fake news they created their own dataset.

[26] created an annotated dataset of COVID 19 Tweets in English, Hindi, Bengali for fake news detection. Features selected were re-tweet count, Bias score, Twitter User Features, Source Tweet Embedding, Fact verification score. They worked on Supervised Learning algorithms techniques and BERT classifier for multiple languages to prove F-Score of 89%. In future, the authors like to add an automatic feature selection module which can take the most informative subset of handcrafted features for the classification task.

Table 1: Literature review

| Author & year | Dataset | Type of input | Language | Domain | Algorithmic techniques | Highest Accuracy |
|---------------|--|---------------|----------|---------------------------------|------------------------------|------------------|
| [4] | twitter | images | - | Natural disaster | Naive Bayes Decision tree | 97% |
| [7] | Twitter & Weibo | | English | | | |
| [8] | DS1-ISOT Fake news Dataset () DS2-Kaggle (all domains) DS3-Kaggle (, politics) | | English | political sports, entertainment | ML and ensemble techniques | 99% |
| [9] | Twitter | text | English | Breaking news. | sequential classifiers | |
| [10] | PHEME | text | English | breaking news. | NLP and data mining Methods | |
| [11] | LIAR | Text | English | | Supervised Learning | +16% |

| Author & year | Dataset | Type of input | Language | Domain | Algorithmic techniques | Highest Accuracy |
|---------------|---|-----------------------------|----------|--|--|------------------|
| | (12,836 records) | | | | algorithms | |
| [12] | Github with 11000 articles from Facebook | Text | English | Politics, science and business | Naive Bayes Classifiers | 91% |
| [5] | | picture, video, text, title | English | | 1. traditional ML methods 2. deep learning methods | |
| [14] | social media posts | Text | English | Healthcare | Bi-LSTM 1D-CNN BERT ROBERTa | 97% |
| [22] | Hindi news articles | Text | Hindi | New articles | | |
| [27] | Buzzfeed (news articles) | News | English | | Supervised Learning algorithms and ML library XGBoost | 86% |
| [6] | Kaggle | Text | English | User accounts and news | deep learning | 98.6% |
| [25] | 1300 news articles | Text | Urdu | (i) Health, (ii) Sports, (iii) Showbiz, (iv) Technology, and (v) Business, | ML, NN and Transformer techniques | 67% |
| [28] | Twitter | Text | English | General | Multiple machine learning and deep learning algorithms | 72% |
| [29] | News articles | Text | English | General | Machine learning and natural language processing | 93% |
| [30] | Fake news datasets on two fake news datasets (ISO and FA-KES), PolitiFact.com | Text | English | Politics | CNN+RNN | |

III. RESEARCH GAPS

1. Very less work is done on fake-news detection for Low resource languages like Marathi and Hindi.
2. (7, 22 -27)
3. Transformer approach of deep learning are still not explored much (13,22,25,26)
4. Fake news detection is not studied much from the view point of sentiment analysis.(3)
5. Multilingual approach is not much explored (26)
6. Ensemble approach of Deep learning models of classification is an area of research(24)
7. Automatic processing of tweets is a challenge (24)
8. Use of LSTM is explored in only few works (13,14,19,23,30)

IV. Research Questions

9. Why do we need automated fake news detection system for Marathi/Hindi Tweets?
10. How to obtain maximum accuracy of the automated system?
11. Does the introduction of multinomial classification affect the accuracy?
12. Does this automated detection system meet social need to offer authentic news to the users?
13. Which type of algorithms would give optimum results?

V. STATEMENT OF PROBLEM

To develop automatic fake news detection system for Marathi and Hindi Tweets based on misinformation about Politics, Health and Entertainment. The task of classifying news manually requires authentic references to the sources stating the facts as true or false. Also one has to have deep knowledge and reading about the domain of text to be classified. In this research, we are going to discuss the problem of classifying fake news articles using Deep learning techniques- Transformer approach, CNN and LSTM. The data we will use in our work is collected from the Twitter and World Wide Web and will contain news articles from the domains of Politics, Health and entertainment. The principal goal of the research is to differentiate fake articles from true news. We will extract different textual features from the text to feed as an input to our model. Learning models will be trained and parameter-tuned to obtain optimal accuracy.

VI. OBJECTIVES OF THE STUDY

1. To identify Twitter fake news in Marathi and Hindi language related to Politics, Health and Entertainment.
2. To find out suitable feature extraction method.
3. To develop suitable classification methodology.
4. To apply appropriate classifier and cross validate the result (with good accuracy).
5. To identify combination of methods in order to accurately detect fake news in social media
6. To classify the fake news into below category thus making it a multiclass classification-
 - a) Rumor
 - b) Fake news
 - c) Hoax
 - d) Clickbait
 - e) Misinformation
7. To prove the results for Hindi and Marathi languages.
8. To develop a tool to automatically detect the input text as fake or authentic.

VII. SCOPE AND LIMITATIONS

We are working on Fake information on Tweeter during last 5 years for the Tweets written in Devnagri Script (Hindi and Marathi). Also we are verifying the probable fake tweets by authenticating the truthfulness of the tweet with dataset of online edition of popular Print media from Maharashtra like Maharashtra Times, Lokmat and Sakal.

VIII. METHODOLOGY, TOOLS & TECHNIQUES

A. Overall methodology

Step1:

1. The first step for our research was to go through the related papers and get the datasets for training the model.
2. We found that very less data was present for Marathi and Hindi.

Step 2:

1. So, we need to start scraping the Marathi and Hindi tweets and labeling them manually as fake or authentic
2. We will run models already present and then provide results
3. Then we will build our own Deep learning model based on Transformer approach, CNN and LSTM and compare the results.

B. Datasets:

- Target Dataset: Tweet based data about Marathi and Hindi language related to Politics, Health and Entertainment
- Available dataset-
 - <https://www.kaggle.com/datasets/disibig/marathi-news-dataset> (scrapped data from Marathi news website) (Classification as State, Entertainment, sports)
 - <https://www.kaggle.com/datasets/atulpatare/abp-live-marathi-news-headlines> (Classification as Maharashtra, Bollywood, sports)
 - A Marathi Tweet-based Sentiment Analysis Dataset (by Atharv Kulkarni, et al)
- Proposed dataset: Need to be created

A. Proposed dataset creation Approach

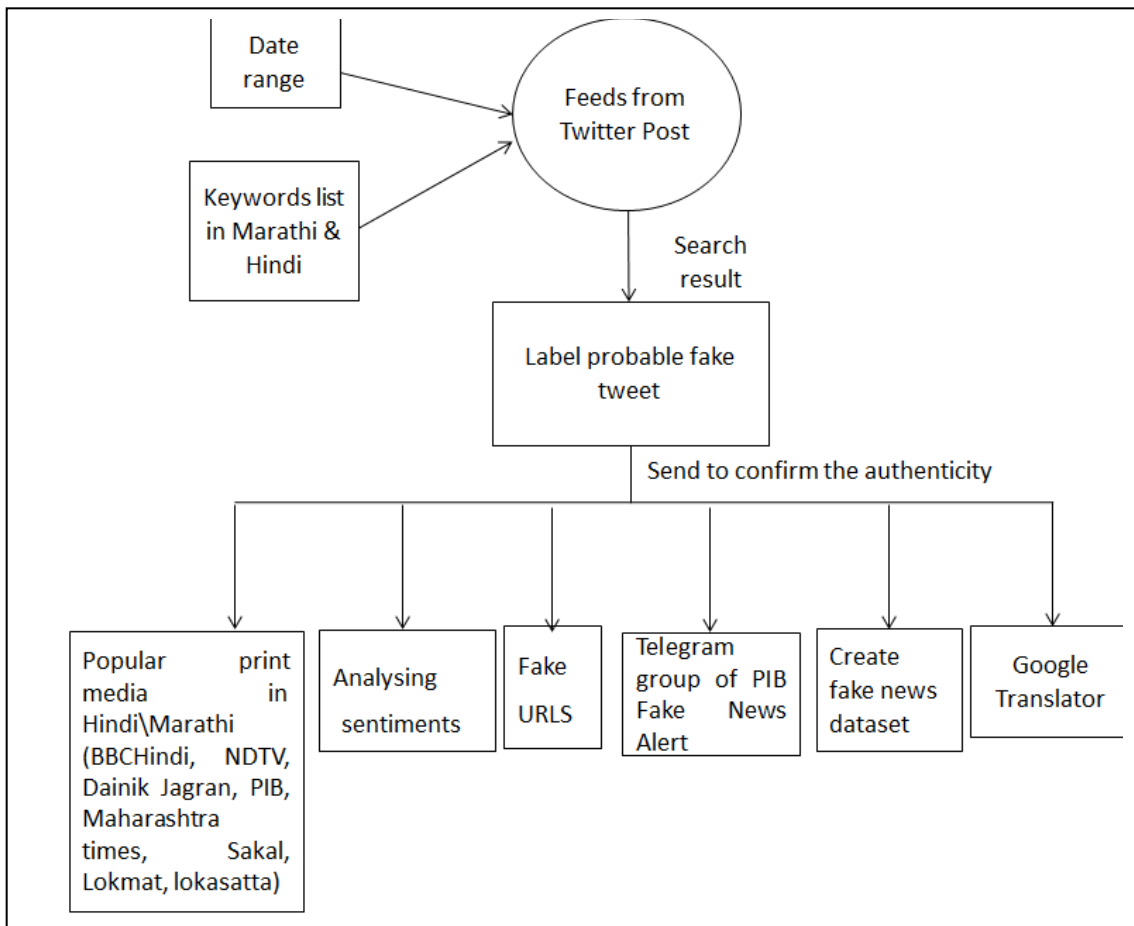
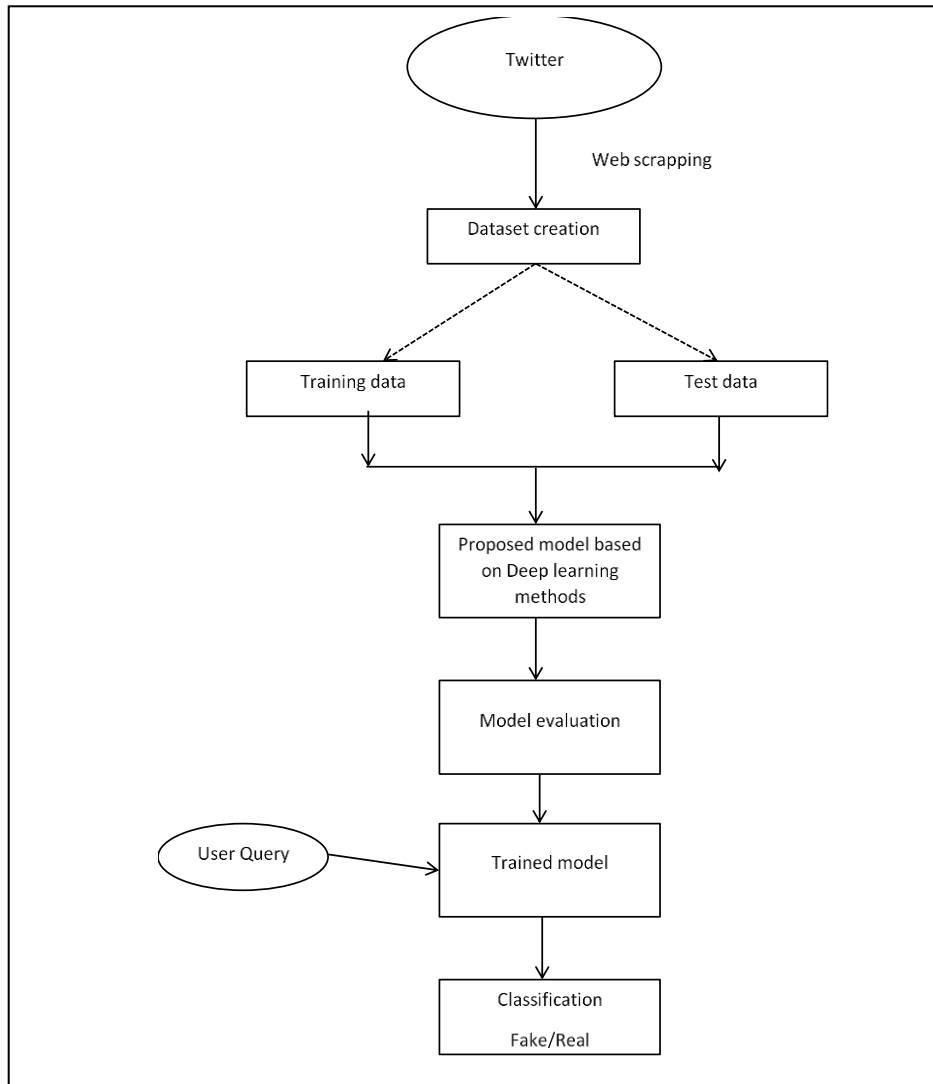


Figure 1: Proposed dataset creation Approach



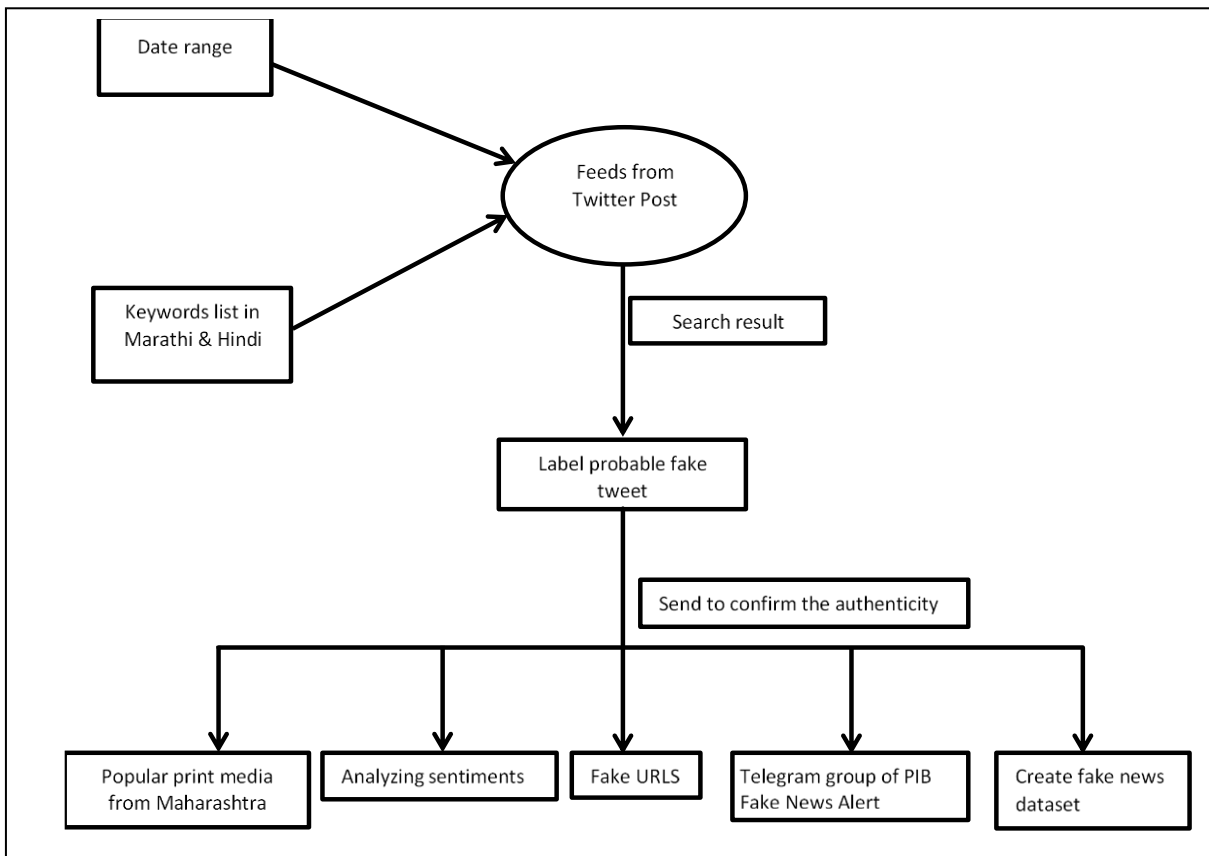


Figure 2: Proposed Architecture

- B. Deep Learning Techniques: Transformer approach, CNN and LSTM
- C. Programming Language: Python libraries for social networking services (SNS)

IX. CONCLUSION

Any kind of public communication medium like news papers, magazines, news channels, etc are the major sources to spread the information about the events happening around. Since last couple of years, the social media platforms like Facebook, Tweeter, Instagram, WhatsApp, etc are becoming widely used as a means of communication by the people belonging to differrent age groups, varied professions and societies. Along with the positive feature to allow users to have conversations, share information and create web content; social media platforms also pose an issue of media manipulation to create and spread fake news. The manipulated information can create biased views among the readers and hence affect the social health by the incidences like crimes, mob violence and riots. Fake news is intentionally written to mislead readers to believe false information, which makes it difficult and nontrivial to detect based on news content It is important to identify news as real or fake so as to stop its spread among citizens.

We have reviewed around 35 papers and concluded our review as per below table.

| Paper Ref. No. | Work done on low resource languages | |
|----------------|-------------------------------------|----------------|
| 7 | Chinese | RL |
| 22 | Hindi | ML and DL |
| 23 | Hindi | ML and LSTM |
| 24 | Arabic | ML |
| 25 | Urdu | DL (CNN, BERT) |
| 26 | English, Hindi, Bengali | ML and BERT |
| 27 | Hindi | ML |

References

1. S. Aphiwongsophon, and P. Chongstitvatana. "Detecting fake news with Machine Learning methods." *15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*. IEEE, 2018. 528-531.
2. Aditi Gupta, Hemank Lamba, Ponnurangam Kumaraguru, and Anupam Joshi. "Faking Sandy: characterizing and identifying fake images on Twitter during Hurricane Sandy." *WWW '13 Companion: Proceedings of the 22nd International Conference on World Wide Web*. ACM, 2013. 729–736.
3. Anjali Jain, Harsh Khatter, and AvinashShakya. "A smart System for Fake News Detection Using Machine Learning." *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*. IEEE, 2019.
4. Antˆonio Diogo Forte Martins, Lucas Cabral, Pedro Jorge Chaves Mourˆao, Jos´e Maria Monteiro, and Javam Machado. "Detection of Misinformation about COVID-19 in Brazilian Portuguese WhatsApp Messages Using Deep Learning." *23rd International Conference on Enterprise Information Systems (ICEIS 2021)*. SCITEPRESS – Science and Technology Publication, 2021. 63-74.
5. Arkaitz Zubiaga, Elena Kochkina, Maria Liakata, Rob N Procter, and Michal Lukasik. "Stance classification in Rumours as a Sequential Task Exploiting the Tree Structure of Social Media Conversations." *Proceedings of the International Conference on Computational Linguistics (COLING)*. 2016.
6. ARKAITZ ZUBIAGA, KALINA BONTCHEVA, AHMET AKER, MARIA LIAKATA, and ROB PROCTER. "Stance classification in Rumours as a Sequential Task Exploiting the Tree Structure of Social Media Conversations." (ACM) 51, no. 2 (2018).
7. Ashima Yadav, and Dinesh Kumar Vishwakarma. "A comparative study on bio-inspired algorithms for sentiment analysis." *Springer Nature*, 2020.
8. Atharva Kulkarni, Meet Mandhane, Manali Likhitar, Gayatri Kshirsagar, and Raviraj Joshi. "L3CubeMahaSent: A Marathi Tweet-based Sentiment Analysis Dataset." *Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, WASSA@EACL 2021*,. 2021.
9. Atharva Kulkarni, Meet Mandhane, Manali Likhitar, Gayatri Kshirsagar, Jayashree Jagdale, and Raviraj Joshi. "Experimental Evaluation of Deep Learning models for Marathi Text." *Proceedings of the 2nd International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications* . Springer Link, 2022. 605–613.
10. Chandra Mouli Madhav Kotteti, Xishuang Dong, Na Li, and Lijun Qian. "Fake News Detection Enhancement with Data Imputation." *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing*. IEEE, 2018. 187-192.
11. Cody Buntain, and Jennifer Golbeck. "Automatically Identifying Fake News in Popular Twitter Threads." *IEEE International Conference on Smart Cloud (SmartCloud)*. IEEE Explore, 2017.
12. Dan S. Nielsen, and Ryan McConville. "MuMiN: A Large-Scale Multilingual Multimodal Fact-Checked Misinformation Social Network Dataset." 2022.
13. Debanjana Kar, Mohit Bhardwaj, Suranjana Samanta, and Amar Prakash Azad. <https://arxiv.org/>. arxiv. 2020. <https://arxiv.org/abs/2010.06906> (accessed 2022).
14. DHIREN ROHERA, et al. "A Taxonomy of Fake News Classification Techniques: Survey and Implementation Aspects." *IEEE Access* 10 (2022): 30367-30394.
15. Eugenio Tacchini, Gabriele Ballarin, Marco L. Della Vedova, Stefano Moret, and Luca de Alfaro. "Some Like it Hoax:Automated Fake News Detection in Social Networks." (Proceedings of the Second Workshop on Data Science for Social Good (SoGood), Skopje, Macedonia) 1960 (2017).
16. Gourab Nath, Gnaneeswar Adhi, Prasanna Kumar Thandul, and Manikandan M. "An Attempt to Detect Fake Messages Circulated on WhatsApp." *Proceedings of 7th International Conference of Business Analytics and Intelligence*. 2020.

17. Hernon, Peter. "Disinformation and misinformation through the internet: Findings of an exploratory study." *Government Information Quarterly* (ELSEVIER) 12, no. 2 (1995): 133-139.
18. Iftikhar Ahmad, Muhammad Yousaf, Suhail Yousaf, and Muhammad Ovais Ahmad. "Fake News Detection Using Machine Learning Ensemble Methods." *Wiley Hindawi*, 2020: 11.
19. Jamal Abdul Nasir, Osama Subhani Khan, and Iraklis Varlamis. "Fake news detection: A hybrid CNN-RNN based deep learning approach." *International Journal of Information Management Data Insights*, 2021.
20. Jathin Badam, Akash Bonagiri, KVLN Raju, and Dipanjan Chakraborty. "Aletheia: A Fake News Detection System for Hindi." 2022: 255-259.
21. Jaynil Gaglani, Yash Gandhi, Shubham Gogate, and Aparna Halbe. "Unsupervised WhatsApp Fake News Detection using Semantic Search." *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020)*. IEEE Xplore, 2020. 285-298.
22. Julio C. S. Reis, André Correia, Fabrício Murai, Adriano Veloso, and Fabrício Benevenuto. "Explainable Machine Learning for Fake News Detection." *WebSci '19: Proceedings of the 10th ACM Conference on Web Science*. ACM, 2019. 17-26.
23. Julio C. S. Reis, André Correia, Fabrício Murai, Adriano Veloso, and Fabrício Benevenuto. "Supervised Learning for Fake News Detection." *IEEE Intelligent Systems* (IEEE) 34, no. 2 (2019): 76 - 81.
24. Kai Shu, Deepak Mahudeswaran, Suhang Wang, Dongwon Lee, and Huan Liu. "FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media." 8, no. 3 (2020).
25. Kai Shu, Deepak Mahudeswaran, Suhang Wang, Dongwon Lee, and Huan Liu. "FakeNewsNet: A Data Repository with News Content, Social Context and Spatiotemporal Information for Studying Fake News on Social Media." (Mary Ann Liebert, Inc., publishers) 8, no. 3 (2020).
26. KARISHMA SHARMA, FENG QIAN, HE JIANG, NATALI R, MING ZHANG, and YAN LIU. "Combating Fake News: A Survey on Identification and Mitigation Techniques." (ACM Transactions on Intelligent Systems and Technology) 10, no. 3 (2019).
27. Kuo-Cheng Chung a, Chun-Hung Chen a, and Hsueh-Hsuan T. "Social media privacy management strategies: A SEM analysis of user privacy behaviors." *Computer Communications* (Elsevier), 2021: 1220133.
28. Liang Wu, and Huan Liu. "Debunking Rumors in Social Networks: A Timely Approach." *11th ACM Conference on Web Science (WebSci '19)*. Boston, MA, USA: ACM, 2019.
29. Lucas Cabral, Jos'e Maria Monteiro, Jos'e Maria Monteiro, C'esar Lincoln Mattos, and Pedro Jorge Chaves Mour'ao. "FakeWhastApp.BR: NLP and Machine Learning Techniques for Misinformation Detection in Brazilian Portuguese WhatsApp Messages." *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021)*. n.d. 63-74.
30. Maaz Amjad, Sabur Butt, Hamza Imam Amjad, Alisa Zhila, Grigori Sidorova, and Alexander Gelbukh. "Overview of the Shared Task on Fake News Detection in Urdu at FIRE 2021." *FIRE 21: Forum for Information Retrieval Evaluation, December 13–17, 2021, India*. CEUR Workshop Proceedings (CEUR-WS.org), 2021.
31. Menglong Lu, Zhen Huang, Binyang Li, Zheng Qin, and DongSheng Li. "SIFTER: A Framework for Robust Rumor Detection." *IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING* 30 (2022): 429-442.
32. Miguel A. Alonso, David Vilares, Carlos Gómez-Rodríguez, and Jesús Vilares. "Sentiment Analysis for Fake News Detection." *Electronics*, 2021: 1-32.
33. N. Kanagavalli, and S. Baghavathi Priya. "Social Networks Fake Account and Fake News Identification with Reliable Deep Learning." *Intelligent Automation & Soft Computing* (Tech Science Press) 33, no. 1 (2022): 191-205.
34. N. Raghavendra Sai, and Satarla SAI Kumar. "Battling Fake News: A Survey on Mitigation Techniques and Identification." *Proceedings of the Fifth International Conference on Trends in Electronics and Informatics (ICOEI)*. IEEE, n.d. 828-834.
35. Pedro Faustini, and Thiago Cov'oes. "Fake News Detection Using One-Class Classification." *8th Brazilian Conference on Intelligent Systems (BRACIS)*. IEEE, 2019. 592-597.

36. Preston, Stephanie , Anthony Anderson, David J. RobertsonID, Mark P. Shephard, and Narisong Huhe. "Detecting fake news on Facebook: The role of emotional intelligence." *PLOS ONE* (Public Library of Science (PLoS)) 16, no. 3 (2021).
37. Priyanga V.T, Sanjanasri J.P, Vijay Krishna Menon, Gopalakrishnan E.A, and Soman K. P. "Exploring fake news identification using word and sentence embeddings." *Journal of Intelligent & Fuzzy Systems* (IOS Press) 41, no. 5 (2021): 5441-5448.
38. Rajshree Varma, and Prathamesh Churi. "A systematic survey on deep learning and machine learning approaches of fake news detection in the pre- and post-COVID-19 pandemic." (International Journal of Intelligent Computing and Cybernetics) n.d.
39. Reema Aswani, Arpan Kumar Kar, and P. Vigneswara Ilavarasan. "Detection of Spammers in Twitter marketing: A Hybrid Approach Using Social Media Analytics and Bio Inspired Computing." *Springer Science+Business Media, LLC*, 2017.
40. Sadiq-Ur-Rahman Shifath, S. M., Mohammad Faiyaz Khan, and Md. Saiful Islam. "A transformer based approach for fighting COVID-19 fake news." (arXiv.org) 2021.
41. Sani Muhamad Isa, Gary Nico, and Mikhael Permana. "INDOBERT FOR INDONESIAN FAKE NEWS DETECTION." *ICIC International*. ICIC Express Letters, 2022. 289-297.
42. Sawinder Kaur, Parteek Kumar, and Ponnurangam Kumar. "Automating fake news detection system using multi-level voting model." *Springer-Verlag GmbH Germany, part of Springer Nature 2019* (SpringerLink), November 2020: 9049–9069.
43. Shailender Kumar, Subodh Kumar, Pooja Yadav, and Meghna Bagri. "A Survey on Analysis of Fake News Detection Techniques." *Proceedings of the International Conference on Artificial Intelligence and Smart Systems (ICAIS-2021)*. IEEE Xplore, 2021. 894-899.
44. —. "Social Networks Fake Account and Fake News Identification with Reliable Deep Learning." *Intelligent Automation & Soft Computing*, 2022.
45. Sudhanshu Kumar, and Thoudam Doren Singh. "Fake News Detection on Hindi News Dataset." *ScienceDirect*, April 2022.
46. Thaer Thaher, Mahmoud Saheb, Hamza Turabieh, and Hamouda Chantar. "Intelligent Detection of False Information in Arabic Tweets Utilizing Hybrid Harris Hawks Based Feature Selection and Machine Learning Models." (Symmetry) 2021.
47. Tracy Jia Shen, Robert Cowell, Aditi Gupta, Thai Le, Amulya Yadav, and Dongwon Lee. "How Gullible Are You? Predicting Susceptibility to Fake News." *WebSci '19: Proceedings of the 10th ACM Conference on Web Science*. Boston, MA, USA: ACM, 2019. 287–288.
48. Usman Saeed, Hammad Fahim, and Dr. Farid Shirazi. "Profiling Fake News Spreaders on Twitter." *11th CLEF conference, Greece*. r Creative Commons License Attribution 4.0 International (CC BY 4.0), 2020.
49. Varshil Mehta, and Wenlin Han. "Fake News Detection in Social Networks Using Machine Learning and Deep Learning: Performance Evaluation." *International Conference on Industrial Internet (ICII)*. IEEE, 2019. 375-380.
50. Vivek Srivastava, and Mayank Singh. "PoliWAM: An Exploration of a Large Scale Corpus of Political Discussions on WhatsApp Messenger." <https://arxiv.org/>. W-NUT 2021, EMNLP. 2021. <https://arxiv.org/abs/2010.13263> (accessed 2022).
51. Z Khanam, B N Alwasel, H Sirafi, and M Rashid. "Fake News Detection Using Machine Learning Approaches." *IOP Conf. Series: Materials Science and Engineering*. n.d.