

A machine learning approach for Fake news detection from Urdu social media posts

Abhinav Kumar¹, Jyoti Kumari²

¹Department of Computer Science & Engineering, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India

²Department of Computer Science & Engineering, National Institute of Technology Patna, India

Abstract

Fake news has the potential to mislead the public, damage social order, undermine government legitimacy, and pose a major danger to societal stability. As a result, early identification of fake news via Internet platforms is critical. The majority of previous research has focused on detecting false news in resource-rich languages like English, Hindi, and Spanish. The current study makes use of an Urdu language dataset to detect fake news. Three different models have been proposed in the paper. The first one is a dense neural network (DNN)-based model, the second one is a Majority voting-based ensemble model, and the third one is the Probability averaging-based ensemble model. The proposed dense neural network-based model performed better with character n-gram TF-IDF features and achieved a macro F_1 -score of 0.59 and an accuracy of 0.72. The code for the proposed models is available at https://github.com/Abhinavkmr/Urdu_Fake_News_Detection.git

Keywords

Fake news, Urdu, Social media post, Machine learning

1. Introduction

People are more inclined to pick an online platform for generating or consuming news because of the ease of access and freedom to distribute Internet content [1, 2]. Several news are initially reported on the Internet before being broadcast on traditional news channels [3, 4, 5]. However, some people misuse the benefits of contemporary technology by broadcasting fake news on these platforms to make fun of a person/society, cause fear, or make money [6, 7, 8, 9]. A piece of false news spreads faster than a piece of factual news due to its high sentimental value. Fake news' extensive propagation has major negative consequences for both individuals and society. Fake news must be recognized and disseminated as quickly as possible to limit the negative implications. Therefore the identification of fake news has emerged as one of the most investigated subjects in natural language processing. The highlighted issue would have been easier to solve if the news on the Internet had only been available in a single language. However, there are over 5000 languages spoken throughout the world. It's virtually hard to create a generalized false news detection system that works in all languages. For resource-rich languages like English, Hindi, Spanish, and others, significant effort has been done.

Verónica et al. [10] extracted lexical, syntactic, and semantic information from English news to detect false news. Duran et al. [11] suggested a model that uses lexical characteristics including

Forum for Information Retrieval Evaluation, December 13-17, 2021, India

✉ abhinavanand05@gmail.com (A. Kumar); j2kumari@gmail.com (J. Kumari)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

bag-of-words, parts of speech, and n-grams to identify false news in Spanish. Giachanou et al. [12] and Ghanem et al. [13] proposed long-short term-memory network-based model for fake news detection. Singh et al. [9] proposed an attention-based LSTM model for the identification of rumour from social media. Anu and Abhinav [14] proposed a deep ensemble-based model for the identification of COVID-19 fake news posted over social media. An extensive survey on fake news detection can be seen in Roy and Chahar [15].

Despite having over 100 million speakers globally, Urdu has a limited number of labeled datasets, making it a resource-poor language in NLP. As a result, only a few efforts for detecting false news in Urdu have been reported. Amjad et al. [16, 17] created a benchmark dataset for Urdu fake news. Kumar et al. [6] extracted character-level features Urdu news articles and proposed a dense neural network for the identification of Urdu fake news. Khilji et al. [18] proposed a generalized autoregressor based model whereas, Reddya et al. [19] proposed a GRU-based model to identify fake news from Urdu news articles. This work proposes three different models: (i) Dense Neural Network (DNN)-based model, (ii) Majority voting-based ensemble model, and (iii) Probability averaging-based ensemble model for the identification of fake news from Urdu news articles. The proposed models are validated with the dataset published in the *UrduFake-FIRE2021* [20, 21] shared task.

The rest of the paper is organized as follows: The following is how the rest of the article is structured: The details of the proposed model, as well as the dataset description and feature extraction, are explained in Section 2. Section 3 details a variety of experiments and their outcomes. Finally, Section 4 brings the article to a close-by presenting the most important finding.

2. Methodology

The overall flow diagram of the proposed models can be seen in Figure 1. Three different models were proposed for the fake news identification from Urdu news: (i) Dense Neural Network (DNN)-based model, (ii) Majority voting-based ensemble model, and (iii) Probability averaging-based ensemble model. The overall data statistic used to validate the proposed system can be seen in Table 1.

Table 1

Overall data statistic used to validate the proposed models

Class	Train	Test
Real	600	200
Fake	438	100
Total	1038	300

2.1. Dense neural network (DNN)-based model

The suggested dense neural network (DNN) architecture is made up of four layers, each with 1,024, 512, 128, and 2-neurons. The top 15,000, uni-gram, bi-gram, and tri-gram character-level

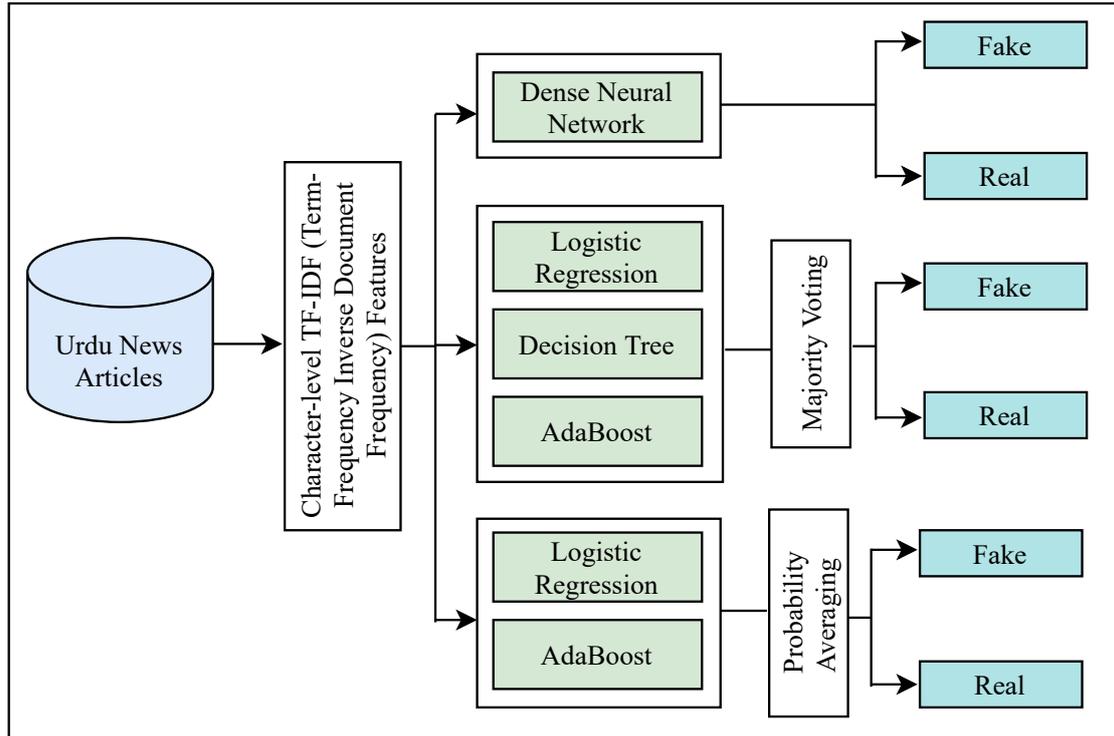


Figure 1: Overall flow diagram for the proposed methodology

TF-IDF features are utilized as input to the DNN model. We conducted extensive experiments to find the best-suited hyper-parameters because the performance of deep learning-based models is sensitive to the hyper-parameters chosen. The best results were obtained using a dropout rate of 0.3, a learning rate of 0.001, a batch size of 16, binary cross-entropy as a loss function, and Adam as the optimizer with 100 epoch training.

2.2. Majority voting-based ensemble model

In the case of the Majority voting-based ensemble model, predictions of Logistic Regression, Decision Tree, and Adaboost classifiers are used to find the final class value. The final class value is decided based on the majority voting. The overall diagram of the model can be seen in Figure 1. To provide input to the classifiers, top 30,000 uni-gram, bi-gram, and tri-gram character-level TF-IDF features were used.

2.3. Probability averaging-based ensemble model

In the case of the Probability averaging-based ensemble model, Logistic Regression and AdaBoost classifiers are used to get the class probability value for fake and real classes. Then the class-wise probability averaging was performed to get the final probability and based on the final

probability final class level is determined. The overall flow diagram of the model can be seen in Figure 1. To provide input to the classifier, top 30,000 uni-gram, bi-gram, and tri-gram character-level TF-IDF features were used. To implement all the classifiers, Sklearn Python library¹ is used with default parameters.

Table 2

Results of different models for fake news detection from Urdu article

Models	Class	Precision	Recall	F_1 -score	Accuracy
Dense neural network (DNN)-based model	Fake	0.79	0.23	0.36	0.72
	Real	0.72	0.97	0.82	
	Macro Avg.	0.75	0.60	0.59	
Majority voting-based ensemble model	Fake	0.87	0.13	0.23	0.70
	Real	0.69	0.99	0.82	
	Macro Avg.	0.78	0.56	0.52	
Probability averaging-based ensemble model	Fake	1.00	0.05	0.10	0.68
	Real	0.68	1.00	0.81	
	Macro Avg.	0.84	0.53	0.45	

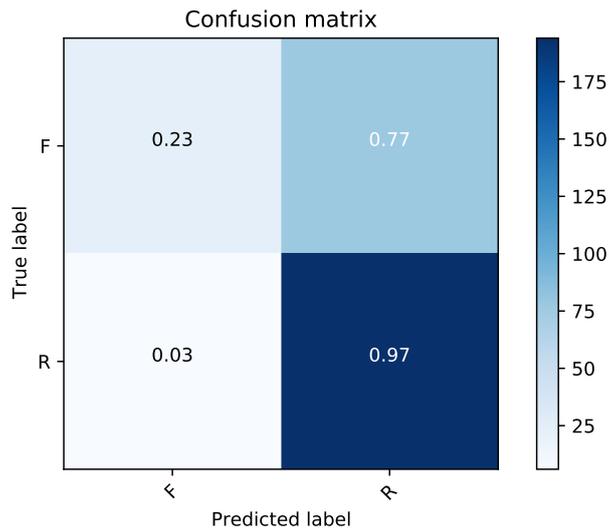


Figure 2: Confusion matrix for dense neural network

3. Results

The performance of the proposed models is measured in terms of precision, recall, F_1 -score, and accuracy. Along with this, the confusion matrix is also plotted to visualize the performance.

¹<https://scikit-learn.org/stable/>

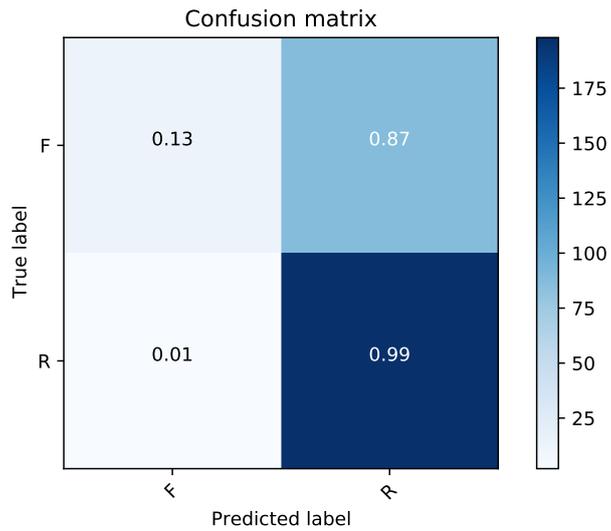


Figure 3: Confusion matrix for majority LR, DT, and AdaBoost

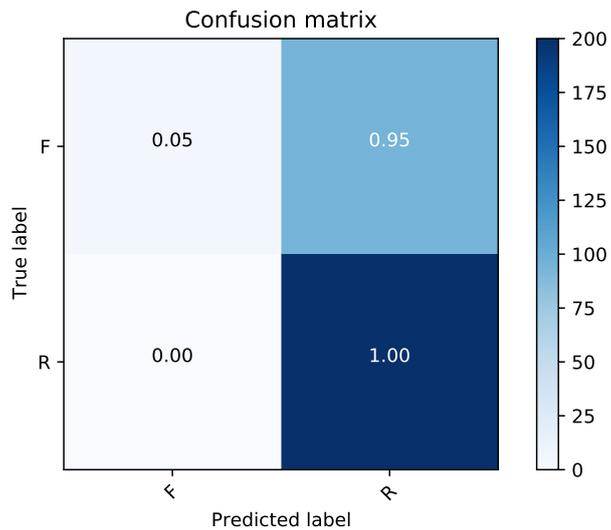


Figure 4: Confusion matrix for averaging LR and DT

The results for different models are listed in Table 2. The accuracy of the suggested DNN model was 0.72 and the macro F_1 -score was 0.59. The suggested method achieves a recall of 0.23 for the fake class. Figure 2 depicts the DNN model's confusion matrix. The suggested Majority voting-based ensemble model has a macro F_1 -score of 0.52 and an accuracy of 0.70. It had a recall of 0.13 for the false article class. Figure 3 shows the confusion matrix for the ensemble model based on majority voting. The proposed Probability averaging-based ensemble model is able to achieve a macro F_1 -score of 0.45 and an accuracy of 0.68. For the fake article class, it

achieved a recall of 0.05. The confusion matrix for the Probability averaging-based ensemble model can be seen in Figure 4.

4. Conclusion

The widespread dissemination of erroneous information has affected both individuals and society. In this paper, we suggest three distinct methods for detecting false news in Urdu news articles. With a macro $F1$ -score of 0.59 and an accuracy of 0.72, the suggested dense neural network-based model fared better. In the future, a more robust ensemble-based model for obtaining classification accuracy might be created. For the identification of fake news from Urdu news articles, a Transformer-based approach can also be investigated.

References

- [1] A. Kumar, J. P. Singh, Location reference identification from tweets during emergencies: A deep learning approach, *International journal of disaster risk reduction* 33 (2019) 365–375.
- [2] A. Kumar, J. P. Singh, S. Saumya, A comparative analysis of machine learning techniques for disaster-related tweet classification, in: *2019 IEEE R10 Humanitarian Technology Conference (R10-HTC)*(47129), IEEE, 2019, pp. 222–227.
- [3] A. Kumar, J. P. Singh, Y. K. Dwivedi, N. P. Rana, A deep multi-modal neural network for informative twitter content classification during emergencies, *Annals of Operations Research* (2020) 1–32.
- [4] J. P. Singh, Y. K. Dwivedi, N. P. Rana, A. Kumar, K. K. Kapoor, Event classification and location prediction from tweets during disasters, *Annals of Operations Research* 283 (2019) 737–757.
- [5] A. Kumar, N. C. Rathore, Relationship strength based access control in online social networks, in: *Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems: Volume 2*, Springer, 2016, pp. 197–206.
- [6] A. Kumar, S. Saumya, J. P. Singh, NITP-AI-NLP@ UrduFake-FIRE2020: Multi-layer dense neural network for fake news detection in Urdu news articles., in: *FIRE (Working Notes)*, 2020, pp. 458–463.
- [7] K. Shu, A. Sliva, S. Wang, J. Tang, H. Liu, Fake news detection on social media: A data mining perspective, *ACM SIGKDD explorations newsletter* 19 (2017) 22–36.
- [8] J. P. Singh, N. P. Rana, Y. K. Dwivedi, Rumour veracity estimation with deep learning for Twitter, in: *International Working Conference on Transfer and Diffusion of IT*, Springer, 2019, pp. 351–363.
- [9] J. P. Singh, A. Kumar, N. P. Rana, Y. K. Dwivedi, Attention-based LSTM network for rumor veracity estimation of tweets, *Information Systems Frontiers* (2020) 1–16. doi:<https://doi.org/10.1007/s10796-020-10040-5>.
- [10] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, R. Mihalcea, Automatic detection of fake news, in: *Proceedings of the 27th International Conference on Computational Linguistics*, Association for Computational Linguistics, 2018, pp. 3391–3401.
- [11] J.-P. Posadas-Durán, H. Gómez-Adorno, G. Sidorov, J. J. M. Escobar, Detection of fake

- news in a new corpus for the spanish language, *Journal of Intelligent & Fuzzy Systems* 36 (2019) 4869–4876.
- [12] A. Giachanou, P. Rosso, F. Crestani, Leveraging emotional signals for credibility detection, in: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2019, pp. 877–880.
 - [13] B. Ghanem, P. Rosso, F. Rangel, An emotional analysis of false information in social media and news articles, *ACM Trans. Internet Technol.* 20 (2020). URL: <https://doi.org/10.1145/3381750>. doi:10.1145/3381750.
 - [14] A. Priya, A. Kumar, Deep ensemble approach for COVID-19 fake news detection from social media, in: *2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, IEEE, 2021, pp. 396–401.
 - [15] P. K. Roy, S. Chahar, Fake profile detection on social networking websites: A comprehensive review, *IEEE Transactions on Artificial Intelligence* 1 (2020) 271–285. doi:10.1109/TAI.2021.3064901.
 - [16] M. Amjad, G. Sidorov, A. Zhila, H. Gómez-Adorno, I. Voronkov, A. Gelbukh, Bend the Truth: A benchmark dataset for fake news detection in Urdu and its evaluation, *Journal of Intelligent & Fuzzy Systems* 39 (2020) 2457–2469. doi:10.3233/JIFS-179905.
 - [17] M. Amjad, G. Sidorov, A. Zhila, Data augmentation using machine translation for fake news detection in the Urdu language, in: *Proceedings of The 12th Language Resources and Evaluation Conference*, 2020, pp. 2537–2542.
 - [18] A. F. U. R. Khiljia, S. R. Laskara, P. Pakraya, S. Bandyopadhyaya, Urdu fake news detection using generalized autoregressors, in: *FIRE (Working Notes)*, 2020.
 - [19] S. M. Reddy, C. Suman, S. Saha, P. Bhattacharyya, A GRU-based fake news prediction system: Working notes for UrduFake-FIRE 2020., in: *FIRE (Working Notes)*, 2020, pp. 464–468.
 - [20] A. Maaz, S. Butt, H. I. Amjad, A. Zhila, G. Sidorov, A. Gelbukh, Overview of the shared task on fake news detection in Urdu at FIRE 2021., in: *In CEUR Workshop Proceedings*, 2021.
 - [21] A. Maaz, S. Butt, H. I. Amjad, A. Zhila, G. Sidorov, A. Gelbukh, UrduFake@ FIRE2021: Shared track on fake news identification in Urdu., in: *In Forum for Information Retrieval Evaluation*, 2021.