Habesha@DravidianLangTech 2024: Detecting Fake News Detection in Dravidian Languages using Deep Learning

Mesay Gemeda Yigezu, Olga Kolesnikova, Grigori Sidorov and **Alexander Gelbukh**
Instituto Politécnico Nacional (IPN),
Centro de Investigación en Computación (CIC), Mexico City, Mexico
Correspondence:mgemedak2022@cic.ipn.mx

Abstract

This research tackles the issue of fake news by utilizing the RNN-LSTM deep learning method with optimized hyperparameters identified through grid search. The model's performance in multi-label classification is hindered by unbalanced data, despite its success in binary classification. We achieved a score of 0.82 in the binary classification task, whereas in the multi-class task, the score was 0.32. We suggest incorporating data balancing techniques for researchers who aim to further this task, aiming to improve results in managing a variety of information.

1 Introduction

Getting updates about what's happening in the world is important for us. It helps us learn more about the different things going on. Many people like to watch the news or read the newspaper in the morning with a cup of tea to stay informed. Fake news is one of the most significant scourges in our digitally interconnected world. It is broadly defined as a subject encompassing news, data, reports, and information that is either wholly or partially false (Kong et al., 2020).

If the news is not true, it can confuse people and spread false information. Sometimes, fake news is used to create rumors or harm the reputation of political leaders. To tackle this issue, we've suggested a system to identify fake news. However, with the enormous amount of data on the internet and social media, it's challenging to manually check if each piece of news is fake or not (Yigezu et al., 2023e; Bade, 2021).

This pervasive phenomenon acts as a wildfire, influencing countless individuals daily. The reach of fake news is extensive and can pose a significant threat to a nation's security, economy, prosperity, and the well-being of its citizens. Unfortunately, many people may not be fully aware of how profoundly fake news can affect the matters that surround them, and they may lack the necessary skills to discern and handle such situations when they arise (Arif et al., 2022; Yigezu et al., 2023c; Bade and Afaro, 2018).

Fake news is crafted with the explicit intent of disseminating information under the guise of propaganda or a hoax, ultimately aimed at achieving financial or political gains (Yigezu et al., 2023d). This deceptive practice manipulates public opinion, steering it towards falsehoods and distortions. The creators of fake news strategically weave narratives designed to exploit vulnerabilities and biases, with the ultimate goal of influencing individuals and, in turn, shaping societal perceptions. This insidious tactic not only undermines the credibility of information but also poses a significant threat to the foundations of democracy and the well-being of communities. Recognizing the nefarious intentions behind fake news is essential for fostering a more discerning public and cultivating a media landscape rooted in truth and integrity (Yigezu et al., 2023b; Shahiki-Tash et al., 2023).

The organizers of a shared task organized this study, which consists of two tasks: identifying multiple labels in Malayalam news and classifying a given social media text as either original or fake (Subramanian et al., 2024).

The organization of this paper is as follows: In Section 2, we thoroughly examine several related research studies. Section 3 focuses on explaining the tasks we are considering, while Section 4 gives a detailed discussion of the chosen methodology and experiments carried out for the task. Section 5 is where we present results and have discussions about them. Lastly, in Section 6, we draw conclusions and discuss possible future trends in research within this field.

2 Related Works

Numerous researchers have delved into the study of fake news detection, employing a variety of ap-

Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pages 156–161

proaches to unravel the complexities of understanding and categorizing fake news. Notably, these approaches encompass machine learning, deep learning, and transformer-based methodologies. Below, we explore a few notable studies that showcase the diversity of techniques employed:

In earlier studies, researchers leveraged on the traditional machine learning techniques (Reis et al., 2019; Shu et al., 2017; Yigezu et al., 2023a; Tash et al., 2022; Liu and Wu, 2018; Singh et al., 2018).

Ahmed et al. (2017) introduces a model for fake news detection utilizing n-gram analysis and machine learning methodologies. The study delves into an examination and comparison of two distinct feature extraction techniques and six varied machine classification methods. The experimental evaluation reveals optimal performance when employing Term Frequency-Inverted Document Frequency (TF-IDF) as the feature extraction technique and Linear Support Vector Machine (LSVM) as the classifier, achieving an impressive accuracy rate of 92%.

Granik and Mesyura (2017) employed a straightforward method for fake news detection, utilizing a naive Bayes classifier. This approach was translated into a software system and evaluated against a dataset comprising Facebook news posts. The achieved classification accuracy on the test set reached approximately 74%, a commendable result given the relatively uncomplicated nature of the model. The article discusses various avenues for potential improvement, indicating that the obtained results offer insights into addressing the fake news detection problem using artificial intelligence methods.

Mahabub (2020) introduces an intelligent detection system for news classification, addressing both real and fake news tasks, utilizing an Ensemble Voting Classifier. The approach involves the incorporation of eleven well-established machine-learning algorithms, including Naïve Bayes, K-NN, SVM, Random Forest, Artificial Neural Network, Logistic Regression, Gradient Boosting, Ada Boosting, among others. Through cross-validation, the top three performing machine-learning algorithms are selected for integration into the Ensemble Voting Classifier. The experimental results validate the efficacy of the proposed framework, achieving an impressive accuracy rate of approximately 94.5%. Additionally, other key metrics such as ROC score, precision, recall, and F1 demonstrate outstanding performance. The proposed detection framework not only attains high accuracy but also effectively identifies crucial features within news data. These identified features hold promise for implementation in other classification techniques, extending the utility of the system to detect fake profiles and messages.

With the advent of deep learning, researchers explored the application of neural networks, (Kong et al., 2020; Kumar et al., 2020; Hiramath and Deshpande, 2019; Yigezu et al., 2022; Tash et al., 2023). These deep learning models demonstrated improved performance in capturing sequential dependencies within textual data.

To tackle the issue of fake news, the Thota et al. (2018) propose a neural network architecture designed to precisely predict the stance between a provided pair of headline and article body. Their model surpasses the performance of existing architectures by 2.5%, achieving an impressive accuracy rate of 94.21% on the test data.

The authors Sahoo and Gupta (2021) present an innovative approach to automatic fake news detection within the Chrome environment, enabling the detection of fake news on Facebook. Their methodology involves the utilization of various features linked to a Facebook account, coupled with certain news content features, to analyze account behavior through deep learning techniques. Through experimental analysis on real-world data, the authors demonstrate that their proposed fake news detection approach attains higher accuracy compared to existing state-of-the-art techniques.

In response to the proliferation of fake news, the imperative for computational methods to detect them has become increasingly apparent. The primary objective of fake news detection is to empower users to discern various forms of fabricated information. The determination of the news veracity hinges on a decision-making process influenced by previously encountered instances of fake or authentic news. Various models can be employed to discern deceptive news circulating on social media platforms. Kaliyar (2018) makes a two fold contribution to the field. Firstly, the authors introduce datasets encompassing both fake and real news, undertaking diverse experiments to design effective fake news detectors. Leveraging Natural Language Processing, Machine Learning, and deep learning techniques, they classify the datasets, providing a comprehensive assessment of fake news detection.

Their contribution extends to encompass fake news categorization, incorporating existing algorithms derived from machine learning techniques.

In this research Chauhan and Palivela (2021), a profound exploration into distinguishing false news from authentic sources is conducted through a deep learning-based approach. The cornerstone of the proposed model is a LSTM neural network. Complementing the neural network, a GloVe word embedding is employed to represent textual words as vectors. Additionally, tokenization is utilized for feature extraction or vectorization, enhancing the model's capacity. The integration of N-grams further refines the proposed approach. A comprehensive comparative analysis of multiple fake news detection techniques is undertaken. The results of the proposed model are meticulously evaluated using accuracy metrics, revealing an exceptional performance with an accuracy rate of 99.88%. This underscores the effectiveness of the LSTM-based model and its ability to discern false news with an exceedingly high level of accuracy.

3 Description of tasks and Dataset

There is an escalating demand for the detection of fake news within social media texts. The dataset for this crucial task has been generously provided by the organizers of the Fake News Detection in Dravidian Languages- DravidianLangTech@EACL 2024 (Subramanian et al., 2023) and our team, Habesha, actively participated in this shared endeavor.

Throughout our involvement in this collaborative initiative, we immersed ourselves in two distinctive tasks. The primary task aimed to classify social media texts as either original or fake. The data sources encompassed various social media platforms such as Twitter and Facebook. Given a social media post, the shared task mandated the classification of the content as either fake or authentic news.

In the second task, namely the Fake News Detection from Malayalam News (FakeDetect-Malayalam) shared task, researchers were provided with a platform to address the formidable challenge of identifying and flagging fake news within the realm of Malayalam-language news articles. Accurate misinformation detection is crucial for fostering trustworthy communication in an era of information overload. The core objective of the FakeDetect-Malayalam shared task was to inspire participants to develop effective models capable of accurately detecting and categorizing fake news articles in the Malayalam language into different categories. In this context, we considered five fake categories - False, Half True, Mostly False, Partly False, and Mostly True.

Through our engagement with these two tasks, our goal was to contribute to a comprehensive understanding of fake news detection, with a specific emphasis on addressing the nuances associated with YouTube comments in both tasks.

We utilized a total of 4,072 data points for training in Task 1, which involves classifying fake and original content. Additionally, 1,019 data points were allocated for evaluating the model's performance. In Task 2, focusing on fake news detection from Malayalam News, we employed 1,669 data points for training and reserved 250 data points for evaluation purposes. For a more comprehensive overview of the statistics related to these specific datasets, kindly refer to Figure 1 for Task 1 and Figure 2 for Task 2.

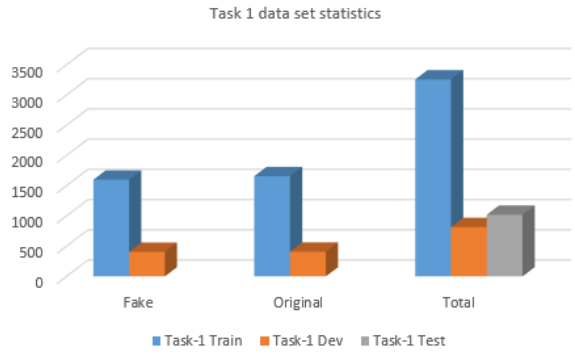

Figure 1: Task 1 data set statistics

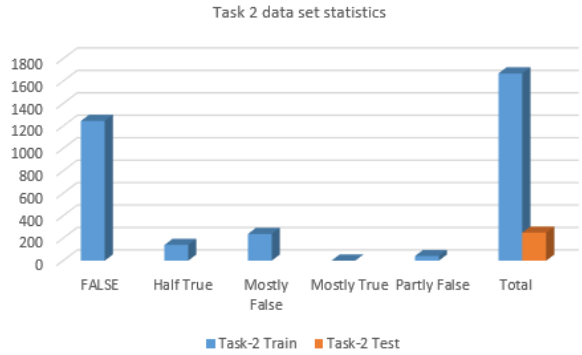

Figure 2: Task 2 data set statistics

In Figure 1, we observe an unbalanced distribution within the two-class dataset, indicating an uneven representation of the classes. In the meantime, Figure 2 shows that Task 2 is a multi-label

assignment with five classes, which shows how the data is spread out in a very different way. This imbalance poses a significant challenge, potentially resulting in biased predictions and impacting the overall performance of the trained model. Recognizing and addressing such imbalances becomes imperative for fostering a more robust and unbiased predictive model (Yigezu et al., 2023b; Tonja et al., 2022; Bade and Seid, 2018).

4 Methodology

This section offers an in-depth insight into the methodology applied in this study, with a specific emphasis on data preprocessing and the adoption of a deep learning approach for fake news detection. Our primary goal was to perform fake news detection for both tasks.

Initiating our analysis, we acknowledged the significance of meticulous data preprocessing. This critical step involved purifying and formatting the raw data to optimize its appropriateness for subsequent stages in our model.

4.1 Experiment

In our methodology for detecting fake news in both shared tasks, we implemented a deep learning approach, specifically utilizing Recurrent Neural Networks (RNN). The process involved several crucial stages to ensure a thorough and accurate analysis. We used Long Short-Term Memory (LSTM) layers to model sequential dependencies in the text, dropout layers to stop overfitting during training, and a sigmoid activation function to tell the difference between fake and real news. This helped us capture semantic relationships.

We used the task organizer's training and development datasets, the Adam Optimizer, the cross-entropy loss function, and test data to evaluate the results of the model training. A grid search was conducted to automate hyperparameter tuning, systematically exploring different configurations, including varying numbers of hidden units and epochs. This rigorous approach aimed to optimize the model's performance and enhance its ability to generalize to diverse datasets.

5 Results and Discussion

As outlined in Section 4.1, we implemented the Recurrent Neural Network (RNN) model for both shared tasks. Our outcomes revealed a macro-F1 score of 0.82 for task 1 and 0.32 for task 2.

In task one, our model exhibited promising performance by effectively distinguishing between fake and original news, resulting in a superior outcome compared to task two. This discrepancy in performance can be attributed to the nature of task two, involving multi-label classification with an unbalanced data set, which presented challenges for achieving satisfactory results. The inherent complexity of addressing multiple classes in task 2 posed difficulties, leading to suboptimal performance in contrast to the binary classification nature of task 1.

6 Conclusion and Future work

The dissemination of fake news or misinformation poses a significant challenge, steering information in undesired directions and impeding the acquisition of reliable and timely information. To address this issue, we employed the deep learning approach, specifically RNN-LSTM. For optimal model training, we utilized grid search methods to fine-tune hyper parameters.

In binary classification, our approach yielded favorable results, showcasing its effectiveness in distinguishing between genuine and fake information. However, the application of multi-label classification suffered from the presence of unbalanced data, which led to less than ideal results.

As a recommendation for researchers aiming to extend this task, we suggest incorporating data balancing techniques. By addressing the imbalance in the dataset, more robust and accurate results can be achieved in the context of multi-label classification, enhancing the model's ability to handle diverse and nuanced information effectively.

Acknowledgments

The work was done with partial support from the Mexican Government through the grant A1-S-47854 of CONACYT, Mexico, grants 20241816, 20241819, and 20240951 of the Secretaría de Investigación y Posgrado of the Instituto Politécnico Nacional, Mexico. The authors thank the CONACYT for the computing resources brought to them through the Plataforma de Aprendizaje Profundo para Tecnologías del Lenguaje of the Laboratorio de Supercómputo of the INAOE, Mexico and acknowledge the support of Microsoft through the Microsoft Latin America PhD Award.

References

Hadeer Ahmed, Issa Traore, and Sherif Saad. 2017. Detection of Online Fake News using N-gram analysis and Machine Learning Techniques. In *Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments: First International Conference, ISDDC 2017, Vancouver, BC, Canada, October 26-28, 2017, Proceedings 1*, pages 127–138. Springer.

Muhammad Arif, Atnafu Lambebo Tonja, Iqra Ameer, Olga Kolesnikova, Alexander Gelbukh, Grigori Sidorov, and Abdul Gafar Manuel Meque. 2022. CIC at CheckThat! 2022: Multi-class and Cross-lingual Fake News Detection. *Working Notes of CLEF*.

Girma Yohannis Bade. 2021. Natural Language Processing and Its Challenges on Omotic Language Group of Ethiopia. *Journal of Computer Science Research*, 3(4):26–30.

Girma Yohannis Bade and Akalu Assefa Afaro. 2018. Object Oriented Software Development for Artificial Intelligence. *American Journal of Software Engineering and Applications*, 7(2):22–24.

Girma Yohannis Bade and Hussien Seid. 2018. Development of Longest-Match Based Stemmer for Texts of Wolaita Language. *vol*, 4:79–83.

Tavishee Chauhan and Hemant Palivela. 2021. Optimization and Improvement of Fake News Detection using Deep Learning Approaches for Societal Benefit. *International Journal of Information Management Data Insights*, 1(2):100051.

Mykhailo Granik and Volodymyr Mesyura. 2017. Fake News Detection using Naive Bayes Classifier. In *2017 IEEE first Ukraine conference on electrical and computer engineering (UKRCON)*, pages 900–903. IEEE.

Chaitra K Hiramath and GC Deshpande. 2019. Fake News Detection using Deep Learning Techniques. In *2019 1st International Conference on Advances in Information Technology (ICAIT)*, pages 411–415. IEEE.

Rohit Kumar Kaliyar. 2018. Fake News Detection using a Deep Neural Network. In *2018 4th international conference on computing communication and automation (ICCCA)*, pages 1–7. IEEE.

Sheng How Kong, Li Mei Tan, Keng Hoon Gan, and Nur Hana Samsudin. 2020. Fake News Detection using Deep Learning. In *2020 IEEE 10th symposium on computer applications & industrial electronics (ISCAIE)*, pages 102–107. IEEE.

Sachin Kumar, Rohan Asthana, Shashwat Upadhyay, Nidhi Upreti, and Mohammad Akbar. 2020. Fake News Detection using Deep Learning Models: A novel approach. *Transactions on Emerging Telecommunications Technologies*, 31(2):e3767.

Yang Liu and Yi-Fang Wu. 2018. Early detection of Fake News on Social Media Through Propagation Path Classification with Recurrent and Convolutional Networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32.

Atik Mahabub. 2020. A Robust Technique of Fake News Detection using Ensemble Voting Classifier and Comparison with Other Classifiers. *SN Applied Sciences*, 2(4):525.

Julio CS Reis, André Correia, Fabrício Murai, Adriano Veloso, and Fabrício Benevenuto. 2019. Supervised Learning for Fake News Detection. *IEEE Intelligent Systems*, 34(2):76–81.

Somya Ranjan Sahoo and Brij B Gupta. 2021. Multiple Features Based Approach for Automatic Fake News Detection on Social Networks using Deep Learning. *Applied Soft Computing*, 100:106983.

Moein Shahiki-Tash, Jesús Armenta-Segura, Zahra Ahani, Olga Kolesnikova, Grigori Sidorov, and Alexander Gelbukh. 2023. Lidoma at HomoMex2023@ Iberlef: Hate Speech Detection Towards the Mexican Spanish-Speaking LGBT+ Population. the importance of preprocessing before using bert-based models. In *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2023)*.

Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake News Detection on Social Media: A Data Mining Perspective. *ACM SIGKDD explorations newsletter*, 19(1):22–36.

Naman Singh, Tushar Sharma, Abha Thakral, and Tanupriya Choudhury. 2018. Detection of Fake Profile in Online Social Networks Using Machine Learning. In *2018 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, pages 231–234. IEEE.

Malliga Subramanian, Bharathi Raja Chakravarthi, Kogilavani Shanmugavadivel, Santhiya Pandiyan, Prasanna Kumar Kumaresan, Balasubramanian Palani, Premjith B, Sandhiya Raja, Vanaja, Mithunajha S, Devika K, Hariprasath S.B, Haripriya B, and Vigneshwar E. 2024. Overview of the Second Shared Task on Fake News Detection in Dravidian Languages. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, Malta. European Chapter of the Association for Computational Linguistics.

Malliga Subramanian, Bharathi Raja Chakravarthi, Kogilavani Shanmugavadivel, Santhiya Pandiyan, Prasanna Kumar Kumaresan, Balasubramanian Palani, Muskaan Singh, Sandhiya Raja, Vanaja, and Mithunajha S. 2023. Overview of the Shared Task on Fake News Detection from Social Media Text. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

M Shahiki Tash, Z Ahani, Al Tonja, M Gemeda, N Hussain, and O Kolesnikova. 2022. Word Level Language Identification in Code-mixed Kannada-English Texts using traditional machine learning algorithms. In *Proceedings of the 19th International Conference on Natural Language Processing (ICON): Shared Task on Word Level Language Identification in Code-mixed Kannada-English Texts*, pages 25–28.

Moein Tash, Jesus Armenta-Segura, Zahra Ahani, Olga Kolesnikova, Grigori Sidorov, and Alexander Gelbukh. 2023. LIDOMA@ DravidianLangTech: Convolutional Neural Networks for Studying Correlation Between Lexical Features and Sentiment Polarity in Tamil and Tulu Languages. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 180–185.

Aswini Thota, Priyanka Tilak, Simrat Ahluwalia, and Nibrat Lohia. 2018. Fake News Detection: a Deep Learning Approach. *SMU Data Science Review*, 1(3):10.

Atnafu Lambebo Tonja, Mesay Gemeda Yigezu, Olga Kolesnikova, Moein Shahiki Tash, Grigori Sidorov, and Alexander Gelbuk. 2022. Transformer-Based Model for Word Level Language Identification in Code-mixed Kannada-English Texts. *arXiv preprint arXiv:2211.14459*.

Mesay Gemeda Yigezu, Girma Yohannis Bade, Olga Kolesnikova, Grigori Sidorov, and Alexander Gelbukh. 2023a. Multilingual Hope Speech Detection using Machine Learning.

Mesay Gemeda Yigezu, Selam Kanta, Olga Kolesnikova, Grigori Sidorov, and Alexander Gelbukh. 2023b. Habesha@ DravidianLangTech: Abusive Comment Detection using Deep Learning Approach. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 244–249.

Mesay Gemeda Yigezu, Tadesse Kebede, Olga Kolesnikova, Grigori Sidorov, and Alexander Gelbukh. 2023c. Habesha@ DravidianLangTech: Utilizing Deep and Transfer Learning Approaches for Sentiment Analysis. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 239–243.

Mesay Gemeda Yigezu, Olga Kolesnikova, Grigori Sidorov, and Alexander Gelbukh. 2023d. Transformer-Based Hate Speech Detection for Multi-Class and Multi-Label Classification.

Mesay Gemeda Yigezu, Moges Ahmed Mehamed, Olga Kolesnikova, Tadesse Kebede Guge, Alexander Gelbukh, and Grigori Sidorov. 2023e. Evaluating the Effectiveness of Hybrid Features in Fake News Detection on Social Media. In *2023 International Conference on Information and Communication Technology for Development for Africa (ICT4DA)*, pages 171–175. IEEE.

Mesay Gemeda Yigezu, Atnafu Lambebo Tonja, Olga Kolesnikova, Moein Shahiki Tash, Grigori Sidorov, and Alexander Gelbukh. 2022. Word Level Language Identification in Code-mixed Kannada-English Texts using Deep Learning Approach. In *Proceedings of the 19th International Conference on Natural Language Processing (ICON): Shared Task on Word Level Language Identification in Code-mixed Kannada-English Texts*, pages 29–33.