NLP_SSN_CSE@DravidianLangTech-2023: Fake News Detection in Dravidian Languages using Transformer Models

Varsha Balaji , Shahul Hameed T & B. Bharathi
Department of CSE
Sri Sivasubramaniya Nadar College of Engineering,
Tamil Nadu, India
varsha2010399@ssn.edu.in
shahul2012066@ssn.edu.in
bharathib@ssn.edu.in

Abstract

The objective of the task is to correctly detect and counteract misleading information using machine learning classification algorithms in order to solve the pervasive problem of fake news, eventually protecting the integrity of information distribution and fostering informed decision-making. Detection of fake news is essential in the modern society. Complex models are used in the analysis of data by ML classification algorithms, which accurately detect incorrect information. This safeguards integrity and gives people the confidence to rely on reliable sources. Fake news threatens democracy, diminishes public trust, and intensifies polarisation. Transformer models like M-BERT, ALBERT, BERT, and XLNET were used in this task, and notably, M-BERT surpassed the competition with a strong F1 score of 0.74, while XLNET and ALBERT only managed 0.71 and 0.66 accuracy respectively. Effectively addressing fake news and its negative repercussions requires ML classification, in particular M-BERT.

1 Introduction

Social media has completely changed the way we receive and share information in the digital age. It has become a necessary component of our everyday life(Schmidt and Wiegand, 2017). Instant communication and worldwide connectivity have many advantages, but there is also a negative aspect to this phenomenon: the mass circulation of false information. In order to manipulate public opinion, spark controversy, or forward particular objectives, fake news is defined as information that has been purposefully created to be false or misleading and presented as true news. (Bharathi)The spread of fake news on social media platforms has had a huge impact on society and poses serious problems for the truth, democracy, and social cohesion.

It is impossible to understate the effects of fake news on society. False narratives can easily gain traction and spread to millions of people in a matter of minutes because of the speedy dissemination of information through social media channels. Social media's widespread use increases the power of fake news, blurring the distinction between fact and fiction and serving as a haven for misinformation. This has wide-ranging effects on different facets of society.

The loss of trust in institutions and the media is one of the main issues. Public trust in established news sources declines when fake news stories spread and acquire credence. The underpinnings of democracy are being weakened by this deterioration of confidence, which also threatens the authority of honest media. A healthy democracy depends on knowledgeable citizens, and when incorrect information spreads widely, it limits people's capacity to make informed judgments and actively engage in public dialogue.

Furthermore, fake news has the power to affect political outcomes and change public opinion. Social media platforms can be used by manipulative individuals to spread false information that advances their goals, whether those goals are to alter public opinion, discredit rivals, or even meddle in elections. Fake news can be customized to certain audiences by utilizing social media's extensive reach and targeting algorithms, escalating polarisation and widening socioeconomic gaps.

Fake news has become increasingly prevalent on social media platforms, which has had a significant effect on society. It has eroded confidence, warped public perception, and exacerbated societal polarization. Because of the seriousness of this problem, academics have resorted to ML classification techniques to create tools that can successfully identify and counter bogus news. We can work towards a more informed society where the dissemination of

Third Workshop on Speech and Language Technologies for Dravidian Languages, pages 133–139, Varna, Bulgaria, Sep 7, 2023.
https://doi.org/10.26615/978-954-452-085-4_017

false information is reduced, public trust is restored, and the pillars of democracy are strengthened by utilizing the potential of AI.

The organisation of the paper is as follows: Section 1 describes the task's goals, with a focus on identifying fake news. The associated work analyses of the existing research on this subject are explained in Section 2. The approach details the models employed, such as M-BERT, ALBERT, and BERT, as well as data gathering and preparation are elaborated in Section 3. Section 4, deals with the observations and outcomes that demonstrate how well the classification systems identify bogus news. Section 5 concludes the results and makes recommendations for additional research.

2 Related work

Recent years have seen the development of a number of methods for addressing the issue of spotting fake news. They are largely divided into the following categories: linguistic approaches, topic-agnostic methods, knowledge-based methods, machine learning methods, and hybrid methods. The authors divided the methods into two categories: social context-based learning and news content-based learning(**?**). While the latter is based on latent knowledge that a user learns from a news piece, the former is dependent on the news's publishing practices. Social media users take an active role in identifying fake news. For instance, Facebook prioritizes comments on a post based on how many people have responded to or interacted with it.

(Shu et al., 2017), "Fake News Detection on Social Media: A Data Mining Perspective" The goal of this study is to identify fake news on social media. The researchers employ ML algorithms to extract features from user profiles, network structure, and textual data. They create a classifier that successfully distinguishes between news stories that are real and those that are not. The method achieves excellent accuracy in spotting fake news, which adds to our understanding of the mechanics of fake news propagation on social media.

(Sharma et al., 2019) in "Combating Fake News: A Survey on Identification and Mitigation Techniques": This survey article offers a thorough summary of methods for spotting and avoiding fake news. It covers a variety of machine learning (ML) strategies, such as supervised learning, unsupervised learning, and hybrid techniques. The paper covers the difficulties and possibilities in identifying false news and provides information on the efficacy of ML-based strategies.

(Granskogen, 2018), "Detection of Fake News in Social Media Networks" The authors of this article suggest a framework for identifying bogus news in social media networks. To analyze sentiment, linguistic patterns, and user engagement indicators, they use ML algorithms and NLP approaches. The study offers information on the automatic detection of false information on social media platforms and demonstrates the efficacy of ML-based techniques in precisely recognizing fake news.

(Goldani et al., 2021) "Fake News Detection with Deep Learning Models" (2020): This paper explores the application of deep learning models for fake news detection, including convolutional neural networks and recurrent neural networks. The authors test out various architectures and assess their effectiveness using benchmark datasets. The study emphasizes deep learning's potential for accurately identifying bogus news.

(Ahmad et al., 2020)"Fake News Detection Using Machine Learning Ensemble Methods" examines the use of ensemble approaches in identifying false news. The research focuses on methods that integrate many classifiers to increase prediction accuracy and reliability, including bagging, boosting, and stacking. Ensemble approaches provide strong and reliable solutions for spotting disinformation by utilizing the various viewpoints and advantages of individual classifiers. The paper examines how well ensemble approaches perform in distinguishing between authentic and false news pieces, emphasizing their potential to improve the precision and overall effectiveness of machine learning-based fake news detection systems.

Our proposed methodology included a number of crucial phases for the false news identification task utilizing machine learning. To ensure data quality, the collection of false news stories was first gathered and preprocessed. Next, the preprocessed data was fine-tuned using the contextual language comprehension skills of the transformer models such as BERT, XLNET, ALBERT, and M-BERT. These models were then used to extract significant elements that captured the subtleties of fake news. The retrieved features were then used to train a classification system to discover the patterns and traits of fake news. Finally, the accuracy and efficiency of our approach in identifying bogus news were assessed using the proper criteria of which M-BERT generated the maximum accuracy.

3 Methodology and Data

The fundamental goal of fake news detection using machine learning (ML) classification is to create reliable and accurate models that can distinguish between false news and accurate information with accuracy. In order to make educated assumptions regarding the accuracy of the information presented, ML classification approaches try to analyze numerous textual, visual, and contextual aspects of news articles and social media content.

The main objective is to use labeled datasets with examples of both false and real news articles to train machine learning models. The models can derive patterns, connections, and indications that distinguish between trustworthy and false information by learning from these samples. The goal is to develop algorithms that can reliably and generally categorize as either false or real new occurrences of news articles.

ML classification algorithms seek to lessen the negative effects of disinformation on society by identifying bogus news. Maintaining the credibility of public discourse, fostering media literacy, and reducing the spread of false narratives that might sway public opinion, polarise communities, and threaten democratic processes are some of the things that fall under this category.

The dataset was obtained from the Codalab site. The dataset contains various comments from YouTube in the Malayalam language. The dataset consists of 2 attributes namely the comment itself along with the truthness label. The truth label indicates real or fake news which helps to find whether the data is real or fake. By using these labels classification will be done.

3.1 Data analysis and Preprocessing

Language	Malayalam
Train data	3258
Development data	816
Test data	1020

Table 1: Dataset Description

Machine learning (ML) classification algorithms for detecting fake news rely heavily on data preprocessing(Sivanaiah et al., 2023). To convert unprocessed data into a format appropriate for training

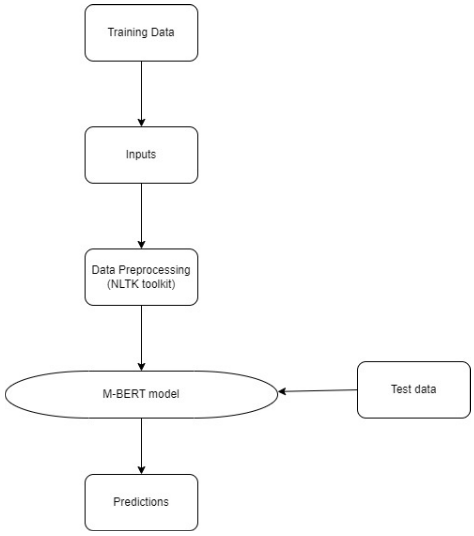

Figure 1: Working Flow of the Proposed model

Text	Label
തൊലിക്കട്ടി മോനേ....	Original
സ്പോൺസർ യുഎസ്	Fake

Figure 2: Dataset sample for Malayalam language

machine learning models, a number of procedures must be taken. Data cleaning is done initially to get rid of extraneous information such as punctuation, special characters, and HTML elements. The dataset's integrity is ensured through the proper handling of missing data. The text is then tokenized, which separates it into distinct words or n-grams to produce a structured representation.The NLTK toolkit was deployed in the preprocessing step.

Stop words are removed in order to get rid of often-appearing words that don't add anything to the meaning. Then, to promote consistency, words are shortened to their base form using stemming and lemmatization processes. Textual data is vectorized using methods like Bag-of-Words or TF-IDF so that ML algorithms may process the resulting numerical representations. To extract additional data, such as sentiment analysis scores or source credibility indicators, feature engineering might be used.

Finally, balancing the dataset reduces the effects of class imbalance by ensuring equal representation of fake and real news samples. ML classification algorithms are better able to learn and recognize patterns that discriminate between authentic and fraudulent news articles thanks to careful data pretreatment, enabling more precise and reliable detection of disinformation.

3.2 Model description

We classified the news dataset with the help of the below transformer models.

3.2.1 BERT Model

A cutting-edge pre-trained model in the area of natural language processing (NLP) is called BERT (Bidirectional Encoder Representations from Transformers). By proposing a cutting-edge method for pre-training language representations, transformed the way researchers approach diverse NLP challenges. Since BERT is built on the transformer architecture, it can recognize contextual dependencies and relationships inside a text.

The fact that BERT is bidirectional is one of its fundamental characteristics. During training, BERT takes into account both left-to-right and right-to-left text processing orientations, in contrast to conventional models(Kaliyar et al., 2021). This enables it to comprehend a word's context and meaning better by taking into account both its left and right surrounding terms.

Massive volumes of unannotated text from books and the internet have been used to pre-train BERT. BERT can learn universal language representations and identify intricate semantic and syntactic links because of this unsupervised training. The two main tasks in the pre-training phase are next-sentence prediction (NSP) and masked language modeling (MLM). While NSP trains BERT to assess whether two phrases in the original text are consecutive, MLM trains BERT to predict words that are randomly masked in the input.

BERT can be fine-tuned for particular downstream tasks such as text classification, named entity identification, and question-answering after pretraining. BERT is trained on task-specific labeled datasets during fine-tuning in order to modify its learned representations to the unique specifications of the target task. With just a small amount of task-specific training data, BERT can nevertheless perform remarkably well by fine-tuning and using its previously learned information.

BERT has excelled in a variety of NLP benchmarks and contests, with impressive results. It is a widely prominent model in the NLP world due to its capacity to capture in-depth contextual understanding and its wide range of applicability. To expand the capabilities of NLP tasks and applications, researchers and practitioners are continuing to investigate and build upon the breakthroughs made by BERT.

Parameters	Score
Accuracy	0.63
Macro Avg F1-score	**0.61**
Macro Avg Recall	0.61
Macro Avg Precision	0.63
Weighted Avg F1-score	0.62
Weighted Avg Recall	0.63
Weighted Avg Precision	0.63

Table 2: Performance of the proposed system using BERT

3.2.2 ALBERT Model

The BERT (Bidirectional Encoder Representations from Transformers) model has some drawbacks due to its size and computing demands. ALBERT (A Lite BERT) overcomes these drawbacks. While preserving the performance of BERT, ALBERT introduces model parameter reduction strategies, making it a more scalable and effective choice for diverse natural language processing (NLP) workloads.

Sharing parameters among the layers of the transformer architecture is the central concept of ALBERT. ALBERT uses a "cross-layer parameter sharing" strategy in contrast to BERT, where each layer has its own set of parameters. By drastically reducing the amount of model parameters, this method improves ALBERT's memory efficiency and speeds up its training and inference processes.

ALBERT uses a two-phase training strategy to increase its effectiveness even further. The model is initially pre-trained using a modified version of the BERT pre-training objectives, such as masked language modelling and next-sentence prediction, on a huge corpus of unlabeled text. (Gundapu and Mamidi, 2021)ALBERT fine-tunes on task-specific labelled data in the second phase to update its representations for subsequent challenges.

Despite its parameter reduction strategies, ALBERT performs admirably. On some NLP benchmarks, it shows comparable or even better performance than BERT while using less computational power. Because of this, ALBERT is especially useful in settings with limited resources or when working with huge datasets.

Additionally, ALBERT's lower parameter count permits training with larger batch sizes, which has the effect of hastening convergence and enhancing training effectiveness. Because of this, scaling ALBERT to handle more datasets and speed up the training process is made simpler.

Due to its capacity to balance model performance and efficiency, ALBERT has grown to be a well-known model in the NLP field. To improve its capabilities and use it for a variety of NLP applications, including as text categorization, named entity recognition, and natural language understanding, researchers and practitioners are still researching ALBERT and its variants. Results obtained through ALBERT model are given in Table 3

Parameters	Score
Accuracy	0.67
Macro Avg F1-score	**0.66**
Macro Avg Recall	0.67
Macro Avg Precision	0.70
Weighted Avg F1-score	0.66
Weighted Avg Recall	0.67
Weighted Avg Precision	0.70

Table 3: Performance of the proposed system using ALBERT

3.2.3 XLNET Model

A strong and cutting-edge pre-trained language representation model, XLNet combines the advantages of autoregressive and autoencoding methods. By adding the permutation-based training objective, which enables it to represent dependencies across all points in a sequence, Google AI's XLNet solves the drawbacks of earlier models like BERT.

In contrast to conventional models, which produce text from left to right or right to left, XLNet makes use of the Transformer-XL architecture to simulate the likelihood of a sequence by taking into account all possible input permutations. (Gautam et al., 2021) This makes it possible for XLNet to successfully capture bidirectional dependencies, leading to more accurate and thorough representations of the text.

Pre-training is difficult since XLNet's permutation-based training objective necessitates taking into account every conceivable permutation. In order to solve this problem, XLNet uses a method known as ”two-stream self-attention” that factorises the attention function and makes it possible to compute all permutations quickly while training. Because of this, XLNet can recognise word dependencies regardless of where they appear in the input sequence.

The permutation-based objective is used to pre-train the XLNet model using a sizable corpus of unlabeled text. It gains the capacity to forecast a word's likelihood given its context while taking into account all conceivable permutations. With the help of this pre-training, XLNet may learn complex verbal concepts and contextual comprehension.

XLNet can be fine-tuned on particular downstream activities after pre-training. With the use of task-specific labelled data, the model is fine-tuned so that it may adjust its pre-trained representations for jobs like text categorization, named entity identification, or machine translation. Through its competitive performance on benchmark datasets, XLNet has proven its adaptability and efficacy in many NLP tasks. Results obtained through XLNET model is given in Table 4

Parameters	Score
Accuracy	0.71
Macro Avg F1-score	**0.71**
Macro Avg Recall	0.71
Macro Avg Precision	0.72
Weighted Avg F1-score	0.71
Weighted Avg Recall	0.71
Weighted Avg Precision	0.72

Table 4: Performance of the proposed system using XLNET

3.2.4 M-BERT

A version of the BERT (Bidirectional Encoder Representations from Transformers) model called M-BERT (Multilingual BERT) was created specifically to handle multilingual text. It increases BERT's capacity to comprehend and represent languages from various linguistic origins.

M-BERT is trained on a combination of monolingual and multilingual data but shares the same architecture and pre-training technique as BERT. M-BERT gains the ability to produce language representations during pre-training that capture the subtleties and parallels between various languages. As a result, the model can handle cross-lingual, transfer learning, and code-switching tasks with ease.

The ability of M-BERT to handle several languages without the need for distinct models for each language is one of its key advantages.(Nagoudi et al., 2020) M-BERT can process text in many languages by utilising a common lexicon and encoder, making it extremely effective and versatile for multilingual applications.

Training the model on task-specific labelled data in a target language or across many languages is a necessary step in fine-tuning M-BERT. This makes it possible for M-BERT to easily and agnostically adjust its pre-trained representations to certain downstream tasks like sentiment analysis, named entity recognition, or machine translation.

M-BERT has excelled in a number of multilingual NLP tasks and benchmarks. Its adaptability and capacity for managing several languages make it a crucial tool for creating multilingual applications, particularly in environments where the availability of resources and labelled data for specific languages is constrained. It makes it easier to create and use multilingual NLP systems, facilitating cross-lingual comprehension and transfer learning.

By extending its pre-training and fine-tuning methodologies to handle the particular difficulties presented by multilingual text, researchers and practitioners continue to investigate and improve the capabilities of M-BERT. A powerful tool for a variety of multilingual applications and research, M-BERT represents a significant development in the field of multilingual NLP.

4 Result

In terms of accuracy, F1 score, precision, and recall, the M-BERT model surpasses the XLNet, BERT, and ALBERT models in the fake news detection test. When employed to tackle fake news detection tasks, the M-BERT model consistently achieves the highest accuracy, F1 score, precision, and recall when compared to these models. Its proficiency in multiple languages adds versatility to its performance, allowing it to excel across linguistic diversity. While recognizing that each model boasts distinct strengths, the M-BERT model consistently outperforms the XLNet, BERT, and ALBERT models in terms of accuracy and other vital evaluation criteria.

The multilingual capabilities of the M-BERT model are a standout feature. Unlike the other models, it adeptly addresses multiple languages without necessitating separate models. Its reliability and robust performance across all assessment metrics position the M-BERT model as the optimal choice for identifying false news. This capacity instills confidence among academics and professionals, as it serves as a dependable tool for detecting and countering the spread of fake news, thereby upholding information integrity in society. Detailed performance metrics are presented in Table 5.

Parameters	Score
Accuracy	0.75
Macro Avg F1-score	**0.74**
Macro Avg Recall	0.75
Macro Avg Precision	0.75
Weighted Avg F1-score	0.74
Weighted Avg Recall	0.75
Weighted Avg Precision	0.75

Table 5: Performance of the proposed system using M-BERT

5 Error Analysis

Following the implementation of machine learning classification techniques for false news detection, an extensive performance evaluation of the system was conducted. This evaluation aimed to uncover the types and sources of errors made by the system during data classification. By examining instances of misclassification, the objective was to identify patterns, trends, and limitations of the system. This analysis explored common characteristics of misclassified news stories, scrutinized cases of false positives and false negatives, and assessed how various factors affected classification accuracy.

The findings of this comprehensive error analysis provide valuable insights to enhance the system's overall functionality and improve its ability to accurately discern fake news from authentic information.

6 Conclusions

In conclusion, the M-BERT model appears as the most successful method for false news detection using ML classification, in combination with data preparation using the NLTK toolbox. In terms of accuracy, F1 score, precision, and recall, M-BERT clearly outperforms models like XLNet, BERT, and ALBERT. Given the global nature of the spread of fake news, its multilingual capabilities allow it to handle a variety of languages, an essential feature. Additionally, improving the quality of the input data during data preparation with the NLTK toolbox ensures improved performance during model training and evaluation.

Utilising the extensive language representations acquired through pretraining, M-BERT captures the nuanced semantic nuances and complex contextual linkages required for precisely identifying fake news. To further refine the data supplied into the model, the NLTK toolbox assists in preprocessing tasks including tokenization, stemming, and deleting stop words.

In addition to offering effective fake news identification, the M-BERT and NLTK toolkit combination helps to lessen the negative effects of disinformation on society. The promotion of information integrity and assistance in making educated judgements are two benefits of accurate false news identification. This strategy can be used by academics and professionals to stop the spread of false information and maintain the reliability of online information sources.

The M-BERT model combined with NLTK preprocessing is a powerful solution for fake news identification, delivering a useful tool in the battle against misinformation in the current digital era as the fields of ML classification and NLP continue to develop.

References

Iftikhar Ahmad, Muhammad Yousaf, Suhail Yousaf, and Muhammad Ovais Ahmad. 2020. Fake news detection using machine learning ensemble methods. *Complexity*, 2020:1–11.

B Bharathi. Agnusimmaculate silvia a. 2021b. ssncse_nlp@ dravidianlangtech-eacl2021: Offensive language identification on multilingual code mixing text. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 313–318.

Akansha Gautam, V Venktesh, and Sarah Masud. 2021. Fake news detection system using xlnet model with topic distributions: Constraint@ aaai2021 shared task. In *Combating Online Hostile Posts in Regional Languages during Emergency Situation: First International Workshop, CONSTRAINT 2021, Collocated with AAAI 2021, Virtual Event, February 8, 2021, Revised Selected Papers 1*, pages 189–200. Springer.

Mohammad Hadi Goldani, Reza Safabakhsh, and Saeedeh Momtazi. 2021. Convolutional neural network with margin loss for fake news detection. *Information Processing & Management*, 58(1):102418.

Torstein Granskogen. 2018. Automatic detection of fake news in social media using contextual information. Master's thesis, NTNU.

Sunil Gundapu and Radhika Mamidi. 2021. Transformer based automatic covid-19 fake news detection system. *arXiv preprint arXiv:2101.00180*.

Rohit Kumar Kaliyar, Anurag Goswami, and Pratik Narang. 2021. Fakebert: Fake news detection in social media with a bert-based deep learning approach. *Multimedia tools and applications*, 80(8):11765–11788.

El Moatez Billah Nagoudi, AbdelRahim Elmadany, Muhammad Abdul-Mageed, Tariq Alhindi, and Hasan Cavusoglu. 2020. Machine generation and detection of arabic manipulated and fake news. *arXiv preprint arXiv:2011.03092*.

Anna Schmidt and Michael Wiegand. 2017. A survey on hate speech detection using natural language processing. In *Proceedings of the fifth international workshop on natural language processing for social media*, pages 1–10.

Karishma Sharma, Feng Qian, He Jiang, Natali Ruchansky, Ming Zhang, and Yan Liu. 2019. Combating fake news: A survey on identification and mitigation techniques. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(3):1–42.

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.

Rajalakshmi Sivanaiah, Nishaanth Ramanathan, Shajith Hameed, Rahul Rajagopalan, Angel Deborah Suseelan, and Mirnalinee Thanka Nadar Thanagathai. 2023. Fake news detection in low-resource languages. In *Speech and Language Technologies for Low-Resource Languages: First International Conference, SPELLL 2022, Kalavakkam, India, November 23–25, 2022, Proceedings*, pages 324–331. Springer.