ML&AI_IIIT Ranchi@DravidianLangTech: Leveraging Transfer Learning for the Discernment of Fake News within the Linguistic Domain of Dravidian Language

Kirti Kumari
IIIT Ranchi / Ranchi India
kirti@iiitranchi.ac.in

Shirish Shekhar Jha
IISER Bhopal / Bhopal India
shirish20@iiserb.ac.in

Zarikunte Kunal Dayanand
IISER Bhopal / Bhopal India
zarikunte20@iiserb.ac.in

Praneesh Sharma
KIIT / Bhubaneswar India
praneeshsharma30@gmail.com

Abstract

The primary focus of this research endeavor lies in detecting and mitigating misinformation within the intricate framework of the Dravidian language. A notable feat was achieved by employing fine-tuning methodologies on the highly acclaimed Indic-BERT model, securing a commendable 4^{th} rank in a prestigious competition organized by DravidianLangTech 2023 while attaining a noteworthy macro F1-Score of 0.78. To facilitate this undertaking, a diverse and comprehensive dataset was meticulously gathered from prominent social media platforms, including but not limited to Facebook and Twitter. The overarching objective of this collaborative initiative was to proficiently discern and categorize news articles into either the realm of veracity or deceit through the astute application of advanced machine learning techniques, coupled with the astute exploitation of the distinctive linguistic idiosyncrasies inherent to the Dravidian language.

1 Introduction

Text classification has emerged as a critical task that has garnered significant attention and contributions from researchers and practitioners worldwide in recent years. The rapid increase in user enrollment on social media platforms has led to an unprecedented surge in the availability of textual data. This abundance of textual data has played a pivotal role in various domains, including business and research, shaping strategies, and driving advancements in software and applications.

Companies rely on textual data, particularly user reviews and comments, to formulate their business strategies and make informed decisions about their products and services. By analyzing the sentiment and content of textual data, businesses gain valuable insights into customer opinions, preferences, and trends. Furthermore, researchers have recognized the potential of textual data analysis in advancing fields such as natural language processing, sentiment analysis, and information retrieval. The exploration of textual data has paved the way for the development of more intelligent software and applications that can understand and process human language more effectively.

The analysis of textual data assumes even greater significance when considering its multilingual nature. With a diverse range of languages spoken worldwide, understanding and analyzing textual data across different languages is crucial. While most contributions in the field of text classification have focused on the English language domain, it is essential to explore the characteristics and challenges posed by other languages. Low-resource languages, such as the Dravidian languages, often face a scarcity of available methodologies and resources for effective text analysis. This poses a unique challenge that requires innovative approaches and solutions.

The impact of textual data goes beyond business and research, particularly in the era of social media and digital communication. A study conducted by MIT [1] revealed a disconcerting trend: fake news spreads faster than accurate information, particularly on platforms like Twitter, where it can propagate at a rate 10 to 20 times faster. This poses a significant challenge to individuals, organizations, and society as a whole, as the proliferation of misinformation can lead to detrimental consequences. Trust in news sources and the ability to discern reliable information become paramount in such an environment.

Addressing this issue becomes even more challenging when dealing with low-resource languages

[1]https://news.mit.edu/2018/study-twitter-false-news-travels-faster-true-stories-0308

Third Workshop on Speech and Language Technologies for Dravidian Languages, pages 198–206, Varna, Bulgaria, Sep 7, 2023.
https://doi.org/10.26615/978-954-452-085-4_028

like the Dravidian languages. The scarcity of methodologies and resources specifically tailored for these languages poses a barrier to effective fake news detection and mitigation. However, recognizing the urgency and importance of finding solutions, competitions and collaborative initiatives have emerged, bringing together great minds from around the world to tackle this pressing problem. These platforms foster innovation, encourage the development of novel methodologies, and promote advancements in technology that can benefit the human race.

With the aim of contributing to this endeavor, we embarked on an experiment to identify fake news in the Malayalam language, a Dravidian language spoken primarily in the Indian state of Kerala. Leveraging various machine learning and deep learning techniques, we sought to devise a suitable solution that could effectively detect and classify fake news in this language. Our experimentation involved training and fine-tuning models using state-of-the-art approaches, with a particular focus on transfer learning. By fine-tuning Indic-BERT, a language model specifically designed for Indian languages, we aimed to leverage its pre-trained knowledge to improve the accuracy and effectiveness of our classification system.

The results of our experiment were promising, with an achieved macro F1-Score of 0.78. This performance placed us in an admirable 4^{th} position in the competition (Subramanian et al., 2023), reflecting the efficacy of our approach and the efforts invested in tackling the challenges specific to the Malayalam language. By harnessing the power of transfer learning and combining it with domain-specific fine-tuning, we enhanced our system's capabilities and achieved commendable results in detecting and classifying fake news.

Rest of the paper summarised in different sections. Related works are presented in Section 2 followed by Task description and Dataset description in Section 3 and Section 4 respectively. In Section 5, we presented our detailed methodology and validation results followed by Section 6, where we presented the results of test data and analysis of the the results. Finally, we summarise our work in Section 7.

2 Related Work

Due to their potential impact on public discourse and the spread of disinformation, the identification and mitigation of fake news on social media have become key fields of research. This part includes an extensive assessment of the pertinent literature and covers significant contributions to the study of how to spot false news.

In their seminal work, the work (Shu et al., 2017) adopt a data mining perspective to address the challenge of detecting fake news on social media. Their research underscores the significance of employing data mining techniques to extract features and patterns from social media content, enabling accurate identification of fake news. The authors propose a framework incorporating textual, visual, and social features to discern instances of fake news. By highlighting the role of data mining in tackling the issue of fake news on social media platforms, this study contributes valuable insights to the field. The article (Wang, 2017) introduces the benchmark dataset "Liar, liar pants on fire," specifically designed for fake news detection. This dataset consists of statements that have been labeled with varying degrees of truthfulness, enabling researchers to develop and evaluate models for fake news detection. The author emphasizes the significance of high-quality datasets in advancing research in this domain and provides a valuable resource for benchmarking and comparing different approaches. The researches **(Mridha et al., 2021)(Li et al., 2021)(Kim and Jeong, 2019)(Aldwairi and Alwahedi, 2018)** explore deep learning approaches for fake news detection. They extensively review a range of deep learning models and techniques, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and attention mechanisms, applied in the context of fake news detection. The survey provides valuable insights into the strengths and limitations of different deep learning approaches and highlights their potential for detecting fake news. Conducting a comparative study of computational fake news detection on social media, the research (Manish et al., 2022) evaluate the performance of various machine learning and deep learning algorithms. The authors assess different features, such as linguistic, stylistic, and temporal characteristics, and examine their effectiveness in distinguishing between fake and real news. This comparative analysis offers valuable insights into the strengths and limitations of different computational approaches for fake news detection.

In conclusion, there are many different viewpoints and methods covered in the literature on detecting fake news on social media. The articles discussed in this part highlight how crucial it is to use data mining methods, benchmark datasets, deep learning models, and comparative analyses to progress the field of false news detection. These studies' findings, benchmark datasets, and suggested frameworks are all useful tools for creating computational techniques that can stop fake news from spreading on social media platforms. We attempted to draw inspiration for our work from these references, and after comparing our dataset with theirs, we attempted to develop some useful approaches.

3 Task Description

DravidianLangTech presents an intriguing competition titled "Fake News Detection in Dravidian Languages." The primary objective of this task is to develop a proficient classification system that can effectively distinguish between original and fake news in social media text written in Dravidian languages. The dataset for this task comprises social media posts extracted from diverse platforms like Twitter, Facebook, and others.

Participants have been provided with social media texts in Dravidian languages. They are tasked with building machine learning or deep learning models to accurately classify the given content as either fake or original news. The focus of the competition revolves around the ability to discern authentic information from fabricated or misleading content within the context of Dravidian languages.The outcomes of this task hold significant implications for enhancing the credibility and trustworthiness of information shared on social media platforms in Dravidian linguistic domains. Furthermore, developing effective classification models will contribute to empowering users with reliable information, foster informed decision-making, and combat fake news dissemination.

4 Dataset Description

The dataset (Kayalvizhi et al., 2022) utilized for the Fake News Detection competition in Dravidian Languages comprises a diverse collection of social media posts sourced from prominent platforms, including Twitter and Facebook. It is segmented into three distinct subsets: a training dataset, a development dataset, and a test dataset. The training dataset furnished to us encompasses a substantial volume of social media posts composed in Dravidian languages in code-mixed format. Each post within the training dataset is annotated with labels indicating its categorization as either an original or fake news post. The total texts in the Training dataset were 3257 in number. These labeled annotations served as the ground truth for training classification models, playing a crucial role in facilitating the development and refinement of machine learning or deep learning models.

The development dataset, also referred to as the validation dataset, augments the labeled data provided. It allowed us to evaluate models' performance and fine-tune the hyperparameters during the development phase. The labeled information within the development dataset aids in the assessment of the models' efficacy in accurately classifying social media text as either original or fake news.The total text in development dataset were 815. Table 1 summarizes the data distribution in training, development, and test dataset.

Table 1: Distribution of the dataset

Class	**Training**	**Development**	**Test**	**Total**
Original	1658	409	512	2579
Fake	1599	406	507	2512
Total	3257	815	1019	5091

5 Methodology and Validation Results

In our study, we employed a range of methodologies on the development dataset with the objective of identifying the most effective approach for making predictions on the test dataset. Different techniques (Chanda et al., 2022) presents a technique where *mBert* has been utilized along with word-level language tag to classify the comments. The article (Varsha et al., 2022) presented an approach where feature extraction has been done in association with transfer learning to perform the classification task. Being inspired by the related work, we have tried to introduce methodologies that could prove effective in this task.

For example, when working with TF-IDF (Term Frequency-Inverse Document Frequency) features and embeddings-based classification, extensive data pre-processing was undertaken to optimize the performance of the models. However, in the case of fine-tuning Indic-BERT, this data pre-processing step was not necessary. Indic-BERT, have been pretrained on a large corpus of social media texts, specifically in Indian languages, incorporates the linguistic nuances and contextual information prevalent in social media language, including the presence of emojis[2]. Consequently, removing emojis in this case could potentially compromise the linguistic semantics captured by IndicBERT.

Data Pre-Processing was conducted specifically for methods such as TF-IDF feature extraction and sentence embedding-based classification. This crucial step encompassed several operations aimed at enhancing the quality and structure of the text data.

The initial phase involved the removal of punctuation marks, Emojis (only for methodologies of section 5.1 to 5.3) and alphanumeric characters to eliminate noise and ensure a cleaner representation of the textual content. Following this, we proceeded with the removal of stop words, which are commonly occurring words that do not contribute significantly to the overall meaning of the text. Additionally, we focused on expanding any contracted words to their full forms, allowing for a more comprehensive analysis of the text.

The final step of the data pre-processing pipeline involved tokenization, which entailed breaking down the text into individual units such as words or sub-words. This process facilitates subsequent analysis and modeling tasks by providing a structured representation of the text data. Furthermore, we applied lemmatization to the tokens, aiming to reduce inflected or variant forms to their base or dictionary form, promoting consistency and coherence within the dataset.

By diligently performing these data pre-processing techniques, we aimed to enhance the quality and reliability of the text data, preparing it for subsequent analysis and classification tasks.

Following the feature extraction stage, we applied several machine learning models and deep learning techniques to perform the classification task. The specific methods utilized are discussed in subsequent subsections of this study, outlining the intricacies and nuances associated with each approach.

By exploring and evaluating multiple methodologies, we aimed to identify the most effective techniques for classifying social media texts as original or fake news in the context of our research.

[2]https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02221/full

5.1 TF-IDF Based Classification

In this research study, we focused on utilizing the TF-IDF technique for classification tasks in machine learning. The primary objective was to explore the effectiveness of TF-IDF-based approaches for text classification.

TF-IDF is a widely used technique in natural language processing that assigns weights to individual terms based on their frequency within a document and their rarity across the entire corpus. By considering both the local and global importance of terms, TF-IDF enables identifying key features that are discriminative for classification.

We employed various machine learning algorithms to implement the TF-IDF-based classification, including the Ridge classifier and Logistic Regression. These algorithms are well-known for their effectiveness in handling text classification tasks. We also explored similar algorithms to assess their performance and compare their results.

The choice of max_df was 0.9, meaning we ignored terms that appeared in more than 90% of documents, and min_df was set to 5, meaning we also ignored words that appeared in less than five documents; it was made to strike a balance between capturing sufficient vocabulary diversity and avoiding computational complexity. By limiting the dictionary to the most frequent and informative terms, we aimed to ensure robust classification performance while managing the dimensionality of the feature space. Table 2 summarizes the results of different machine learning algorithms deployed on training and development datasets.

5.2 Bag-Of-Words Based Classification

Bag-of-Words (BoW) based text classification is a widely adopted approach in natural language processing for representing text documents as numerical feature vectors. In this research study, we also leveraged BoW-based text classification with machine learning algorithms to analyze and classify textual data effectively.

The initial step in BoW-based text classification involved constructing a dictionary or vocabulary of unique words or terms that were utilized to represent the documents. To ensure comprehensive coverage, we created a dictionary consisting of 10,000 terms, encompassing the most frequent and informative terms derived from the training dataset.

Once the dictionary was established, we proceeded to transform each document into a sparse

Table 2: Summary of TF-IDF based Results on Training and Validation Datasets

Classifier	Macro Precision	Macro Recall	Macro F1
Ridge-Classifier	0.76	0.75	0.75
Perceptron-Classifier	0.72	0.71	0.71
SGD-classifier	0.75	0.75	0.74
Passive-AggressiveClassifier	0.74	0.74	0.74
Decision-Tree-Classifier	0.74	0.72	0.72
Random-Forest-Classifier	0.75	0.75	0.75
AdaBoost-Classifier	0.75	0.74	0.73
SVM-Classifier	0.75	0.74	0.75

vector representation. This vector representation captured the presence or absence of dictionary terms within the document, along with their frequency or weighted values using techniques like term frequency-inverse document frequency (TF-IDF).

To train and classify the BoW representations of the documents, we employed a diverse range of machine learning algorithms, including well-known models such as logistic regression, support vector machines, random forests, and decision trees, among others. Each algorithm underwent training using a labeled training dataset comprising documents and their respective class labels.

During the training phase, the machine learning algorithms acquired an understanding of the patterns and relationships between the BoW features and their associated classes. Subsequently, we evaluated the trained models on a development dataset, utilizing metrics such as precision macro, recall macro, and F1-score macro, to assess their performance and generalization capabilities.

We carefully considered the trade-off between capturing sufficient vocabulary diversity and managing computational complexity in selecting the dictionary size. By opting for a dictionary size of 10,000, we aimed to strike an optimal balance.Table 3 summarizes the results of experiments done on training and validation datasets. It was noted that the Bag-Of-Words features were more effective in classifying the text data.

5.3 Sentence Embedding Based Classification

We have employed Sentence-BERT (SBERT) (Reimers and Gurevych, 2019, 2020), a powerful technique for calculating sentence embeddings, to extract meaningful representations from the text data. By leveraging transfer learning, we utilized the pre-trained model xlm-r-100langs-bert-base-nli-stsb-mean-tokens to generate high-quality sentence embeddings.

To perform classification, we fed the sentence embeddings into a Deep Neural Network (DNN) architecture. To enhance the performance of our model, we extracted the features from the last layer of the neural network model. These extracted features were then used as inputs for various machine-learning classifiers. The concept behind approaching such hybrid architecture was to capture the complexity that came with Dravidian languages.

In addition to the DNN-based approach, we also constructed a custom ensemble classifier by combining multiple machine learning algorithms. The ensemble classifier integrated Decision Tree, Gaussian Naive Bayes, Support Vector Machine, Logistic Regression, and Linear Discriminant Analysis (LDA) classifiers. This ensemble classifier leveraged the features extracted from the last layer of the neural network to classify the text data.

To optimize the performance of our models, we carefully tuned the hyperparameters. Details regarding the hyperparameters can be found in Table 6. We aimed to achieve the best possible classification results by fine-tuning these parameters.

The outcomes of our experiments and evaluations are summarized in Table 4. These results provide insights into the effectiveness and performance of the proposed approach.
Overall, this experiment showcased the utilization of SBERT for sentence embedding, combined with DNN-based classification and a custom ensemble classifier. By extracting features from the last layer of the neural network, we were able to leverage the power of transfer learning and machine learning algorithms.

Table 3: Bag-Of-Words bases Results Summary on Training and Validation Datasets

Classifier	Macro Precision	Macro Recall	Macro F1
Ridge-Classifier	0.78	0.76	0.76
Perceptron-Classifier	0.72	0.72	0.72
SGD-classifier	0.79	0.79	0.78
Passive-AggressiveClassifier	0.79	0.78	0.78
Decision-Tree-Classifier	0.76	0.74	0.74
Random-Forest-Classifier	0.80	0.79	0.79
AdaBoost-Classifier	0.76	0.75	0.74
SVM-Classifier	0.78	0.77	0.77

Table 4: Sentence Embedding based Results Summary on Training and Validation Datasets

Classifier	Macro Precision	Macro Recall	Macro F1
SBERT + DNN	0.69	0.69	0.69
SBERT+DNN+Random Forest	0.70	0.69	0.70
SBERT+DNN+AdaBoost	0.68	0.68	0.68
SBERT+DNN+Ensemble	0.67	0.67	0.67

Table 5: Fine Tuned Indic-BERT Results Summary on Training and Validation Datasets

Classifier	Macro Precision	Macro Recall	Macro F1
IndicBERT	0.79	0.78	0.78
IndicBERT+Random Forest	0.75	0.75	0.75
IndicBERTT+AdaBoost	0.78	0.78	0.78
IndicBERT+Extra Trees	0.75	0.75	0.75
IndicBERT+SGD Classifier	0.79	0.78	0.78
IndicBERT+SVM	0.78	0.78	0.78
IndicBERT+Logistic Regression	0.78	0.78	0.78
IndicBERT+Decision Tree	0.75	0.75	0.75

Table 6: Deep Neural Network Architecture and Hyperparameters

Hyperparameters	Values
Number of Layers	4
Activation Function(s)	Tanh and ReLU
Dropout Rate	0.2
Optimizer	Adam
Number of Epochs	50

Table 7: Indic-BERT results on the test dataset

Method	Result
Indic-BERT	**macro F1-Score** 0.78

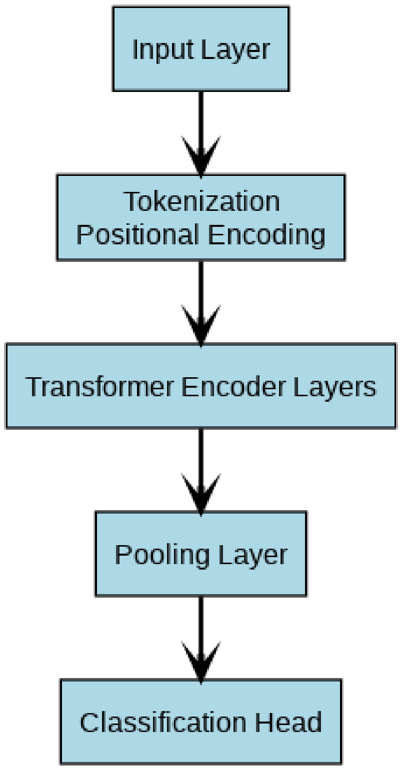
5.4 Fine Tuning Indic-BERT

In the final stage of our research experimentation, we focused on fine-tuning the pre-trained model Indic-BERT, which AI4BHARAT developed[3]. Indic-BERT represents a multilingual ALBERT model that has been trained on extensive corpora comprising 12 major Indian languages. These languages include Assamese, Bengali, English, Gujarati, Hindi, Kannada, Malayalam, Marathi, Oriya, Punjabi, Tamil, and Telugu. Remarkably, Indic-BERT outperforms other publicly available models such as *mBERT* and XLM-R, despite having a significantly lower parameter count. It has demonstrated state-of-the-art performance across a range of tasks (Dabre et al., 2021; Ganesan et al., 2021; Kunchukuttan et al., 2020; Doddapaneni et al., 2021; Kakwani et al., 2020).

To tailor the Indic-BERT model to our specific classification objectives, we engaged in the process of fine-tuning by adjusting the model weights using the development dataset. This step allowed us to leverage the dataset's unique characteristics and linguistic nuances to enhance the performance and adaptability of Indic-BERT. Through fine-tuning, we aimed to optimize the model's ability to capture the relevant features and patterns necessary for accurate classification. We replaced the last layer for classification with a single neuron to make binary classification. We used the Sigmoid function to make classifications and Adam as our optimizer. With backpropagation, we were able to adjust the weights of the model according to the task. The Indic-BERT architecture can be described in the following flowchart 1.

[3]https://ai4bharat.iitm.ac.in/indic-bert

Figure 1: Utilized Indic-BERT Architecture



To handle the varying lengths of text sequences within the datasets, we defined maximum sequence lengths for each subset of data. For the training dataset, the maximum sequence length was set to 97, while for the development and test datasets, the respective max sequence lengths were adjusted to 88 and 89. The specification of sequence lengths plays a vital role in optimizing the tokenization process and generating meaningful tokens. Text sequences shorter than the defined length were padded, while longer sequences were truncated, ensuring consistency and compatibility throughout the analysis.

Subsequently, we extracted the features from the last layer of the fine-tuned Indic-BERT model. These features, representing the rich contextual information captured by the model, were utilized as inputs for the machine learning classifiers. By leveraging the combination of the fine-tuned model's representations and the discriminative power of the classifiers, we aimed to harness the benefits of both approaches and construct a hybrid model. This hybrid model would offer a comprehensive and robust solution for text classification tasks.

To evaluate the effectiveness of our approach, we conducted extensive experiments and obtained results for various evaluation metrics. The performance of the hybrid model, including accuracy, precision, recall, and F1 score, was carefully analyzed and documented. These metrics provide valuable insights into the model's capabilities and enable a comprehensive assessment of its performance across different classification tasks.

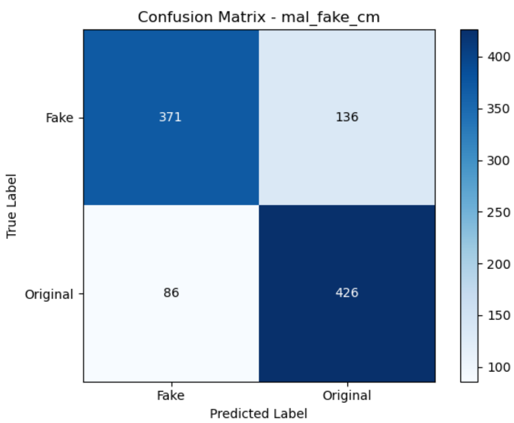
We present the detailed results of our experiments in Table 5, which serves as a comprehensive summary of the findings. The table provides a holistic view of the model's performance and highlights its effectiveness in accurately classifying text based on the features extracted from the fine-tuned Indic-BERT model. Observing the results, we used the fine-tuned model to make final predictions for our task submission.

6 Results and Analysis

In this section, we present the results of our submitted task. After making predictions using the fine-tuned Indic-BERT, we were evaluated using the macro F1-Score, macro Precision, and macro Recall. We obtained a macro F1 Score of 0.78 on the test dataset.

The confusion matrix has been presented in Figure 2 that tells the classification of various classes along with misclassified classes. It is a critical tool for analyzing the performance and efficiency of our model.

Figure 2: Confusion Matrix of Test Predictions



7 Conclusion

To summarize, this research study delved into multiple text classification techniques, encompassing Bag-of-Words based classification, TF-IDF-based classification, Sentence Embedding based classification and the fine-tuning of the pre-trained Indic-BERT model. Each approach exhibited distinct advantages and showcased its effectiveness in accurately categorizing textual data. Our model and experiments performed satisfactorily over task; all the results are summarized in the Tables. We can enhance the performance by fine-tuning the pre-trained model with more data.

Acknowledgements

We are thankful to Indian Institute of Information Technology Ranchi for all the support during our research.

References

Monther Aldwairi and Ali Alwahedi. 2018. Detecting fake news in social media networks. *Procedia Computer Science*, 141:215–222.

Galen Andrew and Jianfeng Gao. 2007. Scalable training of L1-regularized log-linear models. In *Proceedings of the 24th International Conference on Machine Learning*, pages 33–40.

Supriya Chanda, Anshika Mishra, and Sukomal Pal. 2022. Sentiment analysis and homophobia detection of code-mixed dravidian languages leveraging pre-trained model and word-level language tag. In *Working Notes of FIRE 2022-Forum for Information Retrieval Evaluation (Hybrid). CEUR*.

Raj Dabre, Himani Shrotriya, Anoop Kunchukuttan, Ratish Puduppully, Mitesh M Khapra, and Pratyush Kumar. 2021. Indicbart: A pre-trained model for indic natural language generation. *arXiv preprint arXiv:2109.02903*.

Sumanth Doddapaneni, Gowtham Ramesh, Mitesh M Khapra, Anoop Kunchukuttan, and Pratyush Kumar. 2021. A primer on pretrained multilingual language models. *arXiv preprint arXiv:2107.00676*.

Vinod Ganesan, Gowtham Ramesh, and Pratyush Kumar. 2021. Supershaper: Task-agnostic super pre-training of bert models with variable hidden dimensions. *arXiv preprint arXiv:2110.04711*.

Divyanshu Kakwani, Anoop Kunchukuttan, Satish Golla, Gokul N.C., Avik Bhattacharyya, Mitesh M. Khapra, and Pratyush Kumar. 2020. IndicNLPSuite: Monolingual Corpora, Evaluation Benchmarks and Pre-trained Multilingual Language Models for Indian Languages. In *Findings of EMNLP*.

S Kayalvizhi, Thenmozhi Durairaj, Bharathi Raja Chakravarthi, et al. 2022. Findings of the shared task on detecting signs of depression from social media. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 331–338.

Kyeong-Hwan Kim and Chang-Sung Jeong. 2019. Fake news detection system using article abstraction. In *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pages 209–212.

Anoop Kunchukuttan, Divyanshu Kakwani, Satish Golla, Avik Bhattacharyya, Mitesh M Khapra, Pratyush Kumar, et al. 2020. Ai4bharat-indicnlp corpus: Monolingual corpora and word embeddings for indic languages. *arXiv preprint arXiv:2005.00085*.

Xiangyang Li, Yu Xia, Xiang Long, Zheng Li, and Sujian Li. 2021. Exploring text-transformers in aaai 2021 shared task: Covid-19 fake news detection in english. 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 106–115. Springer.

Manish Kumar Singh Manish, Jawed Ahmed Jawed, Mohammad Afshar Alam Alam, Kamlesh Kumar Raghuvanshi Kamlesh, and Sachin Kumar Sachin. 2022. A comparative study of computational fake news detection on social media.

Muhammad F Mridha, Ashfia Jannat Keya, Md Abdul Hamid, Muhammad Mostafa Monowar, and Md Saifur Rahman. 2021. A comprehensive review on fake news detection with deep learning. *IEEE Access*, 9:156151–156170.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Nils Reimers and Iryna Gurevych. 2020. Making monolingual sentence embeddings multilingual using knowledge distillation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

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.

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.

Josephine Varsha, B Bharathi, and A Meenakshi. 2022. Sentiment analysis and homophobia detection of youtube comments in code-mixed dravidian languages using machine learning and transformer models. In *Working Notes of FIRE 2022-Forum for Information Retrieval Evaluation (Hybrid). CEUR*.

William Yang Wang. 2017. " liar, liar pants on fire": A new benchmark dataset for fake news detection. *arXiv preprint arXiv:1705.00648*.