SetFit: A Robust Approach for Offensive Content Detection in Tamil-English Code-Mixed Conversations Using Sentence Transfer Fine-tuning

Kathiravan Pannerselvam[1], **Saranya Rajiakodi**[1], **Sajeetha Thavareesan**[2], **Sathiyaraj Thangasamy**[3], **Kishore Kumar Ponnusamy**[4]

[1]Department of Computer Science, Central University of Tamil Nadu, India.
[2]Eastern University, Sri Lanka.
[3]Sri Krishna Adithya College of Arts and Science, India.
[4]Digital University of Kerala, India.

Abstract

Code-mixed languages are increasingly prevalent on social media and online platforms, presenting significant challenges in offensive content detection for natural language processing (NLP) systems. Our study explores how effectively the Sentence Transfer Fine-tuning (SetFit) method, combined with logistic regression, detects offensive content in a Tamil-English code-mixed dataset. We compare our model's performance with five other NLP models: Multilingual BERT (mBERT), LSTM, BERT, IndicBERT, and Language-agnostic BERT Sentence Embeddings (LaBSE). Our model, SetFit, outperforms these models in accuracy, achieving an impressive 89.72%, significantly higher than other models. These results suggest the sentence transformer model's substantial potential for detecting offensive content in code-mixed languages. Our study provides valuable insights into the sentence transformer model's ability to identify various types of offensive material in Tamil-English online conversations, paving the way for more advanced NLP systems tailored to code-mixed languages.

Keywords

Hate speech, SetFit, Natural language processing, Offensive detection, Code-mixed languages, Tamil-English dataset

1 Introduction

Text classification is a part of Information Extraction (IE), an emerging area of research and application that explores how to discover knowledge (information) from a vast amount of text (Ek et al., 2011). So many data resources are available on the internet, like social media, e-commerce sites, blogs, news portals, personal websites, and others, where

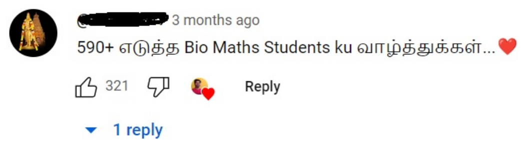

Figure 1: Example Tamil-English code-mixed youtube comment

people share their thoughts and opinions in their native language (Kathiravan and Haridoss, 2018). Code-mixed language is a phenomenon in which speakers interchange two or more languages or language varieties within a single conversation or text. It is common in multilingual societies, particularly in urban areas where people with diverse linguistic backgrounds interact regularly (Ravikiran and Annamalai, 2021) (Singh et al., 2018). Code-mixing is prevalent in many parts of the world, including India, Africa, and the Americas. Fig. 1 is an example YouTube comment with Tamil-English code-mixed languages. In code-mixed language, speakers often borrow words or phrases from one language and incorporate them into another, resulting in a hybrid language combining multiple languages' grammatical structures and vocabulary (Pannerselvam et al., 2023). The prevalence of code-mixed language in social media and online platforms has posed significant challenges for NLP systems. Traditional NLP models are designed to work with monolingual texts and struggle with the complexity of code-mixed language. Consequently, there is an increasing demand for developing NLP systems that can effectively process code-mixed language, especially in identifying offensive language (Li, 2021).

Detecting offensive content in code-mixed lan-

Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pages 35–42

guage is challenging due to the linguistic complexity of code-mixed language and the nuances of language use (Chakravarthi, 2023),(Kumaresan et al., 2023). Offensive language in code-mixed language can involve gender-based language, stereotypes, derogatory terms, and hate speech (Kumaresan et al., 2022). Developing NLP systems that effectively detect offensive language in a code-mixed language is crucial for promoting safe and inclusive online spaces (Ahluwalia et al., 2018),(Alowibdi et al., 2014). This study focuses on detecting offensives in a code-mixed Tamil-English dataset using Sentence Transformer Fine-Tune (SetFit). It is a substitute for few-shot text classification and involves fine-tuning a Sentence Transformer using task-specific data. This approach can be effortlessly implemented using the sentence-transformers library.

The main objective of this research is to evaluate the performance of the SetFit model and compare it with other NLP models in offensive content detection within code-mixed Tamil-English language data on social media and online platforms. The study aims to provide insights into the effectiveness of SetFit in identifying offensive content and its potential applicability in enhancing the robustness of NLP systems for code-mixed languages.

The following research questions are conceived from the above-mentioned objective of this research work.

RQ1 How does the performance of SetFit compare to LSTM, BERT, multilingual BERT (mBERT), IndicBERT, and LaBSE in the task of offensive content detection within a code-mixed Tamil-English dataset?

RQ2 To what extent does the performance of SetFit vary when detecting different types of offensive content (e.g., Offensive targetediInsult group, Offensive untargeted, Not offensive, Not Tamil, Offensive targeted insult individual, Offensive targeted insult others) within code-mixed Tamil-English online discourses?

RQ3 How does the performance of the NLP models (SetFit, mBERT, LSTM, BERT, IndicBERT, LaBSE) differ in terms of precision, recall, and F1-score when identifying offensive content in the code-mixed Tamil-English dataset?
The findings from our research can be beneficial for developing NLP models that are more effective in detecting various offensives in code-mixed languages, which can help mitigate the harmful effects of such language on online platforms. The findings can benefit further research and practitioners developing NLP applications for social media and online platforms.

Additionally, developing NLP models capable of detecting various offensives in code-mixed languages can help mitigate the harmful effects of such language in online platforms, promote safer and more inclusive online spaces, and contribute to building a more equitable and inclusive society.

2 Related works

In this related work section, we review the existing literature on this topic, focusing on studies explicitly addressing the problem of various offensive language detection in code-mixed text. We also discuss the existing approaches to code-mixed text classification, including language-specific and language-agnostic models, and highlight the strengths and weaknesses of each approach.

Offensive language identification is a task that has been receiving considerable attention in recent years. However, identifying offensive spans in long sentences has been a challenging problem, especially for context-dependent ones. To address this issue, a study (Ravikiran and Annamalai, 2021) evaluated several models, including Bi-LSTM CRF, MuRIL, and LIME. The rationale extraction-based approach involving a combination of MuRIL and LIME performed significantly better than the other models. However, the Bi-LSTM CRF model was sensitive toward shorter sentences and performed worse than the random baseline. Extracting offensive spans for long sentences was also found to be complicated. In the future, the researchers plan to re-do the offensive span identification task, where participants will be required to identify offensive spans while simultaneously classifying different types of offensiveness. The researchers also plan to release this study's baseline models and datasets to encourage further research.

Kalaivani et al. (2021) (Kalaivani et al., 2021) propose a model called TOLD (Tamil Offensive Language Detection) that detects offensive language in code-mixed social media comments written in Tamil and English. The study uses Multilingual BERT (mBERT) with feature-based selection to generate contextualized word embeddings and selects relevant features using the chi-square test. The proposed TOLD model outperforms the accuracy, precision, recall, and F1 score of existing models. The study also analyzes the importance of various features and finds that character-based features and part-of-speech tags are highly effective in detecting offensive language. The study demonstrates the effectiveness of using mBERT with feature-based selection in identifying and moderating offensive language in social media platforms, promoting a safer and more inclusive online environment.

Ravikiran and Annamalai (2021) (Ravikiran and Annamalai, 2021) present a new dataset called DOSA (Dravidian Code-mixed Offensive Span Identification Dataset) for identifying the offensive language in code-mixed social media comments in four Dravidian languages: Tamil, Telugu, Kannada, and Malayalam. The authors collected and annotated the dataset with offensive spans, the smallest text segment containing an offensive word or phrase. The dataset contains 7,500 comments and 15,547 offensive spans, and multiple annotators did the annotations to ensure high inter-annotator agreement. The paper also provides baseline experiments on the dataset using various models and shows the effectiveness of the dataset in identifying offensive language in code-mixed comments in Dravidian languages.

Chakravarthi et al. (2021) (Chakravarthi et al., 2021) present a study on sentiment analysis of code-mixed text in four Dravidian languages: Tamil, Telugu, Kannada, and Malayalam. The authors collected a dataset of code-mixed social media comments and used various preprocessing techniques to prepare the data for sentiment analysis. They then used various machine and deep learning models to perform sentiment analysis and evaluated the performance using various metrics. The results showed that deep learning models such as LSTM and GRU outperformed traditional machine learning models in sentiment analysis of code-mixed text in Dravidian languages. The study provides insights into the challenges and opportunities in sentiment analysis of code-mixed text in Dravidian languages. It can help develop practical tools for sentiment analysis in these languages.

Rajalakshmi et al. (2021) (Rajalakshmi et al., 2021) present an approach for offensive language identification in code-mixed Tamil using a transformer-based model. The authors describe their participation in the DLRG shared task at the Dravidian Language Technologies Workshop 2021, which aimed to develop models for identifying the offensive language in code-mixed social media comments. The authors used a transformer-based model called RoBERTa and fine-tuned it on the provided dataset. They also used data augmentation techniques to improve the performance of the model. The results showed that their approach achieved high accuracy in identifying the offensive language in code-mixed Tamil, outperforming the baseline models provided by the shared task. The study demonstrates the effectiveness of transformer-based models for offensive language identification in code-mixed Tamil and provides insights into developing models for similar tasks in other Dravidian languages.

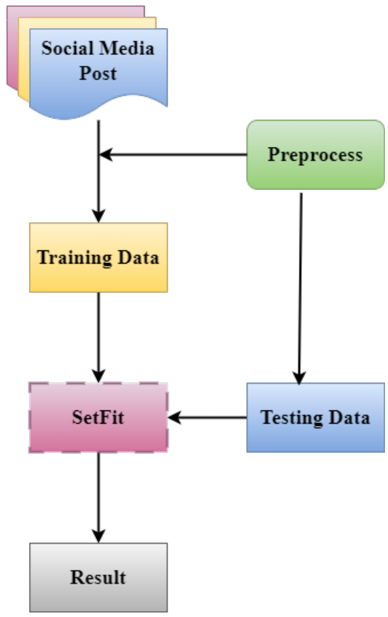

Figure 2: General architecture of proposed model

3 Proposed Methodology:

In this section, we described the experimental setup for our experimental work on classifying different types of offensive content in code-mixed text data in Tamil and English (Chakravarthi et al., 2021). Figure 2 illustrates the workflow of the general architecture of the offensive language detection framework.

3.1 Dataset

The benchmark dataset (Chakravarthi et al., 2022) contains a manually annotated dataset for sentiment analysis and offensive language identification in social media comments of three under-resourced Dravidian languages - Tamil, Kannada, and Malayalam-consisting of over 60,000 YouTube comments. The dataset contained various types of code-mixing and was annotated by volunteer annotators with high inter-annotator agreement. The authors also conducted baseline experiments using machine learning methods to establish benchmarks on the dataset.

We randomly selected 35139 code-mixed comments from the dataset mentioned above. Table 1 and Figure 3 illustrate the description of the dataset. Below are sample offensive YouTube comments along with their corresponding labels.

- **Offensive Targeted Insult Group**
  Dey dey deyyy,, loosu pasangala,, munna pinna jayalalitha amma va pathrukingal-ada gambeeramna ennanu avanga kannula tha paakanum,, amma oda history ah ithu,, soniya gandhi mari iruku chaiiii,,, avanga bio edukanumna a-z pathutu vanthu edungadaa,, dummy akkaathinga

- **Offensive Untargeted**
  Intha maari comments ku like kekuravangala india va vittu veliya annupanum.

- **Not offensive**
  Ellam okay than....but antha ponnu aasa pattu love panni avanaiye kola panna thappu il-laya...yosinga............

- **Not Tamil**
  Abe sale ye toh tatti hai

- **Offensive Targeted Insult Individual**
  Pa Ranjith paru Da unaku sc St quote job government job quarters ellam cancel pannanum

- **Offensive Targeted Insult Other**
  Iron man fans dis like podunga intha vidio kku

3.1.1 Preprocess and Data preparation

We preprocessed the data by removing stopwords, punctuations, and URLs and converted the text to lowercase. We then split the dataset into training and test sets in a 70:30 ratio, and we used various models such as LSTM, mBERT, BERT, IndicBERT, SetFit, and LaBSE.Finally, the accuracy of these models was on the test dataset.

Table 1: Dataset description.

Label	**# of Posts**
Not offensive	25,425
Not Tamil	1,454
Offensive Targeted Insult group	2,557
Offensive Targeted Insult Individual	2,343
Offensive Targeted Insult Other	454
Offensive Untargeted	2,906
Total	**35,139**

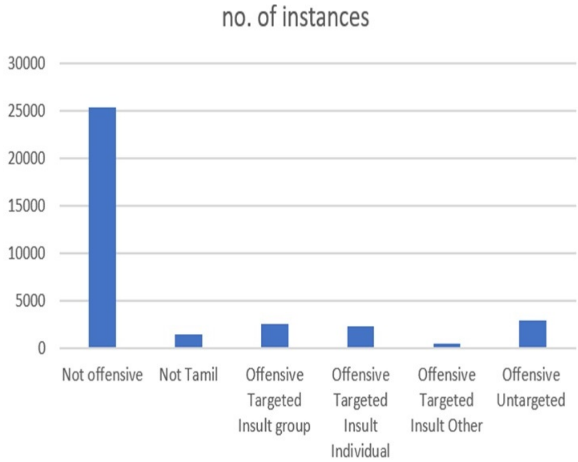

Figure 3: Dataset distribution among various class

3.2 Model

3.2.1 Sentence Transformer Fine-tuning (SetFit)

The Sentence Transformer (ST) method, widely utilized in semantic search, similarity, and clustering of words, encodes sentences into unique vector representations based on semantic content. This encoding process involves adapting a transformer model within a Siamese architecture, as detailed by (Reimers and Gurevych, 2019), (Wasserblat, 2021). The training process actively minimizes the distance between vectors of semantically similar sentences and maximizes it for those that are semantically distant through contrastive training. The performance of Sentence Transformers (ST) excels beyond other embedding representations, yet it does not match the classification capabilities of cross-encoders such as BERT (Reimers and Gurevych, 2019).

The initial stage of the training process involves selecting a sentence-transformers (ST) model from the model hub. Subsequent steps include configuring the training class, populating the training data loader, and fine-tuning. The training dataset consists of positive and negative sentence pairs to

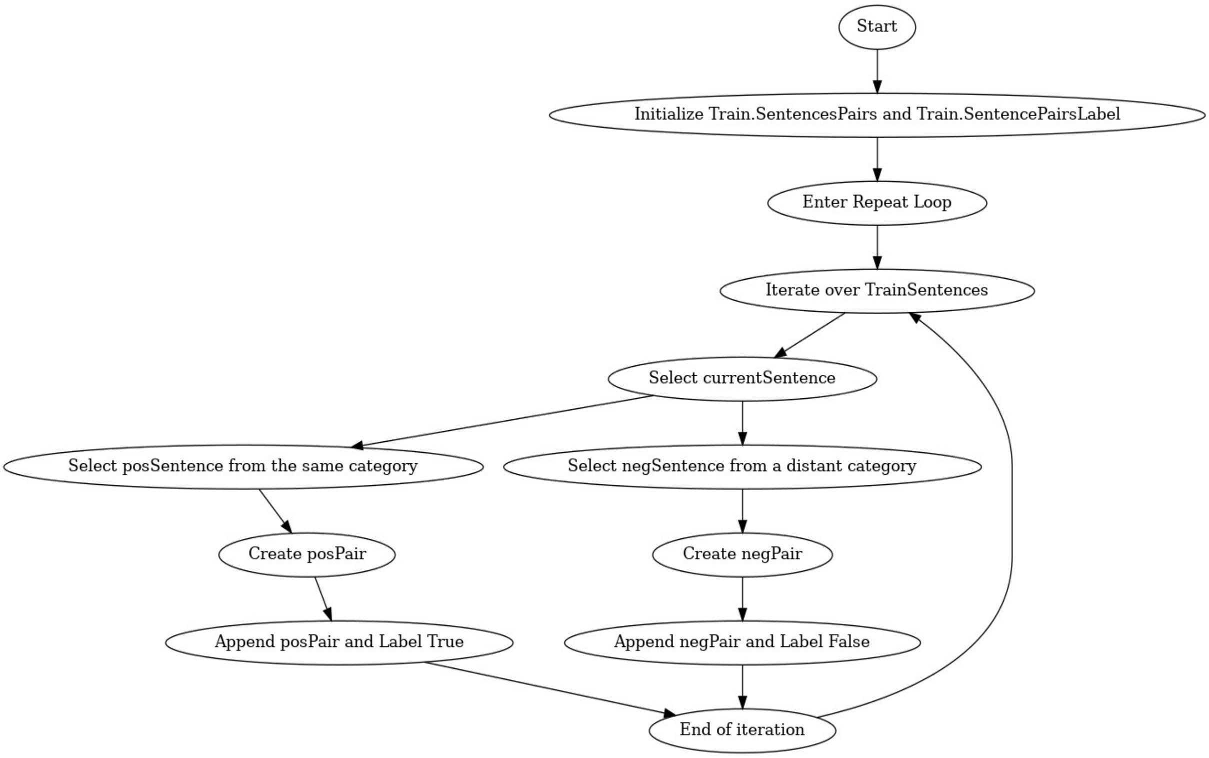

Figure 4: sentence pair selection of SetFit model

effectively address the limited labeled training data in the few-shot scenario. Positive pairs involve two sentences randomly selected from the same class, while antagonistic pairs comprise two sentences randomly chosen from different classes. Each iteration involving sentence pairs generates 2xN training pairs, where N represents the total number of training samples per task—illustrated in Figure 4.

These generated sentence pairs are employed for fine-tuning the ST model. Upon completion of the fine-tuning step, an adapted ST model is generated. The training data sentences are then encoded using the adapted ST, and the encoded data is used to train a Logistic Regression (LR) model for simplicity. Each test sentence undergoes encoding with the adapted ST in the inference phase, and the LR model predicts its category.

The proposed model is compared with the following state-of-the-art algorithms.

Long Short-Term Memory (LSTM) is an improved version of Recurrent Neural Network (RNN) architecture widely used in natural language processing and time-series prediction tasks. It is designed to address the vanishing gradient problem in standard RNNs by introducing memory cells that can selectively forget or retain information from the previous time steps. The LSTM network consists of three gates: the input gate, output gate, and forget gate, which controls the flow of information in and out of the memory cells (Zhang et al., 2018),(Kathiravan and Saranya, 2021). In this experimental work, we used the configuration for the LSTM model to include 100 hidden units, a dropout rate of 0.2, and a recurrent dropout rate of 0.2. The output layer consisted of 6 units, and the activation function used was softmax.

Bidirectional Encoder Representations from Transformers (BERT) is a pre-trained language model developed by Google, based on the Transformer architecture, and is trained on large amounts of text data to learn contextualized word embeddings. BERT is unique because it can learn bidirectional context by looking at the entire sentence, unlike other language models that only look at the previous or next word in the sequence (Alsharef et al., 2022),(Fan et al., 2021),(Li et al., 2022). The pre-trained model can be fine-tuned on a specific task with more minor task-specific data, making it useful for practical applications. We utilized CSE CUDA and trained the model for three epochs. CSE CUDA is a parallel computing platform that enables Graphics Processing Units (GPUs) to accelerate the training of deep learning models, significantly speeding up the training process. The choice of 3 epochs was based on experimentation with the dataset and the model architecture.

Multilingual BERT is a variant of the BERT model pre-trained in multiple languages. Google developed it and trained on massive text data in 104 languages. The goal of mBERT is to learn a shared representation for all languages, enabling the model to perform well on various NLP tasks for different languages without requiring language-specific training data. The model is fine-tuned on a specific task using task-specific data in the target language, and it has been shown to achieve state-of-the-art performance on various cross-lingual benchmarks. mBERT is a useful tool for multilingual applications, as it eliminates the need for separate training models for each language, saving time and resources (Kalaivani et al., 2021).

IndicBERT is a language model using the BERT architecture designed for Indian languages (Rajalakshmi et al., 2021). It is trained on large amounts of various Indian language text data. Researchers at IIT Bombay developed IndicBERT, which has shown promising results in various Indian language benchmarks. It is an open-source project. So, anybody can incorporate state-of-the-art language models into their projects (Kohli et al., 2021).

Language-agnostic BERT Sentence Embedding (LaBSE) is a multilingual language model based on the BERT architecture, trained on a large amount of text data from over 100 languages (Feng et al., 2020). It generates high-quality sentence embeddings for text in any language, making it useful for cross-lingual natural language processing tasks such as classification and machine translation. Developed by Google, LaBSE has shown state-of-the-art performance on benchmark datasets and is an open-source project for developers and researchers working with multilingual text data.

In the next section, we presented the results and analysis of our experiments on sentiment analysis in code-mixed social media comments in Tamil and English.

4 Results and Discussion

In this experimental work, we proposed Sentence Transfer Fine-tuning (SetFit) to detect offensive content in code-mixed languages. It achieved an impressive 89% accuracy on the test set, surpassing other models like LSTM, BERT, mBERT, IndicBERT, and LaBSE, which recorded accuracies

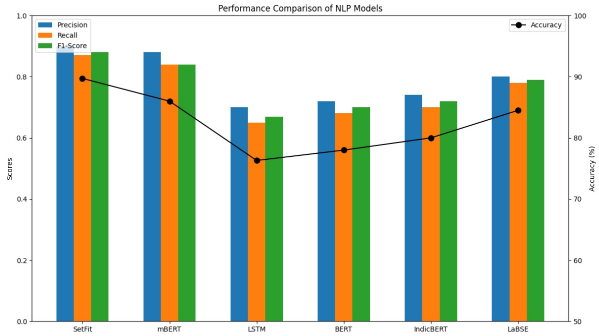

Figure 5: Accuracy of various classifiers

of 76.3%, 78%, 86%, 80%, and 84.5% respectively. Figure 5 and Table 2 depict the detailed evaluation metrics used in our experiment. SetFit's remarkable effectiveness across these metrics, as discussed in Research Questions 1 and 3 (RQ1 and RQ3), showcases its capability to enhance offensive content detection in code-mixed languages.

While a thorough analysis of SetFit's varied performance in identifying different kinds of offensive content is pending, its overall efficiency solidifies its position as a promising approach for improving NLP systems in code-mixed language contexts, aligning with the objectives outlined in Research Question 2 (RQ2).

Further analysis conducted to examine the impact of dataset size on model efficacy revealed that SetFit maintained consistent performance when the dataset size was increased from 10k to 35k samples. This finding suggests that SetFit is adept at handling larger datasets effectively.

These results endorse SetFit as a highly effective method for classifying various forms of offensive code-mixed text. The high accuracy achieved by SetFit likely stems from its ability to capture the complex linguistic patterns within the dataset effectively. Moreover, our findings stress the importance of employing large datasets for training models in this specialized area.

5 Conclusion

In conclusion, our experimental work, which introduced the Sentence Transfer Fine-tuning (SetFit) approach, marks a significant advancement in Natural Language Processing (NLP), particularly in code-mixed languages. The primary objective of our research was to enhance the detection of offensive content in such languages, a challenge that has grown increasingly relevant in today's digital communication landscape. Additionally, our analysis

Table 2: Evaluation metrics of various classifiers.

Model	Precision	Recall	F1-score	Accuracy
SetFit	0.90	0.87	0.88	0.89
mBERT	0.88	0.84	0.84	0.86
LSTM	0.70	0.65	0.67	0.76
BERT	0.72	0.68	0.70	0.78
indicBERT	0.74	0.70	0.72	0.80
LaBSE	0.80	0.78	0.79	0.84

of the impact of dataset size on model performance revealed a key strength of SetFit: its ability to maintain stable performance with the increase in dataset size from 10k to 35k samples. This robustness in handling larger datasets is crucial, especially in code-mixed language processing, where data variability is high. SetFit's high accuracy is attributed to its advanced capability to capture the complex linguistic patterns inherent in the dataset. This highlights the importance of using comprehensive and large datasets in training models for this domain.

Our research indicates that SetFit is a highly effective and reliable method for classifying different types of offensive content in code-mixed text. Its performance provides a new benchmark in the field and opens avenues for future research and development in creating more sophisticated and nuanced NLP systems for multilingual and culturally diverse digital communications.

References

Resham Ahluwalia, Himani Soni, Edward Callow, Anderson Nascimento, and Martine De Cock. 2018. Detecting hate speech against women in english tweets. *EVALITA Evaluation of NLP and Speech Tools for Italian*, 12:194.

Jalal S Alowibdi, Ugo A Buy, S Yu Philip, and Leon Stenneth. 2014. Detecting deception in online social networks. In *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, pages 383–390. IEEE.

Ahmad Alsharef, Karan Aggarwal, Deepika Koundal, Hashem Alyami, Darine Ameyed, et al. 2022. An automated toxicity classification on social media using lstm and word embedding. *Computational Intelligence and Neuroscience*, 2022.

Bharathi Raja Chakravarthi. 2023. Detection of homophobia and transphobia in youtube comments. *International Journal of Data Science and Analytics*, pages 1–20.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Vigneshwaran Muralidaran, Navya Jose, Shardul Suryawanshi, Elizabeth Sherly, and John P. McCrae. 2022. Dravidiancodemix: sentiment analysis and offensive language identification dataset for dravidian languages in code-mixed text. *Language Resources and Evaluation*.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Sajeetha Thavareesan, Dhivya Chinnappa, Durairaj Thenmozhi, Elizabeth Sherly, John P McCrae, Adeep Hande, Rahul Ponnusamy, Shubhanker Banerjee, et al. 2021. Findings of the sentiment analysis of dravidian languages in code-mixed text. *arXiv preprint arXiv:2111.09811*.

Tobias Ek, Camilla Kirkegaard, Håkan Jonsson, and Pierre Nugues. 2011. Named entity recognition for short text messages. *Procedia-Social and Behavioral Sciences*, 27:178–187.

Hong Fan, Wu Du, Abdelghani Dahou, Ahmed A Ewees, Dalia Yousri, Mohamed Abd Elaziz, Ammar H Elsheikh, Laith Abualigah, and Mohammed AA Al-qaness. 2021. Social media toxicity classification using deep learning: real-world application uk brexit. *Electronics*, 10(11):1332.

Fangxiaoyu Feng, Yinfei Yang, Daniel Cer, Naveen Arivazhagan, and Wei Wang. 2020. Language-agnostic bert sentence embedding. *arXiv preprint arXiv:2007.01852*.

Adaikkan Kalaivani, Durairaj Thenmozhi, and Chandrabose Aravindan. 2021. Told: Tamil offensive language detection in code-mixed social media comments using mbert with features based selection.

P Kathiravan and N Haridoss. 2018. Preprocessing for mining the textual data-a review. *International Journal of Scientific Research in Computer Science Applications and Management Studies IJSRCSAMS*, 7(5).

Panner Selvam Kathiravan and Rajiakodi Saranya. 2021. Named entity recognition (ner) for social media tamil posts using deep learning with singular value decomposition. Technical report, EasyChair.

Guneet Kohli, Prabsimran Kaur, and Jatin Bedi. 2021. Arguably at comma@ icon: Detection of multilingual aggressive, gender biased, and communally charged tweets using ensemble and fine-tuned indicbert. In *Proceedings of the 18th International Conference on*

Natural Language Processing: Shared Task on Multilingual Gender Biased and Communal Language Identification, pages 46–52.

Prasanna Kumar Kumaresan, Rahul Ponnusamy, Ruba Priyadharshini, Paul Buitelaar, and Bharathi Raja Chakravarthi. 2023. Homophobia and transphobia detection for low-resourced languages in social media comments. *Natural Language Processing Journal*, 5:100041.

Prasanna Kumar Kumaresan, Rahul Ponnusamy, Elizabeth Sherly, Sangeetha Sivanesan, and Bharathi Raja Chakravarthi. 2022. Transformer based hope speech comment classification in code-mixed text. In *International Conference on Speech and Language Technologies for Low-resource Languages*, pages 120–137. Springer.

Hui Li, Lin Yu, Jie Zhang, and Ming Lyu. 2022. Fusion deep learning and machine learning for heterogeneous military entity recognition. *Wireless Communications and Mobile Computing*, 2022:1–11.

Zichao Li. 2021. Codewithzichao@ dravidianlangtech-eacl2021: Exploring multilingual transformers for offensive language identification on code mixing text. In *Proceedings of the first workshop on speech and language technologies for dravidian languages*, pages 164–168.

Kathiravan Pannerselvam, Saranya Rajiakodi, Rahul Ponnusamy, and Sajeetha Thavareesan. 2023. CSS-CUTN@DravidianLangTech:abusive comments detection in Tamil and Telugu. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 306–312, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Ratnavel Rajalakshmi, Yashwant Reddy, and Lokesh Kumar. 2021. Dlrg@ dravidianlangtech-eacl2021: Transformer based approachfor offensive language identification on code-mixed tamil. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 357–362.

Manikandan Ravikiran and Subbiah Annamalai. 2021. Dosa: Dravidian code-mixed offensive span identification dataset. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 10–17.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.

Vinay Singh, Deepanshu Vijay, Syed Sarfaraz Akhtar, and Manish Shrivastava. 2018. Named entity recognition for hindi-english code-mixed social media text. In *Proceedings of the seventh named entities workshop*, pages 27–35.

Moshe Wasserblat. 2021. Sentence transformer finetuning (setfit) outperforms gpt-3 on few-shot text classification while. Accessed: 2021-12-14.

Ling Zhang, Magie Hall, and Dhundy Bastola. 2018. Utilizing twitter data for analysis of chemotherapy. *International journal of medical informatics*, 120:92–100.