

# A Transformer-based Model for Detecting Multilingual Sarcasm in Social Media Posts

Shraddha Chauhan<sup>1</sup>, Abhinav Kumar<sup>2\*,†</sup>

<sup>1</sup>Department of Electronics and Communication Engineering, Motilal Nehru National Institute of Technology Allahabad, Prayagraj, Uttar Pradesh, 211004

<sup>2</sup>Department of Computer Science and Engineering, Motilal Nehru National Institute of Technology Allahabad, Prayagraj, Uttar Pradesh, 211004

## Abstract

The rapid progress of social media like Facebook, Instagram, LinkedIn, YouTube, and X has enabled people from different linguistic and cultural backgrounds to engage in global conversations. However, this multicultural digital landscape poses various challenges in detecting sarcasm. Sarcasm is characterized by its use of irony to convey mockery and depends on contextual and cultural nuances that can vary dramatically across languages. Sarcastic posts have the potential to invert the overall meaning of phrases, and that's why there is a need to make an accurate sarcasm detection system that detects sarcasm in multilingual languages. Sarcasm detection gained considerable attention, particularly in the English language. However, the exploration of sarcasm detection in Dravidian languages like Tamil and Malayalam remains significantly underdeveloped. These languages present unique challenges due to their morphology, agglutinative nature, and diverse syntactic structures. This paper aims to bridge the gap by exploring sarcasm detection in Dravidian languages, focusing on the challenges posed by code-mixing, dialectal variations. This paper examines the use of three different transformer-based models: (i) Distil-mBERT, (ii) mBERT, and (iii) RoBERTa to effectively capture the nuances of sarcasm in these languages. Basically, these transformer models help in detecting subtle expressions of sarcasm by surmounting the challenges posed by code-switching and incorporating cultural context to eventually enhance the performance of systems targeted at detecting sarcasm. The experimental results show the potential of transformers in achieving promising performance in multilingual sarcasm identification, forming the path for further research in this less-explored domain.

## Keywords

Sarcasm, NLP, Transformers, Multilingual, Deep Learning, Social Media

## 1. Introduction

The arrival of social media has expanded communication across linguistic and cultural boundaries, highlighting the need for advanced NLP tools capable of navigating this multilingual landscape [1, 2, 3]. Among the innumerable challenges determined by this diversity, sarcasm detection stands out due to its reliance on complicated contextual and cultural cues. It employs irony to convey a meaning opposite to the literal interpretation, making it difficult to detect in languages with rich cultural diversity [4, 5]. In the context of Dravidian languages, sarcasm detection becomes very complex as these languages are very different from English. People of southern India mostly speak these languages and have unique linguistic features, structures, and cultural references that are not always straightforwardly interpretable by standard NLP models. As a result, sarcasm expressed in these languages may be misinterpreted by conventional sentiment analysis systems, leading to inaccuracies in understanding and responding to user interactions [6].

Most of the current methods for sarcasm detection depend on statistical and rule-based models, leveraging linguistic and pragmatic features and elements like sentiment shifts, interjections, and punctuation to identify sarcasm in posts. Various machine learning models are also used in the literature to detect sarcasm in various languages effectively [7]. Sarcasm detection gained considerable attention, particularly in the English language, but very little attention is given to code-mixed multilingual

Forum for Information Retrieval Evaluation, December 12-15, India

✉ shraddha76830@gmail.com (S. Chauhan); abhik@mnnit.ac.in (A. Kumar)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Dravidian languages. Sarcasm detection in text poses different challenges compared to other media types like images, videos, and speech due to the absence of contextual information and indicators such as tones and physical gestures [8]. This difficulty is further increased when the text is multilingual, i.e., consists of more than one language. Code-mixed text can mix words, phrases, or clauses from different languages. For example, “*Edutha vecchathu nokki vaayikkunna pole undu dialoges... very bad trailer.*” This comment is a mix of Malayalam and English languages. Its English translation is “*The dialogues feel like they’re just reading what they picked up... very bad trailer.*”. This comment criticizes the quality of the dialogue in the trailer, suggesting that it sounds unnatural or forced, leading to an overall negative impression of the trailer. Humans can easily detect whether the comment’s literal meaning is sarcastic because we can understand the sentence’s context better than machine learning models.

Recent advancements in deep learning techniques [9] have drawn significant attention from researchers due to their remarkable ability to detect sarcasm in social media post. Despite having that much importance of this task, only a few number of studies have explored deep learning models for multilingual sarcasm detection in text. social media plays a vital role in daily communication, with platforms like Facebook and X heavily featuring images and videos. However, text remains the primary mode of communication, despite its inherent limitations, such as the absence of non-verbal cues. Deep learning models can learn and understand hierarchical representations of language data, enabling them to capture complex patterns and relationships that traditional models cannot capture. Therefore, this work uses deep learning techniques for sarcasm detection. Among these deep learning models, Transformers have emerged as a robust architecture, with models like Distil-mBERT, mBERT and RoBERTa setting new benchmarks in different NLP tasks [10]. The Pre-Trained Transformer-based models for detecting multilingual sarcasm in social media posts require careful preprocessing of Tamil and Malayalam text and fine-tuning on our dataset to adapt the pre-trained knowledge to the specific nuances of sarcastic and non-sarcastic expressions in these languages [11]. Therefore, this paper explores three different deep learning-based models- Distil-mBERT, mBERT and RoBERTa to identify sarcastic social media posts from Tamil-English and Malayalam-English posts.

The rest of the paper is organized as follows: Section 2 lists related work for sarcasm identification, Section 3 discusses proposed methodology. The outcome of the proposed model is listed in Section 4, error analysis on test data is listed in 5 and the paper is concluded in Section 6.

## 2. Related Work

As sarcasm detection is gaining significant importance in natural language processing, various approaches have been explored in the field, but Dravidian languages did not get sufficient attention [12, 13, 14]. Most of the work till now has been done in the English language, but the models that are good at detecting sarcasm in English did not work effectively in other languages like Tamil, Malayalam, etc., because of their distinct and unique features and differences. A new method is needed to detect sarcasm in the Dravidian language accurately. Many methods mainly rely on rule-based systems, and lexicons are limited in capturing sarcasm’s complexity effectively. A number of conventional machine learning models uses bag-of-words and syntactic features to identify sarcastic expressions. However, these methods also struggle with the contextual and complex nature of sarcasm.

The introduction of deep learning and transformer-based models marked an important advancement in sarcasm detection [15, 16, 17]. Models like mBERT have demonstrated significant improvements in effectively understanding context and semantics. Its bidirectional attention mechanism allows it better to grasp the irony and contradictions in sarcastic posts. Limited work has been reported in the literature to identify sarcastic posts from code-mixed social media posts [18]. The models trained on English datasets fail to generalize efficiently in multilingual settings due to linguistic and cultural differences. Recent studies have addressed this gap by exploring sarcasm detection in languages like Arabic, Tamil,

Malayalam, and Marathi [19]. These studies highlight the importance of language-specific models and the need for diverse training datasets to capture unique sarcasm patterns across different languages.

The research work for identifying Tamil and Malayalam code-mixed sarcastic posts is limited [20]. Existing work in sentiment analysis for these languages focuses on general sentiment classification rather than sarcasm [21]. Many efforts are made to develop Tamil and Malayalam sentiment analysis tools and resources. However, these resources lack the granularity required for effective sarcasm detection [22]. The use of pre-trained transformer models can offer promising avenues for addressing these challenges [23]. By adapting models like Distil-mBERT, mBERT and RoBERTa to Dravidian languages, we aim to handle contextual information and improve sarcasm detection.

### 3. Methodology

The overall flow diagram of the proposed model can be seen in Figure 1. The detailed description of the tasks, dataset statistics, and methodology can be seen in the subsequent subsections.

#### 3.1. Task Description

Tamil and Malayalam are under resource Dravidian languages, where a few resources are available mainly for multilingual sarcasm detection. Some of the data samples for Tamil-English and Malayalam-English provided by in the FIRE-2024 workshop with their translation in English can be seen in Table 1. The task is to classify Tamil-English and Malayalam-English social media posts into sarcastic and not-sarcastic classes.

**Table 1**

Malayalam and Tamil dataset’s overview with their English translation.

Lang	Text	English Meaning	Label
Mal	Enna oru feel olla trailer	What a trailer with such a great feel!	Non-sar
Mal	Sha rukh Khan nte fan padam pole undallo	It’s like Shah Rukh Khan’s movie Fan	Sar
Tam	90s kids nanga padam eppada release	90s kids, when is the movie releasing?	Non-sar
Tam	10.5 m 3 clock kulla varuma	Will it arrive within 10.5 meters by 3 o’clock?	Sar

#### 3.2. Dataset Description

The dataset used to detect sarcasm is imbalanced and code-mixed in Tamil and Malayalam and is collected from social media [24]. The posts contain more than one sentence, but the average sentence length of the corpora is 1. Each post is labeled with sentiment polarity (Sarcastic or Non-sarcastic). The dataset is organized with a significant amount of data for both training and validation purposes. For training, the Malayalam dataset consists of 13,188 training samples, while the Tamil dataset includes 29,570 training samples. The validation set contains 2,826 samples for Malayalam and 6,336 samples for Tamil. The Test set contains 2,826 samples for Malayalam and 6,338 samples for Tamil. The overall data statistics for the Tamil-English and Malayalam-English can be seen in Table 2.

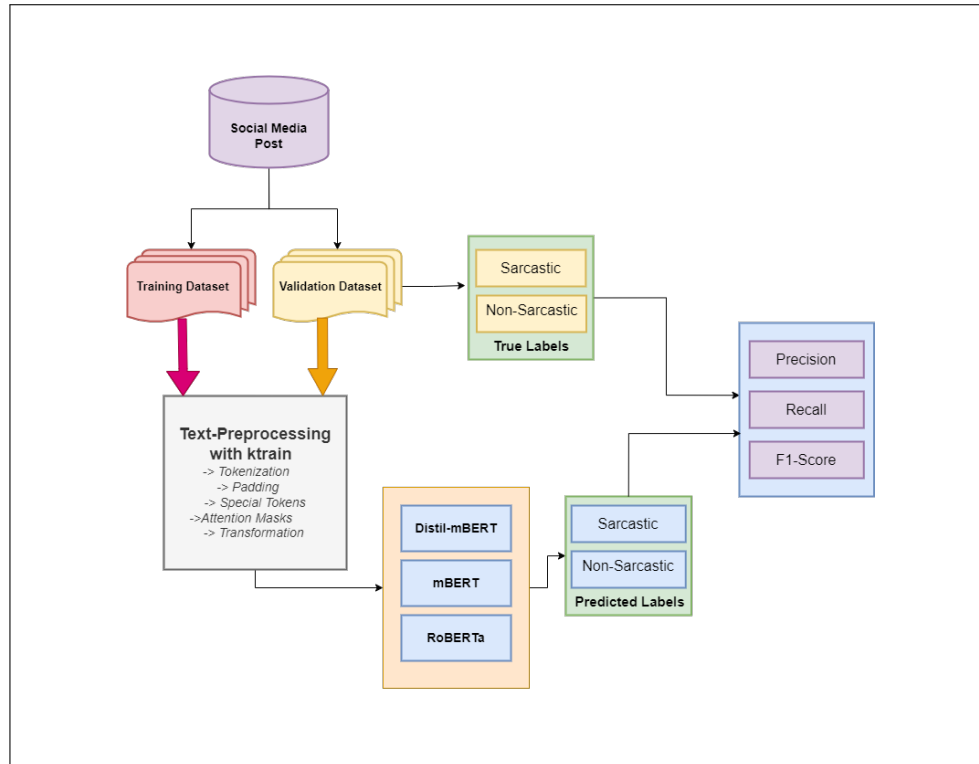
Three different pre-trained multilingual transformer-based models, (i) Distil-mBERT (ii) mBERT and (iii) RoBERTa were utilized to identify sarcastic social media posts. These models are pre-trained on vast text datasets, such as Wikipedia, and have an excellent performance when fine-tuned for a wide range of downstream tasks.

- **Distil-mBERT** is a small, fast, cheap and lighter version of BERT. It is highly effective in NLP tasks, including sarcasm detection, and is trained using the self-supervised learning method, which helps the model understand patterns, context, and word meanings, which are very critical for detecting sarcasm. It uses a task called Masked Language Modeling in which some words

**Table 2**

Overall data statistics for Tamil-English and Malayalam-English social media posts

SNo	Language	Class	Train	Validation	Test
1	Malayalam	Not-Sarcastic	10689	2305	2314
		Sarcastic	2499	521	512
		Total	13188	2826	2826
2	Tamil	Not-Sarcastic	21740	4630	4621
		Sarcastic	7830	1706	1717
		Total	29570	6336	6338

**Figure 1:** Flow chart of our proposed work.

are hidden, and the model is trained to predict those words based on the context [25]. It is bidirectional and captures the contextual information of words in a sentence, which can help to understand when a phrase is sarcastic. Fine-tuning Distil-mBERT on the dataset can help the model learn to identify sarcasm patterns, such as irony, exaggeration, and tone shifts.

- **mBERT** is pre-trained on massive amounts of text data using two key tasks, which are masked language modeling and next sentence prediction.[26]. It processes text in a deeply contextual manner, allowing it to recognize hyperbole, irony, or contrast between expectations and reality. For example, in the sentence "I love being ignored" mBERT can detect that the word 'love' is being used sarcastically because the context provided by "being ignored" negates the usual positive meaning of 'love'. The self-attention mechanism of mBERT allows the model to weigh the importance of different words in a sentence when predicting the overall meaning. This helps mBERT to focus on the keywords or phrases that might point to sarcasm. Its deep contextual understanding helps it to capture the nuances even in short texts. By fine-tuning mBERT on our datasets and leveraging its bidirectional architecture, mBERT can detect sarcasm in diverse forms, from subtle irony to exaggerated statements, even in challenging multilingual and code-mixed settings like social media posts.

- **RoBERTa** stands for robustly optimized BERT approach[27]. It uses self-attention, which helps to focus more on an essential part of the sentence that is important in the identification of sarcasm. It is pre-trained using a self-supervised approach and uses a masking technique during training to make the model more robust. Sarcasm detection needs not just only individual words but their contextual meaning in sentences also. Roberta’s bidirectional nature can enable it to consider both the preceding and following context of each word in a sentence, and it can be crucial for sarcasm detection. It performs better on social media-like text where informal language, slang, and symbols abound. It’s large-scale pretraining helps it adapt to nonstandard inputs like emojis, hashtags, or even informal punctuation. It handles noisy inputs, mixed modalities much better.

Our work is setting up a classification pipeline using Ktrain library [28] and transformer-based models like Distil-mBERT, mBERT and RoBERTa. After that, we perform data preprocessing in which we initialize the Ktrain text preprocessing transformer for the model selected. We set a maximum length of 30 tokens per text input and convert text and labels into a format that the model requires for training. We trained the pre-trained Transformer model with Adam optimizer, set it up for classification, and trained all three models for 50 epochs. The learning rate is  $5e^{-5}$ , and the batch size is 32.

## 4. Result

To determine the effectiveness of the proposed transformer-based model, several evaluation metrics such as confusion matrix, precision, recall,  $F_1$ -score, and AUC (Area under the ROC curve) and ROC (Receiver Operating Characteristic) are used.

**Table 3**

Precision, Recall, F1-scores, and Support for Tamil-English and Malayalam-English using Distil-mBERT Model

Class	Tamil-English				Malayalam-English			
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support
Non-sarcastic	0.84	0.88	0.86	4621	0.88	0.95	0.91	2314
Sarcastic	0.63	0.53	0.58	1717	0.64	0.43	0.51	512
Macro avg	0.73	0.71	0.72	6338	0.76	0.69	0.71	2826
Weighted avg	0.78	0.79	0.78	6338	0.84	0.85	0.84	2826
Accuracy	0.79 (6338)				0.85 (2826)			

**Table 4**

Precision, Recall, F1-scores, and Support for Tamil-English and Malayalam-English using mBERT Model

Class	Tamil-English				Malayalam-English			
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support
Non-sarcastic	0.83	0.90	0.86	4621	0.88	0.95	0.92	2314
Sarcastic	0.65	0.50	0.56	1717	0.67	0.41	0.51	512
Macro avg	0.74	0.70	0.71	6338	0.77	0.68	0.72	2826
Weighted avg	0.78	0.79	0.78	6338	0.84	0.86	0.84	2826
Accuracy	79.62 (6338)				0.86 (2826)			

### 4.1. Evaluation

- **Confusion Matrix:** is used to describe the performance of a classification model by comparing the actual labels with the predicted labels. The rows in the matrix represent the actual classes and the columns in the matrix represent the predicted classes. The elements in diagonal represents the correct predictions. while the off diagonal elements represent the incorrect predictions made by the model (see Table 6).

**Table 5**

Precision, Recall, F1-scores, and Support for Tamil-English and Malayalam-English using RoBERTa Model

Class	Tamil-English				Malayalam-English			
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support
Non-sarcastic	0.85	0.88	0.86	4621	0.84	0.99	0.91	2314
Sarcastic	0.64	0.57	0.60	1717	0.73	0.14	0.23	512
Macro avg	0.74	0.73	0.73	6338	0.78	0.56	0.57	2826
Weighted avg	0.78	0.79	0.79	6338	0.82	0.83	0.78	2826
Accuracy	0.80 (6338)				0.83 (2826)			

**Table 6**

Confusion Matrix with True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) representation.

	Predicted positive	Predicted negative
Actual positive	TP	FN
Actual negative	FP	TN

- **Accuracy:** is the percentage of correct predictions made by a model divided by all predictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Precision:** is the proportion of all true positive predictions out of all positive predictions made by the model.

$$\text{Precision} = \frac{TP}{TP + FP}$$

- **Recall:** is the proportion of all true positive predictions out of all actual positive instances. It is also known as sensitivity or true positive rate.

$$\text{Recall} = \frac{TP}{TP + FN}$$

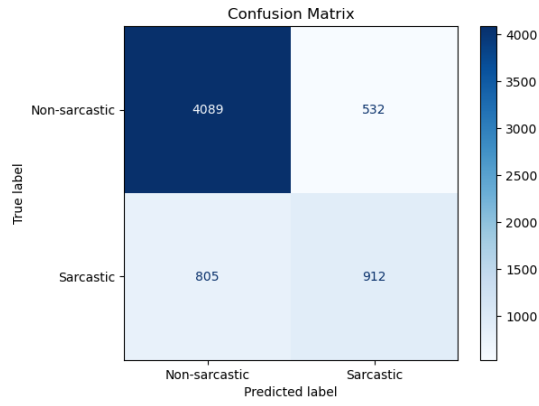
- **F1 Score:** is a score that combines both precision and recall, calculated as the harmonic average of the precision and recall.

$$\text{F1 Score} = 2 \times \frac{P \times R}{P + R}$$

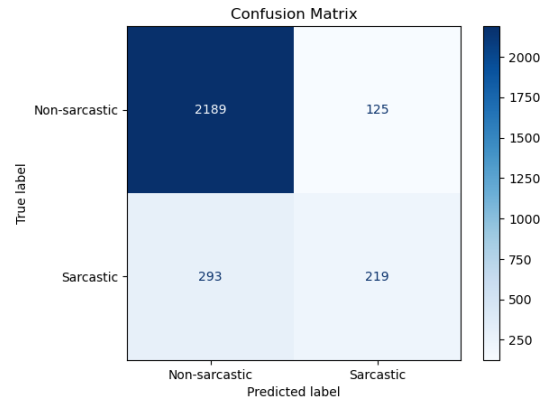
- **AUC-ROC:** is a curve that plots the False Positive Rate (FPR) against the True Positive Rate (TPR). The area under the ROC curve is the AUC. The ROC curve displays how the true positive rate changes with respect to the false positive rate at different threshold levels. Each ROC curve highlights the models ability to distinguish between classes (Sarcastic or Non-sarcastic) with the area under the curve serving as a measure of overall classification performance. A higher AUC reflects better model performance, providing a clear visual insight into the trade-offs involved in prediction accuracy.

$$\text{TPR} = \frac{TP}{TP + FN} \quad \text{FPR} = \frac{FP}{FP + TN}$$

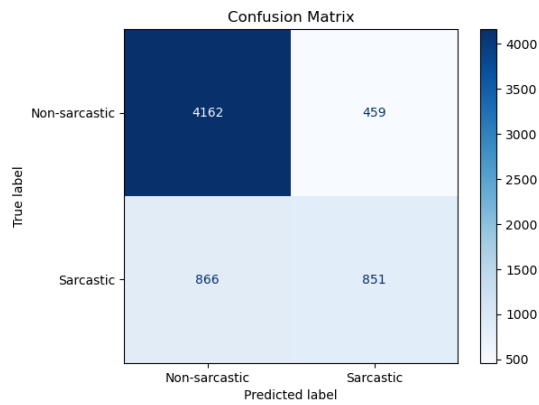
The results of Tamil-English and Malayalam-English for Distil-mBERT, mBERT and RoBERTa can be seen in Tables 3, 4, and 5, respectively. The confusion matrix for the Distil-mBERT model of Tamil-English and Malayalam-English can be seen in Figures 2 and 3, respectively. The confusion matrix for the mBERT model of Tamil-English and Malayalam-English can be seen in Figures 4 and 5, respectively. Similarly, the confusion matrix for the RoBERTa model of Tamil-English and Malayalam-English can be seen in Figures 6 and 7, respectively. The ROC curve for the Distil-mBERT model of Tamil-English and Malayalam-English can be seen in Figures 8 and 9, respectively. The ROC curve for the mBERT



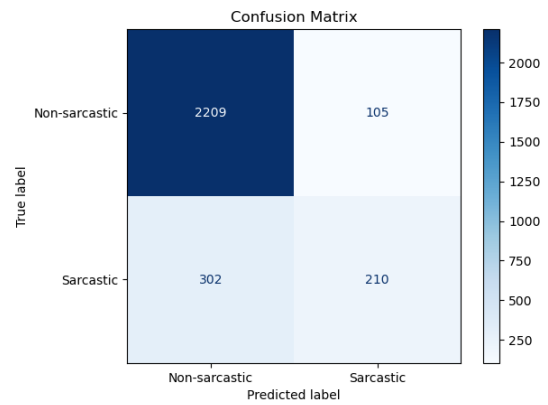
**Figure 2:** Confusion matrix for Tamil (Distil-mBERT)



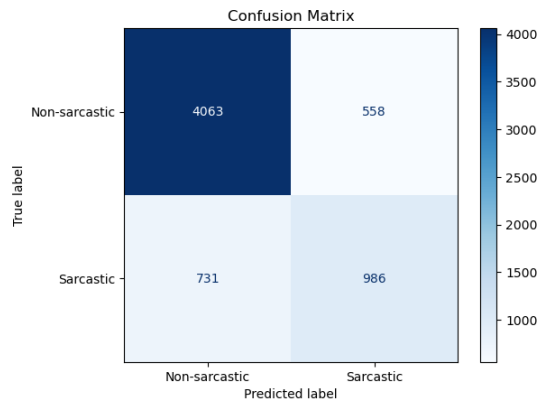
**Figure 3:** Confusion matrix for Malayalam (Distil-mBERT)



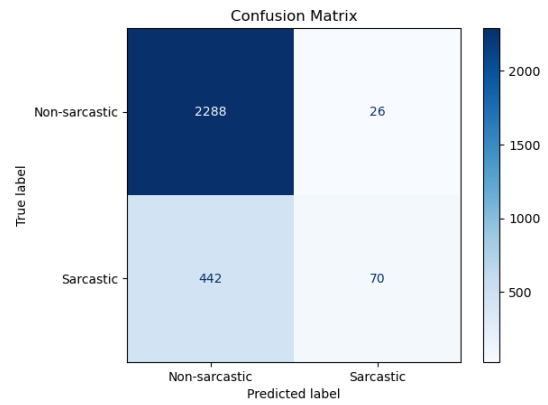
**Figure 4:** Confusion matrix for Tamil (mBERT)



**Figure 5:** Confusion matrix for Malayalam (mBERT)



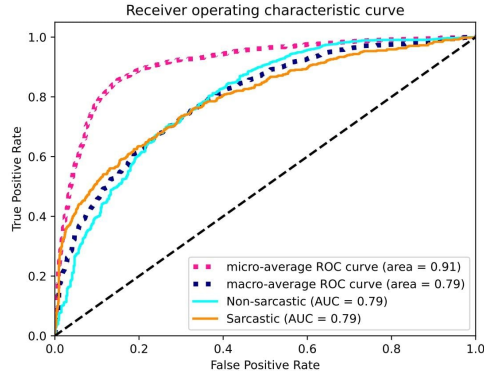
**Figure 6:** Confusion matrix for Tamil (RoBERTa)



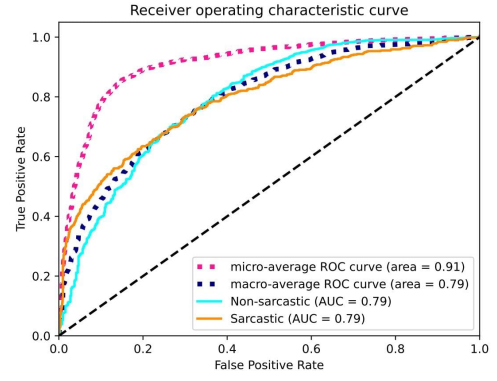
**Figure 7:** Confusion matrix for Malayalam (RoBERTa)

model of Tamil-English and Malayalam-English can be seen in Figures 10 and 11, respectively. The ROC curve for the RoBERTa model of Tamil-English and Malayalam-English can be seen in Figures 12 and 13, respectively. All the results reported here were obtained after the release of the labeled testing dataset to show the class-wise performance of the models.

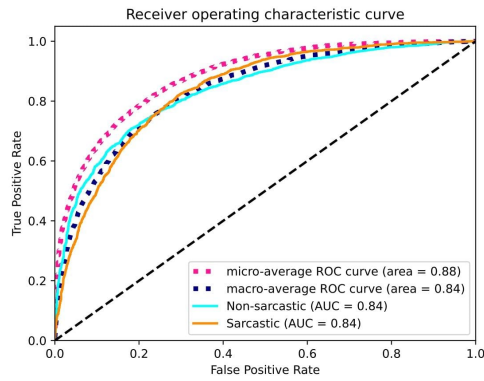




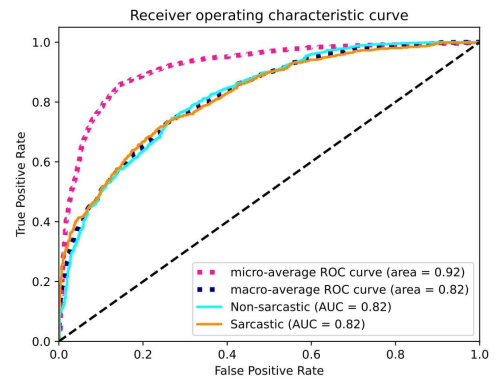
**Figure 8:** ROC curve for Tamil (Distil-mBERT)



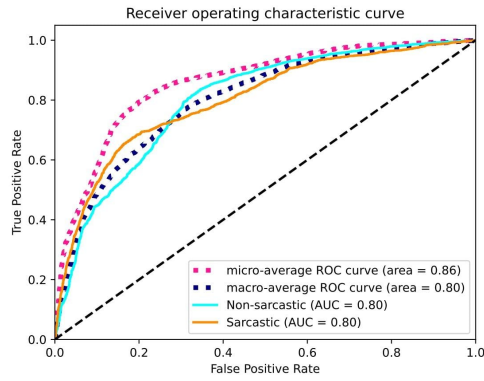
**Figure 9:** ROC curve for Malayalam (Distil-mBERT)



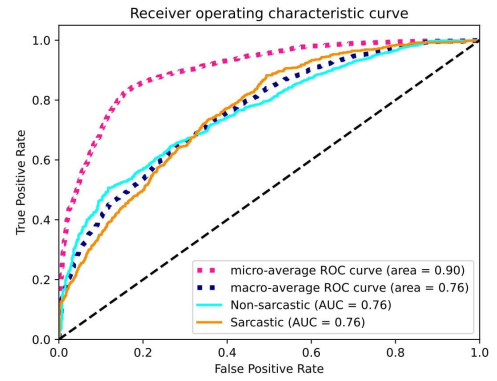
**Figure 10:** ROC curve for Tamil (mBERT)



**Figure 11:** ROC curve for Malayalam (mBERT)



**Figure 12:** ROC curve for Tamil (RoBERTa)



**Figure 13:** ROC curve for Malayalam (RoBERTa)

## 5. Error Analysis

Analyzing code-mixing and language switching within text-data where model predict wrong label can provide deeper insights into the complexities that models face in real-world social media text. **For example:-** Consider the text: *Shavakallarayile Kuzhimaadathile Peril Oru Letter marach pedich lalettan ninnappo Onnu Nadungi*

Tokens : ['Shavakallarayile', 'Kuzhimaadathile', 'Peril', 'Oru', 'Letter', 'marach', 'pedich', 'lalettan',



‘ninnappo’, ‘Onnu’, ‘Nadungi’]

Language Sequence of tokens: [‘tr’, ‘sw’, ‘id’, ‘de’, ‘no’, ‘pl’, ‘it’, ‘it’, ‘fi’, ‘fr’, ‘id’]

It is a Non-Sarcastic text but our model misinterpret it as Sarcastic. As there are 9 language switches in this text and this frequent switching creates a highly dynamic context where the model needs to constantly adjust between different linguistic representations.

In the error analysis of three transformer models for the sarcasm detection task on code-mixed social media posts, mBERT, Distil-mBERT, and RoBERTa perform differently because of their handling capabilities of code-mixing, language switching, and emojis. The mBERT does well in handling language switching but is not robust enough in informal components such as the usage of emojis and slang found typically in social media and X. The RoBERTa performs very well in informal language understanding and social media text but with poor performance regarding code-mixed data due to monolingual pretraining. Though Distil-mBERT was efficient, it lost the capability to handle code-mixing and mixed modalities together, hence bringing less accuracy to sarcasm detection. In summary, the models face varied challenges with the complexities of the social media text, especially with frequent language shifting and the use of emojis.

## 6. Conclusion

Despite the increasing popularity of social media and the limited work on Dravidian code-mixed sarcasm identification, we developed a framework for sarcasm detection in two code-mixed Dravidian corpora, Malayalam and Tamil. Our proposed approach of fine-tuning the multilingual Transformer model yields a comparative performance, achieving competitive scores in DravidianCodeMix@FIRE-2024. This paper shows a comparative study of three transformer-based models—Distil-mBERT, mBERT and RoBERTa in detecting sarcasm in two Dravidian languages: Malayalam and Tamil. For Sarcasm detection in Tamil, RoBERTa performance is better than the other two models with accuracy and macro average precision, recall, and  $F_1$ -score of 80%, 74%, 73%, and 73%, respectively. For Sarcasm detection in Malayalam, mBERT performance is better than the other two models with accuracy and macro average precision, recall, and  $F_1$ -score of 86%, 77%, 68%, and 72%, respectively.

Our finding demonstrates that RoBERTa performs well on the Tamil dataset but struggles with the Malayalam dataset. For Malayalam, mBERT proved to be more effective, outperforming the other models, whereas in Tamil, mBERT achieved a competitive performance. This variation highlights the challenges presented by the intricacies of social media text, where frequent language switching, the use of emojis, and code-mixing add layers of complexity. Notably, our dataset contains predominantly monolingual text, with only a few instances of code-mixing. These findings emphasize the importance of optimizing models for specific languages to handle the nuanced demands of diverse real-world code-mixed datasets.

## Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

## References

- [1] A. Kumar, J. P. Singh, A. K. Singh, Explainable bert-lstm stacking for sentiment analysis of covid-19 vaccination, IEEE Transactions on Computational Social Systems (2023) 1–11. doi:10.1109/TCSS.2023.3329664.
- [2] A. Kumar, J. P. Singh, Deep neural networks for location reference identification from bilingual disaster-related tweets, IEEE Transactions on Computational Social Systems 11 (2024) 880–891. doi:10.1109/TCSS.2022.3213702.

- [3] A. Kumar, J. P. Singh, N. P. Rana, Y. K. Dwivedi, Multi-channel convolutional neural network for the identification of eyewitness tweets of disaster, *Information Systems Frontiers* 25 (2023) 1589–1604.
- [4] R. Pandey, A. Kumar, J. P. Singh, S. Tripathi, Hybrid attention-based long short-term memory network for sarcasm identification, *Applied Soft Computing* 106 (2021) 107348.
- [5] R. Pandey, A. Kumar, J. P. Singh, S. Tripathi, A hybrid convolutional neural network for sarcasm detection from multilingual social media posts, *Multimedia Tools and Applications* (2024) 1–29.
- [6] T. Yue, X. Shi, R. Mao, Z. Hu, E. Cambria, Sarcnet: A multilingual multimodal sarcasm detection dataset, in: *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 2024, pp. 14325–14335.
- [7] Y. Kumar, N. Goel, Ai-based learning techniques for sarcasm detection of social media tweets: State-of-the-art survey, *SN Computer Science* 1 (2020) 318.
- [8] G. H. Aleryani, W. Deabes, K. Albishre, A. E. Abdel-Hakim, Impact of emoji exclusion on the performance of arabic sarcasm detection models, 2024. URL: <https://arxiv.org/abs/2405.02195>. arXiv:2405.02195.
- [9] S. Lakshmi, 10 - sarcasm detection using deep learning in natural language processing, in: D. J. Hemanth (Ed.), *Computational Intelligence Methods for Sentiment Analysis in Natural Language Processing Applications*, Morgan Kaufmann, 2024, pp. 187–205. URL: <https://www.sciencedirect.com/science/article/pii/B9780443220098000136>. doi:<https://doi.org/10.1016/B978-0-443-22009-8.00013-6>.
- [10] M. A. Galal, A. Hassan Yousef, H. H. Zayed, W. Medhat, Arabic sarcasm detection: An enhanced fine-tuned language model approach, *Ain Shams Engineering Journal* 15 (2024) 102736. URL: <https://www.sciencedirect.com/science/article/pii/S2090447924001114>. doi:<https://doi.org/10.1016/j.asej.2024.102736>.
- [11] E. Hashmi, S. Y. Yayilgan, S. Shaikh, Augmenting sentiment prediction capabilities for code-mixed tweets with multilingual transformers, *Social Network Analysis and Mining* 14 (2024) 86.
- [12] R. Bhukya, S. Vodithala, Deep learning based sarcasm detection and classification model, *Journal of Intelligent & Fuzzy Systems* (2024) 1–14.
- [13] Y. Liu, M. Chi, Q. Sun, Sarcasm detection in hotel reviews: a multimodal deep learning approach, *Journal of Hospitality and Tourism Technology* (2024).
- [14] C. Thaokar, J. K. Rout, M. Rout, N. K. Ray, N-gram based sarcasm detection for news and social media text using hybrid deep learning models, *SN Computer Science* 5 (2024) 163.
- [15] A. Nandi, K. Sarkar, A. Mallick, A. De, A survey of hate speech detection in indian languages, *Social Network Analysis and Mining* 14 (2024) 70.
- [16] J. Dai, A bert-based with fuzzy logic sentimental classifier for sarcasm detection, in: *2024 7th International Conference on Advanced Algorithms and Control Engineering (ICAACE)*, 2024, pp. 1275–1280. doi:10.1109/ICAACE61206.2024.10548550.
- [17] M. Amal, R. Boujelbane, M. Ellouze, ANLP RG at StanceEval2024: Comparative evaluation of stance, sentiment and sarcasm detection, in: N. Habash, H. Bouamor, R. Eskander, N. Tomeh, I. Abu Farha, A. Abdelali, S. Touileb, I. Hamed, Y. Onaizan, B. Alhafni, W. Antoun, S. Khalifa, H. Haddad, I. Zitouni, B. AlKhamissi, R. Almatham, K. Mrini (Eds.), *Proceedings of The Second Arabic Natural Language Processing Conference*, Association for Computational Linguistics, Bangkok, Thailand, 2024, pp. 788–793. URL: <https://aclanthology.org/2024.arabicnlp-1.90>. doi:10.18653/v1/2024.arabicnlp-1.90.
- [18] M. E. Hassan, M. Hussain, I. Maab, U. Habib, M. A. Khan, A. Masood, Detection of sarcasm in urdu tweets using deep learning and transformer based hybrid approaches, *IEEE Access* 12 (2024) 61542–61555. doi:10.1109/ACCESS.2024.3393856.
- [19] A. Ameer, S. Hamdi, S. B. Yahia, Domain adaptation approach for arabic sarcasm detection in hotel reviews based on hybrid learning, *Procedia Computer Science* 225 (2023) 3898–3908.
- [20] H. Ghous, M. H. Malik, J. Altaf, S. Nayab, I. Sehrish, S. A. Nawaz, Navigating sarcasm in multilingual text: An in-depth exploration and evaluation, *Journal of Computing & Biomedical Informatics* (2024).

- [21] A. Rawat, S. Kumar, S. S. Samant, Hate speech detection in social media: Techniques, recent trends, and future challenges, *Wiley Interdisciplinary Reviews: Computational Statistics* 16 (2024) e1648.
- [22] L. S. Kumar, A. Hegde, B. R. Chakravarthi, H. Shashirekha, R. Natarajan, S. Thavareesan, R. Sakuntharaj, T. Durairaj, P. K. Kumaresan, C. Rajkumar, Overview of second shared task on sentiment analysis in code-mixed tamil and tulu, in: *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, 2024, pp. 62–70.
- [23] O. Nimase, S. Hong, When do "more contexts" help with sarcasm recognition?, 2024. URL: <https://arxiv.org/abs/2403.12469>. arXiv:2403.12469.
- [24] B. R. Chakravarthi, S. N. B. B. N. K. T. Durairaj, R. Ponnusamy, P. K. Kumaresan, K. K. Ponnusamy, C. Rajkumar, Overview of sarcasm identification of dravidian languages in DravidianCodeMix@FIRE-2024, in: *Forum of Information Retrieval and Evaluation FIRE - 2024, DAIICT*, Gandhinagar, 2024.
- [25] V. Sanh, L. Debut, J. Chaumond, T. Wolf, Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter, *ArXiv abs/1910.01108* (2019).
- [26] J. Devlin, M. Chang, K. Lee, K. Toutanova, BERT: pre-training of deep bidirectional transformers for language understanding, *CoRR abs/1810.04805* (2018). URL: <http://arxiv.org/abs/1810.04805>. arXiv:1810.04805.
- [27] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, Roberta: A robustly optimized BERT pretraining approach, *CoRR abs/1907.11692* (2019). URL: <http://arxiv.org/abs/1907.11692>. arXiv:1907.11692.
- [28] A. S. Maiya, ktrain: A low-code library for augmented machine learning, 2022. URL: <https://arxiv.org/abs/2004.10703>. arXiv:2004.10703.