PrimeLine@DravidianLangTech 2026: Hope Speech Detection in Tulu Using XLM-RoBERTa for Coarse and Fine-Grained Classification

Rithikaa V **Sanjay Krishnan K** **Nithya Varshini C N R** **S. Sumathi**
Department of Information Technology, St. Joseph's College of Engineering
https://github.com/Rithikaa7/PrimeLine-Hope_Speech_Detection

Abstract

Hope speech detection in low-resource, code-mixed languages presents a genuine challenge for natural language processing. Tulu, a Dravidian language spoken along the coastal regions of Karnataka and Kerala, is one such language where social media content is deeply code-mixed, blending Tulu, Kannada script, and English within a single comment. Two classification tasks are addressed: a four-class coarse-grained setting (Track 1) and a five-class fine-grained setting (Track 2). XLM-RoBERTa, a cross-lingual transformer pre-trained on more than 100 languages, is fine-tuned on the task-provided datasets using Google Colab with an NVIDIA T4 GPU. The system achieves a Macro F1-score of 0.34 on Track 1 and 0.19 on Track 2 on the official Codabench evaluation, establishing the first transformer-based baseline for hope speech classification in Tulu.

1 Introduction

Hope speech refers to positive, forward-looking online content that encourages, supports, or inspires others. As NLP research has expanded its scope, hope speech detection has grown into a meaningful area of study, offering a constructive counterpart to the widely explored tasks of hate speech and toxicity detection. While considerable effort has been directed at identifying harmful content, the computational identification of positive and hopeful discourse has received comparatively less attention.

Identifying hopeful language in regional and minority languages is particularly valuable, as it gives visibility to communities rarely represented in large-scale language studies. For speakers of languages such as Tulu, computational tools can open up possibilities for social listening and content moderation tailored to their specific needs.

Tulu is spoken by roughly four to five million people in the Tulu Nadu region of coastal Karnataka and Kerala. Despite its significant speaker population, Tulu remains almost entirely absent from NLP research. There is no dedicated pre-trained language model for Tulu and no established NLP pipeline for the language. Its presence on social media is marked by heavy code-mixing, where native Tulu, Kannada script, English, and emoji are blended freely within a single comment, making automated classification considerably harder than equivalent tasks in monolingual settings.

The DravidianLangTech shared task on Hope Speech Detection in Tulu, organized as part of ACL 2026, provides the first structured benchmark for this problem. Track 1 covers four coarse-grained categories: Encouraging Hope, Discouraging Hope, Blended Hope, and Uninvolved. Track 2 covers five fine-grained categories: Inspiring Hope, Hopelessness, Realistic Hope, Optimistic Hope, and Fading Hope. The transition to fine-grained classification makes the task progressively more demanding, as Track 2 categories differ in degree rather than in kind, with genuinely ambiguous boundaries.

The approach taken in this work centers on fine-tuning XLM-RoBERTa [6], a multilingual transformer pre-trained on more than 100 languages. Its cross-lingual representations allow it to handle Tulu's mixed scripts without any language identification or transliteration steps, making it a practical and principled choice for this low-resource, code-mixed setting. The system is trained end-to-end with a linear classification head on top of the encoder, and predictions are generated by taking the argmax of the softmax output.

2 Related Work

Hope speech detection as a formal NLP task was first introduced by Chakravarthi [2] through the HopeEDI dataset, covering English, Tamil, and Malayalam. The dataset and baselines demonstrated that transformer-based models clearly outperform classical approaches even on small mul-

Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pages 336–340
July 3, 2026

tilingual corpora, setting the stage for the shared tasks that followed.

Chakravarthi and Muralidaran [4] reported findings from a shared task on hope speech detection across English, Tamil, and Malayalam, confirming that multilingual transformers consistently outperform classical machine learning approaches on code-mixed text.

Several systems from those shared tasks demonstrated the strength of XLM-RoBERTa. Ziehe et al. [16] showed that fine-tuning XLM-RoBERTa achieved strong results across English, Malayalam, and Tamil without any task-specific feature engineering. Mahajan et al. [11] ranked first in both English and Tamil by fine-tuning RoBERTa and XLM-RoBERTa respectively. Balouchzahi et al. [1] built an XLM-RoBERTa ensemble specifically for code-mixed Dravidian text. Huang and Bai [9] combined XLM-RoBERTa with TF-IDF features and ranked first in both English and Malayalam. Singh et al. [13] explored multilingual transformer models pre-trained on Indian languages and achieved second place in both English and Malayalam. These results collectively established XLM-RoBERTa as the dominant baseline for Dravidian hope speech classification.

On the model side, Vaswani et al. [15] introduced the Transformer architecture, which set the foundation for the modern pre-trained language model paradigm. Devlin et al. [7] built on this with BERT, demonstrating the power of masked language modelling for downstream NLP tasks. Liu et al. [10] followed with RoBERTa, showing that more robust pre-training strategies yield better performance. Conneau et al. [6] extended this to the multilingual setting with XLM-RoBERTa, pre-trained on over 100 languages including several Indic scripts.

Dravidian language NLP has received growing attention in recent years. Chakravarthi et al. [5] reported a comprehensive shared task overview on hope speech across five languages. Chakravarthi [3] further presented a multilingual hope speech dataset covering English and Dravidian languages. Sundar et al. [14] proposed a stacked encoder architecture using cross-lingual embeddings and demonstrated improvements over standard fine-tuning baselines.

Research on Tulu specifically is extremely sparse. Shetty [12] examined sentiment analysis in code-mixed Tulu and Tamil text, noting the severe scarcity of annotated data. Ehsan et al. [8] tackled Tulu sentiment classification using BiLSTM models with ELMo embeddings, identifying the lack of Tulu-specific language resources as the primary bottleneck. Neither work addresses hope speech detection in Tulu; the current work is the first to do so.

3 Dataset and Preprocessing

3.1 Dataset

The datasets for both tracks were prepared by the shared task organizers and drawn from social media platforms used by Tulu speakers. The text reflects real-world Tulu online writing, mixing Tulu, Kannada script, English, and emoji freely within individual comments. This code-mixed, multi-script nature is a defining feature of the data and makes classification non-trivial even for human annotators.

Track 1 contains 5,991 training samples across four coarse categories. The distribution is imbalanced, with Uninvolved accounting for 2,490 samples (41.6%) and Discouraging Hope having the fewest at 711 (11.9%). Track 2 has 3,185 training samples across five fine-grained labels, with Inspiring Hope as the majority class at 35.4% and Fading Hope as the minority at 7.4%. Tables 1 and 2 show the full distributions.

Label	Count	%
Encouraging Hope	1895	31.6%
Discouraging Hope	711	11.9%
Blended Hope	895	14.9%
Uninvolved	2490	41.6%
Total	5991	100%

Table 1: Class distribution – Track 1 (Coarse-Grained).

Label	Count	%
Inspiring Hope	1129	35.4%
Hopelessness	937	29.4%
Realistic Hope	503	15.8%
Optimistic Hope	380	11.9%
Fading Hope	236	7.4%
Total	3185	100%

Table 2: Class distribution – Track 2 (Fine-Grained).

3.2 Preprocessing

Preprocessing was kept minimal by design. Text samples were passed directly to the XLM-RoBERTa tokenizer, which handles diverse scripts through a SentencePiece-based subword vocabulary. Padding and truncation were applied at a maximum length of 128 tokens. Transliteration, language identification, and stopword removal were deliberately skipped, as these steps risk erasing code-switching cues that carry genuine meaning. No normalization of Kannada or Roman script was applied, preserving the natural variation present in user-generated content. This minimal preprocessing strategy is consistent with prior work on code-mixed Dravidian NLP, where aggressive preprocessing has been shown to hurt rather than help performance.

4 Methodology

4.1 System Pipeline

The system follows a standard transformer fine-tuning pipeline. Each raw text sample is first passed to the XLM-RoBERTa tokenizer, which applies SentencePiece subword segmentation and encodes the input as a sequence of token IDs with a special `[CLS]` token prepended. The encoded sequence is fed into the XLM-RoBERTa encoder, which produces contextualized representations through twelve layers of multi-headed self-attention. The hidden state of the `[CLS]` token at the final encoder layer is extracted and passed through a linear classification head. Output logits are converted to class probabilities via softmax, and the predicted label is the class with the highest probability. For Track 1, the output layer has four neurons; for Track 2, five. Figure 1 illustrates the pipeline.

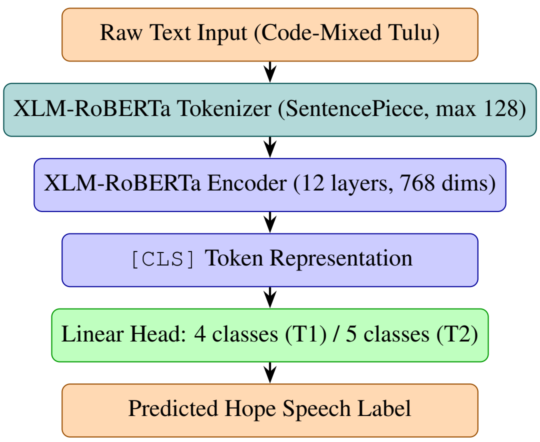

Figure 1: End-to-end system pipeline for hope speech classification in Tulu.

4.2 Model Formulation

The core of the system is `xlm-roberta-base`, a transformer encoder built on the architecture of Vaswani et al. [15] and trained using the robust pre-training strategy of Liu et al. [10] extended to 100+ languages by Conneau et al. [6]. The model contains approximately 270 million parameters and was pre-trained on filtered Common Crawl data spanning 100 languages, including Indic scripts.

Within each transformer layer, the self-attention mechanism computes attention scores over all token pairs. Given query matrix $\mathbf{Q}$, key matrix $\mathbf{K}$, and value matrix $\mathbf{V}$, scaled dot-product attention is:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{d_k}}\right)\mathbf{V} \quad (1)$$

where d_k is the key vector dimensionality. This allows the model to capture long-range dependencies across mixed scripts and languages simultaneously.

The `[CLS]` representation is fed into a linear head, and logits $z_1, \ldots, z_C$ are normalized via softmax to produce class probabilities:

$$P(y = c \mid x) = \frac{\exp(z_c)}{\sum_{j=1}^{C} \exp(z_j)} \quad (2)$$

The model is trained end-to-end by minimizing categorical cross-entropy loss:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic} \log P(y = c \mid x_i) \quad (3)$$

where N is the number of training samples, C the number of classes, and $y_{ic} \in \{0, 1\}$ the one-hot ground truth indicator. Gradients are backpropagated through the full encoder, adapting pre-trained representations to the specific vocabulary and style of Tulu social media text.

4.3 Training Setup

Both models were fine-tuned on Google Colab with an NVIDIA T4 GPU for 2 epochs with batch size 8 and AdamW optimizer (learning rate 2×10^{-5}, weight decay 0.01). A linear warmup was applied over the first 6% of training steps, followed by linear decay to zero. The development set served as the validation split throughout, and the best checkpoint by validation loss was selected for inference. Predictions on test sets were obtained by argmax over output logits, converted to label strings, and submitted to Codabench. No ensembling was applied.

5 Results and Discussion

5.1 Official Results

Table 3 reports official Codabench evaluation scores. Accuracy and macro-averaged precision, recall, and F1 are reported. Macro-averaging is the primary metric as it weights all classes equally, making it sensitive to performance on minority categories regardless of their frequency.

Track	Acc	P	R	F1
Track 1 – Coarse	0.60	0.30	0.41	0.34
Track 2 – Fine	0.24	0.19	0.19	0.19

Table 3: Macro-averaged results on Codabench (P = Precision, R = Recall).

On Track 1, an accuracy of 0.60 is achieved with a Macro F1 of 0.34. The gap between these two figures reflects a model that performs well on the majority Uninvolved class but struggles with less frequent categories. Track 2 is harder, with accuracy at 0.24 and Macro F1 at 0.19, reflecting the semantic closeness of its five fine-grained categories and severe class imbalance.

5.2 Error Analysis

A detailed look at predictions reveals that the accuracy-F1 gap on Track 1 is driven primarily by over-prediction of the Uninvolved class (41.6% of training data). Samples from Blended Hope and Discouraging Hope are prone to being absorbed into this dominant category, as they share surface-level vocabulary with neutral comments. The decision boundary between Encouraging Hope and Blended Hope is especially difficult, since a comment can simultaneously express genuine positivity and realistic acknowledgment of hardship, making clean categorical assignment ambiguous.

Track 2 exhibits a more extreme version of the same problem. The five fine-grained categories are semantically adjacent and differ in degree rather than kind. A comment expressing uncertainty about the future could plausibly belong to Fading Hope, Hopelessness, or Realistic Hope depending on subtle contextual cues difficult to capture with only two epochs of fine-tuning. The scarcity of Fading Hope samples (236 of 3,185) further limits learning for this class.

5.3 Discussion and Future Directions

The results demonstrate that XLM-RoBERTa provides a viable starting point for hope speech detection in Tulu even without Tulu-specific pre-training. Future work should address class imbalance via weighted loss functions or back-translation augmentation from Kannada. Training for more epochs with early stopping based on macro F1 could help balance performance across classes. Larger multilingual models or Dravidian-family-specific pre-trained models could yield better representations for Tulu's lexical patterns. Contrastive learning objectives that explicitly separate semantically similar hope categories could also resolve the fine-grained boundary confusion observed in Track 2.

6 Conclusion

This work presents the first transformer-based system for hope speech detection in code-mixed Tulu, fine-tuned on datasets from the DravidianLangTech@ACL 2026 shared task. XLM-RoBERTa is applied to both tracks, achieving Macro F1 scores of 0.34 and 0.19 on Track 1 and Track 2 respectively. The results establish a meaningful baseline for future work and highlight the dual challenge of class imbalance and fine-grained semantic similarity. Despite modest absolute scores, the cross-lingual transfer capabilities of XLM-RoBERTa prove sufficient to produce non-trivial predictions even for a language with no dedicated NLP resources. It is hoped that this contribution encourages further computational research on Tulu and other low-resource Dravidian languages that remain largely absent from the NLP literature.

Limitations

The primary limitation is the use of only two training epochs, restricting the model's ability to learn reliable decision boundaries for minority classes. The absence of Tulu-specific pre-trained representations means cross-lingual transfer may not fully capture the language's phonological and morphological patterns. Class imbalance was not addressed through resampling or loss-weighting, suppressing performance on rare categories. The system does not perform language identification or script normalization, both avenues for future improvement.

Acknowledgments

The organizers of the DravidianLangTech@ACL 2026 shared task are thanked for providing the datasets and the Codabench evaluation platform. Sincere gratitude is expressed to the project guide, S. Sumathi, Department of Information Technology, St. Joseph's College of Engineering, for constant guidance throughout this work.

References

[1] Fazlourrahman Balouchzahi, Aparna B K, and H. L. Shashirekha. MUCS@LT-EDI-EACL2021: CoHope – hope speech detection for equality, diversity, and inclusion in code-mixed texts. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 180–187, 2021. URL https://aclanthology.org/2021.ltedi-1.25.

[2] Bharathi Raja Chakravarthi. HopeEDI: A multilingual hope speech detection dataset for equality, diversity, and inclusion. In *Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotion's in Social Media*, pages 41–53, 2020. URL https://aclanthology.org/2020.peoples-1.5.

[3] Bharathi Raja Chakravarthi. Multilingual hope speech detection in English and Dravidian languages. *International Journal of Data Science and Analytics*, 14 (4):389–406, 2022. URL https://doi.org/10.1007/s41060-022-00324-5.

[4] Bharathi Raja Chakravarthi and Vigneshwaran Muralidaran. Findings of the shared task on hope speech detection for equality, diversity, and inclusion. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 61–72, 2021. URL https://aclanthology.org/2021.ltedi-1.8.

[5] Bharathi Raja Chakravarthi et al. Overview of the shared task on hope speech detection for equality, diversity, and inclusion. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 378–388, 2022. URL https://aclanthology.org/2022.ltedi-1.53.

[6] Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451, 2020. URL https://aclanthology.org/2020.acl-main.747.

[7] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT 2019*, pages 4171–4186, 2019. URL https://aclanthology.org/N19-1423.

[8] Toqeer Ehsan, Amina Tehseen, Kengatharaiyer Sarveswaran, and Amjad Ali. AlphaBrains@DravidianLangTech: Sentiment analysis of code-mixed Tamil and Tulu using ELMo embeddings. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 152–159, 2023. URL https://aclanthology.org/2023.dravidianlangtech-1.22.

[9] Bo Huang and Yang Bai. TEAM HUB@LT-EDI-EACL2021: Hope speech detection based on pre-trained language model. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 122–127, 2021. URL https://aclanthology.org/2021.ltedi-1.17.

[10] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. RoBERTa: A robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019. URL https://arxiv.org/abs/1907.11692.

[11] Khyati Mahajan, Erfan Al-Hossami, and Samira Shaikh. TeamUNCC@LT-EDI-EACL2021: Hope speech detection using transfer learning with transformers. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 136–142, 2021. URL https://aclanthology.org/2021.ltedi-1.19.

[12] Poorvi Shetty. Poorvi@DravidianLangTech: Sentiment analysis on code-mixed Tulu and Tamil corpus. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 124–132, 2023. URL https://aclanthology.org/2023.dravidianlangtech-1.19.

[13] Pankaj Singh, Prince Kumar, and Pushpak Bhattacharyya. CFILT IIT Bombay@LT-EDI-EACL2021: Hope speech detection using multilingual representation from transformers. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 193–196, 2021. URL https://aclanthology.org/2021.ltedi-1.27.

[14] Arunima Sundar, Akshay Ramakrishnan, Avantika Balaji, and Thenmozhi Durairaj. Hope speech detection for Dravidian languages using cross-lingual embeddings with stacked encoder architecture. *SN Computer Science*, 3:67, 2022. URL https://doi.org/10.1007/s42979-021-00948-3.

[15] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, pages 5998–6008, 2017. URL https://papers.nips.cc/paper/2017/has 3f5ee243547dee91fbd053c1c4a845a html.

[16] Stefan Ziehe, Franziska Pannach, and Aravind Krishnan. GCDH@LT-EDI-EACL2021: XLM-RoBERTa for hope speech detection in English, Malayalam, and Tamil. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 132–135, 2021. URL https://aclanthology.org/2021.ltedi-1.18.