SJM_MINDS@DravidianLangTech@ACL2026: Machine Learning Approaches for Hope Speech Detection in Code-Mixed Tulu

Hosahalli Lakshmaiah Shashirekha, Manjula Suvarna and Jayashree Krishna
Department of Computer Science, Mangalore University, Mangalore, India
{hlsrekha, manjulashivamu, jayashreekmu}@mangaloreuniversity.ac.in

Abstract

Hope speech detection is an important task in understanding emotionally constructive communication in online platforms, especially in low-resource and code-mixed languages. This paper describes our system submitted to the first shared task on Hope Speech Detection in Code-Mixed Tulu, organized by DravidianLangTech@ACL 2026. The shared task consists of two tasks: Task 1 - Coarse-Grained Hope Tone Classification and Task 2 - Fine-Grained Hope Type Classification, with the objective of detecting and classifying the tone and type of hope expressed in code-mixed Tulu texts. We experimented with Logistic Regression (LR) and Linear Support Vector Classifier (LinearSVC) - classical Machine Learning (ML) approaches, trained with Term Frequency and Inverse Document Frequency (TF-IDF) of word ngrams (n = 1, 2). For Task 1, we employed both models, whereas for Task 2, we employed only the LR model. Linear SVC obtained a macro F1-score of 0.51 in Task 1 and secured 4th rank, while the LR model obtained a macro F1-score of 0.37 in Task 2 and secured 5th rank. The results demonstrate that traditional ML approaches remain effective for low-resource code-mixed language scenarios.

1 Introduction

Digital technologies have enabled instant information sharing and interaction through social networks such as Facebook, Instagram, and YouTube. This has led to the increased volume of user-generated content in regional and code-mixed languages. Users are taking advantage of social media and are sharing hate speech, hope speech, in addition to other types of content. Detecting constructive emotional content, such as hope speech, is essential for promoting positive online interactions and reducing harmful communication.

Hope can be described as an openness of spirit toward the future—a desire, expectation, or wish for something to occur or be true—that significantly influences a person's mindset, emotions, behavior, and decision-making (Balouchzahi et al., 2023). It is a powerful force that motivates individuals to persist despite life's uncertainties (Kumaresan et al., 2023). Detecting hope speech focuses on identifying whether a piece of text expresses hope or not. However, detecting it manually is both tedious and time-consuming because of the rising number of users and the expanding volume of social media content (Divakaran et al., 2024).

While sentiment analysis and emotion detection have been extensively studied, hope speech detection in low-resource and code-mixed languages such as Tulu remains underexplored. The detection of hope speech has emerged as a significant task, driven by the need to recognize motivational expressions and goal-oriented behavior of users on social media platforms (Ramos et al., 2025). As there are no guidelines to create any post/comment on social media, users usually combine words and sub-words belonging to more than one language they know, leading to code-mixed text. Code-mixed Tulu text usually combines Tulu with Kannada and English words, introducing challenges such as spelling variations, transliteration inconsistencies, and informal grammar.

Hope Speech Detection in Code-Mixed Tulu Language shared task[1], organized by DravidianLangTech@ACL 2026, aims to promote research in low-resource and code-mixed Tulu language in the context of social media text (Thenmozhi et al., 2026). The primary objective of this task is to automatically identify and categorize hope-related expressions in user-generated content. The shared task consists of two tasks:

- **Task 1: Coarse-Grained Hope Tone Classification** - consists of four classes - *Encouraging*, *Discouraging*, *Uninvolved*, and *Blended*

[1] https://www.codabench.org/competitions/11328/

Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pages 359–365

Table 1: Sample text with their English Translation and Corresponding Label

Sample Code-mixed Text	**English Translation**	**Classes**
Nett comdey olu unndu ...?	Where's the comedy in this?	Blended Hope
Mast samaryd boka..benner batter	We had a guest after a very long time	Uninvolved
ಬಾರಿಷೋಕುಅತ್ಂಡ್	It was very nice	Encouraging Hope
Anchor n onji change malple	Kindly change the anchor	Discouraging Hope

Tone, with the goal to identify the overall emotional tone of a given social media post (tweet, comment, etc.) based on the presence, absence, or ambiguity of hope-related expressions.

- **Task 2: Fine-Grained Hope Type Classification** - consists of five classes - *Optimistic Hope*, *Realistic Hope*, *Inspiring Hope*, *Fading Hope*, and *Hopelessness*, with the goal of refining the classification by identifying specific types of hope expressed in the text.

Few representative samples from the shared task dataset (Thenmozhi et al., 2026) along with their English translations, are shown in Table 1.

To address the challenges of the shared task, in this paper, we describe our proposed system for hope speech detection in code-mixed Tulu. We explored classical ML models: LR and LinearSVC, trained with TF-IDF of words n-grams (n = 1, 2) for hope speech detection.

The rest of the paper is organized as follows: Section 2 presents the related work, Section 3 describes the methodology, Section 4 discusses the experiments, results and error analysis and Section 5 concludes the paper with future work.

2 Related Work

Compared to high-resource languages, low-resource languages face issues such as limited labeled data, high dialectal variation, and lack of pretrained language models. Despite these limitations, hope speech detection in low-resource languages has gained attention with the increase in the number of social media users and also the increasing volume of user-generated text. A few studies relevant in this direction are discussed below:

Mesay Gemeda Yigezu et al. (Yigezu et al., 2023) focused on detecting hope speech in social media posts in English and Spanish as part of IberLEF 2023 shared task. Support Vector Machine (SMV)-based model trained with TF-IDF of words achieved macro F1-scores of 0.489 for English and 0.481 for Spanish. Aggarwal et al. (Aggarwal et al., 2023) examined ML methods for classifying English YouTube comments into *Hope Speech*, *Non-Hope Speech*, or *Neutral*, using the EACL-2021: Hope Speech Detection for Equality, Diversity, and Inclusion" shared task dataset. They identified label inconsistencies and relabeled the dataset, testing traditional models (Naïve Bayes, LR, SVM) and BERT. BERT achieved the best results, with an F1-score of 0.6966, and improved to 0.85 when the neutral class was excluded for binary classification.

Hegde et al. (Hegde et al., 2023b) addressed the identification of hope speech in English and code-mixed texts in Bulgarian, Spanish, and Hindi, in a shared task on Hope Speech Detection for Equality, Diversity, and Inclusion (LT-EDI) at RANLP 2023. The authors proposed models that combined transformer-based embeddings, including BERT and multilingual BERT, with TF-IDF character n-gram features, and employed LinearSVC for classification. Their approaches achieved strong performance, securing top ranks across multiple language tracks. Uzor et al. (Uzor et al., 2025) compared SVM, deep learning models, and transformer-based approaches for hope speech detection. SBERT achieved the highest accuracy in both binary (92.81%) and multiclass (84.29%) tasks, outperforming BERT, GPT-2, and traditional models. Overall, transformer models proved most effective at capturing nuanced categories, though challenges remain for underrepresented classes in low-resource settings.

Abiola et al. (Abiola et al., 2026) evaluated traditional ML models against fine-tuned transformer architectures for hope speech detection. The LaBSE (Language-agnostic BERT Sentence Embedding) model consistently outperformed the Ridge Classifier baseline, achieving macro-F1 scores of 0.861 (English), 0.868 (Spanish), 0.948 (Urdu), and 0.839 (German). These results show that while traditional models remain competitive, transformer models more effectively capture subtle semantic patterns in hope speech. Similarly, Sharma et al. (Sharma et al., 2025) proposed an ensemble framework for low-resource languages (English, Kannada, Malayalam, Tamil), combining LSTM, mBERT, and XLM-RoBERTa. Their ensemble outperformed individual models, with weighted F1-scores of 0.93 (English), 0.74 (Kannada), 0.82 (Malayalam), and 0.60 (Tamil), highlighting the strength of ensemble methods for multilingual hope speech detection in low-resource settings.

Research on Tulu NLP is steadily expanding through shared tasks and resource development initiatives led by multiple researchers. These efforts—ranging from sentiment analysis to parallel corpus construction—are gradually building the foundation for more advanced low-resource language processing in Tulu.

DravidianLangTech@RANLP 2023 introduced sentiment analysis for code-mixed Tamil and Tulu, marking the first public-domain study of Tulu sentiment analysis, where systems were evaluated using macro F1-score (Hegde et al., 2023a). To advance neural machine translation for low-resource Dravidian languages, a Kannada–Tulu parallel corpus was developed, with transformer-based models outperforming traditional encoder–decoder approaches (Hegde and Shashirekha, 2023). LT-EDI@EACL 2024 addressed the detection of homophobia and transphobia in multilingual social media comments, including Tulu (Chakravarthi et al., 2024). Morphological segmentation studies for Kannada and Tulu further demonstrated the effectiveness of CRF-based models with syllable and character features for morpheme boundary detection (Hegde and Shashirekha, 2024). The shared task at DravidianLangTech@NAACL 2025 addressed multi-class sentiment classification using data from platforms such as YouTube and Twitter (Thenmozhi et al., 2025). Narayanan et al. (Narayanan and Aepli, 2024) developed an English–Tulu parallel corpus and machine translation system for the low-resource Tulu language.

While most of the work focuses on English, low-resource languages, and code-mixed texts in low-resource languages continue to face issues such as lexical sparsity, limited annotated data, imbalanced data, and lack of pretrained models. This motivated us to explore traditional ML models for hope speech detection in code-mixed Tulu language.

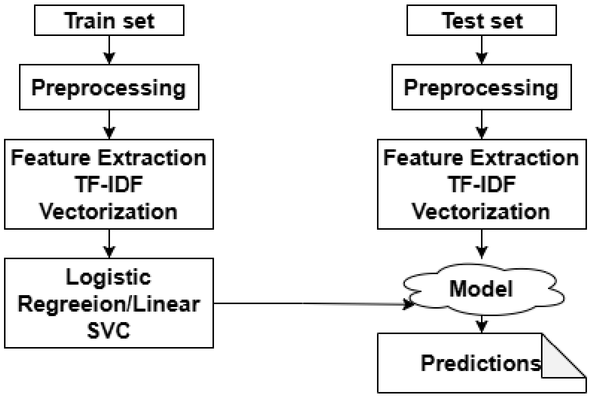

Figure 1: Framework of the Proposed Methodology for Hope Speech Detection in Code-mixed Tulu

3 Methodology

The workflow of the proposed methodology is depicted in Figure 1 and the steps involved are explained in the following subsections:

3.1 Preprocessing

Text data is subjected to a standardized preprocessing pipeline to ensure consistency and reduce noise. The preprocessing steps included lower-casing, removal of punctuation and special characters, and tokenization. These steps help normalize textual input and reduce sparsity in feature space, thereby improving the effectiveness of downstream tasks.

3.2 Feature Extraction

To transform textual data into numerical form, we extracted word n-grams (n = 1, 2) and vectorized them using TfidfVectorizer[2]. TF-IDF is widely used in text classification as it effectively captures the importance of terms relative to their frequency within individual documents and across the corpus.

3.3 Model Description

We employed the classical ML models - LR and LinearSVC, to detect hope speech in code-mixed Tulu text.

- **LR** classifier is a strong baseline for text classification due to its robustness, interpretability, and efficiency in high-dimensional feature spaces. LR model is trained with TF-IDF of word ngrams (n = 1, 2) to capture both individual words and short contextual dependencies with a maximum feature size of 5,000. Limiting the feature size helps control model complexity and reduce overfitting. To ensure

[2]https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html

proper convergence during optimization, the maximum number of training iterations was set to 1000.

- **LinearSVC** aims to find an optimal hyperplane that maximizes the margin between classes. It is particularly effective for sparse text representations and often performs competitively in text classification tasks. LinearSVC is trained with TF-IDF of word ngrams (n = 1, 2) to capture both individual words and short contextual dependencies. The maximum feature size is set to 8,000 to allow a richer representation. Additionally, a minimum document frequency (min_df) of 2 is applied to remove rare terms that may introduce noise. Since the dataset exhibited class imbalance, we set the class_weight parameter to balanced. This automatically adjusts class weights inversely proportional to class frequencies, thereby mitigating bias toward majority classes and improving generalization on minority categories.

These classifiers are selected due to their effectiveness in handling high-dimensional sparse feature representations. We explored both LR and LinearSVC for Task 1, whereas only LR model for Task 2.

4 Experiments and Results

The dataset provided for the shared task consists of code-mixed Tulu comments collected from social media platforms. The distribution of data across Train and Development sets for Task 1 and 2 are shown in Tables 2 and 3, respectively. The Test sets for Task 1 and 2 consists of 1,284 and 682 samples respectively.

Model performance is evaluated by the organizers using Precision, Recall, and F1-score, and the models are ranked based on macro F1-score. The performances of the proposed models on Development and Test sets are reported in Tables 4 and 5, respectively. In Task 1, Linear SVC achieved a macro F1-score of 0.51 and secured 4th rank in the shared task, whereas for Task 2, LR model obtained a macro F1-score of 0.37 and secured 5th rank. The comparison of macro F1-scores of the participating teams for Task 1 and 2 are shown in Figures 2 and 3, respectively.

Table 2: Distribution of data in Train and Development Sets for Task 1

Class	Train Set	Development Set
Encouraging	1,895	406
Discouraging	711	153
Uninvolved	2,490	534
Blended Tone	895	191

Table 3: Distribution of data in Train and Development Sets for Task 2

Class	Train Set	Development Set
Optimistic Hope	380	81
Realistic Hope	503	108
Inspiring Hope	1,129	242
Fading Hope	236	51
Hopelessness	937	200

4.1 Error Analysis

To gain deeper insights into the classification behavior of the proposed models, the confusion matrices for Task 1 and Task 2 shown in Figures 4 and 5, are analyzed.

For Task 1, the confusion matrix shows that the model is able to correctly identify a significant portion of the *Encouraging* and *Uninvolved* categories However, the model fails to identify *Discouraging* and *Blended Tone*, as these categories often contain overlapping linguistic expressions that combine both supportive and discouraging sentiments. Additionally, several *Encouraging* and *Discouraging* comments are also incorrectly predicted as *Uninvolved*. This may be due to implicit sentiment expressions in code-mixed Tulu comments, where supportive or discouraging intent is not always explicitly conveyed by strong sentiment indicators.

For Task 2, the model demonstrates strong performance in identifying the *Inspiring Hope* and *Hopelessness* categories. However, the instances of *Optimistic Hope* and *Realistic Hope* are frequently misclassified as *Inspiring Hope*. This suggests that the classifier tends to associate general positive sentiment with inspirational expressions. Furthermore, the *Fading Hope* category shows very few correct predictions, indicating that the model struggles to capture the subtle linguistic cues of gradually diminishing hope. These misclassifications highlight the complexity of fine-grained hope classification, particularly in code-mixed social media text, where contextual and semantic nuances are crucial.

The relatively small dataset and the presence of

Table 4: Results for Development Sets of Task 1 and Task 2

Task	Model	Precision	Recall	F1-score	Accuracy
Task 1	LR	0.62	0.51	0.52	0.67
	LinearSVC	0.53	0.53	0.53	0.62
Task 2	LR	0.32	0.34	0.31	0.50

Table 5: Results for Test Sets of Task 1 and Task 2

Task	Model	Precision	Recall	F1-score	Accuracy	Rank
Task 1	LinearSVC	0.52	0.51	0.51	0.62	4
Task 2	LR	0.64	0.39	0.37	0.56	5

code-mixed Tulu text with informal spellings and linguistic variations, collectively contribute to the lower macro F1-score achieved by the models. Few misclassified instances are shown in Table 6.

5 Conclusion

This paper presents the models submitted for Hope Speech Detection in the Code-Mixed Tulu shared task organized by DravidianLangTech@ACL 2026. We experimented with LR and LinearSVC models using word n-grams (n = 1, 2) features. Linear SVC with class-weight balance achieved the macro F1-score of 0.51 in Task 1 and secured the 4th rank, while the LR model achieved the macro F1-score of 0.37 and secured the 5th rank in Task 2. The findings demonstrate that classical ML approaches remain competitive for low-resource and code-mixed language classification tasks. More feature engineering and other ML models will be explored further.

Acknowledgments

The authors acknowledge the use of OpenAI's ChatGPT for assistance in grammar checking, language refinement, and improving the presentation of the manuscript. All research contributions, interpretations, and conclusions are developed by the authors.

Limitations

Despite the promising results obtained in this study, several limitations remain.

- The dataset used for the experiments is relatively small, which may limit the model's ability to learn diverse linguistic patterns and generalize effectively to unseen data.

- The presence of code-mixed Tulu text introduces additional challenges due to informal spellings, transliteration variations, and inconsistent grammatical structures commonly found in social media comments. These characteristics make it difficult for traditional ML models to capture the full contextual meaning of the text.

- Class imbalance across categories, particularly in Task 2, where some classes, such as *Fading Hope* have significantly fewer instances. This imbalance may negatively affect the model's ability to accurately learn and predict minority classes.

- The use of TF-IDF for feature representation fails to capture deeper semantic relationships in the text.

Overall, the study is limited by the small dataset size, code-mixed text complexities, class imbalance and the limitations of the ML models to generalize.

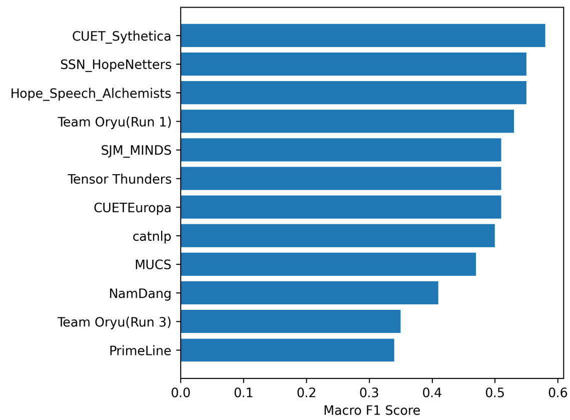

Figure 2: Comparison of Macro F1 scores of the Participating Teams for Task 1

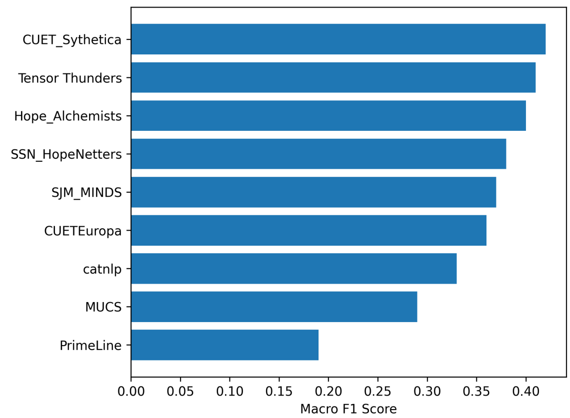

Figure 3: Comparison of Macro F1-scores of the Participating Teams for Task 2

Table 6: Sample Misclassified Instances in Task 1 and Task 2 along with their True and Predicted Labels

Task	Text	True	Predicted
Task 1	Ini pura picture baruva	Blended Tone	Uninvolved
	ಸತ್ಯೋದ ಬೆರಿಯ ಸತ್ಯೋಲು ಉಲ್ಲಾ, ಪಂಡ್ ತೋಜಾದ್ ಕೊರಿಯರ್.	Encouraging	Uninvolved
	20 ne sala thupuni	Discouraging	Uninvolved
Task 2	ಅಜ್ಜ ಭಕ್ತಿ ಉಂಡು	Optimistic	Inspiring
	Tulunaad n oripuga... tulu bhashen oripaaga.	Realistic	Inspiring
	Yakshaganada choukattuve gotthijji.	Fading	Hopelessness
	ಈಈಈಈರ್ ಭಯಂಕರ ಮಾರ್ಲೆ 😄😄😁😁😆😆👌👌👌	Inspiring	Hopelessness

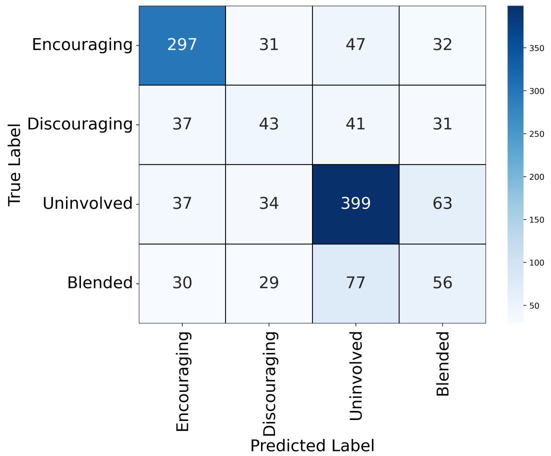

Figure 4: Confusion matrix for Task 1 - LinearSVC Model

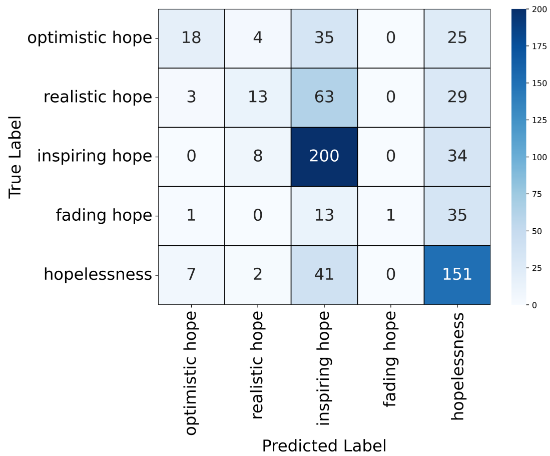

Figure 5: Confusion matrix for Task 2 - LR Model

References

Oluwatobi Joseph Abiola, Tolulope Olalekan Abiola, Obdulia Pichardo Lagunas, and Grigori Sidorov. 2026. Hope Speech Detection: A Comparative Study across Four Languages. *International Journal of Combinatorial Optimization Problems and Informatics*, 17(2):215–230.

Pranjal Aggarwal, Pasupuleti Chandana, Jagrut Nemade, Shubham Sharma, Sunil Saumya, and Shankar Biradar. 2023. Hope Speech Detection on Social Media Platforms. In *Cybercrime in Social Media*, pages 67–84. Chapman and Hall/CRC.

Fazlourrahman Balouchzahi, Grigori Sidorov, and Alexander Gelbukh. 2023. Polyhope: Two-Level Hope Speech Detection from Tweets. *Expert Systems with Applications*, 225:120078.

Bharathi Raja Chakravarthi, Prasanna Kumar Kumaresan, Ruba Priyadharshini, Paul Buitelaar, Asha Hegde, Hosahalli Lakshmaiah Shashirekha, Saranya Rajiakodi, Miguel Ángel García-Cumbreras, Salud María Jiménez-Zafra, José Antonio García-Díaz, Rafael Valencia-García, Kishore Kumar Ponnusamy, Poorvi Shetty, and Daniel García-Baena. 2024. Overview of Third Shared Task on Homophobia and Transphobia Detection in Social Media Comments. In *Proceedings of the Fourth Workshop on Language Technology for Equality, Diversity, Inclusion*, pages 124–132, St. Julian's, Malta. Association for Computational Linguistics.

S. Divakaran, K. Girish, and H. L. Shashirekha. 2024. Hope on the Horizon: Experiments with Learning Models for Hope Speech Detection in Spanish and English. In *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2024), Co-located with the 40th Conference of the Spanish Society for Natural Language Processing (SEPLN 2024)*. CEUR Workshop Proceedings.

Asha Hegde, Bharathi Raja Chakravarthi, Hosahalli Lakshmaiah Shashirekha, Rahul Ponnusamy, Subalalitha Chinnaudayar Navaneethakrishnan, Lavanya Sambath Kumar, Durairaj Thenmozhi, Martha Karunakar, Shreya Sriram, and Sarah Aymen. 2023a. Findings of the Shared Task on Sentiment Analysis in Tamil and Tulu Code-Mixed Text. pages 64–71, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Asha Hegde, Kavya G, Sharal Coelho, and Hosahalli Lakshmaiah Shashirekha. 2023b. MUCS@LT-EDI2023: Learning Approaches for Hope Speech Detection in Social Media Text. In *Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 279–286, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Asha Hegde and Hosahalli Lakshmaiah Shashirekha. 2023. KT2: Kannada–Tulu Parallel Corpus Construction for Neural Machine Translation. In *Proceedings of the 20th International Conference on Natural Language Processing (ICON)*, pages 743–753, Goa University, Goa, India. NLP Association of India (NLPAI).

Asha Hegde and Hosahalli Lakshmaiah Shashirekha. 2024. *Syllable-Level Morphological Segmentation of Kannada and Tulu Words*, pages 113–133. Wiley, Hoboken, NJ, USA.

Prasanna Kumar Kumaresan, Bharathi Raja Chakravarthi, Subalalitha Cn, Miguel Ángel García-Cumbreras, Salud María Jiménez Zafra, José Antonio García-Díaz, Rafael Valencia-García, Momchil Hardalov, Ivan Koychev, Preslav Nakov, Daniel García-Baena, and Kishore Kumar Ponnusamy. 2023. Overview of the Shared Task on Hope Speech Detection for Equality, Diversity, and Inclusion. In *Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 47–53, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Manu Narayanan and Noëmi Aepli. 2024. A Tulu Resource for Machine Translation. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 1756–1767, Torino, Italia. ELRA and ICCL.

Luis Ramos, Hiram Calvo, and Olga Kolesnikova. 2025. Hope Speech Detection in Social Media English Corpora: Performance of Traditional and Transformer Models. *Preprint*, arXiv:2510.23585.

Deepawali Sharma, Vedika Gupta, Vivek Kumar Singh, and Bharathi Raja Chakravarthi. 2025. Stop the Hate, Spread the Hope: An Ensemble Model for Hope Speech Detection in English and Dravidian Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 24(3):1–18.

Durairaj Thenmozhi, Bharathi Raja Chakravarthi, Asha Hegde, Hosahalli Lakshmaiah Shashirekha, Rajeswari Natarajan, Sajeetha Thavareesan, Ratnasingam Sakuntharaj, Krishnakumari Kalyanasundaram, Charmathi Rajkumar, Poorvi Shetty, and Harshitha S Kumar. 2025. Overview of the Shared Task on Sentiment Analysis in Tamil and Tulu. In *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 732–738, Acoma, The Albuquerque Convention Center, Albuquerque, New Mexico. Association for Computational Linguistics.

Durairaj Thenmozhi, Rathnakar Shetty P, Parameshwar R. Hegde, M D Anusha, Raksha Adyanthaya, Mohammed Fadhel Aljunid, Prasanna Kumar Kumaresan, and Bharathi Raja Chakravarthi. 2026. Findings of the Shared Task on Hope Speech Detection in Tulu. In *Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 45–55, Melbourne, Australia. Association for Computational Linguistics.

GodsGift Uzor, Mohammadsaleh Pajuhanfard, Michael Beebe, Maaz Amjad, and Victor Sheng. 2025. Hope Speech Detection: A Comparative Study of Machine Learning, Deep Learning, and Transformer-Based Models. *Journal of Artificial Intelligence Research*, 74:123–145.

Mesay Gemeda Yigezu, Girma Yohannis Bade, Olga Kolesnikova, Grigori Sidorov, and Alexander F Gelbukh. 2023. Multilingual Hope Speech Detection Using Machine Learning. In *Proceedings of the IberLEF 2023 Workshop at SEPLN*, volume 3498, pages 123–134. CEUR Workshop Proceedings.