cantnlp@DravidianLangTech 2026: organic domain adaptation improves multi-class hope speech detection in Tulu

Andrew Li
Lake Washington School District
landrewi@hotmail.com

Sidney Wong
University of Otago, New Zealand
Te Pūnaha Matatini, New Zealand
sidney.wong@otago.ac.nz

Abstract

This paper presents our systems and results for the Hope Speech Detection in Code-Mixed Tulu Language shared task at the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages (DravidianLangTech-2026). We trained an XLM-RoBERTa-based text classification system for detecting hope speech in code-mixed Tulu social media comments. We compared this organically adapted hope speech detection model with our baseline model. On the development set, the organically adapted model outperformed the baseline system. While our submitted systems performed more modestly on the official test set, these results suggest that further adapting XLM-RoBERTa on organically collected Tulu social media text containing code-mixed and mixed-script variation can improve hope speech detection in code-mixed Tulu.

1 Introduction

There has been increasing interest in detecting different forms of language on social media, including hate speech, offensive language, and other emotionally meaningful content. More recently, researchers have also begun to study hope speech, which includes supportive, encouraging, and empathetic expressions in online communication (Chhabra and Vishwakarma, 2023; Ghosal et al., 2026). Detecting hope speech is useful for promoting healthier online discourse, identifying constructive engagement, and supporting research on emotional well-being in low-resource communities (Chakravarthi et al., 2022b). These goals become more difficult in under-resourced languages, where limited labeled data and strong linguistic variation can reduce the effectiveness of standard text classification systems. This challenge is especially apparent in code-mixed settings, where multiple languages and scripts may appear within the same comment.

In this paper, we describe our system for the Hope Speech Detection in Code-Mixed Tulu shared task at DravidianLangTech-2026 (Thenmozhi et al., 2026). The purpose of this shared task is to develop a multi-level hope speech classification system in code-mixed Tulu. Our approach used XLM-RoBERTa as a transformer-based text classifier (Conneau et al., 2020). Furthermore, we extended a similar fine-tuning strategy used in prior shared task work (Wong and Durward, 2024). Our primary motivation was to account for the script-switching phenomena observed in the training data to ensure that we account for the orthographic (and linguistic) context. We further adapt the model using additional organic Tulu social media text containing code-mixed and mixed-script variation to better match the linguistic patterns found in real Tulu social media text.

2 Related Works

The first shared task to explored hope speech detection in YouTube comments across three language conditions: English, Tamil, and Malayalam (Chakravarthi and Muralidaran, 2021). In descending order, the best performing hope speech detection models based on weighted average F_1-score were for English at 0.93, Malayalam at 0.85 (Hossain et al., 2021), and Tamil at 0.61 (Sharma and Arora, 2021). Unsurprisingly, the best performing hope speech detection models utilized variants of the Bidirectional Encoder Representations from Transformers (BERT) model architecture (Devlin et al., 2019). This contrasts earlier detection systems which relied on shallow classifiers (Chakravarthi, 2022).

Successive shared tasks focusing on hope speech detection would expand to other language conditions. The second shared task, also with YouTube comments, included Kannada and Spanish (Chakravarthi et al., 2022a). Meanwhile, the

Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pages 169–175

third shared task focused on three Indo-European languages: Bulgarian, English, Hindi, and Spanish (Kumaresan et al., 2023). Using data from X, formerly known as Twitter, the best performing model for English achieved a weighted average F_1-score of 0.50 using shallow classifiers (Kumari et al. 2023; Thandavamurthi et al. 2023). Related shared tasks, such as Hope 2024 (García-Baena et al., 2024) and PolyHope (Butt et al., 2025), have explored hope speech detection in English and Spanish across binary and multi-class classification conditions. Some common challenges observed across the systems include class imbalance, code-mixing and script-switching, and sociolinguistic differences (Butt et al., 2025).

It is clear from existing literature that hope speech detection is treated as a text classification problem (Kowsari et al., 2019). With a focus on optimizing model architectures, Sharma et al. (2025) found that an ensemble approach offered the best model performance by combining the architectures of Long Short-Term Memory (LSTM; Hochreiter and Schmidhuber 1997), mBERT (Devlin et al., 2019), and XLM-RoBERTa (Conneau et al., 2020) achieving an impressive weighted average F_1-score of 0.93 for English. However, solely focusing on model architectures fails to account for linguistic contexts beyond lexical structures (Li, 2022). We are now only observing approaches that account for the linguistic context, such as translation-based techniques (Ahmad et al., 2025), but research within this area remain limited.

3 Methodology

With approximately 2.5 million speakers, Tulu is an under-resourced language (Shetty, 2024). In contrast to other Dravidian languages, such as Kannada, Malayalam, and Tamil, resources for natural language processing remain scarce. Therefore, the primary purpose of *Hope Speech Detection in Code-Mixed Tulu Language* shared task was to develop multi-level hope speech classification systems in code-mixed Tulu. We treat this shared task as a classification problem (Kowsari et al., 2019). We used XLM-RoBERTa as the base pre-trained language model (PLM) for our system (Conneau et al., 2020). To summarize, XLM-RoBERTa embeddings were trained on 2.5 TB of filtered web-crawled data containing 100 languages. However, it is important to note that Tulu was not one of these languages. Encoder-decoder transformer-based PLMs were used because of their ability to support domain adaptation, meaning that we can continue training a pretrained language model on additional in-domain data without the need to train a new PLM from scratch.

Table 1: Coarse-Grained Training Class Labels

Class	Count
Blended Tone	895
Discouraging	711
Encouraging	1895
Uninvolved	2490

Table 2: Fine-Grained Training Class Labels

Class	Count
Fading Hope	236
Hopelessness	937
Inspiring Hope	1129
Optimistic Hope	380
Realistic Hope	503

The shared task was broken down into two subtasks which we refer to as Task 1 and Task 2 (Thenmozhi et al., 2026). For Task 1, the goal was to classify each comment into one of four coarse-grained categories: *encouraging*, *discouraging*, *uninvolved*, and *blended tone*. For Task 2, the goal was to classify hope-related content into five fine-grained categories: *optimistic hope*, *realistic hope*, *inspiring hope*, *fading hope*, and *hopelessness*. For each task, we first developed a baseline system and then developed organically adapted candidate systems by further training XLM-RoBERTa on organically collected Tulu social media text containing code-mixed and mixed-script variation before downstream fine-tuning on the labeled shared-task data. For submission, we selected the best performing candidate system in each task based on development set macro average F_1-scores.

3.1 Training Data

The organizers of the *Hope Speech Detection in Code-Mixed Tulu* shared task provided labeled training, development data, and a test set consisting of social media comments in code-mixed Tulu (Thenmozhi et al., 2026). The training data was split into coarse-grained hope tone classification (Task 1) and fine-grained hope type classification (Task 2). These datasets formed the basis of our baseline system development and evaluation. The labeled training data was used to train the classification models, the development set was used for validation and model comparison, and the official test set was used only for final shared-task submission and evaluation.

The distribution of target class labels in the training data is shown in Table 1 for the coarse-grained task and Table 2 for the fine-grained task. There is noticeable class imbalance across labels in both tracks. For the coarse-grained task, *uninvolved* and *encouraging* account for most of the training examples, while *discouraging* and *blended tone* appear less frequently. For the fine-grained task, *inspiring hope* and *hopelessness* are the most common labels, while *fading hope*, *optimistic hope*, and *realistic hope* are less represented.

3.2 Domain Adaptation

The first stage in developing our system was domain adaptation, where we further adapted a PLM using organically collected Tulu social media text containing code-mixed and mixed-script variation before downstream supervised classification. Liu et al. (2022) showed that domain adaptation can improve the performance of transformer-based language models in downstream tasks. We incorporated additional organic Tulu social media text for domain adaptation. This corpus consisted of naturally occurring Tulu social media comments and was intended to better reflect real-world code-mixed and mixed-script usage.

Unlike the labeled shared-task comments, this organic data consisted of naturally occurring Tulu social media comments collected to better reflect the linguistic variation found in real-world usage, including mixed-script and code-mixed forms. Following the data development approach described in Wong and Durward (2024), the organic Tulu data was retrieved using a Tulu language identification pipeline trained with Wikipedia-based data, which was used to identify candidate Tulu social media text for continued language-model adaptation. We used this organically collected Tulu social media text from the *Global Corpus of Language Use* (Dunn, 2020), which contains code-mixed and mixed-script variation, as unlabeled in-domain data for further adapting XLM-RoBERTa before downstream classification.

We adapted XLM-RoBERTa using a masked language modeling objective and then trained an `AutoModelForMaskedLM` model on this corpus. We trained the masked language model for two epochs with a per-device batch size of 16 and a learning rate of 5×10^{-5}. We used the AdamW optimization setup provided in the training framework (Loshchilov and Hutter, 2019). After training, we saved the adapted checkpoint and tokenizer for use in downstream supervised fine-tuning.

3.3 Supervised Classification

As discussed in Section 3.2, we developed our classification systems using both the original XLM-RoBERTa checkpoint and the organically adapted checkpoint produced during the domain adaptation stage. We then fine-tuned these models on the labeled shared-task data for the two Tulu classification settings: Task 1 (coarse-grained hope tone classification); and, Task 2 (fine-grained hope type classification). All model training was conducted in a shared Google Colab environment. Note that the relevant code notebook and scripts are available on GitHub[1]. For each task, we trained a sequence classification model using the appropriate number of output labels. We used `AutoModelForSequenceClassification` with the XLM-RoBERTa tokenizer and then trained the classifier for up to four epochs with a learning rate of 2×10^{-5}, using the same fine-tuning configuration for both the baseline and organically adapted systems. Four candidate systems overall were produced: a baseline and an organically adapted model for both Task 1 and Task 2 each. The performance of these candidate systems on the development set is reported in Section 4. Based on the development set macro average F_1-scores, we selected the organically adapted models as our submitted systems for both Task 1 and Task 2.

4 Results

The macro average F_1-scores of our candidate models on the development set are presented in Table 3 and Table 4. Both organic models performed better than the baseline models on the development set. In Task 1, the best performing candidate model was the organic model which yielded a macro average F_1-score of 0.5238 compared to 0.5227 for the baseline model. In Task 2, the best performing candidate model was also the organic model which yielded a macro average F_1-score of 0.3416 compared to 0.3171 for the baseline model. Based

[1]https://github.com/landrewi/cantnlp-HopeSpeechDetection-DravidianLangTech2026

Table 3: Validation results for Task 1: Coarse-Grained Hope Speech Classification. The best performing model based on macro average F_1-score is in **bold**.

Metric	Baseline	Organic
Accuracy	0.6269	0.6869
Macro F_1	0.5227	**0.5238**
Weighted F_1	0.6364	0.6545

Table 4: Validation results for Task 2: Fine-Grained Hope Speech Classification. The best performing model based on macro average F_1-score is in **bold**.

Metric	Baseline	Organic
Accuracy	0.4135	0.5337
Macro F_1	0.3171	**0.3416**
Weighted F_1	0.3753	0.4732

on these model performance metrics, we therefore selected the organically adapted models as our submitted systems for both Task 1 and Task 2.

5 Discussion

Based on the evaluation metrics alone, our organically adapted systems performed modestly better than the baseline systems on the development set for both tasks. With reference to Table 3 and Table 4, we see that the organically adapted model improved model performance over the baseline model in both the coarse-grained and fine-grained conditions. In Task 1, the macro average F_1-score increased from 0.5227 to 0.5238, while in Task 2 the macro average F_1-score increased from 0.3171 to 0.3416, suggesting that further adapting XLM-RoBERTa on organically collected Tulu social media text containing code-mixed and mixed-script variation provided a measurable benefit for hope speech classification, particularly for the fine-grained task (Task 2).

While the development set results indicate consistent improvement, they were not uniform across metrics and tasks. In Task 1, the macro average F_1 improvement was small, but accuracy and weighted F_1 increased more clearly, possibly reflecting that the organically adapted model better captured the majority-class patterns while still struggling with minority-class distinctions. As for Task 2, the organically adapted model improved across all reported metrics, suggesting that exposing the model to organically collected in-domain text containing code-mixed and mixed-script variation may be specifically helpful when the label space was larger and the distinctions between classes were more subtle.

Table 5: Macro average F_1-score from official test evaluation and rank for both tasks.

Task	F_1-score	Rank
Coarse-Grained	0.50	5
Fine-Grained	0.33	7

When comparing our submitted systems with the official test evaluation results found in Table 5, our organically adapted systems achieved a macro average F_1-score of 0.50 for Task 1 and 0.33 for Task 2, ranking fifth and seventh overall, respectively. The best performing team in Task 1 was Team `CUET_Synthetica` which yielded a macro average F_1-score of 0.58, while we yielded a macro average F_1-score of 0.50 and ranked 5th overall. The best performing team in Task 2 was also Team `CUET_Synthetica` which yielded a macro average F_1-score of 0.42, while we yielded a macro average F_1-score of 0.33 and ranked 7th overall. Despite this mid-ranged performance, these results indicate that the approach of adapting multilingual pretrained models using organically collected unlabeled Tulu social media text for further language-model adaptation is viable in this low-resource code-mixed setting, but the complexity of code-mixed Tulu requires more robust data or architectural innovations.

6 Conclusion

While our submitted systems did not achieve the best results, our organically adapted approach produced improvements over the baseline models on the development set for both Task 1 and Task 2. These results suggest that further adapting XLM-RoBERTa on organically collected Tulu social media text containing code-mixed and mixed-script variation can improve hope speech detection in code-mixed Tulu. At the same time, the official test results indicate that this task remains challenging, especially under class imbalance and fine-grained label distinctions.

Limitations

One limitation of the current study is that we only evaluated one domain adaptation strategy based on organically collected Tulu social media text and did not compare it directly with a synthetic script-switching variant such as that explored in (Wong and Durward, 2024). This prior work explored both synthetic and organic script-switching for domain adaptation, whereas our current system focused only on further adapting XLM-RoBERTa using organically collected Tulu social media text containing code-mixed and mixed-script variation. As a result, we cannot determine whether the gains observed on the development set were specific to the organic adaptation pipeline or whether a synthetic approach would have produced similar improvements.

Another limitation was class imbalance in both tasks, as shown in Table 1 and Table 2. In Task 1, the training set was dominated by the *uninvolved* and *encouraging* classes, while *discouraging* and *blended tone* were underrepresented. In Task 2, *inspiring hope* and *hopelessness* appeared most frequently, while *fading hope*, *optimistic hope*, and *realistic hope* were underrepresented. We did not apply any explicit class-imbalance mitigation techniques such as class-weighted loss, focal loss, or over/under-sampling during training. Future work could test whether these strategies improve performance on minority classes, particularly for Task 2. This imbalance may reduce the model's ability to learn reliable distinctions for less frequent hope-related categories and may have contributed to the modest official test performance.

Finally, beyond the upstream and downstream impacts of bias in multilingual PLMs, we also recognize that incorporating externally collected digital language data for language-model adaptation may introduce additional biases not fully addressed in this paper, such as sampling bias, geographic bias, and variation tied to platform-specific language-use (Goldhahn et al., 2012; Dunn, 2020). These limitations suggest that future work should compare multiple adaptation strategies more directly and further examine the effects of data balance and external-data bias in code-mixed Tulu hope speech detection.

Ethical Considerations

Our system contributes to much needed research on under-resourced languages such as Tulu (Shetty, 2024). The results suggest that existing PLMs, such as XLM-RoBERTa (Conneau et al., 2020), are ill-equipped to meet the needs of Tulu language speakers. Therefore, the findings from our paper provides offers an insightful benchmark for Tulu within the area of hope speech detection. However, we need to think beyond development, but also applied to benefit Tulu speaking communities. Much like hate speech detection (Wong, 2024), we should also adopt a similar view to determine the social impact of hope speech detection research. Moreover, we should consider how systems developed can be deployed to detect hope speech in real-world social media contexts.

References

Muhammad Ahmad, Iqra Ameer, Wareesa Sharif, Sardar Usman, Muhammad Muzamil, Ameer Hamza, Muhammad Jalal, Ildar Batyrshin, and Grigori Sidorov. 2025. Multilingual hope speech detection from tweets using transfer learning models. *Scientific Reports*, 15(1):9005.

Sabur Butt, Fazlourrahman Balouchzahi, Maaz Amjad, Salud María Jiménez-Zafra, Hector G. Ceballos, and Grigori Sidorov. 2025. Overview of PolyHope at IberLEF 2025: Optimism, Expectation or Sarcasm? *Procesamiento del Lenguaje Natural*, 75(0):461–474.

Bharathi Raja Chakravarthi. 2022. Hope speech detection in YouTube comments. *Social Network Analysis and Mining*, 12(1):75.

Bharathi Raja Chakravarthi and Vigneshwaran Muralidaran. 2021. 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, Kyiv, Ukraine. Association for Computational Linguistics.

Bharathi Raja Chakravarthi, Vigneshwaran Muralidaran, Ruba Priyadharshini, Subalalitha Chinnaudayar Navaneethakrishnan, John Philip McCrae, Miguel Ángel García-Cumbreras, Salud María Jiménez-Zafra, Rafael Valencia-García, Prasanna Kumar Kumaresan, Rahul Ponnusamy, Daniel García-Baena, and José Antonio García-Díaz. 2022a. 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, Dublin, Ireland. Association for Computational Linguistics.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Thenmozhi Durairaj, John McCrae, Paul Buitelaar, Prasanna Kumaresan, and Rahul Ponnusamy. 2022b. Overview of The Shared Task on Homophobia and Transphobia Detection in Social Media Comments. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 369–377, Dublin, Ireland. Association for Computational Linguistics.

Anusha Chhabra and Dinesh Kumar Vishwakarma. 2023. A literature survey on multimodal and multilingual automatic hate speech identification. *Multimedia Systems*, 29(3):1203–1230.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised Cross-lingual Representation Learning at Scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451, Online. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of NAACL-HLT*, pages 4171–4186.

Jonathan Dunn. 2020. Mapping languages: the Corpus of Global Language Use. *Language Resources and Evaluation*, 54(4):999–1018. Https://doi.org/10.1007/s10579-020-09489-2.

Daniel García-Baena, Fazlourrahman Balouchzahi, Sabur Butt, Miguel Ángel García-Cumbreras, Atnafu Lambebo Tonja, José Antonio García-Díaz, Selen Bozkurt, Bharathi Raja Chakravarthi, Hector G. Ceballos, Rafael Valencia-García, Grigori Sidorov, L. Alfonso Ureña-López, Alexander Gelbukh, and Salud María Jiménez-Zafra. 2024. Overview of HOPE at IberLEF 2024: Approaching Hope Speech Detection in Social Media from Two Perspectives, for Equality, Diversity and Inclusion and as Expectations. *Procesamiento del Lenguaje Natural*, 73(0):407–419.

Sayani Ghosal, Devendra Kumar Tayal, and Amita Jain. 2026. Hope WVBL: Improving Automatic Hope Speech Detection Using Word Embedding and Bidirectional LSTM. In *Proceedings of International Conference on Information Technology and Intelligence*, pages 131–142, Singapore. Springer Nature.

Dirk Goldhahn, Thomas Eckart, and Uwe Quasthoff. 2012. Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 759–765, Istanbul, Turkey. European Language Resources Association (ELRA).

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780.

Eftekhar Hossain, Omar Sharif, and Mohammed Moshiul Hoque. 2021. NLP-CUET@LT-EDI-EACL2021: Multilingual Code-Mixed Hope Speech Detection using Cross-lingual Representation Learner. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 168–174, Kyiv. Association for Computational Linguistics.

Kamran Kowsari, Kiana Jafari Meimandi, Mojtaba Heidarysafa, Sanjana Mendu, Laura Barnes, and Donald Brown. 2019. Text Classification Algorithms: A Survey. *Information*, 10(4):150.

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.

Kirti Kumari, Shirish Shekhar Jha, Zarikunte Kunal Dayanand, and Praneesh Sharma. 2023. ML&AI_iiitranchi@LT-EDI-2023: Identification of Hope Speech of YouTube comments in Mixed Languages. In *Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 214–222, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Xianbo Li. 2022. Text Classification of Conversational Implicatures Based on Lexical Features. *Applied Artificial Intelligence*, 36(1):2127598.

Xiaofeng Liu, Chaehwa Yoo, Fangxu Xing, Hyejin Oh, Georges El Fakhri, Je-Won Kang, and Jonghye Woo. 2022. Deep Unsupervised Domain Adaptation: A Review of Recent Advances and Perspectives. *APSIPA Transactions on Signal and Information Processing*, 11(1).

Ilya Loshchilov and Frank Hutter. 2019. Decoupled Weight Decay Regularization. In *Proceedings in the Seventh International Conference on Learning Representations*, New Orleans, LA.

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 Trans. Asian Low-Resour. Lang. Inf. Process.*

Megha Sharma and Gaurav Arora. 2021. Spartans@LT-EDI-EACL2021: Inclusive Speech Detection using Pretrained Language Models. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 188–192, Kyiv. Association for Computational Linguistics.

Poorvi Shetty. 2024. Natural Language Processing for Tulu: Challenges, Review and Future Scope. In *Speech and Language Technologies for Low-Resource Languages*, pages 93–109, Cham. Springer Nature Switzerland.

Priyadharshini Thandavamurthi, Samyuktaa Sivakumar, Shwetha Sureshnathan, Thenmozhi D., Bharathi B, and Gayathri Gl. 2023. TERCET@LT-EDI-2023: Hope Speech Detection for Equality, Diversity, and Inclusion. In *Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 257–261, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Durairaj Thenmozhi, Rathnakar Shetty P, Parameshwar R. Hegde, Anusha M D, 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*, San Diego, CA. Association for Computational Linguistics.

Sidney Wong and Matthew Durward. 2024. cantnlp@LT-EDI-2024: Automatic Detection of Anti-LGBTQ+ Hate Speech in Under-resourced Languages. In *Proceedings of the Fourth Workshop on Language Technology for Equality, Diversity, Inclusion*, pages 177–183, St. Julian's, Malta. Association for Computational Linguistics.

Sidney Gig-Jan Wong. 2024. What is the social benefit of hate speech detection research? A Systematic Review. In *Proceedings of the Third Workshop on NLP for Positive Impact*, pages 1–12, Miami, Florida, USA. Association for Computational Linguistics.