Findings of the Shared Task on Emotion Analysis in Tamil

Anbukkarasi Sampath [1], **Durairaj Thenmozhi** [2], **Bharathi Raja Chakravarthi,**[3]
Ruba Priyadharshini[4]**, Subalalitha Chinnaudayar Navaneethakrishnan**[5]**,**
Kogilavani Shanmugavadivel[1]**, Sajeetha Thavareesan**[6]**, Sathiyaraj Thangasamy**[7]**,**
Parameswari Krishnamurthy[8]**, Adeep Hande**[9] **,Sean Benhur**[10]**,**
Kishore Kumar Ponnusamy[11]**, Santhiya Pandiyan**[1]

[1] Kongu Engineering College, India, [2] SSN College of Engineering, India
[3]Insight SFI Research Centre for Data Analytics,National University of Ireland Galway, Ireland
[4]Madurai Kamaraj University, India, [5]SRM Institute Of Science And Technology, Tamil Nadu,India
[6]Eastern University, Sri Lanka,[7]Sri Krishna Adithya College of Arts and Science, India
[8] University of Hyderabad, India
[9] Indian Institute of Information Technology Tiruchirappalli, India
[10] PSG College of Arts and Science, India, [11] Guru Nanak College, India.
anbu.1318@gmail.com

Abstract

This paper presents the overview of the shared task on emotional analysis in Tamil at DravidianLangTech-ACL 2022. This overview paper presents the dataset used in the shared task, task description, the methodologies used by the participants and the evaluation results of the submissions. Emotion analysis in Tamil shared task consists of two sub tasks. Task A aims to categorize the social media comments in Tamil to 11 emotions and Task B aims to categorize the comments into 31 fine-grained emotions. For conducting experiments, training and development datasets were provided to the participants and results are evaluated for the unseen data. In total, we have received around 24 submissions from 13 teams. For evaluating the models, Precision, Recall, micro average metrics are used.

1 Introduction

Emotional analysis is the process of mining emotions in texts for categorization, which is used in a variety of natural language applications such as e-commerce review analysis, public opinion analysis, comprehensive search, customized suggestion, health-care, and online teaching. Emotions are mental states that are frequently represented through behavior or words. Emotional analysis assists in analyzing the text and extracting the emotions portrayed in the text. There are a variety of monolingual datasets available for Dravidian languages, which may be used for a variety of research purposes (Priyadharshini et al., 2021;

Kumaresan et al., 2021; Chakravarthi and Muralidaran, 2021; Chakravarthi et al., 2020b). There have, however, been few attempts to create code-mixed datasets for Tamil, Kannada, and Malayalam (Chakravarthi et al., 2021a,b, 2020a). We will assist Tamil in overcoming this resource barrier by producing this dataset, which will provide Tamil with cost-effective and quick natural language processing support in emotional analysis. We gathered comments on several Tamil movie trailers and teasers from YouTube to develop resources for a Tamil scenario.

Data on social media platforms such as Twitter, Facebook, and YouTube is rapidly changing and may have a significant impact on an individual's or community's perception or reputation (Ravikiran et al., 2022; Chakravarthi et al., 2022; Bharathi et al., 2022; Priyadharshini et al., 2022). This highlights the need of automation in emotional analysis. Because of its numerous content offerings, such as movies, trailers, music, tutorials, product reviews, and so on, YouTube is a popular social media network in the Indian subcontinent (Priyadharshini et al., 2020; Chakravarthi, 2020). YouTube allows users to create content as well as participate and interact with it through activities such as like and commenting. As a result, more user-generated content is available in underdeveloped languages (Chakravarthi et al., 2019, 2018; Yasaswini et al., 2021).

In this paper the results of the shared task on emotional analysis in Tamil are presented and held at DravidianLangTech. This task consists of two

sub tasks, namely Task A and Task B. The first task is focused on analysing the emotion in social media Tamil comments with 11 emotions. The second task is focused on analysing 31 fine-grained emotions in social media Tamil comments. The task is focused on analysing emotions on Tamil language which is a low resource language that belongs to Dravidian family. All the data including training data, testing data, development data along with results are made publicly available. We strongly believe that the curated dataset will contribute more in advancing the research in the field of emotional analysis field of Tamil.

2 Related Work

As the availability of text data is more today, there are many good works carried out in the field of Natural Language Processing (NLP). Identifying the emotions in a given text data helps in various applications, including online teaching, recommendation systems, public opinion analysis. Even though many researchers contribute in emotion analysis domain, very few works have been made in Dravidian languages like Tamil. In recent days, deep learning based techniques yield good results in this domain. With social media content, emotion detection from code mixed Hindi and English language (Sasidhar et al., 2020) is performed. They created an emotion annotated corpus using 12000 code-mixed text sentences from different social media sites as the data availability in their field of interest is less. It is reported that the CNN-BiLSTM model provides good accuracy. Xu et al. (2020) performed emotion analysis using the proposed CNN_Text_Word2vec model in Chinese microblog. All the words in the microblog are trained using word2vec neural network model. The trained feature vectors are fed as input feature for attention-based convolution neural network model which classifies the text emotion into positive and negative. Sina Weibo microblog is used for experimental analysis in which 40000 training and 40000 test microblog samples were considered.

A combination of cognitive linguistic and deep learning features are used for performing emotion analysis on Persian text (Sadeghi et al., 2021). These models help to identify emotions including anger, happiness, sadness, surprise and fear. For performance measurement, a self-labeled 23,000 Persian documents are taken. They used Word2Vec embedding technique for vector conversion. The

deep learning based architectures (Tocoglu et al., 2019) such as Artificial Neural Network, Convolutional Neural Network, and Long Short-Term Memory based Recurrent Neural Network are examined for Turkish tweets emotional prediction (S.Anbukkarasi and S.Varadhaganapathy, 2020). To curate a dataset, the lexicon-based approach is proposed for automatic annotation of Turkish tweet text and compared the deep learning based architectures that outperformed the traditional machine learning techniques for emotion recognition in Turkish. Hybrid model is proposed to identify the emotions in Arabic text (Alswaidan and Menai, 2020). Human-Engineered feature-based model and deep feature-based model are investigated in which Human-Engineered Feature-based model is used to select the features based on various kinds of the text like lexical, stylistics, and semantic. Deep Feature based model contains stacked neural network that reinserts the embedding layer several times to delay the learning process. From the study, it is clear that there is a research gap in emotional analysis field for Tamil language with advanced techniques.

3 Tamil

Tamil is a classical Dravidian language spoken by Tamil people of South Asia. It is the official language of the Indian state, Tamil Nadu, Singapore and Sri lanka. Tamil is one of the longest surviving classical languages and it was the first to be listed as classical language of India. It is one of the 22 scheduled languages in Indian constitution. The history of Tamil language is categorized into Old Tamil, Middle Tamil and Modern Tamil. Tamil has 247 alphabets including 12 vowels, 18 consonants, 216 compound letters and one special character known as ayudham (S.Anbukkarasi and S.Varadhaganapathy, 2020; Chakravarthi et al., 2018; Ghanghor et al., 2021a,b)

4 Task Description

4.1 Task

Two sub tasks are carried out as part of this shared task: Emotional analysis with 11 emotions annotated data for social media comments in Tamil and Task B were organized with 31 fine-grained emotions annotated data for social media comments in Tamil. The dates followed for releasing the training and testing data is provided in Table 1. Participants are encouraged to participate in any one of

the tasks or both the tasks. For this task, training, testing development datasets with the comments and corresponding emotion label is given to all the participants. The task was to classify the given comment with corresponding emotion.

4.2 Dataset

The data has been collected from various comments of Youtube videos. The data for Task A consists of around 22,200 Tamil Youtube comments with 11 emotions and Task B (Vasantharajan et al., 2022) consists of around 46,000 YouTube comments with 31 emotions. The YouTube Comment Scraper tool is used to collect Tamil comments from various domains such as sports, news and movies, and most of the sentences contain the English text, Tamil text and code-mixed Tamil-English text. For identifying the Tamil language, the language detection library known as langdetect [1] is used to find the Tamil comments from the sentences which are written fully in Tamil and discarded the other language comments. For annotation, the guidelines are given in both English and Tamil and each sentence in the dataset is annotated by minimum three annotators. The personal information of the annotators are gathered to understand about them and informed them the reason for collecting the information. Those who accepted to annotate are given with the instructions and data. They were given the freedom of relieving from the task at any point of time. This process was done for annotating data for both Task A and Task B.

4.3 System Description

We received around 16 submissions for Task A and 8 submissions for Task B, with total of 8 teams for Task A and 7 teams for Task B. For the teams that have submitted multiple submissions, we have considered the highest scoring submission. The system descriptions of each teams are given in this section.

1. **MUCS** (Hegde et al., 2022)- This team participated in Task A and used an ensemble of logistic regression models with three penalties, L1, L2 and Elasticnet. The team ranked 4th with macro average F1 score of 0.04.

2. **Judith Jeyafreeda** (Andrew Jayafreeda, 2022) - has used pretrained word embeddings

with CNN model for classifying emotions, the system achieved 0.094 Macro F1 for Task A and 0.057 for Task B and achieved 6th place in Task B.

3. **pandas** - This team has participated in Task A and have used LaBSE (Feng et al., 2020) feature extraction method with SVM classifier. To handle class imbalance problem, they have oversampled the dataset with SMOTE.

4. **CUET-NLP** (Mustakim et al., 2022) - The team has participated in Task A, and has experimented with multiple classical Machine Learning models such as logistic regression,naive bayes, decision tree, SVM and Pretrained multilingual transformer models, mBERT(Devlin et al., 2019) and XLM-R (Lample and Conneau, 2019) base model, The XLM-R model achieved the highest Macro F1 score of 0.210 and were ranked 2nd in the Task A.

5. **Varsini And Kirthanna** (S et al., 2022) - Participated in Task A and have used a combination of Keyword spotting and Lexical affinity based methods, using emojis and external datasets. The system ranked 5th place with Macro F1 score of 0.030.

6. **GJG** (Prasad et al., 2022) - Used pretrained transformers XLM-R and Deberta. They participated in both of the Tasks A and B, in Task B they achieved the Macro F1 score of 0.45 in Task A and 0.26 in Task B. The participants achieved Rank 1st in Task A.

7. **Optimize_Prime** (Gokhale et al., 2022) - Have experimented with multiple methods, an ensemble of class machine learning models, RNN based models such as LSTM and ULMFIT (Howard and Ruder, 2018), and pretrained transformer models such as XLM-R, Indic-BERT and MuRIL. They have participated in both subtasks, In Task A XLM-R achieved best result of 0.030 Macro F1 ranking 5th place and in Task B, MuRIL achieved best result of 0.125 ranking second place.

8. **UMUTeam** (García-Díaz and Valencia-García, 2022) - Have participated in both Tasks A and B, used a Model that combines both Linguistic Features such as psychological features, POS tags and contextual em-

---

[1] https://github.com/Mimino666/langdetect

Event	Date
Training Set Release	Nov 20, 2021
Test Set Release	Jan 14, 2022
Submission Deadline	Jan 30, 2022
Results Announcement	Feb 10, 2022
Paper Submission	March 10, 2022

Table 1: Emotional Analysis Task Schedule

beddings from pretrained models and Fasttext. They have achived Rank 1st in Task B with the score of 0.151.

9. **IIITSurat** - Have participated in both Task A and B. They have used an ensemble of deep learning models with oversampling the minority classes. They were able to rank 4 with Macro F1 score of 0.090 in Task B and rank 7th with Macro F1 score of 0.020 in Task A.

10. **MSD** - The team has participated in the Task A and have submitted submissions on Support Vector Machines, BiLSTM and an ensembles. Out of the three submissions, BiLSTM provided higher results than other models.
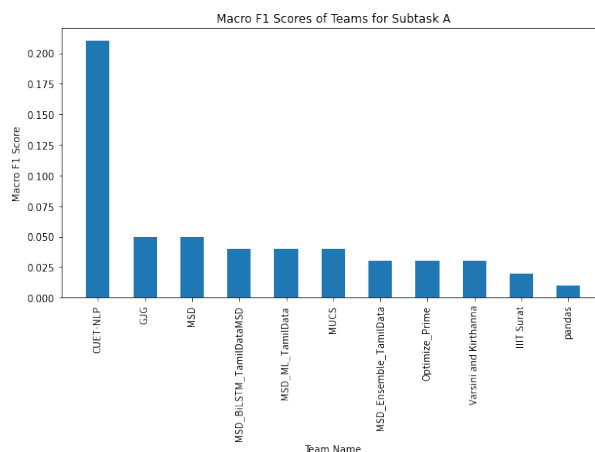
5 Evaluation

In this section we describe about the evaluation method that were used for the both Tasks A and B. We primarily evaluated the submissions using major classification metrics such as Macro Averaged and Weighted Average Precision, Recall and F1-Score. For ranking the teams, we primarily used Macro Averaged F1 Score as it finds f1 score for each label and find their unweighted mean.

6 Results and Discussion

Fine-grained Emotion Detection is a hard problem, it is even harder for low resource languages such as Tamil, In this shared task teams have submitted results for both Tasks A and B. Figures 1 and 2 depicts the pictorial representation of scores of each team for both Task A and Task B. The results are discussed in Table 2 and Table 3.

In Task A, Team GJG has achieved Rank 1 with Macro F1 Score of 0.310 and UMUTeam has achieved Rank 1 with 0.151 F1-score in Task B. Task A received more number of submissions compared to Task B, we hypothesize that this is due to the fact that Task B is relatively harder than Task A since it contains 31 emotions which are more
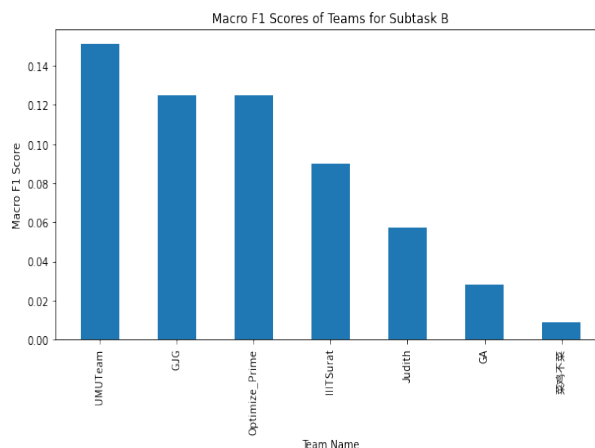

Figure 1: Results for Task A

Figure 2: Results for Task B

Team	Precision	Recall	F1-Score	Rank
GJG	0.320	0.310	0.310	1
CUET-NLP	0.220	0.250	0.210	2
MSD	0.090	0.080	0.050	3
MUCS	0.110	0.130	0.040	4
Optimize_Prime	0.090	0.080	0.030	5
Varsini and Kirthanna	0.090	0.110	0.030	5
IIITSurat	0.090	0.090	0.020	7
pandas_tamil	0.080	0.070	0.010	8

Table 2: Results of the Submissions for Task A

Team	Precision	Recall	F1-Score	Rank
UMUTeam	0.150	0.171	0.151	1
GJG	0.142	0.144	0.125	2
Optimize_Prime	0.132	0.140	0.125	2
IIITSurat	0.156	0.099	0.090	4
Judith Jeyafreeda	0.094	0.068	0.057	5
GA	0.033	0.031	0.028	6
菜鸡不菜	0.005	0.032	0.009	7

Table 3: Results of the Submissions for Task B

similar in nature and harder for Machine Learning models to classify.

7 Conclusion and Future Work

Two new datasets with fine-grained emotions for classifying the emotions in Tamil language is created for this shared task. In this paper, various approaches used for analysing emotions from Tamil comments and their results are analyzed. The task has been carried out as two sub tasks. In the Task A, 8 teams have participated and submitted their results. For Task B, 7 teams submitted their results. The XLM-R model produces the highest F1-score for Task A and the combination of techniques including linguistic features, POS tags, contextual embeddings yields remarkable results for Task B. As a future work, we planned to increase the size of the dataset and include more fine grained emotional labels to increase the performance of the system.

Acknowdgements

(CARDAMOM-Comparative Deep Models of Language for Minority and Historical Languages).

References

Nourah Alswaidan and Mohamed El Bachir Menai. 2020. Hybrid feature model for emotion recognition in Arabic text. *IEEE Access*, 8:37843–37854.

Judith Andrew Jayafreeda. 2022. JudithJeyafreedaAndrew@TamilNLP-ACL2022:CNN for Emotion Analysis in Tamil. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

B Bharathi, Bharathi Raja Chakravarthi, Subalalitha Chinnaudayar Navaneethakrishnan, N Sripriya, Arunaggiri Pandian, and Swetha Valli. 2022. Findings of the shared task on Speech Recognition for Vulnerable Individuals in Tamil. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*. Association for Computational Linguistics.

Bharathi Raja Chakravarthi. 2020. 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, Barcelona, Spain (Online). Association for Computational Linguistics.

Bharathi Raja Chakravarthi, Mihael Arcan, and John P. McCrae. 2018. Improving wordnets for under-

resourced languages using machine translation. In *Proceedings of the 9th Global Wordnet Conference*, pages 77–86, Nanyang Technological University (NTU), Singapore. Global Wordnet Association.

Bharathi Raja Chakravarthi, Mihael Arcan, and John P. McCrae. 2019. WordNet gloss translation for under-resourced languages using multilingual neural machine translation. In *Proceedings of the Second Workshop on Multilingualism at the Intersection of Knowledge Bases and Machine Translation*, pages 1–7, Dublin, Ireland. European Association for Machine Translation.

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. Association for Computational Linguistics.

Bharathi Raja Chakravarthi, Vigneshwaran Muralidaran, Ruba Priyadharshini, and John Philip McCrae. 2020a. Corpus creation for sentiment analysis in code-mixed Tamil-English text. In *Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL)*, pages 202–210, Marseille, France. European Language Resources association.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Thenmozhi Durairaj, John Phillip McCrae, Paul Buitaleer, Prasanna Kumar Kumaresan, and Rahul Ponnusamy. 2022. Findings of the shared task on Homophobia Transphobia Detection in Social Media Comments. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*. Association for Computational Linguistics.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Vigneshwaran Muralidaran, Shardul Suryawanshi, Navya Jose, Elizabeth Sherly, and John P McCrae. 2020b. Overview of the track on sentiment analysis for Dravidian languages in code-mixed text. In *Forum for Information Retrieval Evaluation*, pages 21–24.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Rahul Ponnusamy, Prasanna Kumar Kumaresan, Kayalvizhi Sampath, Durairaj Thenmozhi, Sathiyaraj Thangasamy, Rajendran Nallathambi, and John Phillip McCrae. 2021a. Dataset for identification of homophobia and transophobia in multilingual YouTube comments. *arXiv preprint arXiv:2109.00227*.

Bharathi Raja Chakravarthi, Priya Rani, Mihael Arcan, and John P McCrae. 2021b. A survey of orthographic information in machine translation. *SN Computer Science*, 2(4):1–19.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. *ArXiv*, abs/1810.04805.

Fangxiaoyu Feng, Yinfei Yang, Daniel Cer, Naveen Arivazhagan, and Wei Wang. 2020. Language-agnostic BERT sentence embedding.

José Antonio García-Díaz and Rafael Valencia-García. 2022. UMUTeam@TamilNLP-ACL2022: Emotional Analysis in Tamil. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Nikhil Ghanghor, Parameswari Krishnamurthy, Sajeetha Thavareesan, Ruba Priyadharshini, and Bharathi Raja Chakravarthi. 2021a. IIITK@DravidianLangTech-EACL2021: Offensive language identification and meme classification in Tamil, Malayalam and Kannada. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 222–229, Kyiv. Association for Computational Linguistics.

Nikhil Ghanghor, Rahul Ponnusamy, Prasanna Kumar Kumaresan, Ruba Priyadharshini, Sajeetha Thavareesan, and Bharathi Raja Chakravarthi. 2021b. IIITK@LT-EDI-EACL2021: Hope speech detection for equality, diversity, and inclusion in Tamil, Malayalam and English. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 197–203, Kyiv. Association for Computational Linguistics.

Omar Gokhale, Shantanu Patankar, Onkar Litake, Aditya Mandke, and Dipali Kadam. 2022. Optimize_Prime@TamilNLP-ACL2022: Emotion Analysis in Tamil. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Asha Hegde, Sharat Coelho, and Shashrekha H.L. 2022. MUCS@DravidianLangTech@ACL2022: ensemble of logistic regression penalties to identify Emotions in Tamil Text. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Jeremy Howard and Sebastian Ruder. 2018. Universal language model fine-tuning for text classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 328–339, Melbourne, Australia. Association for Computational Linguistics.

Prasanna Kumar Kumaresan, Ratnasingam Sakuntharaj, Sajeetha Thavareesan, Subalalitha Navaneethakrishnan, Anand Kumar Madasamy, Bharathi Raja Chakravarthi, and John P McCrae. 2021. Findings of shared task on offensive language identification in Tamil and Malayalam. In *Forum for Information Retrieval Evaluation*, pages 16–18.

Guillaume Lample and Alexis Conneau. 2019. Cross-lingual language model pretraining. *Advances in Neural Information Processing Systems (NeurIPS).*

Nasehatul Mustakim, Rabeya Rabu Aktar, Golam Sarwar Md Mursalin, Eftekhar Hossain, Omar Sheriff, and Mohammed Hoque Moshiui. 2022. CUET-NLP@TamilNLP-ACL2022 multi-class textual emotion detection from social media using transformers. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Janvi Prasad, Gaurang Prasad, and Gunavathi Chellamuthu. 2022. GJG@TamilNLP-ACL2022: emotion analysis and classification in Tamil using Transformers. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Ruba Priyadharshini, Bharathi Raja Chakravarthi, Subalalitha Chinnaudayar Navaneethakrishnan, Thenmozhi Durairaj, Malliga Subramanian, Kogilavani Shanmugavadivel, Siddhanth U Hegde, and Prasanna Kumar Kumaresan. 2022. Findings of the shared task on Abusive Comment Detection in Tamil. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Ruba Priyadharshini, Bharathi Raja Chakravarthi, Sajeetha Thavareesan, Dhivya Chinnappa, Durairaj Thenmozhi, and Rahul Ponnusamy. 2021. Overview of the DravidianCodeMix 2021 shared task on sentiment detection in Tamil, Malayalam, and Kannada. In *Forum for Information Retrieval Evaluation*, pages 4–6.

Ruba Priyadharshini, Bharathi Raja Chakravarthi, Mani Vegupatti, and John P McCrae. 2020. Named entity recognition for code-mixed Indian corpus using meta embedding. In *2020 6th international conference on advanced computing and communication systems (ICACCS)*, pages 68–72. IEEE.

Manikandan Ravikiran, Bharathi Raja Chakravarthi, Anand Kumar Madasamy, Sangeetha Sivanesan, Ratnavel Rajalakshmi, Sajeetha Thavareesan, Rahul Ponnusamy, and Shankar Mahadevan. 2022. Findings of the shared task on Offensive Span Identification in code-mixed Tamil-English comments. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Varsini S, Kirthana S Ranjan, Angel S Deborah, S Rajalakshmi, R.S Milton, and T.T Mirnalinae. 2022. Varsini_and_Kirthanna@DravidianLangTech-ACL2022-Emotional Analysis in Tamil. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Seyed Soheil Sadeghi, Hassan Khotanlou, and Mohammad Rasekh Mahand. 2021. Automatic persian text emotion detection using cognitive linguistic and deep learning. *Journal of AI and Data Mining*.

S.Anbukkarasi and S.Varadhaganapathy. 2020. Analyzing sentiment in Tamil tweets using deep neural network. In *2020 Fourth International Conference on Computing Methodologies and Communication*. IEEE.

T Tulasi Sasidhar, Premjith B, and Soman K P. 2020. Emotion detection in Hinglish(Hindi+English) code-mixed social media text. *Procedia Computer Science*, 171:1346–1352. Third International Conference on Computing and Network Communications (CoCoNet'19).

Mansur Alp Tocoglu, Okan Ozturkmenoglu, and Adil Alpkocak. 2019. Emotion analysis from Turkish tweets using deep neural networks. *IEEE Access*, 7:183061–183069.

Charangan Vasantharajan, Sean Benhur, Prasanna Kumar Kumarasen, Rahul Ponnusamy, S. Thangasamy, Ruba Priyadharshini, Thenmozhi Durairaj, Kanchana Sivanraju, Anbukkarasi Sampath, Bharathi Raja Chakravarthi, and John P. McCrae. 2022. TamilEmo: finegrained emotion detection dataset for Tamil. *ArXiv*, abs/2202.04725.

Dongliang Xu, Zhihong Tian, Rufeng Lai, Xiangtao Kong, Zhiyuan Tan, and Wei Shi. 2020. Deep learning based emotion analysis of microblog texts. *Information Fusion*, 64:1–11.

Konthala Yasaswini, Karthik Puranik, Adeep Hande, Ruba Priyadharshini, Sajeetha Thavareesan, and Bharathi Raja Chakravarthi. 2021. IIITT@DravidianLangTech-EACL2021: Transfer learning for offensive language detection in Dravidian languages. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 187–194, Kyiv. Association for Computational Linguistics.