# TamilEcho_Political@DravidianLangTech 2026: Hybrid XLM-RoBERTa with Sarcasm-Aware Feature Fusion for Political Multiclass Sentiment Analysis in Tamil X


**Kanimozhi Selvi C S[1], Inigashree N S[1], Moneissh A G[1], Kavinraj J[1]**

[1]Department of Computer Science and Engineering
Kongu Engineering College, Perundurai, India
kanimozhi@kongu.ac.in
inigashreens.22cse@kongu.edu
moneisshag.22cse@kongu.edu
kavinrajj.22cse@kongu.edu



## Abstract

Political sentiment analysis of Tamil tweets is challenging due to the informal use of language, sarcasm, emoji-driven sentiment shifts, and a high degree of class imbalance. This paper introduces our system for the Shared Task on Political Multiclass Sentiment Analysis of Tamil X (Twitter) Comments at DravidianLangTech@ACL 2026. We employ a hybrid design comprising contextual representations from XLM-RoBERTa in combination with lexical TF-IDF features and sarcasm-aware emoji indicators. We also use domain-aware hashtag expansion to provide additional political context and apply class-weighted label smoothing to address data imbalance. On the official test set, our system achieves a Macro-F1 score of 0.3559, ranking 10th among participating teams. The findings indicate that combining semantic, lexical, and pragmatic features enhances fine-grained classification of political sentiments in Tamil.



**Keywords:** Political Comments, Tamil Tweets, Natural Language Processing, XLM-RoBERTa, TF-IDF, Emoji Features, Sentiment Analysis.


## 1 Introduction

Social media platforms such as X (Twitter) have taken center stage in political discourse, where users are able to share their views and communicate with political actors in real time. Voter attitudes and campaign plans have been analyzed using automated political sentiment analysis (Tumasjan et al., 2010). The effectiveness of social media analysis in terms of modeling political discourse patterns has also been proven with the help of computational methods (Stieglitz and Dang-Xuan, 2012).

Tamil is a Dravidian language which is mainly spoken in Tamil Nadu, India and some in Sri Lanka, Singapore and Malaysia. Despite the abundant literary tradition Tamil is relatively under-reserved in terms of annotated corpora and pre-trained language models. Tweets about political activities usually have spelling differences, code-mixing in English, sarcasm, emojis, and colloquialism that greatly complicate classification. Emerging research on Tamil and other languages in India has emphasized the effect of code-mixing and informal text written in social media in sentiment classification effectiveness (Sreelakshmi et al., 2024). These properties necessitate models that can be used to represent contextual semantics as well as pragmatic signs.

The shared task on Political Multiclass Sentiment Analysis of Tamil X (Twitter) Comments was held toward this purpose as a part of DravidianLangTech@ACL 2026 (Vegupatti et al., 2026). The task includes the categorization of tweets into seven sentiment classes and the performance of systems is evaluated based on Macro-F1 because there is class imbalance. We suggest a hybrid model in this work that combines contextual embedding with transformers, lexical and emoji-based features to contextualize the Tamil political discourse.

## 2 Related Work

The study of political sentiment analysis has been a major research area in recent years especially in the identification and comprehension of the public opinion and biasness in social media. Prathvi et al. (Prathvi et al., 2024) reported that the Macro-F1 score of ensemble models, which were run using TF-IDF with n-grams, was 0.260 and 0.550, respectively, on code-mixed Tamil and Tulu sentiment analysis. The correctly extraction of lexical features methods in low-resource and code-mixed environment was proven. Kumar et al. (Kumar et al., 2017) developed a sentiment analysis BiLSTM-CNN model to analyze Malayalam tweets and reached an accuracy of 0.9824, which has made it possible to adopt deep neural archi-




*Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 387–391

tecture on Dravidian languages. Chakravarthi et al. (Chakravarthi et al., 2021) obtained a weighted F1 score of 0.711 in Tamil-English sentiment analysis, which indicates the problems of multilingual and informal political speech. Rajalakshmi et al. (Rajalakshmi et al., 2022) demonstrated that MuRIL through emoji based sentiment representations showed that explicit emoji model increases code-mixed Tamil text classification. Finally, Angdresey et al. (Angdresey et al., 2025) proposed a hybrid approach incorporating tagging models with BERT representations, random oversampling, and Multinomial Naïve Bayes, achieving an accuracy of 85.155 % and AUC score of 96.80 on political sentiment analysis of YouTube comments.

## 3 Dataset

The dataset given by the organizers of the shared task entails 5,440 Tamil political tweets, which were gathered in X (Twitter). All tweets are coded into one of seven categories of sentiments, namely Substantiated, Sarcastic, Opinionated, Positive, Negative, Neutral, and None of the Above. The data is divided into the train (4,352 tweets, 80%), development (544 tweets, 10%), and test (544 tweets, 10%) groups. The large percentage of the dataset is constituted by Opinionated and Sarcastic tweets, with the Substantiated and None of the Above classes being relatively low. The imbalance between classes in the training data is depicted in Figure 1. This tradeoff encourages the application of Macro-F1 to be the evaluation metric.

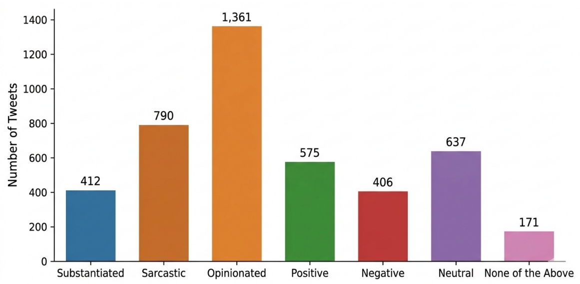


Figure 1: Class distribution in the training dataset.

## 4 Methodology

### 4.1 Pre-processing

Tamil political tweets are informal and they are also full of spelling variants, short forms, emojis, and politically charged hashtags. These attributes introduce noise and hidden contextual cues that may adversely influence downstream sentiment classification. So, a specific preprocessing approach is used to enhance the readability of the text and background knowledge. The general process of the preprocessing and feature engineering is shown in Figure 2.

**Hashtag Knowledge Expansion:**

Political hashtags can also have implicit contextual information, not explicitly found in the text of the tweet. An example would be the use of the hashtag #DMK, #BJP,#NTK and so on to refer to a particular political party or leader. In our method, we augment regularly reoccurring political hashtags with a manually defined dictionary that is written in code. As an example, the #DMK can be mapped to Dravida Munnetra Kazhagam ruling party Tamil Nadu. This expansion which is rule-based assists in including more semantic context and enhances contextual representation in transformer-based models.

**Identification of Sarcasm with the help of emojis:**

Emojis is important in conveying sarcasm, mockery, and emotional tone in the Tamil politics. Implicit sentiment cues are commonly done with emojis like laughing, angry, folded-hands emojis, and thinking emojis. In order to express this phenomenon explicitly, we derive two emoji-related features: the number of emojis in the tweet in total and the number of emojis denoting sarcasm. These features are pragmatic sentiment predictors and assist the classifier to differentiate among sarcastic and literal sentiment.

### 4.2 Feature Extraction

Once preprocessed, every tweet is coded into various numerical descriptions of complementary linguistic characteristics. We choose a hybrid method of representation combining transformer embeddings, TF-IDF features, and emoji-based features as shown in Figure 3.

**TF-IDF Representation of features:**

The importance of lexical meaning in comparison to the corpus is captured with the help of Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF puts more emphasis on words that occur more in a document than in the entire dataset, and therefore focus on discriminative political words. TF-IDF is applied to identify lexical significance. In the case of semantic representation, we use

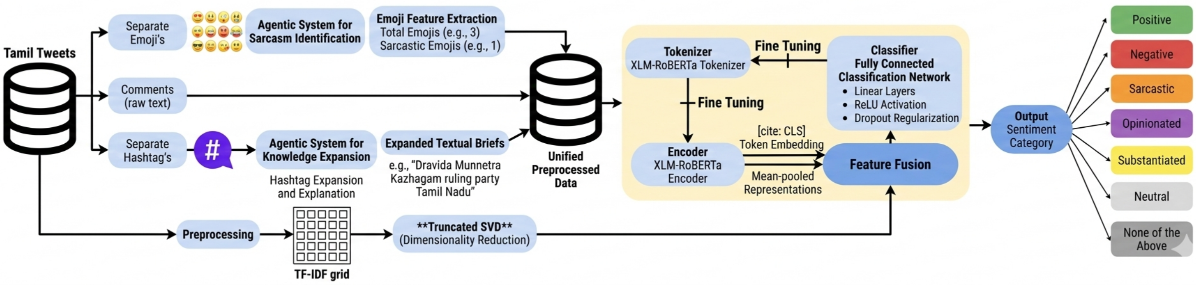


Figure 2: Proposed workflow for political sentiment analysis in Tamil X (Twitter) comments.

XLM-RoBERTa. TF-IDF of a term $t$ in document $d$ is calculated as:

$$TF{-}IDF(t,d) = TF(t,d) \times \log\left(\frac{N}{DF(t)}\right) \tag{1}$$

TF(t,d) is the frequency of a term in a document $d$, $N$ is the number of documents and DF(t) is the number of documents that contain term $t$. TF-IDF especially is useful in extracting domain-specific political expressions that the transformer models might not necessarily highlight.

**SVD as a form of dimensionality Reduction:** As TF-IDF vectors are both high-dimensional and sparse, Truncated Singular Value Decomposition (SVD) is used to trim the dimensionality without compromising the most informative ones. This step increases noise reduction, computation efficiency and generates a small dense form that can be utilized by neural feature fusion.

**Contextual Embeddings based on XLM-RoBERTa:** To achieve semantic presentation, we use XLM-RoBERTa, which is a multilingual transformer that has been trained on multilingual cross-linguistic corpora. Two complementary contextual representations are obtained, specifically that (i) token embedding [CLS], which is a representation of global sentence-level semantics, and (ii) mean-pooled representational embeddings of all tokens, which represent distributed contextual representations. These representations enable the model to exhibit holistic and fine-grained cues in the form of semantics in the tweet.

**Feature Design Justification:** We combine XLM-RoBERTa (semantic), TF-IDF (lexical), and emoji features (pragmatic) as they provide complementary information, which is more robust to noisy social media text.

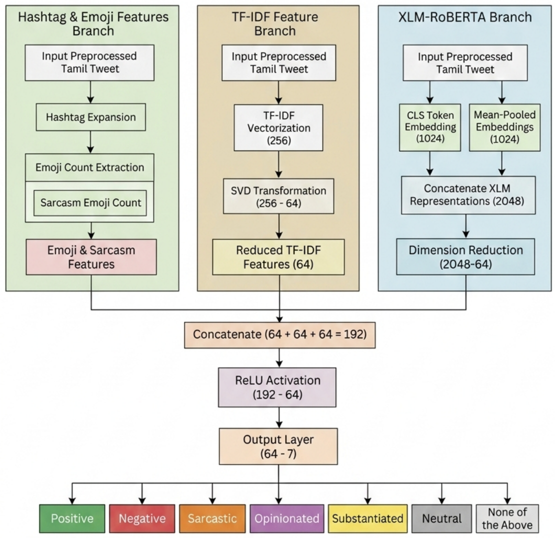


Figure 3: Architecture of the hybrid XLM-RoBERTa-based classification model.

## 5 Model and Training

We suggest a hybrid architecture to combine contextual embeddings in both XLM-RoBERTa and TF-IDF and emoji-based sarcasm. This architecture allows this design to obtain semantic information of transformer-based representations as well as the lexical significance and the pragmatic cues. The general system design is depicted in Figure 3. The system is comprised of a transformer-based encoder and statistical feature module. The preprocessed tweets are fed to XLM-RoBERTa to get the [CLS] embedding and mean-pooled token representations, which represents both global and contextual sentence data. These embeddings are added to the truncated TF-IDF representation and emojis-based features to create a feature representation. The merged vector is then passed through a completely connected classification network comprising of linear layers, ReLU activation, and dropout regularization. There is then a final softmax layer that estimates one of the seven possible categories

| Model | Macro-F1 |
|---|---|
| XLM-RoBERTa Only | 0.331 |
| XLM-RoBERTa + TF-IDF | 0.346 |
| XLM-RoBERTa + TF-IDF + Emoji | 0.356 |

Table 1: Performance of model configurations.

of sentiment: Positive, Negative, Sarcastic, Opinionated, Substantiated, Neutral and None of the Above. We use AdamW optimizer with the learning rate of $2 \times 10^{-5}$ to train. In order to deal with the imbalance in classes, weighted CrossEntropyLoss computed on inverse class frequencies is employed. Generalization is enhanced by the use of label smoothing. The methods of progressive fine-tuning are followed, i.e. encoder layers are unfreezed across epochs and gradient clipping is utilized to maintain stable optimization.

## 6 Results and Analysis

### 6.1 Macro Average F1-Score

The F1 macro average that we employ as needed in the shared task, is the harmonic mean of precision and recall of each class, and then averages all classes in an imbalanced dataset whereby each class is given equal importance.

$$F1_{macro} = \frac{1}{C} \sum \left( \frac{2 \times P_i \times R_i}{P_i + R_i} \right) \quad (2)$$

In which $C$ stands out as the amount of classes and where $P_i$ and $R_i$ represent precision and recall of class $i$ respectively.

### 6.2 Results

We experimented on three model configurations of our XLM-RoBERTa model. The performance comparison across different model configurations is presented in Table 1. Hybrid feature fusion was effective as the XLM-RoBERTa + TF-IDF + Emoji model did it best. The final model class-wise precision, recall and F1-score is listed in Table 2.

**Ablation Analysis:**

The addition of TF-IDF features boosts Macro-F1 from 0.331 to 0.346, and emoji features further improve Macro-F1 to 0.356.

The model works well on the class of none and less well on the alternative classes of negative and substantiated. In the shared task, our last system scored 0.3559 in Macro-F1, which ranked 10.

| Class | Precision | Recall | F1-score |
|---|---|---|---|
| Negative | 0.18 | 0.22 | 0.20 |
| Neutral | 0.26 | 0.31 | 0.28 |
| None | 0.91 | 0.88 | 0.89 |
| Opinionated | 0.45 | 0.39 | 0.42 |
| Positive | 0.29 | 0.33 | 0.31 |
| Sarcastic | 0.47 | 0.44 | 0.45 |
| Substantiated | 0.23 | 0.26 | 0.24 |

Table 2: Class-wise performance of the final model.

### 6.3 Error Analysis

The model performs well on "None" (F1 = 0.89) but struggles with "Negative" (0.20) and "Substantiated" (0.24). The latter involves factual reasoning that the model is not able to perform, whereas negative sentiment tends to cover the area of sarcasm. Also, sarcasm devoid of emojis and code-mixed text all add to errors. External knowledge and better modeling of sarcasm should be included in future work.

## 7 Conclusion

We introduce a hybrid approach for Tamil political sentiment analysis, achieving Rank 10 in the shared task. The strategy is a combination of XLM-RoBERTa, TF-IDF, and emoji-based characteristics to manage informal political speech. The model is effective in capturing both semantic and pragmatic cues to multiclass sentiment classification and this reflects the advantage of hybrid feature fusion. The code files of the project can be found on GitHub.[1]

## 8 Limitations

The dataset is small and imbalanced, limiting the generalizability and influencing the level of performance on minority classes like *Negative* and *Substantiated*. The ambiguities of the political rhetoric and the implicit group mentions also make it more difficult to detect sentiment in political rhetoric (van Dijk, 2002; Coulthard and Johnson, 2007). The model also struggles with implicit sarcasm, where emojis are not available. Although XLM-RoBERTa can offer very strong multilingual representations, other models like MuRIL that are specific to Tamil may be more effective in depicting linguistic nuances.

[1] https://github.com/Inigashree/PoliticalSentiment_DravidianLangTech_2026

## References


A. Angdresey, L. Sitanayah, and I. L. H. Tangka. 2025. Sentiment analysis for political debates on youtube comments using bert labeling, random oversampling, and multinomial naïve bayes. *Journal of Computing Theories and Applications*, 2(3):342–354.

Bharathi Raja Chakravarthi, R. Priyadharshini, S. Thavareesan, D. Chinnappa, D. Thenmozhi, E. Sherly, and C. Vasantharajan. 2021. Findings of the sentiment analysis of dravidian languages in code-mixed text. *arXiv preprint arXiv:2111.09811*.

Malcolm Coulthard and Alison Johnson. 2007. Forensic linguistic analysis of political communication. *Journal of Language and Politics*, 6(3):426–431.

S. Sachin Kumar, M. Anand Kumar, and K. P. Soman. 2017. Sentiment analysis of tweets in malayalam using long short-term memory units and convolutional neural nets. In *Proceedings of the 5th International Conference on Mining Intelligence and Knowledge Exploration (MIKE 2017)*, pages 320–334. Springer.

B. Prathvi, K. Manavi, K. Subrahmanyapoojary, Asha Hegde, G. Kavya, and Hosahalli Shashirekha. 2024. MUCS@dravidianlangtech-2024: A grid search approach to explore sentiment analysis in code-mixed tamil and tulu. In *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, pages 257–261, St. Julian's, Malta. Association for Computational Linguistics.

R. Rajalakshmi, S. Selvaraj, A. Shibani, and B. R. Chakravarthi. 2022. Understanding the role of emojis for emotion detection in tamil. In *Proceedings of the First Workshop on Multimodal Machine Learning in Low-resource Languages*, pages 9–17.

K. Sreelakshmi, B. Premjith, Bharathi Raja Chakravarthi, and K. P. Soman. 2024. Detection of hate speech and offensive language codemix text in dravidian languages using cost-sensitive learning approach. *IEEE Access*, 12:20064–20090.

Stefan Stieglitz and Linh Dang-Xuan. 2012. Political communication and influence through microblogging—an empirical analysis of sentiment in twitter messages. *Journal of e-Democracy*.

Andranik Tumasjan, Timm O. Sprenger, Philipp G. Sandner, and Isabell M. Welpe. 2010. Predicting elections with twitter: What 140 characters reveal about political sentiment. *Journal of Information Technology*.

Teun A. van Dijk. 2002. Political discourse and ideology. *Discourse & Society*, 13(2):203–221.

Mani Vegupatti, Kishore Kumar Ponnusamy, Bharathi Raja Chakravarthi, Saranya Rajiakodi, Durairaj Thenmozhi, Prasanna Kumar Kumaresan, and Sathiyaraj Thangasamy. 2026. TamilPoliSent 2026: A Shared Task report on Multiclass Political Sentiment Analysis in Tamil. In *Proceedings of the Sixth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.