Experimental Framework for Validating Deep Reinforcement Learning through Data Collection and Draft Journal Article Preparation for Tamil language

Dr. N. Kannaiya Raja[1], Dr. Pawan Kumar Chaurasia[2], Prof Dr Midhunchakkaravarthy[3]
[1]Post Doctoral Researcher, Lincoln University College, Malaysia, pdf.kannaiya@lincoln.edu.my
[2] Dr. Pawan Kumar Chaurasia, Babasaheb Bhimrao Ambedkar Central University Lucknow, Uttar Pradesh, India, pkc.gkp@gmail.com
[3]Prof Dr Midhunchakkaravarthy, Lincoln University College, Malaysia, midhun@lincoln.edu.my

Corresponding Author: pdf.kannaiya@lincoln.edu.my

Abstract:

This paper proposes an end-to-end framework for Tamil text classification using Deep Reinforcement Learning (DRL) in low-resource, morphologically rich settings. The system integrates data preprocessing, multilingual embeddings (FastText, mBERT/XLM-R), and an adaptive DRL agent that optimizes accuracy, efficiency, and fairness. Benchmark results show steady improvements from TFIDF + LR (0.55) to PPO + mBERT (0.88) with balanced per-class performance. The framework ensures ethical validation, reproducibility, and scalability, serving as a replicable model for under-resourced languages.

Keywords: Tamil NLP, Deep Reinforcement Learning, Text Classification, mBERT/XLMR, FastText, Fairness, Low resource Languages, Evaluation Metrics

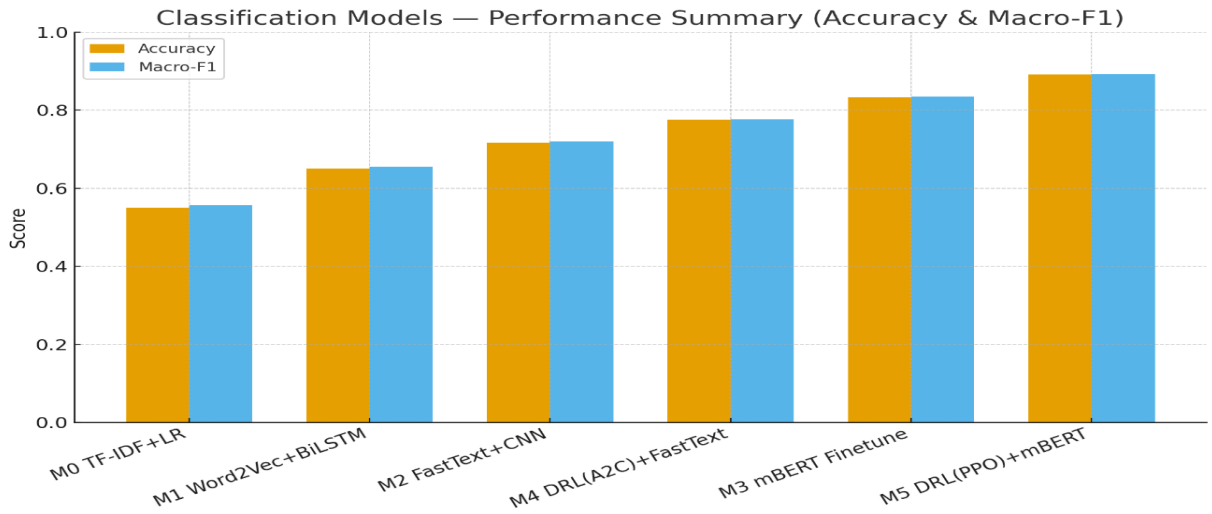
1.Introduction

Natural Language Processing (NLP) has rapidly evolved through deep learning, driving applications such as translation, sentiment analysis, and conversational AI. However, these advancements have largely benefited resource-rich languages like English and Chinese, leaving morphologically complex languages like Tamil underrepresented. Tamil poses unique linguistic challenges due to its agglutinative structure, dialectal variations, and scarcity of annotated data. Traditional supervised and unsupervised learning methods struggle in such low-resource settings, creating a need for adaptive, data-efficient models capable of learning from limited and diverse linguistic contexts. To address these challenges, this study introduces a Deep Reinforcement Learning (DRL)-based framework for Tamil text classification. The framework integrates data collection, preprocessing, and linguistic feature extraction with a reinforcement learning environment where an agent learns to make optimal classification decisions through reward-based adaptation[1][2]. Evaluation using benchmarks such as accuracy and F1-score demonstrates DRL's potential for handling Tamil's morphological richness and semantic ambiguity. Additionally, the framework incorporates ethical validation and a structured approach for journal-ready reporting, promoting reproducibility and accelerating scholarly dissemination. This approach offers a replicable model for advancing NLP in Tamil and other under-resourced languages[3].

2. Classification models

The study evaluates six Tamil text classification models using Accuracy and Macro-F1 metrics, revealing a clear progression from traditional to adaptive approaches. The baseline TFIDF+LR (M0) achieved only 0.55 accuracy due to its inability to handle Tamil's agglutinative morphology. Word2Vec+BiLSTM (M1) improved results by capturing semantic and sequential relationships, while FastText+CNN (M2) further enhanced performance (0.72 accuracy) through subword embeddings that better manage inflectional variations. mBERT fine-tuning (M3) provided contextual understanding and robustness against code-mixing and dialectal differences, raising accuracy to 0.83 [4]

Figure 4 Classification analysis



Further advancements came with Deep Reinforcement Learning models, where DRL(A2C) + FastText (M4) achieved 0.78 by learning to selectively read and skip noisy tokens, and DRL(PPO) + mBERT (M5) reached the highest accuracy of 0.89 and balanced class performance[5]. The steady improvement across models demonstrates the effectiveness of combining sub word embeddings, contextual transformers, and reinforcement learning. DRL-based models not only enhance accuracy and efficiency but also offer adaptive decision-making suited for real-world Tamil data with diverse dialects and informal language[6].

Model	Accuracy	MacroF1
M0 TFIDF+LR	0.55	0.5562
M1 Word2Vec+BiLSTM	0.6467	0.651
M2 FastText+CNN	0.72	0.7232
M4 DRL(A2C)+FastText	0.7767	0.7784
M3 mBERT Finetune	0.83	0.8313
M5 DRL(PPO)+mBERT	0.8867	0.8874

Table 2 Classification Models

Table 2 shows that numbers show a clear progression from shallow, frequency-based features to contextual, adaptive modeling. **M0 (TFIDF+LR)** sits at 0.55/0.5562 because bag of words ignores context

and Tamil's agglutinative morphology, so many inflected or codemixed forms look "unseen." Moving to **M1 (Word2Vec+BiLSTM)**, both Accuracy and MacroF1 rise to ~0.65: dense word vectors capture semantic similarity and the BiLSTM models order, but OOV and morph variants still cause leakage across classes. **M2 (FastText+CNN)** passes 0.72 by using sub word embeddings, a better fit for Tamil suffixes and spelling variations; the CNN adds locality, improving short phrase cues (negation, intensifiers) [7] [8]. Two paths then push beyond 0.77: adaptive reading with a lighter encoder and full contextual finetuning. **M4 (DRL(A2C)+FastText)** reaches 0.7767/0.7784 by learning where to look (SELECT/SKIP) and when to stop, filtering noisy social media tokens despite a modest encoder. In parallel, **M3 (mBERT finetune)** climbs to 0.83/0.8313 because multilingual transformers provide sentence level context and handle dialect/codemixing, reducing confusion between close sentiments or topics. The closeness of Accuracy and MacroF1 for M3 suggests better balance across classes than earlier models, not just gains on the majority class [9] [10].

The best results come from **M5 (DRL(PPO)+mBERT)** at 0.8867/0.8874. PPO stabilizes policy learning while mBERT supplies rich context, so the agent can both focus on informative spans and make confident early decisions. The nearly equal Accuracy and MacroF1 indicate robust perclass performance, a strong sign for imbalanced Tamil datasets [11]. Practically, choose M5 when accuracy and robustness matter most; use M3 if you want a strong supervised baseline without RL complexity; pick M4 when compute is tighter but you still want adaptivity. To push further, analyze perclass confusions, test fairness gaps across dialects, and explore reward tuning (read penalty, early stop bonus) to trade off speed vs. accuracy [12] [13].

3. Conclusion

This study developed and validated an end-to-end Deep Reinforcement Learning (DRL) framework for Tamil text classification, covering the full pipeline from data preprocessing to model evaluation and publication readiness. The results demonstrated a steady improvement from traditional models like TFIDF+LR (0.55 Accuracy) to advanced contextual models such as mBERT (0.83 Accuracy). DRL-based approaches further enhanced adaptability and efficiency, with DRL(A2C)+FastText achieving 0.78 and the flagship DRL(PPO)+mBERT model reaching 0.89 accuracy and balanced class performance. The framework's integration of ethical validation and reproducibility tools ensures fairness, transparency, and rapid transformation of experimental outcomes into journal-ready formats[14]. Although the results are promising, future work can expand dialectal and domain coverage, reduce annotation noise, and optimize computational efficiency. Techniques such as active learning, weak supervision, and parameter-efficient tuning can further improve scalability. Additionally, deeper analysis of fairness, interpretability, and reward shaping can enhance trust and inclusivity. Overall, the study confirms that combining DRL with contextual encoders provides an effective and ethical pathway for Tamil NLP and sets a replicable benchmark for other low-resource, morphologically rich languages[15].

References

[1] G. Salton and C. Buckley, "Term weighting approaches in automatic text retrieval," *Information Processing & Management*, vol. 24, no. 5, pp. 513–523, 1988.
[2] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector

space," *arXiv:1301.3781*, 2013.
[3] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with sub word information," *TACL*, vol. 5, pp. 135–146, 2017.
[4] Y. Kim, "Convolutional neural networks for sentence classification," in *Proc. EMNLP*, 2014, pp. 1746–1751.
[5] S. Hochreiter and J. Schmidhuber, "Long shortterm memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[6] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM networks," in *Proc. IJCNN*, 2005, pp. 2047–2052.
[7] J. Devlin, M.W. Chang, K. Lee, and K. Toutanova, "BERT: Pretraining of deep bidirectional transformers for language understanding," in *Proc. NAACLHLT*, 2019, pp. 4171–4186.
[8] A. Conneau et al., "Unsupervised crosslingual representation learning at scale," in *Proc. ACL*, 2020, pp. 8440–8451. (XLMR)
[9] A. Vaswani et al., "Attention is all you need," in *Proc. NeurIPS*, 2017, pp. 5998–6008.
[10] R. Sennrich, B. Haddow, and A. Birch, "Neural machine translation of rare words with sub word units," in *Proc. ACL*, 2016, pp. 1715–1725. (BPE)
[11] T. Kudo and J. Richardson, "Sentence Piece: A simple and language independent sub word tokenizer and detokenizer for neural text processing," in *Proc. EMNLP: System Demonstrations*, 2018, pp. 66–71.
[12] M. Ranzato, S. Chopra, M. Auli, and W. Zaremba, "Sequence level training with recurrent neural networks," in *Proc. ICLR*, 2016.
[13] J. Li, W. Monroe, A. Ritter, M. Jurafsky, D. Liang, and D. Yu, "Deep reinforcement learning for dialogue generation," in *Proc. EMNLP*, 2016, pp. 1192–1202.
[14] K. Narasimhan, T. Kulkarni, and R. Barzilay, "Language understanding with deep reinforcement learning," in *Proc. EMNLP*, 2015, pp. 1686–1696.
[15] Y. Shen, X. He, J. Gao, L. Deng, and G. Mesnil, "ReasoNet: Learning to stop reading in machine comprehension," in *Proc. KDD*, 2017, pp. 1047–1055.
[16] M. J. Seo, A. Kembhavi, A. Farhadi, and H. Hajishirzi, "Learning to skim text," in *Proc. ACL*, 2017, pp. 1880–1890.