Researching Deep Reinforcement Learning Algorithms for Sentiment Analysis

Nhat-Minh Hoang
Faculty of Electrical and Electronic Engineering, Phenikaa University,
Yen Nghia, Hanoi, 12116, Vietnam
21013299@st.phenikaa-uni.edu.vn

July 1, 2024

Abstract

This study explores the potential of deep reinforcement learning (DRL) for enhancing sentiment analysis in Vietnamese text. We propose a novel Reinforcement Learning for Sentence Classification (RLFSC) model that incorporates a word-level selection mechanism guided by a policy network. This network learns to identify and retain sentiment-bearing words while filtering out less informative ones. We evaluate our approach on a dataset collected from Vietnamese e-commerce platforms, comparing its performance to state-of-the-art methods, including traditional machine learning and deep learning techniques. Our results demonstrate that while the RLFSC model achieves comparable or slightly improved accuracy over baseline methods, the contribution of the reinforcement learning component remains limited. We discuss potential reasons for this observation, including the model's tendency to retain most words during training, and suggest directions for future research to further leverage the power of DRL in sentiment analysis.

Keywords: Machine Learning; Reinforcement Learning; Deep Reinforcement Learning and Sentiment Analysis

1 Introduction

Sentiment analysis plays a crucial role in understanding user opinions and emotions expressed in text data. While traditional machine learning approaches and deep learning models like RNNs and GNNs have made significant progress in sentiment analysis, challenges remain in effectively capturing complex linguistic nuances and context-dependent sentiment expressions.

Inspired by recent successes in applying deep reinforcement learning (DRL) to natural language processing tasks, we investigate its potential for enhancing sentiment analysis, particularly in the context of Vietnamese text. Our work builds upon previous studies like Chen et al. [2] and Mengxin [3], which demonstrated the promise of DRL for word-level sentiment analysis and word selection in English sentiment classification tasks.

In this paper, we introduce the RLFSC model, a novel DRL-based approach for Vietnamese sentiment classification. RLFSC employs a policy network to learn a word-level selection policy, aiming to identify and retain sentiment-bearing words while filtering out less informative ones. The filtered sentence representation is then passed to a sentiment classifier for final prediction. We evaluate RLFSC on a dataset collected from Vietnamese e-commerce websites, comparing its performance to existing state-of-the-art methods.

2 Literature Review

Several studies have explored the integration of deep learning models, particularly recurrent neural networks (RNNs) like LSTMs, and graph neural networks (GNNs) for sentiment analysis tasks. Chen et al. propose a novel framework called Word-level Sentiment LSTM (WS-LSTM) [2] that utilizes reinforcement learning for word-level sentiment analysis. Their approach defines actions corresponding to sentiment polarities (positive, neutral, negative) and employs separate LSTM tunnels for each action. This work demonstrates the potential of reinforcement learning in capturing fine-grained sentiment information within sentences. Or the approach of keeping/deleting word(s) in an English sentence of Mengxin [3] is giving a promising result for applying Reinforcement Learning to the Sentiment task.

Further building on the application of deep learning to sentiment analysis, Zhang et al. provide a comprehensive overview of aspect-based sentiment analysis (ABSA) in their survey [9]. They categorize ABSA tasks based on the targeted sentiment elements and highlight the rising importance of compound ABSA tasks that involve predicting multiple interconnected elements, such as aspect-opinion pairs or triplets. The survey underscores the impact of pre-trained language models (PLMs) like BERT and RoBERTa in advancing ABSA performance and emphasizes the need for addressing challenges related to cross-domain and cross-lingual settings.

The use of reinforcement learning specifically for aspect-based sentiment classification is explored by Wang et al. (2020) in their paper [8]. They introduce SentRL, a framework that leverages reinforcement learning to explore dependency graphs of text, enabling an agent to discover sentiment-bearing paths related to specific aspects. By employing a policy network for action selection and a language understanding module for state updates, SentRL aims to pinpoint the most relevant sentiment information while minimizing the influence of task-irrelevant words. This work, along with Cao et al. [1], further highlights the potential of combining reinforcement learning with external knowledge sources like knowledge graphs to address the inherent sparsity and ambiguity present in text data for sentiment analysis.

3 Methodology

The target of this project is to apply Reinforcement Learning to improve the Sentiment prediction of the sentence. The Actor-based model provides the probability of keeping or discarding the word in the sentence. This acts as the filter for the feature of the sentence. After that, the adjusted feature of the sentence is fed into the classifier to get the final reward which is necessary for updating the Reinforcement learning model.

3.1 Policy Network

For the policy network, the goal is to guide the agent to find which words in the sentence are important to the classification task. The state that is fed into the policy model is the embedding matrix of a word in the sentence. It can be received by a pre-trained BERT model which in this project is the PhoBERT model, the pre-trained BERT for Vietnamese [5]. However, the feature of one word at a time is not enough information for the model to provide the output. It needs an architecture that can remember and deal with sequence data which is ideal for using BRNN or Bi-LSTM [7]. The state now is extracted more clearly for the model to get the reward.

$$S_t = BiLSTM(A_t, S_{t-1}), \tag{1}$$

Where S_t is the current state containing the feature of the word in a sentence and the A_t is the word feature of the corresponding t-th actions.

3.2 Sentiment Classifier

After the agent goes through all the words in the sentence and provides the filtered sentence which is the input of the classification model:

$$S'_t = S_t * A_t \tag{2}$$

$$p = \delta(W_p S'_t + b_p) \tag{3}$$

The equation (3) has Softmax activation to get the probability of the sentiment, which affects the reward. The W_t is the weight of the model and b_p is the bias.

3.3 Reward

The reward is a guide to the model to know whether it is on the right way to the goal or not. In this project, the ultimate goal is the classification performance and depends on multiple steps to get the final reward. For the reward, we define it as the minus of the Kullback - Leibler divergence loss which is suitable for probability output:

$$r_i = -D_{KL}(p||Y) \tag{4}$$

Where p is the prediction probability and Y is the one-hot vector of the ground true label.

4 Experiment

4.1 Dataset

The dataset collected from 3 big E-commercial websites in Vietnam has been used. Which have the statistical information in the Figure 1 and Figure 5. Also, the UIT-ViSFD [4] dataset and UIT-VSFC [6] dataset were also used to have a further test for the model.

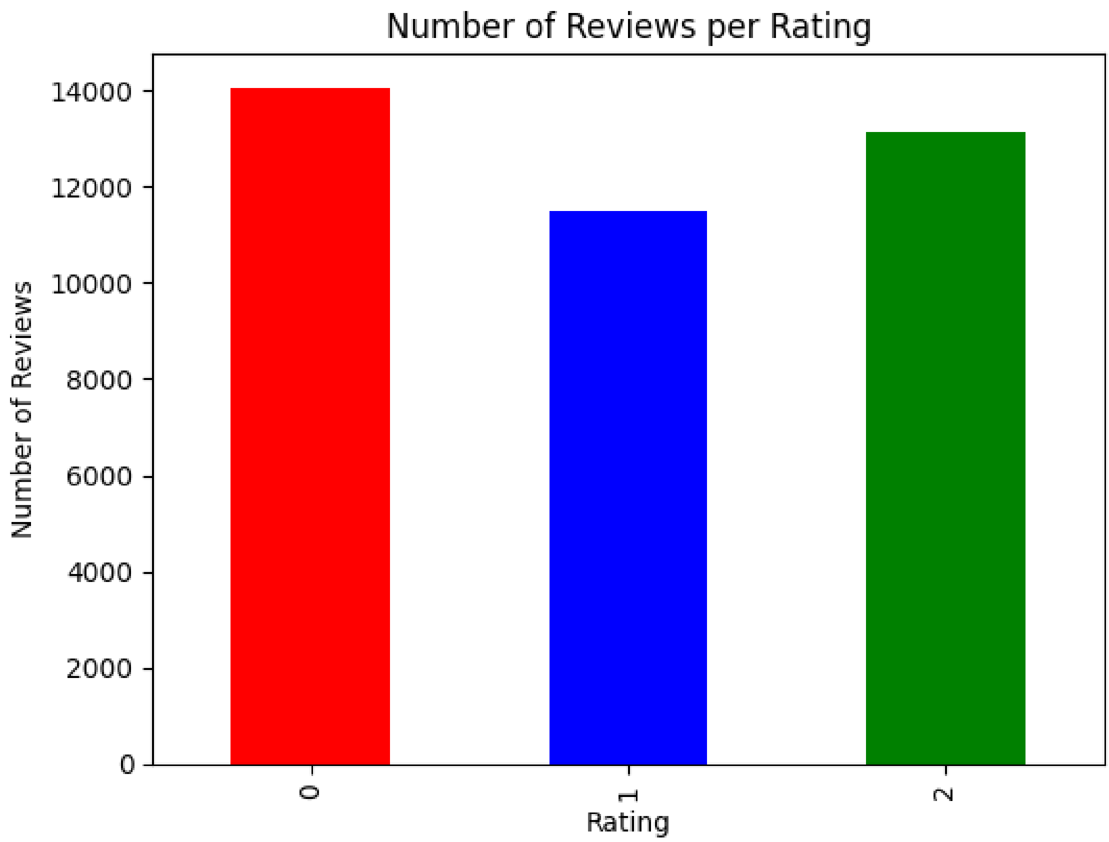

Figure 1: STATISTICS OF THE DATASET

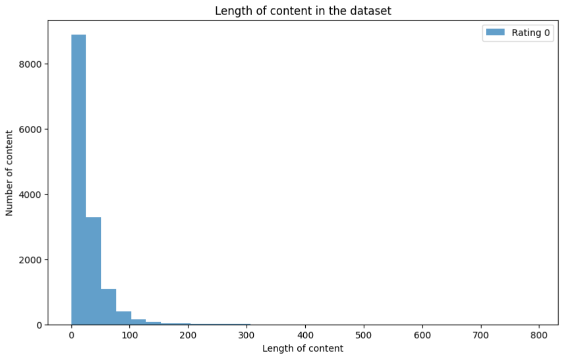

Figure 2: Length of sentence in label 0

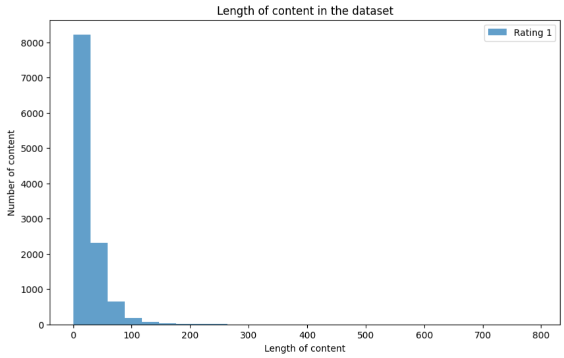

Figure 3: Length of sentence in label 1

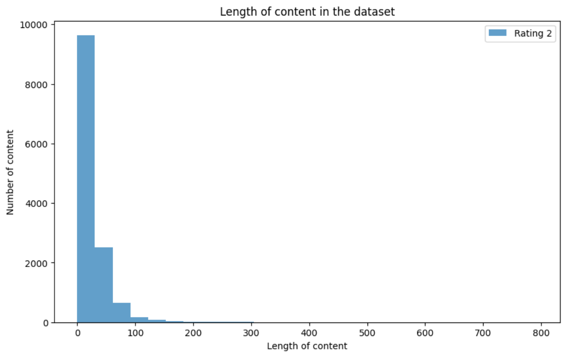

Figure 4: Length of sentence in label 2

Figure 5: STATISTICS OF THE SENTENCE LENGTH IN DATASET

	Precision	Recall	F1-score
UIT-VSFC MaxEnt	0.88	**0.89**	0.88
UIT-VSFC RLFSC	**0.89**	**0.89**	**0.89**
UIT-ViSFD LSTM	**0.57**	0.48	0.52
UIT-ViSFD RLFSC	**0.57**	**0.61**	**0.59**

Table 1: UIT datasets performance

4.2 Performance Results

The accuracy often gets over 60% on the dataset. It's not a big improvement but still can be adjusted to have a better performance. After that, the model easily gets overfitted during training, so the hyper-parameters must be more tuned. However, after optimizing the parameters and applying the regularization, the model RLFSC (Reinforcement Learning For Sentence Classification) model performs 5% better and the training step is more stable which can be seen in Figure 10.

Table 1 gives the overall results on the sentiment-based task in a previous paper of 2 UIT's datasets. In detail, the reinforcement learning classifier model performs better compared to the related work. The most reasonable changes are in the UIT-ViFSD dataset, the Recall and F1-score of the reinforcement learning approach are noticeably

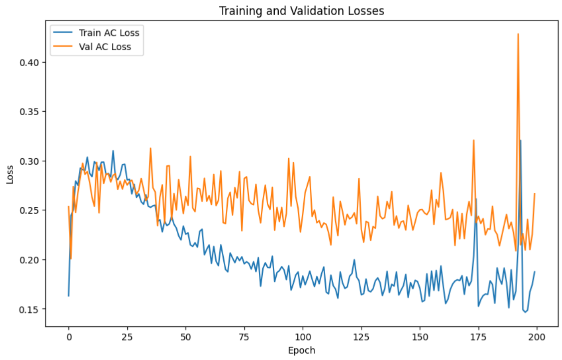

Figure 6: A2C loss each epoch

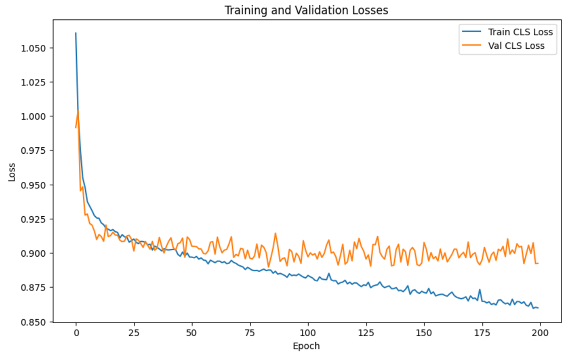

Figure 7: Classifier loss each epoch

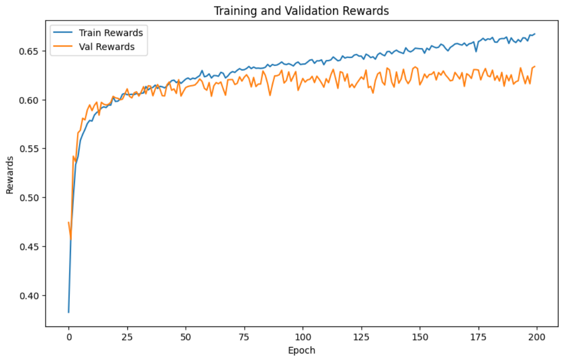

Figure 8: Rewards each epoch

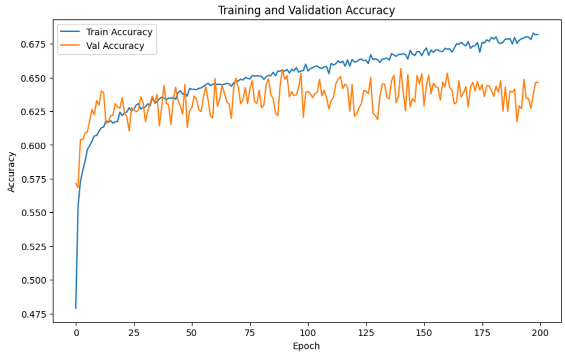

Figure 9: Accuracy each epoch

Figure 10: RLFSC Sentiment Training process's performance

higher.

5 Discussion

Although the performance of the RLFSC model is better than the State-of-the-art works, the contribution of Reinforcement Learning in this RLFSC approach is negligible. For example, with the UIT-VSFC dataset, the action 'keep' distribution is more than 93% of the words in the training data. This a reasonable choice because the more words in the sentence, the more features the classifier model gets to propose the output better.

Another approach is forcing Reinforcement Learning to find the main subject that makes the sentence's sentiment which is much less word consumption and less dependency in the classification model. This can be solved by establishing a different environment to solve the issue.

6 Conclusion

This research investigated the application of deep reinforcement learning for sentiment analysis in Vietnamese text. Our proposed RLFSC model, incorporating a word-level selection mechanism, achieved comparable or slightly improved accuracy compared to state-of-the-art methods on benchmark datasets. However, analysis revealed that the contribution of the reinforcement learning component was limited, as the model tended to retain most words during training. This suggests that while the framework holds promise, further exploration is needed to fully harness the power of DRL in capturing salient sentiment information.

Future research directions include:

Refining the reward function: Designing a reward mechanism that encourages the model to focus on truly discriminative words for sentiment classification.

Exploring different DRL algorithms: Experiment with more advanced DRL algorithms (e.g., A3C, PPO) that might be better suited for handling the sequential nature of text data.

Incorporating external knowledge: Enhancing the model's understanding of sentiment by integrating external knowledge sources, such as sentiment lexicons or knowledge graphs.

By addressing these challenges, we aim to unlock the full potential of DRL in developing more robust and accurate sentiment analysis systems for Vietnamese and other languages.

7 References

References

[1] Yukun Cao, Yijia Tang, Haizhou Du, Feifei Xu, Ziyue Wei, and Chengkun Jin. Heterogeneous reinforcement learning network for aspect-based sentiment classification with external knowledge. *IEEE Transactions on Affective Computing*, 2023.

[2] Ruiqi Chen, Yanquan Zhou, Liujie Zhang, and Xiuyu Duan. Word-level sentiment analysis with reinforcement learning. In *IOP conference series: materials science and engineering*, volume 490, page 062063. IOP Publishing, 2019.

[3] Helen Mengxin Ji. Sentiment analysis using rl.

[4] Luong Luc Phan, Phuc Huynh Pham, Kim Thi-Thanh Nguyen, Sieu Khai Huynh, Tham Thi Nguyen, Luan Thanh Nguyen, Tin Van Huynh, and Kiet Van Nguyen. Sa2sl: From aspect-based sentiment analysis to social listening system for business

intelligence. In Han Qiu, Cheng Zhang, Zongming Fei, Meikang Qiu, and Sun-Yuan Kung, editors, *Knowledge Science, Engineering and Management*, pages 647–658, Cham, 2021. Springer International Publishing.

[5] Dat Quoc Nguyen and Anh Tuan Nguyen. PhoBERT: Pre-trained language models for Vietnamese. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 1037–1042, 2020.

[6] Kiet Van Nguyen, Duc-Vu Nguyen, Phu X. V. Nguyen, Tham T. H. Truong, and Ngan Luu-Thuy Nguyen. Uit-vsfc: Vietnamese students' feedback corpus for sentiment analysis. *2018 10th International Conference on Knowledge and Systems Engineering (KSE)*, pages 19–24, 2018.

[7] Mike Schuster and Kuldip K Paliwal. Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, 45(11):2673–2681, 1997.

[8] Lichen Wang, Bo Zong, Yunyu Liu, Can Qin, Wei Cheng, Wenchao Yu, Xuchao Zhang, Haifeng Chen, and Yun Fu. Aspect-based sentiment classification via reinforcement learning. In *2021 IEEE International Conference on Data Mining (ICDM)*, pages 1391–1396. IEEE, 2021.

[9] Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam. A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. *IEEE Transactions on Knowledge and Data Engineering*, 2022.