

# Reinforcement Learning-Enhanced Chatbot for Natural Language Processing

Sanidhaya Sharma  
Purdue University Fort Wayne  
Fort Wayne, Indiana  
shars11@pfw.edu

## Abstract

The integration of Reinforcement Learning (RL) with Natural Language Processing (NLP) has shown promising results in improving dialogue systems. Traditional chatbots often lack the ability to adapt and improve based on user interactions. By leveraging RL, we aim to develop a chatbot that can learn from user feedback and enhance its conversational abilities over time. This project will explore novel RL methods to optimize chatbot responses, providing a more engaging and effective user experience.

## 1 Introduction

The primary problem we aim to address is the static nature of conventional chatbots, which do not adapt or improve based on user interactions. We will focus on developing a chatbot that can dynamically refine its responses using RL, thereby improving user satisfaction and conversation quality.

## 2 Model/Algorithm

We propose to use a combination of pre-trained language models (e.g., GPT-2 or GPT-3) and RL algorithms (e.g., Proximal Policy Optimization - PPO or Deep Q-Learning). The chatbot will use the conversation history as the state, generate responses as actions, and receive rewards based on user feedback or conversation coherence.

## 3 Datasets

We plan to use the following datasets for training and evaluation:

- **Cornell Movie Dialogues:** For initial training of the base chatbot.
- **OpenSubtitles:** To provide diverse conversational contexts.
- **DailyDialog:** For fine-tuning and evaluating the chatbot's performance in everyday conversations.

## 4 Measurement of Success

The success of the project will be measured using several metrics:

- **Accuracy:** The correctness of the chatbot's responses.
- **F1 Score:** To evaluate the balance between precision and recall in response generation.
- **User Satisfaction:** Qualitative feedback from users to assess the chatbot's performance.
- **BLEU Score:** To measure the quality of generated text against reference responses.

## 5 Possible Analysis

Throughout the project, we will conduct various analyses to ensure the effectiveness of our approach:

- **Training Curves:** Monitoring the reward over time to assess the learning process.
- **Comparative Analysis:** Comparing the RL-enhanced chatbot with the base version to evaluate improvements.
- **Hyperparameter Tuning:** Experimenting with different learning rates and discount factors to optimize performance.

## Limitations

While our approach shows promise, there are several limitations to consider:

- The method may work best for languages with limited morphology, like English.
- Scalability to long text and the requirement of large GPU resources could be challenging.
- Further investigation is needed to address potential overfitting and reward function design.

## Ethics Statement

We acknowledge the importance of ethical considerations in developing AI systems. Our chatbot will be designed to respect user privacy and ensure that the data collected is used responsibly. We will adhere to the ACL Ethics Policy and encourage transparency in our methodology and results.

## Acknowledgements

We would like to thank the authors of the ACL 2023 LaTeX template for providing a clear and comprehensive guide for formatting our proposal.

## References

### A Project Timeline

- **Week 1:** Project Planning and Literature Review
- **Week 2:** Design the Chatbot Architecture
- **Week 3:** Implement the Base Chatbot
- **Week 4:** Integrate Reinforcement Learning Framework
- **Week 5:** Train the RL-Enhanced Chatbot
- **Week 6:** Evaluate and Analyze Results
- **Week 7:** Iterate and Improve
- **Week 8:** Prepare for Research Paper
- **Week 9:** Finalize the Chatbot and Experiments
- **Week 10:** Write and Submit the Research Paper