Benefits of the reward system in text generation

Yedarm Seong\textsuperscript{1} and Myungjoo Kang\textsuperscript{2}

\textsuperscript{1) Department of Computational Science and Technology, Seoul National University, Seoul 151-742, KOREA}
\textsuperscript{2) Department of Mathematical Sciences, Seoul National University, Seoul 151-742, KOREA}

Corresponding Author: Myungjoo Kang, mkang@snu.ac.kr

ABSTRACT

Generative Pre-trained Transformer 2 (GPT-2) is one of the powerful language models, which records high scores in various Natural Language Processing (NLP) tasks. Especially in the text generation field, it performs like a human for writing novels and news. We introduce a GPT-2 based text generation algorithm, which by adding an optimization process using the Reinforcement Learning (RL) algorithm, Proxy policy optimization (PPO), and show that the reward system in RL can contribute to NLP and generative models for text.

REFERENCES