

ABSTRACT

A NOVEL DEEP LEARNING APPROACH FOR CONTROLLED MULTI-TOPIC TEXT GENERATION

Çağlayan, Cansen

M.S., Department of Computer Engineering

Supervisor : Assoc. Prof. Dr. Kasım Murat Karakaya

September 2022, 95 pages

One of the most important tasks in the Controllable Text Generation (CTG) domain is to create topic-controlled texts. In this study, we propose and design three different approaches, and conduct extensive experiments on them to observe the performance of the controlled multi-topic reviews generated in Turkish. In the first approach, we generate controlled multi-topic text using a single-layer GPT language model by incorporating several control techniques. To control the language model, we first add topic information to the sequential input, as a second technique we add the automatically extracted keywords for each topic to the sequential input in addition to the first technique. The last technique that we propose is a novel sampling strategy. We propose to use a topic selection classifier that enables the next token selection according to the probability of the selected tokens being on the desired topic. Then, we apply these approaches to a more advanced language model, the multi-layer GPT, and interpret the results. In addition to these experiments, we compare three different deep learning text classification models in order to create a reliable multi-topic review classifier.

Keywords: Controllable text generation, text generation, text classification, nlp.