Automatic Generation of Research Objectives and Citation Sentences in Academic Papers
学術論文の研究目的と引用文の自動生成

Student No.: 201821629

氏名: 前野　翔平

Name: MAENO SHOHEI

The purpose of this study is to automatically generate the objective sentences and citation sentences included in the introduction secetion of academic papers. The introduction of an academic paper is the most important section in order to explain the importance of research. Therefore, it is a challenge to write an introduction that will help the reader understand the significance of the research. In addition, the introduction is a section that always contains sentences that state research objectives and citations of related researches, regardless of the research field. Therefore, we believe that the automatic generation of objective sentences and citation sentences can make the writing of the introduction section easier.

To build the academic article dataset used in this study, we manually collected open access articles. For the automatic generation of objective sentences, we first built the objective sentence extraction model to collect automatically the objective sentences from articles. Second, we used the Method, Results, Discussion and similar sections of articles, and trained the model to generate the objective sentence. For the automatic generation of citation sentences, we first collected citation sentences from the introduction section of colleted papers and extracted abstracts of the corresponding citation sentences from PubMed (`https://pubmed.ncbi.nlm.nih.gov/`) to build the dataset. Secondly, we trained the model to generate citation sentences from the abstract of the cited paper and the Method, Results, Discussion, and similar sections of the collected paper.

We implemented the ROUGE-TITLE and ROUGE-ABSTRACT metrics to determine whether the quality of human-written or model-generated sentences is better. ROUGE-TITLE calculates the ROUGE-1 of the title of the paper and the sentence, and ROUGE-ABSTRACT calculates the ROUGE-1 of the abstract of the paper and the sentence. In the objective sentence generation task, the title and abstract of collected papers are used to calculate the two metrics. On the other hand, the title and abstract of the cited papers are used in the citation sentence generation task. ROUGE-TITLE, ROUGE-ABSTRACT of human-written sentences and model-generated sentences are computed to evaluate the models.

We succeeded in generating equally or higher quality sentences than humans in the objective sentence generation task and the citation sentence generation task with a probability of 17.5% and 11.8%, respectively. The automatic generation of the objective sentences and citation sentences in the introduction section will help people who have difficulties in writing papers in English.

Principal Academic Advisor: Tetsuya MAESHIRO

Secondary Academic Advisor: Hideo JOHO