

# Graduate School of Science and Technology Master's Thesis Abstract

Laboratory name (Supervisor)	Natural Language Processing (WATANABE TARO (Professor ))		
Student ID	2111433	Submission date	2024 / 1 / 18
Name	VASSELLI JUSTIN RAY		
Thesis title	Retrieval-Based Grammatical Error Correction		
Abstract			
<p>In various natural language processing tasks, such as named entity recognition and machine translation, example-based approaches have been used to improve performance by leveraging existing knowledge. However, the effectiveness of this approach for Grammatical Error Correction (GEC) is unclear. In this work, we explore how an example retrieval-based approach affects the accuracy and interpretability of the output of GEC systems and the trade-offs involved. The approach we investigate has shown great promise in machine translation by using the <math>k</math> nearest translation examples to improve the results of a pretrained Transformer model. We find that using this technique increases precision by reducing the number of false positives, but recall suffers as the model becomes more conservative overall. Increasing the number of example sentences in the datastore does lead to better performing systems, but with diminishing returns and a high decoding cost. Synthetic data can be used as examples, but the effectiveness varies depending on the base model. Finally, we find that finetuning on a set of data may be more effective than using that data during decoding as examples.</p>			