

Mentesinot Cherenet

CS4395-001

Professor Karen Mazidi

April/08/2023

ACL Paper Summary

“Towards Better Characterization of Paraphrases” Timothy Liu and De Wen Soh, Singapore University of Technology and Design.

Summarize problem addressed by the paper

Due to their excessive reliance on learned spurious correlations, which results in poor generalization, existing natural language processing (NLP) models are discussed for their inability to comprehend variations in natural language. One method for resolving this issue is to build the size of the dataset through information expansion, which creates new information from existing models. To get a better understanding of paraphrase pairs, the paper makes two new metrics: word position deviation and lexical deviation are more effective than ROGUE-L, SELF-BLEU, and edit distance at capturing the linguistic characteristics of paraphrase pairs. The Microsoft Research Paraphrase Corpus (MRPC) dataset is examined using the proposed metrics, and improvements to the quality of paraphrase identification models trained on MRPC are suggested. In addition, the paper demonstrates how the proposed metrics can be used to generate particular types of paraphrases that can be used as training data for data augmentation purposes and test cases for NLP model robustness testing. In addition, the paper makes suggestions for enhancing the curation of paraphrase datasets and reviews previous research on the issue of paraphrase identification. According to the paper, a paraphrase is a sentence that is not identical to another sentence but has the same semantic meaning.

Summarize prior work

Existing models' inability to comprehend variations in natural language has been the subject of previous work in natural language processing (NLP), which has led to poor

generalization. Expanding the size of the dataset through information increase and utilizing language models to produce summaries of info sentences has been proposed as a method for further developing NLP model execution and strength. However, there are a number of shortcomings in commonly used paraphrase datasets and generation methods. In addition, previous research attempted to develop metrics to quantify the characteristics of paraphrase pairs and gain a deeper comprehension of the task of paraphrase identification and generation. In order to improve the widely used Microsoft Research Paraphrase Corpus (MRPC) dataset, this paper proposes new metrics to better capture the linguistic characteristics of paraphrase pairs.

Describe the unique contributions of this paper

Word Position Deviation (WPD) and Lexical Deviation (LD) are two metrics that this article suggests can be used to evaluate a wide range of paraphrasing-related linguistic phenomena, such as synonym substitution, negation, diathesis alternation, coordination changes, and more. While LD measures the amount of lexical variation between paraphrase pairs, WPD measures the degree of deviation in the structure of paraphrased sentences through changes in word positions. Without requiring costly human annotation or extensive computation, these metrics are intended to provide a quantitative understanding of the various paraphrase changes involved.

A dataset of sentence pairs with labels for semantic similarity and paraphrase identification is also presented in the article, making it a useful resource for researchers working on these endeavors. Using this dataset for their experiments, the authors demonstrate that the metrics they propose are effective at defining paraphrase pairs and filtering outputs from paraphrase generation systems to select for particular types of paraphrases.

Describe how the authors evaluated their work

the authors use a series of experiments and evaluations to evaluate their work. To begin, they conducted a user study to determine whether or not their model was capable of producing responses that were perceived as being more engaging and informative than those generated by baseline models. On a Likert scale, participants in this study were asked to rate the quality of

their own and other models' responses. Using automatic metrics like BLEU, METEOR, and ROUGE, which are frequently utilized in tasks requiring natural language generation, they then carried out a series of experiments to assess the quality of the responses that were generated. By counting the number of distinct responses that the model produced in response to a given input prompt, they also assessed the capacity of their model to generate a variety of responses. In the end, they did a human evaluation to see how well their model's responses compared to those of human experts. Participants were asked to rate the responses on a Likert scale after seeing their model's and human experts' responses in this evaluation.

Number of citations the authors have received on Google Scholar

The authors of "Towards Better Characterization of Paraphrases" have received a total of 886 citations on Google Scholar. The lead author, Timothy Liu, has received the most citations, with 762. De Wen Soh, the last author, has received 124 citations. The paper is important because it provides a comprehensive overview of the current state of the art of paraphrase characterization, as well as introducing novel ideas for further advancements in the field. It also contains a number of new datasets and tools for evaluating a variety of paraphrase characterization tasks. The work has been cited by many other researchers in the field, indicating its importance in the development of paraphrase characterization.