

Automatic Citation Screening Using Pattern-Exploiting Training and Paraphrasing

Alice Romagnoli¹, Wojciech Kusa² and Gabriella Pasi¹

¹Università degli Studi di Milano-Bicocca, Piazza dell'Ateneo Nuovo, 1, 20126 Milano MI, Italy

²TU Wien, Favoritenstraße 9-11, 1040 Wien, Austria

Abstract

Systematic Reviews (SRs) are a technique of presenting and synthesizing data and results related to a specific topic. Creating SRs is a time-consuming process primarily due to the manual work involved and the constant publication of new results. The SR process includes several steps, one of which is citation screening, which involves extracting titles and abstracts from the documents and evaluating them to assess the document's quality and relevance. This work explores using prompt-based learning to automate citation screening and tests it on seven systematic reviews from the domain of computer science. The screening is seen as a binary classification problem and is implemented using Pattern-Exploiting Training (PET). Given the imbalanced review datasets, we demonstrate how two data augmentation techniques (oversampling and paraphrasing) improve the quality of automatic screening. Our results show that the PET approach alone yields poor results. However, when evaluating the approach at an intermediate phase, (i.e., combining prompts with an ensemble of language models) leads to more promising results.

Keywords

Systematic reviews, Citation screening, Prompt-based learning, Pattern-Exploiting Training


2nd Workshop on Augmented Intelligence for Technology-Assisted Reviews Systems: Evaluation Metrics and Protocols for eDiscovery and Systematic Review Systems. 2nd April 2023

✉ a.romagnoli7@campus.unimib.it (A. Romagnoli); wojciech.kusa@tuwien.ac.at (W. Kusa);
gabriella.pasi@unimib.it (G. Pasi)

ORCID 0000-0003-4420-4147 (W. Kusa); 0000-0002-6080-8170 (G. Pasi)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)