Improving Medical Cases Retrieval using an Online Fact Database

Ivan Kitanovski, Katarina Trojacanec, Ivica Dimitrovski, Suzana Loshkovska

Faculty of Computer Science and Engineering, University "Ss. Cyril and Methodius", Skopje, Macedonia
{ivan.kitanovski,katarina.trojacanec,ivica.dimitrovski,suzana.loshkovska}@finki.ukim.mk
http://www.finki.ukim.mk

Abstract. This paper presents an approach for retrieval of medical cases using a novel query expansion method. The approach relies purely on the text data in the medical cases. The cases are indexed with Terrier IR search engine based on their text content including the caption of the figure contained within them. Furthermore, in the retrieval phase there is an input consisted of a long text query in a narrative form. The input query is expanded by using on-line fact databases, such as Freebase, with the aim that this will add more terms relevant to the concepts mentioned in the text. The goal is to provide a way of query expansion, so that the query is more defined, which should provide more narrowed and precise results in the retrieval. The retrieval is done with the BM25 weighting model. Our approach shows that expanding the input text query in this fashion can provide a boost in the retrieval performance.

Keywords: Medical articles retrieval, Query expansion, PubMed articles, Terrier IR, Freebase.