

Question Answering with Transformers and Few-Shot Learning to find Inventive Solutions for IDM-TRIZ Problems and Contradictions in Patents

Stefan Trapp^{1(✉)}, Nathanael Großer¹, and Joachim Warschat¹

¹ University of Hagen, Universitätsstraße 47, 58097 Hagen, Germany
stefan.trapp@fernuni-hagen.de

Abstract. In this paper, a question-answering (QA) system based on BERT Transformer models is introduced to identify innovative solutions to problems and contradictions from patents. As a first step, a classification system based on Sentence-Transformer Fine-Tuning (SetFit) is developed to classify patents sentence-wise into IDM-TRIZ entities and to add the identified problems and contradictions to a search index. The second step is the creation of an end-to-end QA system. It is shown that a RoBERTa model trained with the SQuAD 2.0 dataset is better suited for the extraction of problem solutions than a BERT model fine-tuned with the same data set. Comparing different patent sections and search strategies, it turns out that more solutions can be extracted from the "State-of-the-Art" than from the "Claims" and "Abstract" sections. The presented QA system achieves an F1 score of 0.82, outperforming a state-of-the-art QA approach for IDM-TRIZ solution extraction.

Keywords: BERT, Bi-Encoder, Classification, Dense Retrieval, Dual-Encoder, Few-Shot Learning, Fine-Tuning, IDM-TRIZ, Information Retrieval, Question-Answering, Retriever-Reader Architecture, RoBERTa, SBERT, Sentence-Transformer, SetFit, SQuAD, Transformer.