

The BiLSTM baseline was implemented in PyTorch 1.13.1, with the required NLP pipeline and word embeddings built using Keras 2.11.0 and GloVe~\cite{Pennington2014Glove}. We used the same approach for hyperparameter optimization of the BiLSTM baseline as we did for BERT variants.

Our Keyword Search baseline employs the Python-based implementation of Arora et al.'s keyword extraction approach~\cite{Arora2016Automated}, available as part of the WikiDoMiner tool~\cite{Ezzini2022WikiDoMiner}. This baseline further uses scikit-learn 1.2.0 to implement both its Count Vectorizer and Random Forest classifier. In addition, we use scikit-learn 1.2.0 for dataset splitting and metrics calculation throughout our evaluation. All our implementation artifacts are available [online](#)~\cite{Code}.