Question Answering

Approaches

Fine-Tuning Sentence-BERT for Question Answering

CapitalOne produced a tutorial (mirror) about using sentence-transformers for Question Answering.

- They use SBERT because it is optimised for fast compute on individual sentence and has good general performance on a number of NLP tasks. They are most interested in the STS capability of SBERT
- They fine-tuned the model on 7k utterances and saw a huge improvement in performance:
  - 52% match rate on 200 test Q-A pairs with no fine tuning
  - 79% match rate on the 200 sample dataset after fine tuning
- They use triplet loss (minimise distance between alike sentences, maximise distance between different sentences)
- This is a nice simple approach where you want to present the answer, as-written, back to the user but its not appropriate for factoid type use cases where we want to highlight the exact word or phrase in the answer text.

Haystack

Haystack is an open source NLP framework for use cases involving large collections of documents. It could be used for searching and ranking type use cases and question answering type use cases.

Haystack works flexibly with existing document stores including DB systems and ElasticSearch.

Revision #4
Created 30 March 2022 19:37:26 by James
Updated 8 February 2024 21:05:14 by James