IBM @ TREC CT-Track 2021

Laura Biester, Venkata Joopudi, Bharath Dandala
lbiester@umich.edu, {vnjoopud,bdand}@us.ibm.com

Introduction

- Clinical trials often fail because they do not meet recruitment targets
- AI system to link patient descriptions to clinical trials

Impact

- Advance science and medicine by allowing clinical trials to succeed
- Help patients who may benefit from participation in a clinical trial

Data

With limited training data availability, we leverage information in both the MIMIC III corpus and clinical trials to create a large-scale silver-standard dataset

Dataset	Patient Description	Clinical Trials	Labeled Pairs
TREC	75	375K	0
SIGIR	60	204K	3870
AutoGT	18k	375K	700K+

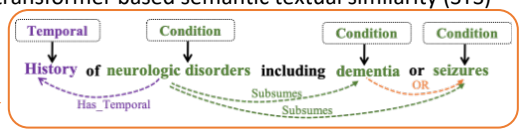
Results

System	NDCG@10	PREC@10	Reciporical Rank
IBMLucene	**.32**	**.20**	**.39**
IBMSTS	.22	.15	.27
IBM SIGIR	.14	.09	.19
IBMAUTOGT	.13	.09	.14
IBMSIGIRACT		.06	.13

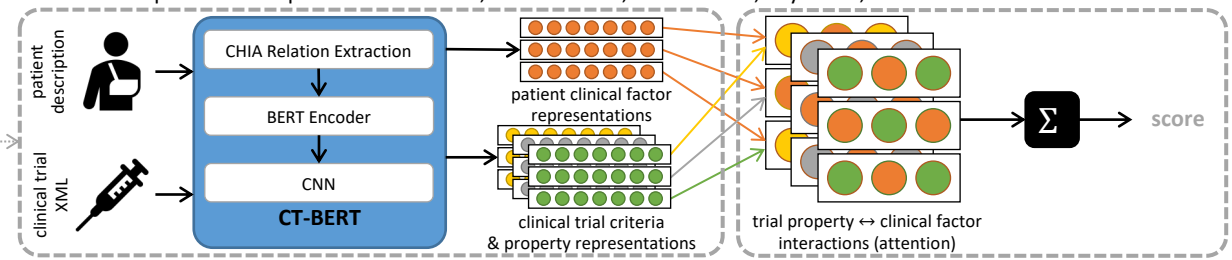
- Our base retrieval system (IBMLucene) achieved the best results across three metrics.
- Existing SOTA information retrieval methods lead to poor performance on development data, and we did not submit runs using them.
- Large-scale training data for this problem is not available, and the proposed silver-standard training data approaches led to worse performance by our deep learning systems.
- We plan to perform deeper analysis to identify the causes once the relevance judgements are publicly released.

Methodology

1. **Query generation module** annotates topics using IBM Watson Annotator for Clinical Data. Features specific to the topics were ranked using MIMIC, PubMed, and a rare lexicon disease to prioritize uncommon conditions.
2. **Retrieval modules** using a BM25 retrieval on a Lucene index and a transformer based semantic textual similarity (STS) model retrieve candidate trials.
3. **Neural BERT-Based rerankers** rank the top 2k trails per topic. See a diagram of our reranker architecture below. Entities are extracted from topics and trials using a model trained on CHIA



Text is encoded using a **BERT-Based model (CT-BERT)**, and **attention mechanisms** are used to compute alignment between spans in the topics and the criteria, interventions, MeSH terms, keywords, and condition.



Citations

Koopman, B., & Zuccon, G. (2016, July). A test collection for matching patients to clinical trials. In Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval (pp. 669-672).

Kury, F., Butler, A., Yuan, C., Fu, L. H., Sun, Y., Liu, H., ... & Weng, C. (2020). Chia, a large annotated corpus of clinical trial eligibility criteria. Scientific data, 7(1), 1-11.

Zhang, X., Xiao, C., Glass, L. M., & Sun, J. (2020, April). Deepenroll: Patient-trial matching with deep embedding and entailment prediction. In Proceedings of The Web Conference 2020 (pp. 1029-1037).