Utilizing Passage Similarity Metric for Passage Ranking

Sumanta Kashyapi[1], Dr. Laura Dietz[1]
[1]University of New Hampshire

Objective

- To find out how to improve passage ranking using a Wikipedia passage similarity metric.
- To figure out the best passage similarity metric suitable for ordering passages in a ranking.

Unsupervised Similarity Metric

For a pair of passage p_1,p_2

- **UNH-bm25-[type]**

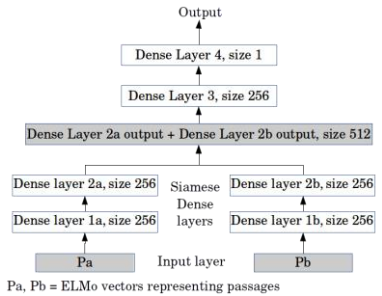
$$score = \frac{BM25(p_1, p_2) + BM25(p_2, p_1)}{2}$$

- **UNH-tfidf-[type]**

$$score = cosine(v_1, v_2)$$

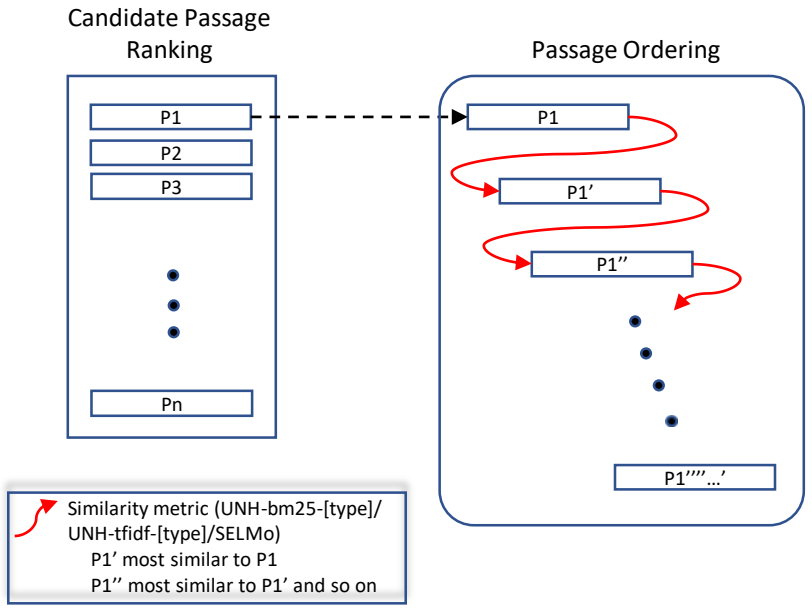
BM25(q,d) = BM25 score between query q and document d
v_i = TFIDF vector of passage p_i
type = Pre-processing method; stemming (stem) / lemmatization (lem) / no pre-processing (pt)

Siamese ELMo (SELMo) Similarity Metric

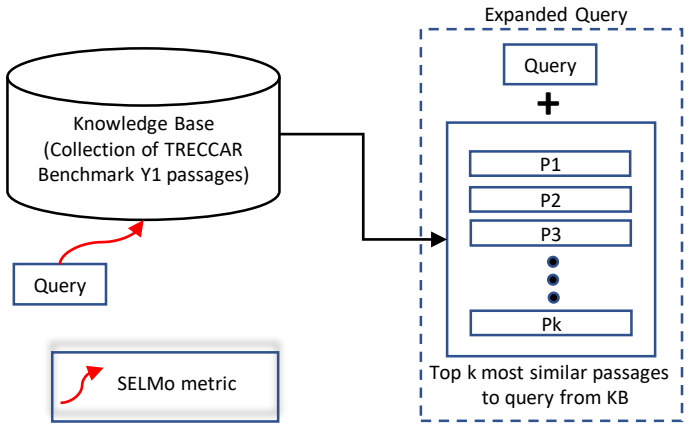


Siamese architecture used to model the similarity metric used for **UNH-dl**

TRECCAR track: Daisy Chain Passage Ordering



Query Expansion using SELMo (UNH-exDL)



Results

TRECCAR track

Method	Facet Overlap	Relevance
UNH-bm25-stem	0.0622 ± 0.0096	**0.1297 ± 0.0170**
UNH-tfidf-ptsim	**0.0756 ± 0.0115**	0.1230 ± 0.0174
UNH-tfidf-lem	0.0686 ± 0.0105	0.1150 ± 0.0165
UNH-tfidf-stem	0.0674 ± 0.0105	0.1168 ± 0.0165
UNH-dl100	0.0403 ± 0.0072	0.1134 ± 0.0165
UNH-dl300	0.0335 ± 0.0065	0.1093 ± 0.0159

DL track

Method	MAP	Mean NDCG	Mean P@10
UNH-bm25 (BM25 baseline)	0.2565	0.5546	0.3465
UNH-exDL	0.0364	0.1400	0.0605

Conclusion

- Current results show that unsupervised similarity metrics outperform SELMo for passage ordering task. In future we plan to improve the similarity metric using an ensemble of all the metrics.
- We observe that our query expansion model performs poorly on the re-ranking task. This suggests that our choice of knowledge base for the query expansion (TRECCAR benchmark Y1) is not suitable for the task.