

Information Retrieval - Assignment 1 Peer Review

Anonymous

June 12, 2019

1 Summary

The assignment should show us 3 methods for using ad-hoc information retrieval. Starting off with BM25 itself (which is a ranking algorithm) and additionally extend BM25 with query extension (by using word embeddings) and as a last method by using word embeddings directly (saliency-weighted semantic network). The results of all 3 methods then can be compared and the results discussed.

2 Peer-review

I will first give some general remarks about the assignment and then give feedback on each task.

3 General Remarks

Well first great structure as I would expect from a research paper. And every section seems to be very detailed described to make it easier for a reader to get into it and get an understanding of the methods used. The only thing missing as you stated is the abstract, but I guess that's a easy thing to add in the end.

4 Report feedback

4.1 Introduction

Very detailed introduction and greatly done to get the reader on the level needed for the texts to come. Also clearly discuss what will be the baseline throughout the document.

4.2 Method

4.2.1 BM25

Mathematical formulations are all included and described well in detail. Not much else to say, well done.

4.2.2 Bm25 Query Expansion

Also here clear structured, both "methods" explained in detail and what Query expansion is, quite like it. Also here you already added that softmax normalization is missing, so I guess will be added later.

4.3 Experiments

Do like your in detail explanation of MAP and you also the explaining of the cross-validation is fine and since implementation is coming then I think it's alright.

4.4 Results

Results are missing even though part of it (without cross-validation) is done in the code, but as soon as the rest is implemented then just adding the table with the results will be sufficient.

4.5 Discuss

Only part of the discussion is done so far, but without the proper results with cross-validation also kind of hard at the moment to discuss much about results. Especially when discussing about the word embeddings since the results should show it best if the word embeddings are actually useful or not.

5 Final Remarks & Conclusion

Overall a very nice report and good structured, Code is also fine. As you yourself also know is that some things are missing but you noted it as todo's inside your report, so I think it will be quite alright as soon as you add the missing parts.