Project Proposal for CS 410

October 25, 2020

1. **Name, NetId and Captain**
   This is a one member team
   Name: Sai Srujan Gudibandi
   NetId: ssg7
   Captain: Sai Srujan Gudibandi

2. **Competition selection**
   I choose to participate in the Information Retrieval Competition

3. **Implementation and ideas for the project**

   i) I plan to use a supervised learning approach owing to the fact that IR is more amenable to empirical tuning over a completely black-box approach of constructing an unsupervised mathematical model

   ii) From the experience of constructing an IR system for MP2.4, I have realized that Okapi-BM25 performed best for the cranfield dataset. While constructing different rankers, I have cycled through various models like Pivoted Length Normalization, INL2 ranking, Zhai's BM25+, Jelinek-Mercer, and Dirichlet smoothing models. After extensive and exhaustive tuning on each and every model, Okapi-BM25 emerged victorious over the others. I, however, didn't have the chance to implement learning-to-rank methods by combining more than one ranking models to obtain a superior performance. I would start with implementing one of the learning-to-rank methodologies described in the optional module of the course to pitch it against individual models to compare their performances.

   iii) In implementing learing-to-rank, I wish to choose a classification based learning to rank from among the different approaches. For this purpose, I'm currently exploring SVM^{map}. I may change my choice of LTRs and choose another model if I'm not satisfied with the performance of a previously chosen one. In addition to working on the datasets provided for the competition, I plan on using LETOR as a benchmark dataset for my model.

4. **Programming language option**
   I wish to implement the majority of the project in Python. However, I may use C++ and/or R in the due course of the project if need be the details of which I will thoroughly document in the user guide.

References

Liu,"Learning to rank for Information Retrieval", *Foundations and Trends® in Information Retrieval*, Vol. 3 No. 3 (2009) pp. 225–331, 2009

http://projects.yisongyue.com/svmmap/

https://www.microsoft.com/en-us/research/publication/letor-benchmark-collection-research-learning-rank-information-retrieval/