Query Understanding

Venkatesh Vinayakarao
Chennai Mathematical Institute
venkateshv@cmi.ac.in

1 INTRODUCTION

Given that the query boxes are typically short in a web search context, how to accurately predict what does the user want? This happens to be the objective of query understanding. Short queries may carry ambiguities. For example, the query "Taj" may refer to the entities Taj Mahal, Taj Tea or Taj Hotel. Researchers over the last decade have used a variety of approaches to improve query understanding. Query Segmentation, Spelling Correction, and Query rewriting are just some such efforts to improve query understanding.

2 CHALLENGES

The natural language ambiguities translate into query understanding challenges. A typical example is "plane". It is unclear whether we are referring to a flying machine or a geometric object. Queries may have multiple intents. For example, consider "US Presidents". Variety of query types may need different ways of handling them. For example, we see navigational, transactional, local ("*where can I find the best coffee in chennai?*" or "*coffee shop near me*"), informational and even QA queries. Also as search engines evolve, the way users query has also changed. This cycle of query understanding and retrieval techniques demands regular re-look at this topic.

Here, we focus on three aspects of query understanding namely conceptualization, capturing temporal context in queries and query intent classification. Embedding *intent* instead of query *terms* have also proven to give good results.

3 CONCEPTUALIZATION

One approach to query understanding lies in understanding the concepts relevant to the query intent. For instance, "*watch harry potter*" is most likely relevant to the concept *movie*. For short queries, simple parsing or even topic modelling do not perform well. Wang et al. [3] argue that external knowledge-bases and our ability to disambiguate are key to achieve this acceptable conceptualization.

4 TEMPORAL CONTEXT

Context plays a significant role in query understanding. A specific type of context is the temporal information contained in the query. About 1.5% of all explicit queries (*fifa 2014*) and 7% (*history of fi*fa) of implicit queries contain temporal information. Hasanuzzaman et al. [2] present an ensemble learning framework defined as a multi-objective optimization problem to attack this issue. One simple idea is to predict the query's temporal category class $c \in \{past, recent, future, atemporal\}$. In spite of having only four categories, such additional clues improve retrieval performance significantly.

5 QUERY ANNOTATION

Glator et al. [1] propose an approach to retrieval that first classifies the query intent and then uses an intent aware learning to rank models. For such query intent classification, they use standard classification algorithms such as AdaBoost and SVM. Retrieval model shows better results when the intent for each query is known. A popular way to categorize the intent is as follows:

(1) Entities (*florida*)
(2) Types (*coastal cities*)
(3) Question (*where is the nearest coffee shop?*)
(4) Others (*chennai to bangalore*)

With high quality knowledge-bases and human judgments as training data, the standard classifiers could achieve up to 77% accuracy. Further, the use of such annotations help the retrieval systems in achieving better precision, NDCG and MAP scores.

6 GENERIC INTENT REPRESENTATION

Queries with same results must be similar. This intuition allows us to embed queries in a distributed space such that similar queries are nearer to each other. Microsoft researchers [4] exploit this idea to build a GEneric iNtent Encoder (GEN Encoder). Authors use the example of queries "cheap cars" and "low-cost autos". Note that even though there are no overlapping terms, the intent could be the same. They use Bing logs to learn the embedding. An *ablation study* studies the performance of an AI system by removing certain components, to understand the contribution of the component to the overall system. Ablation studies reveal the crucial role of learning from implicit user feedback in representing user intent.

REFERENCES

[1] Rafael Glater, Rodrygo L.T. Santos, and Nivio Ziviani. 2017. Intent-Aware Semantic Query Annotation. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Shinjuku, Tokyo, Japan) *(SIGIR '17)*. Association for Computing Machinery, New York, NY, USA, 485–494. https://doi.org/10.1145/3077136.3080825

[2] Mohammed Hasanuzzaman, Sriparna Saha, Gaël Dias, and Stéphane Ferrari. 2015. Understanding Temporal Query Intent. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Santiago, Chile) *(SIGIR '15)*. Association for Computing Machinery, New York, NY, USA, 823–826. https://doi.org/10.1145/2766462.2767792

[3] Zhongyuan Wang, Kejun Zhao, Haixun Wang, Xiaofeng Meng, and Ji-Rong Wen. 2015. Query Understanding through Knowledge-Based Conceptualization. In *Proceedings of the 24th International Conference on Artificial Intelligence* (Buenos Aires, Argentina) *(IJCAI'15)*. AAAI Press, 3264–3270.

[4] Hongfei Zhang, Xia Song, Chenyan Xiong, Corby Rosset, Paul N. Bennett, Nick Craswell, and Saurabh Tiwary. 2019. Generic Intent Representation in Web Search. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (Paris, France) *(SIGIR'19)*. Association for Computing Machinery, New York, NY, USA, 65–74. https://doi.org/10.1145/3331184.3331198