Re-ranking Search Results via Interactive Disambiguation

Malcolm J. Mashig
mjm6jy@virginia.edu
University of Virginia
Charlottesville, Virginia, USA

Ryan Kann
rak3me@virginia.edu
University of Virginia
Charlottesville, Virginia, USA

ABSTRACT

In this paper, we examine the shortcomings of Google's search result rankings when particularly ambiguous queries are issued and we propose a process – requiring minimal user interaction – that clarifies queries before results are shown. We find that this process is a reliable and simple way to elevate the rankings of relevant search results. This process intuitively expands upon how Google's search engine responds to ambiguous queries with two key improvements. First, the process relies on the search engine – rather than the user – to determine if a query is too ambiguous. Second, the process prevents a diversion of the user's attention away from their original information retrieval intent. To illustrate the proposed process, we provide a prototype application. After users issue a query and clarify their query if deemed necessary, the application displays a re-ranked list of Google's search results and a log of how the rankings differ before and after the clarification process. *Please note that this draft of the paper is a proposal for the CS 6501 Information Retrieval final project at the University of Virginia.*

KEYWORDS

Information Retrieval, Natural Language Processing, Named-Entity Disambiguation, Unsupervised Clustering, Fair Ranking

1 INTRODUCTION

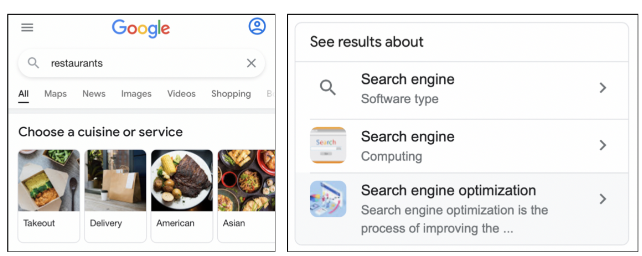
In this section, we will introduce the paper with a brief summary of the motivation, methodology, and results.

2 MOTIVATION

To demonstrate the need for a query-clarification process, please consider this simple motivating example: "columbia" as a search query. If you google this, you will find the first page of results contains websites for the outerwear company, the university in New York, the college in Chicago and Missouri, the city in South Carolina and Missouri, and the record label – all of which bear the "columbia" name. Websites for Columbia Bank, other cities bearing the name, and the twitter account for Columbia University all find themselves on pages beyond the first. Yet a user intending to reach any of these sites likely only issues the one-word query expecting it to be sufficient to find what they are looking for and he/she likely does not venture beyond the first page. The ranking of these search results would fail this user's attempt at retrieving relevant information, simply because the search engine was supplied an ambiguous query. Rather than relying on the user to determine that their query was unclear, a more intuitive resolution is for the search engine to determine this and to actively seek clarification from the user when necessary. In this particular situation, the search engine might present categories such as "location", "clothing", "education", and "other" for the user to select from, before displaying the search results. The search engine might next present "South Carolina", "Missouri", and "other" if the user selects "location" or it might show results (without further clarification) if "clothing" is selected. This is notably an ideal clarification sequence that might be difficult to systematically implement, but it nonetheless motivates the extensive integration of interactive disambiguation as a search engine feature.

To demonstrate how a query-clarification process of this nature leads to more fair rankings, please consider another simple motivating example: "restaurants" as a search query. If you google this, the top search results include a few lucky local places to eat, some likely serving cuisine that is of no interest to you, but nonetheless receiving your attention and perhaps distracting you from your original intent – to explore restaurants serving your preferred cuisine. This is implicitly unfair, because the display of higher-ranked but undesired results implies that lower-ranked but desired results are not being displayed and thus are not being brought to your attention. A simple query-clarification process whereby the user selects from different cuisine categories before seeing results might ensure more fair rankings. When issuing this query on a mobile device, Google actually does allow users to choose the cuisine type at the top of the page. For other queries, Google offers similar features for the purpose of disambiguation. Searching "search engine", for example, leads to a side-bar display allowing the user to see results about search engines as they relate to "Software type", "Computing", or "Search Engine Optimization".

Figure 1: Examples of Interactive Disambiguation



Google's adoption of such features adds credibility to the usefulness of interactive disambiguation, but it also motivates an improvement and expansion of the process. While Google allows users to clarify simple and frequently-searched queries that are known to be ambiguous, it does not allow for clarification on informal or infrequently-issued queries, even if they are ambiguous. The

CS 6501 Information Retrieval, University of Virginia, Spring 2021

ultimate motivation for this paper is to explore a wider adoption (an extension) of informal, interactive query-clarification. The secondary motivation is to demonstrate a more fair implementation of this process, whereby search results are only displayed after queries are sufficiently clarified. These motivations notably highlight the novelty of this work.

3 METHODOLOGY

In this section, we will discuss the main components behind the proposed query-clarification process. These components include:

(1) Text tokenization as a means to extract the subject matters from search results
(2) Multi-round unsupervised clustering as a means to populate clarification categories
(3) Online reinforcement learning as a means to refine clarification categories over time

3.1 Text Tokenization

We begin the process with a list of search results (URLs, titles, descriptions) for a certain query, as ranked by Google, and the first task is to efficiently process text at the title-level, description-level, and website-level. In this first stage, we will need to tokenize the text in a fashion that removes stopwords, performs stemming, and summarizes the key words / terms found in the text – outside of those in the query itself. Words in the title should be weighted more than words in the description and words in the description should likewise be weighted more than words within the website's content. This stage will result in a list of key words / terms and their degree of importance for each search result.

3.2 Multi-round Unsupervised Clustering

In this stage, we conduct a cluster analysis of key words / terms determined in the previous stage in order to identify those that are most helpful in distinguishing different search results. This will require the identification of words / terms that are important in a significant portion of search results while also being unimportant in another significant portion of search results. If there are no such words / terms, then the level of ambiguity will be deemed negligible and further clarification will be deemed unnecessary. Ultimately, this stage will provide a small set of different clarification categories that may be presented to the user. These are the words / terms that terminate as centroids for the identified clusters. Multiple rounds of cluster analysis will be conducted for each path of clarification that a user might follow.

3.3 Online Reinforcement Learning

In the last stage of the process, we organize a reinforcement learning framework that will decide the exact set of categories that are presented to users. The framework will decide categories based on their utility and an element of randomness. The framework will keep track of how users interact with the categories in order to identify how useful each of the categories are. The integration of a partially-random / probability-based selection of categories will assist in the exploration of potentially helpful categories. Meanwhile, categories that are selected more often by users will be exploited (presented with higher probability). This stage will thus actively refine the best subset of categories for each query.

4 EVALUATION

In this section, we will evaluate the results of the interactive query-clarification process. Specifically, we will evaluate 1) how helpful the clarifying categories are and, assuming they are helpful, 2) how well the relevant search results are re-ranked. To determine 1) the general utility of the clarifying categories presented to the user, we will measure similarity between the categories identified via unsupervised clustering and the categories listed by Google Trends [1] in the "Top Related Topics" section. These topics can be extracted via a Google Trends API wrapper [6]. To determine 2) the improvement achieved by re-ranking, we will measure similarity between the re-ranked search results and the search results listed for a query combining the original query and the clarifying category. Beyond this, we will collect metrics that indicate the overall influence of the query-clarification process. For instance, one such metric might be the average sum of change in rankings of the top n search results. We might determine this by simulating a large number of random queries and subsequent clarification processes. Another metric that we might determine in a similar way is the average number of clarifying categories and stages of clarification that are presented to users. This metric would indicate how burdensome the clarification process is for users.

5 PROTOTYPE APPLICATION

In this section, we will provide a link to the prototype application and a more in-depth explanation of the application's features. This will include a discussion of the tool used to scrape Google search results (URLs, titles, and descriptions) [3] as well as the tool used to scrape text from websites listed in the search results [2].

6 RELATED WORK

In this section, we will summarize related research papers in the area of interactive / conversational recommendation such as *Asking Clarifying Questions in Open-Domain Information-Seeking Conversations* [4] and *Recommendation as a Communication Game: Self-Supervised Bot-Play for Goal-oriented Dialogue* [5]. We will then discuss how the research presented in this paper expands upon, improves upon, or differs from related work.

7 CONCLUSION

In this section, we will summarize the paper and discuss potential future directions for the research.

REFERENCES

[1] 2021. Google Trends. (2021). https://trends.google.com/trends
[2] Hugo Bowne-Anderson. 2017. *Web Scraping NLP in Python.* (2017). https://www.datacamp.com/community/tutorials/web-scraping-python-nlp
[3] Ping Shiuan Chua. 2018. *Scraping Search Results from Google Search.* (2018). https://www.pingshiuanchua.com/blog/post/scraping-search-results-from-google-search
[4] Aliannejadi et al. 2019. *Asking Clarifying Questions in Open-Domain Information-Seeking Conversations.* (2019). https://dl.acm.org/doi/pdf/10.1145/3331184.3331265
[5] Kang et al. 2019. *Recommendation as a Communication Game: Self-Supervised Bot-Play for Goal-oriented Dialogue.* (2019). https://arxiv.org/pdf/1909.03922v1.pdf
[6] Philippe Massicotte. 2021. {gtrendsR}. (2021). https://github.com/PMassicotte/gtrendsR