Technology Review:
Google's multitask ranking system
CS 410
Shih-Chiang Lee

We learned from lesson that ranking is very important for accessing, whether the accessing is pull mode like search engines or push mode like recommender system. Google, one of the best searching engine and recommendation system in the word, has a powerful multitask ranking system to process data and users from all over the world. This article is going to discuss how google multitask ranking system works well and my opinions.

For get the most relative results, google must have a set of algorithms is specifically used to determine which results should rank which search terms. To maintain a leading position in the industry, its algorithms are always evaluating. Originally, Google was using page rank algorithm, but it changed to use Hummingbird algorithm in 2013 [1]. Google did not reveal too many technical details about the Hummingbird algorithm. Instead of parsing and searching word by word, Hummingbird algorithm can parse the entire problem more quickly. However, in this set of algorithms, page rank still plays a big role, but it is no longer the only ranking mechanism.

In both searching engines and recommendation system, ranking systems are used to provide the most relative results. The network environment is more and more complex so limited keywords are now not enough to do an effective ranking.

There are several websites post the same article. Which one should get higher score in the ranking system? For example, BBC posts a news, and there is also a private blog reprint the news. When users search about the news, they usually prefer to read the news form the BBC's official website. Besides, the websites which are cited by BBC, will also get higher ranking scores. Therefore, content & functionality of the site and the web page's authority [2] are also significant for the ranking.

In addition, time and location are also needed to be considered. Most information has timeliness, so users usually want to get the last information [2]. When search about weather, most people only want to know the weather today or future, but not past. And location is very important in both searching engine and recommendation system. It facilitates the search and improve the effectiveness of recommendations [3]. For example, when you search just "lunch" in google map, only the local restaurants showed up.

The ranking system should use feedback to optimize its results. In searching engine, the optimization of ranking for a unique user is based on the user's history clicking. The user-clicked websites are assumed to be relevant. The users' visit history, preference and comments [4] could also be a kind of feedback. In recommendation system, it can quickly do the filtering to get the relevant pages.

There are so many elements can affect a page's ranking score. These elements make the algorithm more correct and efficient. However, they may also make the objective function conflict [5]. Besides, the user may click the page just because it has high ranking score when the ranking queue without any other reason. It does not mean that it is relevant to user, and this would cause implicit bias, like position biases [5].

How to solve the objective function conflict and implicit bias? There is a great idea using Multi-gate Mixture-of-Experts (MMoE) to avoid objective function conflict, and using a proposed model architecture which is similar to the Wide & Deep model architecture to dual with implicit bias [5].

The reason why using Mixture-of-Experts (MoE) is that the model is get bigger and the training samples are increasing. Each sample needs to go through all the calculations of the model, which leads to a quadratic increase in training cost. But MOE can split the large model into multiple small models [6]. And for each sample, it is not necessary to go through all the small models to calculate, but only to activate a part of the small models for calculation, which saves computing resources. "MMoE is a soft-parameter sharing model structure designed to model task conflicts and relations." MMoE uses many gates to control if an objective function shares the same experts with other objective function when calculate different samples during the training, which solve the conflict while reducing the training cost[5].

In the model architecture, a shallow tower is trained with features contributing to implicit bias [5]. The bias output will become a part of sample to train the model remove the bias. However, I consider that this can only work on recommending system, not search engine. In the recommending system, users can return feedback to let the system know the recommendations are not relevant, but in searching engine is hard. There is no relevance feedback. That a page is relevant or not, may need user to read most of it. The user may spend same time on both relevant and irrelevant pages so time also cannot be feedback to judge the bias. I suggest that google can set weekly feedback for user on chrome. Let the users to rate how relevance the pages they get in the past week, and return the searching history, browsing records and rates. The bias model can focus on training the sample which users rate it low.

In conclusion, Google's multitask ranking system is very efficient. For a ranking system, the more factors which affect the ranking result does not mean that the result become more relevant, because of objective function conflict. However, Google's ranking system can combine so many objective functions with less conflict. In addition, although it is working on solving bias problems, it is still not enough and can be better. I suggest that more relevance feedback can be added to the ranking system. This may increase the training cost and slightly affect user experience, but I think that removing the bias is worth.

Reference

[1] Moz, Inc., "Google Algorithm Update History: A History of Major Google Algorithm Updates from 2000–Present", July 1, 2021, https://moz.com/google-algorithm-change

[2] B. Guelbert, "How Google Determines Your Ranking on Search Results Pages", October 28, 2019, https://clutch.co/seo-firms/resources/how-google-determines-ranking-search-results-pages

[3] Google Search Help, "Understand & manage your location when you search on Google", https://support.google.com/websearch/answer/179386?hl=en-GB

[4] Google Ads Help, "About recommendations", https://support.google.com/google-ads/answer/3448398?hl=en

[5] Z. Zhao, L. Hong, L. Wei, etc, "Recommending What Video to Watch Next: A Multitask Ranking System", Association for Computing Machinery, New York, NY, United States: pp. 43-51

[6] T Baldacchino,E. Cross, K. Worden, J. Rowson, "Variational Bayesian mixture of experts models and sensitivity analysis for nonlinear dynamical systems". Mechanical Systems and Signal Processing. 66–67: 178–200.