<u>Tech Review</u>

Video Ranking Recommendations: Challenges and Solutions

Ranking recommendation systems are extremely important, especially in recent years given the sharp rise in online video services that are used worldwide. Businesses that offer streaming services such as Netflix and Amazon have grown their online offerings in the past decade. As online video medium businesses surge, they need to adapt to make the user experience as easy and appealing as possible, and businesses want to do that through recommendation systems in order to keep the user engaged and make it easy for them to quickly identify the content that they enjoy watching. However, in order to make video recommendation systems effective, one must overcome the challenges. The main challenges include the following: multiple competing ranking objectives, biases in feedback, along with other factors to consider such as efficiency and scale. Luckily, researchers have identified ways to address these challenges as best as possible.

First, what are ranking objectives and how can they be addressed? There can be multiple objectives or targets to consider when ranking videos for billions of users. The objectives or targets may include such things as recommending videos for a user to watch based on watch time, videos that are rated highly, or click rate. If you optimize best on watch time, click rate decreases [2]. To solve this challenge, the researchers used Multi-gate Mixture-of-Experts, a multi-task learning approach that explicitly learns to model task relationships from data [3]. **It's efficient [3].** It's essentially a complex machine learning world. The MMoE approach estimates multiple types of user behaviors [1, pg 1]. It uses MMoE to automatically learn parameters to share across potentially conflicting objectives. Then by utilizing multiple gating networks, each of the objectives can choose experts to share or not share with others [1, pg 1]. The gaiting allows the model to learn a per-task and per-sample weighting of each of the expert networks, which works well with multiple competing tasks [4]. According to [1], engagement and satisfaction objectives were grouped and those metrics were shown to be greatly improved using the MMoE model [1]. As a result, this method of ranking objectives worked really well and helped address this challenge.

The other main challenge in ranking systems is selection bias, particularly position bias [1]. For example, if a system provides the top ten results for a user query, a user may simply choose the top one, thinking the system has chosen the best match, when in reality this is not the case. The ranking system logs the user has clicked on the video and logged that as a binary relevance judgement. This is fed to the system in future similar queries and it enters an erroneous feedback loop for the user[1]. To address this challenge, researchers found that adopting a Wide & Deep framework helped control or mitigate bias. In other words, it was proposed to add a shallow tower to the main model MMoE model. It corrects ranking selection bias with a side-tower [1]. In detail, it mitigates bias by directly learning the shallow tower together with the main model, which enables learning the selection bias [1]. Basically, they get a user-utility component and factor in learning a bias component -- together it prevents the overall model from leaning too much on position for ranking (avoiding position bias) [1, pg6]. As a result they observed improved reduction in bias in a ranking system.

Lastly, other challenges that come into the mix are challenges of efficiency and scale. There are billions of users for a large scale video platform such as netflix and youtube. These systems much serve ranked recommendations to billions of users across a multitude of video content in a moment, which could be a matter of seconds for a good user experience. There is a trade off of quality and efficiency and scale. Companies must perform the balancing act and decide where to move the needle when judging between the two. The group were able to manage efficiency and scaling by using x, y, z in their experiment. They saw good results and tested with Youtube, a large scale platform to get a real idea of performance of such a large system. Their results were a,b,c in using their system configuration and deep learning models. It is known in the community that quality, efficiency, and scale is still an ongoing topic to further research and dive into for future improvements [cite]. So while a,b may be good to solve the efficiency and scaling factors, there may be other solutions available to ensure the proper ranking along with satisfying many users quickly.

In conclusion, video ranking recommendation systems have challenges which need to be addressed for success. There are challenges of multiple objectives that compete with each other when finding candidates, and there are challenges of bias that are introduced in systems during the ranking phase when results are presented to the users. All the while, there are concerns of managing efficiency and scale, something that businesses today must keep in mind to ensure customer retention to achieve revenues. A ranking system that uses the configuration system of a,b,c addresses the multiple objective concerns and challenges, while the configuration of f,g,h addresses the concerns for selection bias, according to the aforementioned experiments run on Youtube, a large scale video recommending platform -- all the while maintaining a level of efficiency and scale. The group addressed this in their conclusion. Therefore, with the provided results we see a way to mitigate the challenges and create a ideal video ranking system that is efficient and scalable; however, there’s room for improvement with further research to ensure efficiency and scalability.

Works cited

[1] Recommending What Video to Watch Next: A Multitask Ranking System. *Zhe Zhao, Lichan Hong, Li Wei, Jilin Chen, Aniruddh Nath, Shawn Andrews, Aditee Kumthekar, Maheswaran Sathiamoorthy, Xinyang Yi, Ed Chi*. Google, Inc.

[2] Online Article Ranking as a Constrained, Dynamic, Multi-Objective Optimization Problem *Jeya Balaji Balasubramanian, Akshay Soni, Yashar Mehdad, Nikolay Laptev*

[3] Modeling Task Relationships in Multi-task Learning with Multi-gate Mixture-of-Experts *Jiaqi Ma1, Zhe Zhao, Xinyang Yi, Jilin Chen, Lichan Hong, Ed H. Chi*

[4] https://towardsdatascience.com/multi-task-learning-with-multi-gate-mixture-of-experts-b46efac3268