Yunchang Pang

An overview on Google's multitask ranking system

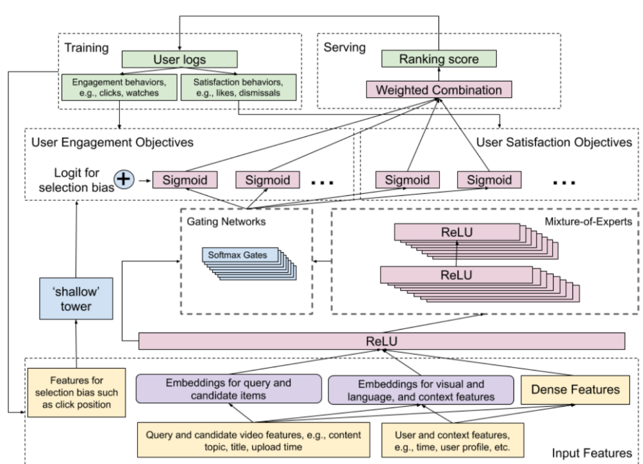
Intro

Video platforms like YouTube play an important role in our lives and it sometimes can be bizarre when we look at a large amount of time we have spent on them. This is partly because these video platforms can always recommend people with high relative content on the sidebar that intrigues people to see further into the topic.

However, what criteria should we base our recommendations on? There are so many aspects we can take into consideration, like users' clicks and users' likes, but is there a way we can put them all into a model for the recommendation? This overview will review Recommending What Video to Watch Next: A Multitask Ranking System. Zhao et al. gave us a glimpse into their ranking system for recommending videos to users.

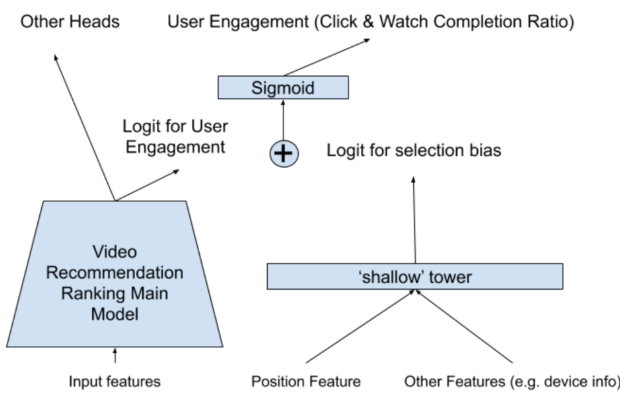
Main Challenges and Solutions

- Conflict objectives caused by events like popular content recommendations and shared content clicked by users. For example, users could click on a shared link that their friends find interesting but are not users' normal preferences [1].
- Implicit bias exists when users tend to click videos that have a higher ranking. For example, a user could possibly click on a video because the video is at the top of the recommendation list [1].

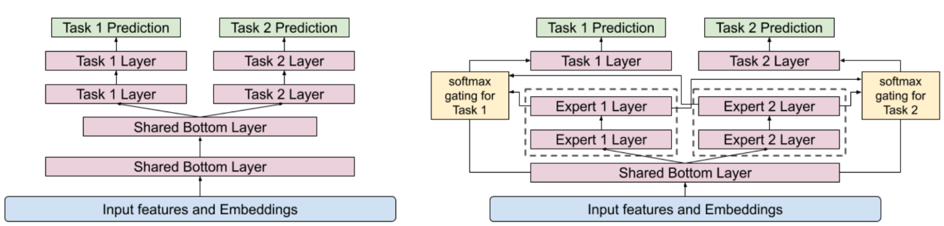


The solution here is Zhao et al. group behaviors into two categories, of which one is users' engagement objectives like clicking on a video and the other is users' satisfaction objectives, for example, promoting a video. Then adopting Multi-gate Mixture-of-Experts (MMOE) for multitask learning [1]. The MMOE here is a model proposed by Ma et al. that could learn multiple tasks in a single model by explicitly learning the model relationships from data[2]. To reduce the implicit bias caused by different positions, they add a "shallow tower," which consists of the bias created by the current system. The architecture is shown above.

Application



As the figure shown above, the idea of removing selection bias created by position is realized by dividing model prediction into the user-utility component from the main tower and a bias component from the shallow tower. The bias component, for example position selection bias, will be sent into the shallow tower training the tower to identify if the log contains selection bias. Then the output from the shallow tower will be used as a parameter with original data into the main tower and will drop out those with selection biases to prevent them from impacting the ranking model [1].



(a) Shared-Bottom Model with shared bottom hidden layers and separate towers for two tasks.

(b) Multi-gate Mixture-of-Expert Model with one shared bottom layer and separate hidden layers for two tasks.

The setup includes training and building models with TensorFlow and sequentially training the proposed model and baseline model [1]. The main difference between the baseline model and the newly proposed model is that the new model used MMOE with one shared bottom layer and separate layers for multiple tasks. The differences between the two models are shown in the above figure. The engagement metrics the paper used include time spent on YouTube and satisfaction matrices used indicators such as survey responses [1].

Model Architecture	Number of Multiplications	Engagement Metric	Satisfaction Metric
Shared-Bottom	3.7M	/	/
Shared-Bottom	6.1M	+0.1%	+ 1.89%
MMoE (4 experts)	3.7M	+0.20%	+ 1.22%
MMoE (8 Experts)	6.1M	+0.45%	+ 3.07%

The above table shows the performances proposed model which included 4 experts and 8 experts. We can see from the data that MMOE using 4 experts and 8 experts are doing better with both engagement metric and satisfaction metrics compared to the base model.

Conclusion

The article provided a new thought on how to recommend videos to users accurately. The use of MMOE increased the model's recommending ability and the use of the "shallow tower" is also a method to reduce selection bias efficiently.

References

[1] Zhe Zhao, Lichan Hong, Li Wei, Jilin Chen, Aniruddh Nath, Shawn Andrews, Aditee Kumthekar, Maheswaran Sathiamoorthy, Xinyang Yi, and Ed Chi. 2019. Recommending what video to watch next: a multitask ranking system. In Proceedings of the 13th ACM Conference on Recommender Systems (RecSys '19). Association for Computing Machinery, New York, NY, USA, 43–51. https://doi.org/10.1145/3298689.3346997

[2] Jiaqi Ma, Zhe Zhao, Xinyang Yi, Jilin Chen, Lichan Hong, and Ed H. Chi. 2018. Modeling Task Relationships in Multi-task Learning with Multi-gate Mixture-of-Experts. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18). Association for Computing Machinery, New York, NY, USA, 1930–1939. https://doi.org/10.1145/3219819.3220007