An Introduction of Learning to Rank Approaches

Ryo Takaki

1. Introduction

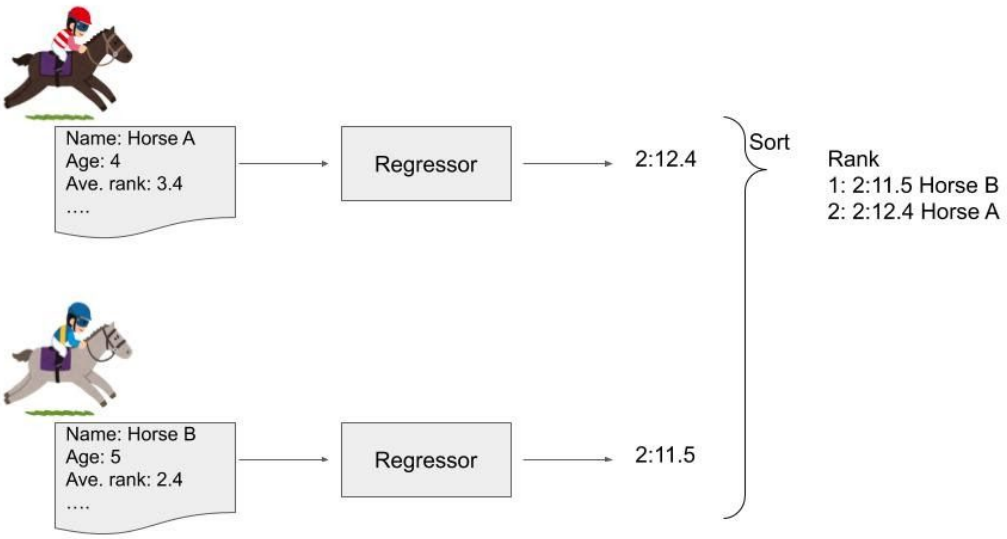
In this technology review, we will introduce major approaches to Learning to Rank (LTR) by using an analogy of a result prediction of a horse-race. In addition to introducing the basic approaches, practical methods widely used for Information Retrieval (IR), Natural Language Processing (NLP) and Data Mining (DM) will be introduced, so that main readers(CS410 students) not only can understand the basic concepts of LTR intuitively but also can dive into more practical techniques.

2. What is Learning to Rank

Learning to Rank is one of the approaches to solve ranking problems. The main difference between the traditional prediction problem(classification or regression) and the ranking problem is that the prediction problem only cares about a predicted class or score itself, but the ranking problem cares about an order of those scores. To take an example of the prediction of a horse-race, the traditional approach predicts a time record of each horse but the ranking problem only cares about the ranking (order of the time records).

3. Pointwise approaches

Pointwise approaches take a single input (e.g. one horse's information) to predict a score (e.g. time record) . A classifier or regressor is trained to minimize a loss function (e.g. time difference between the predicted time and an actual time). The final ranking is calculated by sorting the each scores (e.g. the ranking of the horse race is achieved by sorting the time of each horse). As only one input is treated to train/predict, group structure of ranking is not taken into account (e.g. a horse A's information does not affect a prediction of a horse B's time).

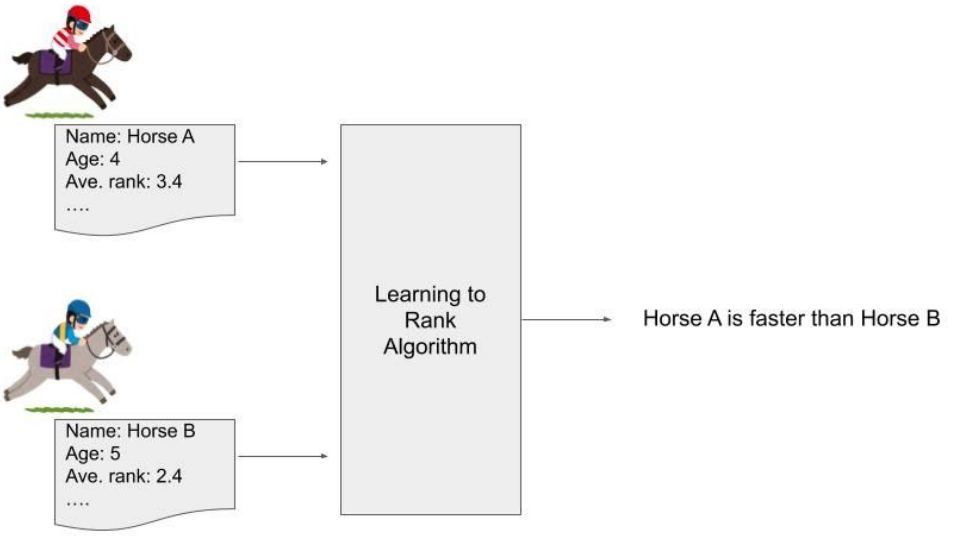


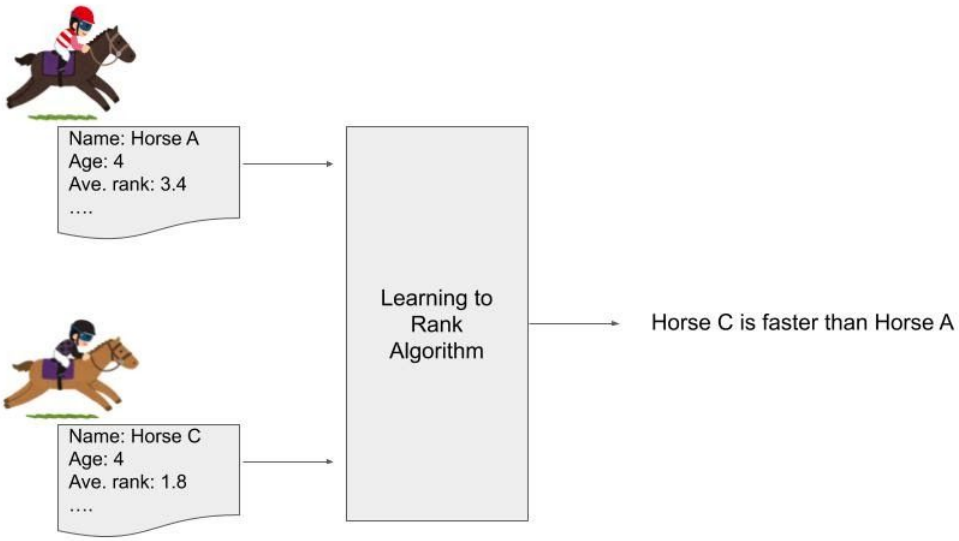
Practical Methods for IR, NLP and DM:

- Subset Ranking
- McRank
- Prank
- OC SVM

4. Pairwise approaches

Pairwise approaches take not a single input but a pair of inputs (e.g. two horses information) to predict a relative order of pairs (e.g. which horse is faster than another horse). In general, the pairwise approaches are better than the pointwise approaches, however they do not always optimize the entire ranking as they optimize the relative order of the pair but entire orders.



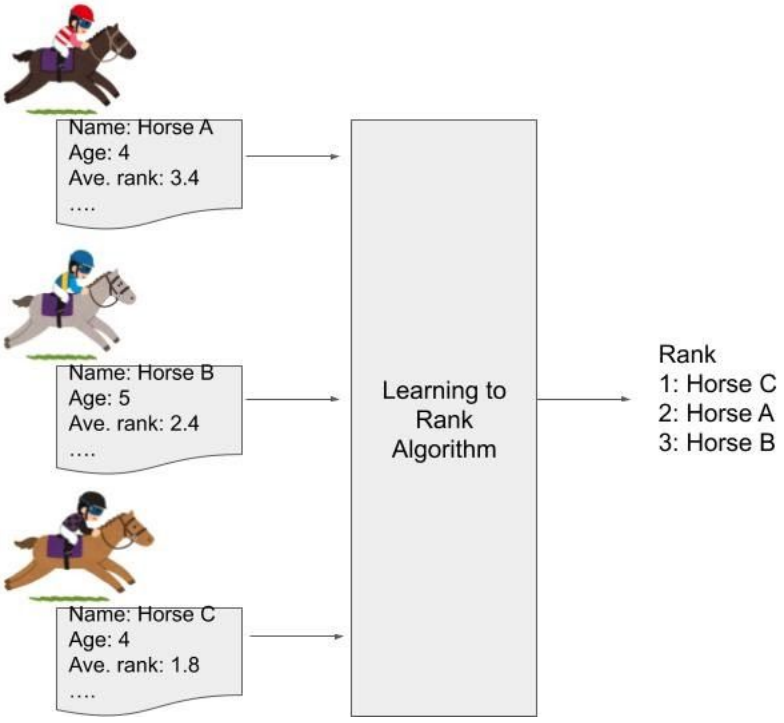


Practical Methods for IR, NLP and DM:

- Ranking SVM
- RankBoost
- RankNet
- GBRank
- IR SVM
- Lambda Rank
- LambdaMART

5. Listwise approaches

Listwise approaches also output the order of the scores, however the input is not just a pair but an entire list. This approach can handle the ranking problem in a more straightforward way by inputting all inputs at once and optimizing the final ranking with certain criteria same as an evaluation method. However, a loss function for this approach tends to be non-continuous. Mathematical techniques such as modifying the non-continuous function to continuous function is used to optimize the loss function.



Practical Methods for IR, NLP and DM::

- ListNet
- ListMLE
- AdaRank
- SVM MAP
- SoftRank

6. Conclusion

In this technology review, pointwise approaches, pairwise approaches and listwise approaches are introduced to handle the ranking problem. Typically, the pairwise approaches are better than pointwise approaches, however, the pairwise approaches do not always optimize the entire ranking. Compared to the pairwise approaches, the listwise approaches can handle the ranking problem in a more straightforward way, however some mathematical techniques such as modifying the non-continuous function to continuous function to optimize the loss function is needed.

7. References

- http://times.cs.uiuc.edu/course/598f14/l2r.pdf
- https://medium.com/@nikhilbd/intuitive-explanation-of-learning-to-rank-and-ranknet-lambdarank-and-lambdamart-fe1e17fac418
- https://medium.com/@nikhilbd/pointwise-vs-pairwise-vs-listwise-learning-to-rank-80a8fe8fadfd
- https://mc.ai/pointwise-pairwise-and-listwise-learning-to-rank/
- https://www.cl.cam.ac.uk/teaching/1516/R222/l2r-overview.pdf
- https://www.cs.cornell.edu/people/tj/svm_light/svm_rank.html