Does LambdaMART Do What You Expect?

Discussion Paper

Claudio Lucchese¹, Federico Marcuzzi¹* and Salvatore Orlando¹

¹Università Ca’ Foscari Venezia, Via Torino, 155, 30170 Mestre, Venezia VE, Italy

Keywords
learning to rank, information retrieval, machine learning

Optimising information retrieval metrics is crucial in Learning-to-Rank tasks. One of the most effective approaches to do this is the well-known LambdaMART [1] algorithm. LambdaMART smoothly approximates for each document a gradient with respect to the IR metric to be optimised, the so-called lambdas. Intuitively, each lambda describes, with some degree of approximation, how much a document score should be pushed up or down to improve the ranking.

In this work, we show that lambdas can be incoherent with respect to document relevance: a document with high relevance can receive a downward push larger than a document with lower relevance. This behaviour goes far beyond the expected degree of approximation. In addition, despite the gain in training efficiency, directly optimising truncated evaluation metrics can exacerbate this discrepancy due to less number of lambdas contributions each document gradient receives. As a consequence each document gradient is only partially computed and this lead to worse model learning.

We analyse the idiosyncrasies of LambdaMART gradients and we introduce some strategies to remove or reduce gradient incoherencies. Specifically, we designed three selection strategies to compute the full gradient for only those documents that should be ranked in the top-\(k\) positions of the ranking. We empirically demonstrate on publicly available datasets that the proposed approach leads to models that can achieve statistically significant improvements in terms of NDCG [2] while maintaining the same training efficiency as optimising truncated metrics.

References