RankBoost on LETOR

Introduction to RankBoost

The basic idea of RankBoost is to formalize learning to rank as a problem of binary classification on instance pairs, and then to adopt boosting approach. Like all boosting algorithms, RankBoost trains one weak ranker at each round of iteration, and combines these weak rankers as the final ranking function. After each round, the document pairs are re-weighted: it decreases the weight of correctly ranked pairs and increases the weight of wrongly ranked pairs.

The details of RankBoost can be found from this JMLR paper.

Learning Parameters

We define each weak ranker on the basis of a feature. With a proper threshold, the weak ranker has binary output, i.e., it takes values from {0, 1}. For each round, we select the best weark ranker from (# of features) x (255 thresholds) candidates.

The number of weak rankers is determined by cross validation.

Dataset	# of weak rankers (from Fold1 to Fold5)
OHSUMED	300, 100, 50, 50, 50, 300
TD2003	300, 150, 250, 100, 300
TD2004	100, 300, 150, 150, 150
HP2003	50, 150, 250, 50, 50
HP2004	50, 100, 200, 100, 300
NP2003	50, 50, 100, 50, 100
NP2004	50, 100, 200, 100, 300

Papers & Docs

Y. Freund, R. Iyer, R. E. Schapire, and Y. Singer. An efficient boosting algorithm for combining preferences. J. Mach. Learn. Res., 4:933-969, 2003.

BibTex

```
@article{964285,
   author = {Yoav Freund and Raj Iyer and Robert E. Schapire and Yoram Singer  },
   title = {An efficient boosting algorithm for combining preferences},
   journal = {J. Mach. Learn. Res.},
   volume = {4},
   year = {2003},
   issn = {1533-7928},
   pages = {933--969},
   publisher = {MIT Press},
   address = {Cambridge, MA, USA},
}
```

Notes

This document was written by Yong-Deok Kim, and the experiments were conducted by Yong-Deok Kim. If any problem, please contact letor@microsoft.com