Improve Ranking Efficiency
by Optimizing Tree Ensembles

Claudio Lucchese[1,3], Franco Maria Nardini[1,3], Salvatore Orlando[2],
Raffaele Perego[1,3], Fabrizio Silvestri[4], and Salvatore Trani[1,5]

[1] ISTI-CNR, Pisa, [2] University Ca' Foscari of Venice,
[3] Istella Srl, [4] Yahoo London, [5] University of Pisa.

Abstract. Learning to Rank (LtR) is the machine learning method of
choice for producing highly effective ranking functions. However, effi-
ciency and effectiveness are two competing forces and trading off effec-
tiveness for meeting efficiency constraints typical of production systems is
one of the most urgent issues. This extended abstract shortly summarizes
the work in [4] proposing CLEaVER, a new framework for optimizing
LtR models based on ensembles of regression trees. We summarize the
results of a comprehensive evaluation showing that CLEaVER is able
to prune up to 80% of the trees and provides an efficiency speed-up up
to 2.6x without affecting the effectiveness of the model.

Modern search engines are expected to return highly relevant results in a
fractions of seconds to satisfy efficiency constraints. Learning-to-Rank (LtR) [1]
methodologies are nowadays pervasively used as effective solutions to ranking
problems. However, efficiency and effectiveness are intertwined concepts than
often counteract each other. In this extended abstract we shortly summarize
the work in [4] where we introduce CLEaVER, a framework developed on top
of QuickRank [5], for the optimization of LtR models based on ensembles of
regression trees *after* the learning phase has completed. Since document scoring
cost by using a tree ensemble model is linear in its size, CLEaVER first removes
a subset of the trees, and then fine-tunes the weights of the remaining ones
according to a given quality measure. Results of a comprehensive evaluation
using QuickScorer [2, 3], a state-of-the-art algorithm for efficient scoring, show
that CLEaVER is able to improve the efficiency of a given ranking ensemble
up to a 2.6x speed-up factor without affecting the effectiveness of the model.

References

1. T.Y. Liu. Learning to rank for information retrieval. *Foundations and Trends in
   IR.* 2009.
2. C. Lucchese, F. M. Nardini, S. Orlando, R. Perego, N. Tonellotto, and R. Ven-
   turini. QuickScorer: A fast algorithm to rank documents with additive ensembles
   of regression trees. In *ACM SIGIR.* 2015.
3. C. Lucchese, F. M. Nardini, S. Orlando, R. Perego, N. Tonellotto, and R. Ven-
   turini. Exploiting CPU SIMD Extensions to Speed-up Document Scoring with
   Tree Ensembles. In *ACM SIGIR.* 2016.
4. C. Lucchese, F. M. Nardini, S. Orlando, R. Perego, F. Silvestri, S. Trani. Post-
   Learning Optimization of Tree Ensembles for Efficient Ranking. In *ACM SIGIR.*
   2016.
5. G. Capannini, C. Lucchese, F. M. Nardini, S. Orlando, R. Perego, N. Tonellotto.
   Quality versus efficiency in document scoring with learning-to-rank models. In
   Information Processing & Management. 2016.