Final Report - P4

Boosting

Shangwen Yan | N17091204 | sy2160

Is Boosting amenable to parallel processing? Justify or show proof.

yes.

The basic concept of boosting is: using a set of weak learners to create a single strong learner. A weak learner is defined to be a classifier that is only slightly correlated with the true classification (it can label examples better than random guessing). In contrast, a strong learner is a classifier that is arbitrarily well-correlated with the true classification

So the training process could be seen as training several models respectively. For each model, we could use several nodes to do the parallel processing just like not using boosting. And then, after getting the optimal parameters of each model (training process is done), we could gather them together, and make it a stronger classifier.

List 3 ways in which you will improve this course from student perspective.

1. add specific TF-IDF demos on NLP topic
2. mongo is not that useful? Maybe it could be removed.
3. the slides or reading materials could be more tidy and better ordered?