
Addendum for “Estimating Accuracy from Unlabeled Data”

Emmanouil Antonios Platanios
Machine Learning Department
Carnegie Mellon University
Pittsburgh, PA 15213
e.a.platanios@cs.cmu.edu

Avrim Blum
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213
avrim@cs.cmu.edu

Tom Mitchell
Machine Learning Department
Carnegie Mellon University
Pittsburgh, PA 15213
tom.mitchell@cs.cmu.edu

1 EXPERIMENTS

In the experiments section of our paper we mention that for the AR method we are solving a convex optimization problem. In fact, we can see that the objective function is convex with respect to each variable separately, but it is not convex with respect to all variables together. We did notice, however, that the solution of the optimization problem was always the same, no matter what initial point we chose, as long as it satisfied the inequality constraints. This led us to believe that the objective function may be convex in the feasible set. However, we have not managed to formally confirm that result and so the corresponding statement in our paper is not necessarily true.

Also, in the experiments section of our paper we accidentally omitted to discuss one important detail. The NELL data set was “filtered” before running the experiments so that the two potential binary classification outputs were equally balanced. The reason for that was that the two classes were highly imbalanced (i.e., it is much more likely that a particular noun phrase does not belong to a specific category, rather than that it does). The “filtering” was performed by keeping all positive-labeled examples for each category and sampling an equal number of negative-labeled examples to consider (the rest of the negative-labeled examples were simply ignored). The results of the experiments performed when using all of the data samples and without “filtering” the data set in that manner are shown in the following table:

Table 1: Mean absolute deviation (MAD) of individual (Ind.), pairwise (Pair.) and all function error rates for the NELL data set, for all three proposed methods and for the cases where we use all of the available data samples.

$\times 10^{-2}$	All Data Samples		
	Ind.	Pair.	All
AR	0.92	0.69	0.66
MLE	58.43	58.27	47.03
MAP	5.96	5.75	5.59