Design and Analysis of Fast Algorithms for Interactive Machine Learning

Bhatia, Jagdeep (School: Watchung Hills Regional High School)

Interactive learning is a data efficient sub-field of machine learning where instruction happens through a series of interactions between a learner and a teacher. Recent research has shown that interactive learning in the Learning from Random Counter-examples (LRC) model is possible in O(log|H|) optimal average learning time for any concept class H, however, the Max-Min algorithm involved is complex. Moreover, an open question remains as to whether there exists an efficient randomized interactive learning algorithm. In addition, the learning rate of an arbitrary interactive learner remains unknown. This research addresses all these challenges by designing novel LRC algorithms and analyzing their performance using mathematical tools from probability, statistics, and computational learning theory. First, it shows a simple LRC algorithm based on majority vote that not only learns in O(log|H|) optimal average time, but also generalizes to non-binary concepts. Second, it solves the open problem with a randomized LRC algorithm and establishes that its performance is also optimal. Finally, it proves that the expected learning rate of any arbitrary LRC algorithm can be upper bounded by O(log(|H|/delta)* 1/epsilon), where epsilon and delta are the allowed learning error and failure probability respectively. While the above results are theoretical, a simulation conducted in this research found that the performance of these algorithms was in fact better than the upper bounds mentioned above. This research therefore establishes that interactive learning in the LRC model is at least as efficient as non-interactive Probably Approximately Correct (PAC) learning.

Awards Won:

Association for Computing Machinery: First Award of $4,000