The Method of Reduction in Machine Learning

"Machine Learning" as a field contains a large set of formalized problems. There is classification, regression, independent component analysis, data visualization, clustering, reinforcement learning, feature extraction, manifold extraction, and many others. Are these problems all fundamentally different? Or are all (some?) of these problems fundamentally related, the solution of one problem being useful in solving another?

This proposal investigates the method of reduction between machine learning problems. Reduction is a standard technique in complexity theory and the crypto community, where solving problem P is shown to be as hard as solving problem Q by showing that a P-solver can be turned into a Q-solver. Reduction in these communities are typically about hardness preservation. In machine learning, it appears more interesting to investigate "ease preservation" instead.

Suppose you find a magic box one day with the property that it can always give you an "edge": for any classification problem, it provides a classifier which is slightly better than random. The first thing to note about this box is that can't possibly exist, yet let's ignore this important detail.

Given this magic box, we can solve "strong classification", producing a classifier which is almost perfect in a small number of steps. One algorithm to do this is Adaboost by Schapire and Freund, which has seen significant practical use. With this magic box, you can also solve "cost sensitive learning" (theoretically and practically as well as KDD-cup champions), "reinforcement learning" (given a generative model, theoretically and practically), and "outlier detection" (theoretically and practically as well as KDD-cup champions).[1]

The magic box can not exist, and yet we consistently observe that algorithms which efficiently use this magic box are amongst the best learning algorithms. Perhaps natural problems (those we encounter, and those we define based on those that we encounter) *are* weak learnable?

The objective of this proposal is to construct a theoretically rigorous and practically useful understanding of the reduction relationships between machine learning problems.

The theoretical side has the intellectual merit. Is weak classification capable of solving all machine learning problems? Or are there machine learning problems not solvable by reduction to classification? Is there some minimal set of magic boxes from which all learning problems can be solved? The goal is to answer these questions, and consequently consolidate the field of machine learning.

The practical side has the broader impact. From a practical standpoint, a system of reduction to classification makes algorithms (and theory) for classification applicable to a larger set of problems. When new problems are encountered, instead of searching for a direct solution to the problem, a (hopefully much simpler) reduction to a known problem could be used. This advance is a step on the path from a field of research to an engineering discipline.

[1] A workshop was recently held on this subject at: http://www.tti-c.org/workshop/machine_learning_reductions.html