Machine Learning, 27, 5 (1997)

Guest Editor's Introduction

The papers in this issue were selected from among those presented at the *Eighth Annual Conference on Computational Learning Theory*, which was held in Santa Cruz, California on July 5-8, 1995. The authors submitted expanded versions of their conference papers which were carefully refereed before appearing here.

The first paper, "An experimental and theoretical comparison of model selection methods," by Kearns, Mansour, Ng and Ron, addresses the problem of trading between the complexity of a hypothesis and how well it fits the data. They compare "penalty-based" methods, including Vapnik's Guaranteed Risk Minimization and Rissanen's Minimum Description Length Principle, with Cross Validation. They describe a variety of results illustrating the strengths and limitations of these algorithms.

Helmbold and Schapire, in "Predicting nearly as well as the best pruning of a decision tree," analyze, using the on-line model, the pruning stage of decision tree learning. They describe an efficient algorithm which, given a decision tree $\mathcal{T}$, makes an expected number of mistakes in classifying new data which, for any pruning $\mathcal{P}$ of $\mathcal{T}$, is bounded as a function of the number of mistakes made by $\mathcal{P}$ on the new data, and by the size of $\mathcal{P}$.

In "Exactly learning automata of small cover time", Ron and Rubinfeld describe an algorithm for learning DFAs in a model where the algorithm does not receive counterexamples to proposed hypotheses. Polynomial-time learning of arbitrary DFAs in this model is known to be impossible, but they show that their algorithm runs in polynomial time for DFAs for which a random walk, with probability at least 1/2, quickly reaches any state.

Finally, Helmbold, Schapire, Singer and Warmuth, in "A comparison of new and old algorithms for a mixture estimation problem," consider the problem of learning a mixture of a number of known distributions. They place the standard EM algorithm, a variant of it with a different learning rate, and a new algorithm, in a common framework. They analyze the new algorithm in the on-line model, and compare the algorithms experimentally.

Thanks to Karen S. Cullen of Kluwer Academic Publishers for her excellent work handling the logistics for this issue. I am also very grateful to the referees. Finally, I would like to thank the authors for contributing their papers.

Philip M. Long
Department of Information Systems and Computer Science
National University of Singapore
Singapore 119260, Republic of Singapore