Machine Learning 4, 113-115 (1989)

Editorial

Scientific success stories make good journal articles: a theoretical innovation backed up by a proof or a running system, the integration of diverse ideas into a more general framework, empirical data testifying to the value of an improved learning technique, and so on. Failed attempts, on the other hand, receive little publicity. Researchers are eager to move on to new ideas, journals and conferences do not seek negative results, and it can be downright embarrassing to admit in print that a new promising method does not work, or is outperformed by simpler well-documented methods in the literature. Research, by its very nature, produces unpredictable outcomes, many of which we classify as failures and relegate them to the proverbial dustbin. The question I wish to raise here is whether some failures can prove instructive, whether they should indeed be published and the field be wiser for them, or whether they should never grace the printed page. Machine learning, and artificial intelligence as a whole, is a sufficiently new science that attention has focused almost exclusively on successes, rather than sharing the limelight with some of the less sexy but equally instructive failures.

When we refer to instructive failures, we can make some clear distinctions as to the cause of the failure, ranging from theoretical limitations and practical impasses to successful implementations that are outperformed by simpler existing methods.

Limitation Proofs

While trying unsuccessfully to work out a method, or build a system based on a published method, researchers occasionally pose the question of whether the method suffers from inherent theoretical limitations. The Minsky and Papert result proving the inability of linear perceptrons to discriminate among some simple patterns is a case in point; it demonstrated the futility of an entire line of research. Some theoreticians, in fact, actively seek to prove limitation theorems, independent of frustrations with the method in question.

Proven theoretical limitations do indeed appear in the literature; they are an accepted mode of communicating hard failures of particular methods or assumptions. Moreover, some theoretical results are based on complexity and tractability arguments (such as proving a problem NP-complete), rather than demonstrating complete impossibility. The only danger comes from divorcing theory and practice, in that proofs pertaining to an unrealistic simplification of a method or based on irrelevant computational models contribute little to the field. Many people would not classify limitation proofs as documented failures, but they do indeed spell doom, or at least temper enthusiasm, for the methods or models addressed.

Application Failures

Sometimes failure comes in the attempt to reduce a theoretically-sound and well-described method to practice in concrete applications. Although the method may be sound *in principle*, it cannot be applied successfully, at least not in its pristine, logically-clean form.

Pure explanation-based learning (EBL) provides a clear illustration. Applying pure EBL to acquire search-control rules in problem solving, as demonstrated in Minton's work for MORRIS and PRODIGY, may lead to inferior performance, just as easily as superior performance. Although this negative result has been known for some time by Minton and his colleagues, it was published only after refinements to the basic EBL method were found to guarantee performance increases (e.g., utility estimation and logical transformations to simplify the new knowledge). Had these improvements not been found, it still would have been an instructive failure worthy of (brief) publication.

Competitive Failures

Failures of a particular method can be absolute (as in the perceptron failure to recognize certain forms) or relative to other more effective methods. A new inductive learning algorithm must be shown superior to existing ones, at least on certain classes of data, in order to be judged a success. It does not suffice to show that it too is capable of learning, but at higher computational cost or lower performance than other well understood or otherwise simpler algorithms.

A weaker form of a competitive failure is narrower applicability of a certain technique as compared to the original claims. That is, the method may apply successfully but only to more circumscribed classes of problems. To cite an example from my own work not in machine learning proper, I published the Δ-MIN search algorithm in the proceedings of the first AAAI conference, claiming it to be a general method to unify multiple sources of knowledge in information-gathering problems such as speech recognition. Future work on applying Δ-MIN indicated that although it may be a good method to guide search in complex sensor-fusion problems, it is outperformed by simpler search methods in speech recognition. The added flexibility is not worth the incremental computational cost. Useful as this practical negative result may be, it is not what we normally publish in the field. A two page technical note quantifying relative performance of a method proven impractical for a given class of applications may save much wasted future effort.

Replication Failures

Experimental results in the physical sciences become established only after they have been replicated multiple times, often with different instrumentation and different data sources. Such a practice insures against artifacts of the experimental setup, peculiarities of a data set, and simple errors committed in the execution or analysis of the experiment. Recently, the field of machine learning is starting to follow this practice, setting up common data bases for inductive methods. However, the practice should be followed more rigorously, and the field should accept short notes confirming a previous result, or just as importantly, notes demonstrating where an attempted replication of an experiment is at variance with the original.

Action Item

Machine Learning wishes to encourage the submission of short technical notes, as discussed in my previous editorial. These notes may range from novel applications and improvements of existing techniques to concise reports of replicated experiments and instructive failures. In general, I wish to encourage the dissemination of information that will increase the scientific productivity of the field, communicating what ought to work but does not, as well as what really does work—and thus save on much needless replication of effort. Some results are best communicated in full formal articles, and others in shorter technical notes, also subject to review but with faster turnaround time.

—Jaime G. Carbonell
Carnegie Mellon University