

# On the Value of Applied Research in Machine Learning

Foster Provost (foster@basit.com)

Tom Fawcett (fawcett@basit.com)

Andrea Danyluk (andrea@cs.williams.edu)

Patricia Riddle (riddle@redwood.rt.cs.boeing.com)

In Proceedings of the First International Workshop on Machine Learning in Practice (ICML 1996)

This is an essay on the topic of applied research and applications papers at IMLC from the point of view of folks working on applied research in machine learning. We hope to minimize the "sour grapes" tone (our papers were not accepted); if some has slipped in, please try to ignore it.

Our goal, given the continued dearth of applications-oriented papers at IMLC, is to stimulate a discussion on the value of applied research to the machine learning community and on why/if IMLC should even be interested in including "applications" papers.

As background, at IMLC-95 in Tahoe we were involved in a large-scale discussion on the ML literature's relative lack of direction from real applications. The consensus among the applications-oriented folks was that there was a perceived bias against applications papers in the community. The response from many members of the MLJ editorial board was that they would welcome applications papers with general lessons to be learned. We were delighted when the call for papers for IMLC-96 included a specific call for applications papers. We were disappointed when our papers were rejected, although we concede that the rejections may have been for reasons other than their applications nature. The rejections were particularly disturbing because the individual reviews were moderately to strongly positive, yet the papers were rejected. While this is not an uncommon occurrence when there is a large number of high quality submissions, it is somewhat strange given that the call for papers explicitly requested applications papers. If the accepted paper titles are any indication, applications work remains poorly represented.

Our purpose here is to initiate a discussion on what the purpose of applications papers should be in the machine learning community, so that in the future we can do a better job (or not waste our time). This essay was stimulated by a comment from a review that was heartening to us regarding one of our papers, but indicates at best a misunderstanding on our part as to what is expected of an applications paper, and at worst a sad state of affairs in the community. In a nutshell, the paper reported on a high-impact, real-world problem for which the literature provided no applicable solution (and little related work), so we made up an ad hoc combination of three existing ML techniques that performed well. The reviewer wrote:

I really wish I could wholeheartedly recommend acceptance for this paper, because it would easily get into any KDD conference and I don't want to lose such papers, but I can't see the average ML researcher getting more out of it than a satiation of curiosity about applications. This is a troubling problem that I hope the program committee and ML as a whole will address; I can't blame the authors for it.

The reviews of a second applications paper caused similar concern. We decided to submit it this year mainly because of the indicated desire for applications papers. We believe that through this work we have learned a great deal about the difference between academic machine learning and real-world machine learning. The reviewers seemed to agree strongly. What we take from the program committee's decision, given the reviews, is that the important lessons from the real world are not as important as the ability to bundle up the complexity of a real-world problem into a nice, neat conference paper; applied work is interesting only if it looks like good academic work. (The reader can draw his/her own conclusions; the submitted paper is available at <http://www.croftj.net/~fawcett/papers/Provost-Danyluk-96.ps.gz>).

Thus, we would like to initiate a discussion as to the value of applications papers in our field. We will offer two related areas of value, at either end of the spectrum of significance.

First, an applications paper can sate the researcher's curiosity about applications: What types of problems are actually being addressed with our technology? Do the real-world results provide support for the results we see in our academic work? In addition, academic grant proposals often make the point that machine learning research can have an impact on real problems. Applications papers provide references for such claims. One might argue that the community's yearly meeting is the appropriate place to present such information (rather than waiting for it to be published in CACM).

At the other end of the spectrum are applications that indicate interesting new directions for the field. Those of us who work in applied machine learning research grumble amongst ourselves about how most academic work seems to ignore important complexities and concentrates on relatively less important problems. As an example, a favorite topic is classification accuracy. There is a consensus among many folks who try to apply machine learning to real-world tasks that very seldom is classification accuracy of primary importance. Yet for the past ten years, the community has produced reams of studies on how to increase classification accuracy. On the other hand, real-world tasks indicate that minimizing cost is often very important. However, the literature contains only a handful of relatively inconclusive studies on cost-sensitivity. When in our work we have wanted to produce classifiers with sensitivity to error costs, the research community has had little to say. In our opinion, real-world work is begging for thorough studies of cost-sensitive learning; showing that one system can beat another in classification accuracy is nearly useless.

One reason for the lack of direction from applications may be the UCI database repository. While we laud the tremendous efforts of the contributors and custodians, we must point out that there is a danger in the assumption that operating on the UCI databases is doing "real-world" machine learning. Often the assumption is tacit--a paper will claim to have used machine learning on "real-world problems," and the reader finds that the author did cross-validation over a set of UCI databases. For academic researchers, the repository is a wonderful resource for benchmarking. It is important that the repository not become a narrow window through which academic researchers view the world.

In conclusion, we view the academic/applied research dichotomy as comprising an important cycle, in which the applied world supplies significant problems and initial (probably ad hoc) solutions, which feed the academic work. The academic work supplies more elegant and well-tested approaches, which feed back to the applied community for use and reality checking. We fear that this cycle may be broken in the machine learning community, although it may be functioning well within individual research programs. Applied researchers feel that the academic world only pays lip service to real problems. The academic world seems to view applied work as interesting only if it looks like good academic work. We, as a community, should strive to define the roles of applied research in machine learning, recognize the differences from academic work, and provide forums for the two to communicate. To that end, we thought it would be valuable to initiate a discussion on the value to the community of applied machine learning research. At the very least, we hope such a discussion will help applied researchers to present their work in a way that the community will find interesting and valuable.