
Machine Learning: Theory, Applications, Experiences—A Workshop for Women in Machine Learning

*San Diego, California
October 4, 2006*

*Organizers: Lisa Wainer, University College London; Hanna Wallach, University of Cambridge; Jennifer Wortman, University of Pennsylvania.
Faculty advisor: Amy Greenwald, Brown University*

Summarized by Lisa Wainer

The workshop was a one-day event offering a showcase of work by women involved in machine learning research. The main objective of the workshop was to offer female faculty, research scientists, and students in the machine learning community an opportunity to meet, exchange ideas, and learn from each other. It also gave women in other areas of computer science the opportunity to learn about cutting-edge research in a growing field. The workshop was open to anyone, male or female, to attend free of charge and was co-located with the Grace Hopper Celebration of Women in Computing. The workshop succeeded in bringing together women from different stages of their careers, from established researchers to Ph.D. candidates and even undergraduate students. It provided an opportunity for established researchers to act as mentors and for students to find much needed role models. There were ninety-six registered participants in all, and quite a few unregistered attendees (two of whom were male).

The invited faculty talks covered a diverse set of topics, were well received by the audience, and inspired much discussion during the session breaks. The student

presentations were divided into short talks, spotlights, and poster presentations. Talks were generally of high quality: Roughly half fell into the category of theory and half were on applications of machine learning. The general reaction of participants was extremely positive. There was a vigorous discussion at the end of the workshop about issues for women in the machine learning community and in computer science in general. The lively discussions covered the differences in working in theoretical as opposed to application-based disciplines, the importance of role models and mentoring, how to raise profiles of women in machine learning, and whether the workshop should be run again in the future.

A few general conclusions were reached. The first was that role models are an important aspect of encouraging more women into computer science and machine learning as well as retaining women once in the field. The second was that events such as this workshop are seen as being very important to help women network, to produce collaborative work, and to meet socially. There was a unanimous response to the last point, in that the participants wanted the workshop to run again next year. Many suggested that they would like it to run alongside a machine learning conference. Many attendees commented that the invited talks and student presentations were of exceptionally high quality. Participants were eager to interact and take part by being active during the poster sessions and asking the speakers questions. Overall, the organizers and the participants felt that the workshop was very valuable, and they are planning to hold the event again next year.

Grace Hopper Celebration of Women in Computing 2006, Making Waves

*San Diego, California
October 4–7, 2006*

Summarized by Rae Harbird

This conference was the sixth in a series designed to bring the research and career interests of women in computing to the forefront. The presenters, from industrial, academic, and government communities, presented their current work while special sessions focused on the role of women in today's technology fields. From my perspective it was a rare and special opportunity to meet women from and hear talks on a diverse range of subject areas spanning the entire breadth of computer science. The atmosphere was truly celebratory, with a strong emphasis on women's achievements and the excitement of working in such a dynamic and fascinating field. The benefits of collaboration and networking for success underpinned the fabric of the conference, reflecting the skills at which women traditionally excel. The conference organizers made great efforts to ensure that the social events were just as rewarding as the technical sessions. I am looking forward to GHC '07.

■ *Dasher: Information-Efficient Text Entry*

Hanna Wallach, University of Cambridge

The objective of the presentation was to introduce the audience to Dasher, a novel information-efficient text-entry system, driven by continuous pointing gestures. Keyboards, despite their ubiquity, are inefficient for two reasons: They do not exploit the predictability of normal language, and they waste the fine analog

capabilities of the user's muscles. Gestural alphabets, such as those used on a Palm Pilot, use fine motor movements but are often unreliable. Devices with limited keyboards, such as mobile phones, use prediction techniques for text entry but this is also clumsy and requires two modes: word completion and disambiguation. Four important things are missing: the ability to take advantage of fine motor movements, exploitation of the redundancy of language, language independence, and single-mode operation allowing users to write and disambiguate at the same time. Dasher is based on principles of machine learning and information theory and is intended to rectify these inefficiencies. Incorporating an adaptive language model of the sort also used in speech recognition, handwriting recognition, and text compression, Dasher offers helpful predictions to the user without constraining the range of words that can be written.

Comparing Dasher with use of a standard keyboard showed that although Dasher users typed fewer characters per minute, error rates (percentage of incorrectly typed words) were lower. Experiments showed that using Dasher with an eyetracker is much faster than using an on-screen keyboard and has a significantly lower error rate. Dasher is designed to be a competitive text-entry system used wherever a full-size keyboard is not possible, such as with wearable and palmtop computers, as well as for disabled users. Anyone can use it, as no training is needed, and it is fast and fun to learn. In a recent case study conducted by Mick Donegan at the ACE Center, Paul, who suffers from cerebral palsy, used Dasher to write his thesis for a degree in Business and IT. Paul reported that Dasher required less head movement,

generated fewer spelling mistakes, and was about four times faster than an on-screen keyboard.

■ **Panel: Building and Managing a Strong Research Group**

Nancy Amato, Texas A & M University; Tracy Camp, Colorado School of Mines; Elizabeth Royer, University of California; Violet Syrotiuk, Arizona State University

As a Ph.D. student who is starting to think about my research career, I was pleased to attend this panel led by women who are experts in their chosen field, which happens to coincide with mine. Although the focus was academia in the United States, the points made were easily translated to a UK context. The talk had two principal strands: first, describing how to get funding for your research group and, second, covering tips on advising students.

As far as obtaining funding goes, some background work is necessary to get your foot in the door: The importance of networking cannot be overemphasized, as you need to establish mentors and partners for your research. Publicize your work and ideas by giving as many talks as possible, both within your institution and outside. The main recommendations for successful proposal writing were to communicate your ideas to as many people as possible and to ensure that your proposal fits the specification provided by the funding body. The panel advised researchers to read and familiarize themselves with proposals that have succeeded already. The pros and cons of working on single or multiple Principal Investigator proposals were discussed. Both are viewed as important to your career, the latter because it provides the opportunity to show the specific expertise that you can bring to a project and the op-

portunity to do it really well. It may also offer the opportunity of working with more senior staff. It is important to remember that proposals are often rejected, and the panel emphasized the importance of giving careful consideration to feedback received and being persistent. Acting as a panelist on funding bodies is also an important learning experience. Some of the panelists admitted to not giving enough talks as new researchers; making such presentations can be hard, but they are a great way to increase your confidence.

Finding good students for your research group was covered next; panelists advised teaching graduate classes and being proactive in recruiting students. There are distinct advantages in implementing schemes that allow undergraduates to gain short periods of experience. The techniques covered for getting the best out of your research students reflected best practice used in industry. Primarily, it is important to set clear expectations and establish clear goals and milestones; preferably, these things will be written down. Tracy Camp has a document outlining student expectations on her Web pages. Establishing a mentoring hierarchy for students not only takes some of the pressure off advisors but also gives newer researchers experience in mentoring for themselves. The visibility and reputation of your research group are, of course, important, and recommendations mirrored those given for publicizing your own research. Network whenever you have the opportunity, volunteer for activities, and be willing to host talks and give talk tours.

■ *On Program Security*

Hongxia Jin, IBM Almaden Research Center

In an email exchange with Hongxia, she said that she aimed to give the audience a quick technical overview of program security and, in particular, to give those who are interested in finding out more a good starting point. Hongxia described some elementary design principles for achieving software security based on her years of experience working in the area. She says that even though everybody's application context may be different, the same design principles should be applicable.

In describing the motivation for her research, Hongxia explained that hackers can reverse-engineer programs to understand or even modify existing programs quite easily. Consequently, competitors may learn trade secrets or copy algorithms to reuse in competing products. They can also remove protections and redistribute the pirated program for a profit. The general problem of program protection is widely thought to be impossible, implying that program protection is an important, wide-open problem. Hongxia described two methods used to defend against attacks. Preserving code integrity by guarding against tampering is one technique. Static integrity can be enforced using cryptographic hashing of all or part of a program, although this can impose a high processing overhead. But what about run-time integrity? You can detect the presence of a debugger, for example, by measuring execution times between particular sections of the code. The actions taken if tampering is detected also need to be given some thought. On the one hand, you could abort execution; on the other, it may be attractive to delay failure until a later time,

inserting plausible yet misleading operations.

Finally, Hongxia gave an example of her research in this area. Bearing in mind that eventually a determined attacker will succeed, it is advantageous to detect on-going tampering as early as possible. In this way you may be able to restrict the extent of the eventual damage. One way in which you can do this is to use an auditing log to record hacking activities. A smart hacker can tamper with the log, but to counteract this you can make the log itself tamper-resistant. Hongxia described a scheme in which log entries are encrypted and transmitted back to a central clearing point. The source of the log entries and the clearing point share a key which is continually but independently evolved in both places. Even if the hackers succeed in subverting the encryption scheme, they cannot go backward, so some evidence of the attack will still be recorded.

■ *Shifting the Tide of Network Security: Being Safe, Being Aware, and Being Active*

Nicole A. Pauls, TriGeo Network Security

In an email exchange with Nicole, she commented that she viewed the conference as an opportunity to talk about improving things for the women of tomorrow and an opportunity for everyone to see what women are doing today. Even though the focus of the conference is on women in academia, it was equally important to see women in industry talking about their research. In Nicole's presentation, she described pragmatic strategies for improving security. Many of us think of network security tools as necessary evils of prevention, but the truth is that we can't protect ourselves from everything. Software holes can be exploited before patches can

be deployed; "trusted" users can become dangerous by opening the wrong email at the wrong time, and everything happens so quickly that we might not even know until it's too late. Nicole explored a defense in-depth network security strategy, covering architecture, monitoring, and active defenses. In terms of architecture, security must be something that we think about from the outset in the design of our networks and systems. We have to think about security on many levels, protecting each element so that it can still operate securely even if, say, the firewall fails. Detection is an essential weapon in our armory, which is where logging and auditing come in. Good organization and management are really the keys to success here; centralizing logging with a syslog server coupled with automated log analysis tools is the way to go. Deciding how you are going to respond to a security incident is equally important. First and foremost, you need to have a set of clearly documented policies clarifying the risks and responses. Your biggest problem here might be wrestling with the internal politics in an organization. On a practical level, you should not expect to do everything at once: prioritize and implement what you can, when you can.

■ *Wireless Sensor Networks and Real-World Applications*

Nirupama Bulusu, Portland State University

In her talk Nirupama encouraged us to "come learn about the opportunities and computing challenges in wireless sensor networks." Although fairly broad, this talk illustrated how sensor technology is being applied in areas such as digitized health care, energy management, condition-based maintenance, and habitat monitoring. The projects

Nirupama chose to illustrate the applications of sensor technology were very exciting; the material would make a great educational pack, encouraging children and young adults to consider a career in computer science or engineering.

Most sensor applications involve monitoring either space or objects or both. In terms of monitoring space, sensors might be used for such things as environmental and habitat monitoring and precision agriculture. For example, a vineyard in Oregon uses sensors to monitor temperature and moisture; Roger the dog collects the data using a wireless collection device in his collar. The ZebraNet project monitoring zebra movement in Kenya was a somewhat more challenging application. Nodes (or zebras) are highly mobile and only sparsely populate the environment. Transferring data to a collection point (which may also be mobile) requires nodes to self-organize and actively route data.

The final example, an application to detect the presence of cane toads in the Australian outback, had a clever twist. The problem of how you detect a cane toad is interesting. It transpires that sensors can be used to hear when a cane toad is nearby by analyzing the acoustic features of the toad call, since cane toad calls have a completely different signature from those of other amphibians. Nirupama gave an "under the bonnet" view of some of the challenges presented by sensor-based applications. For example, localization, determining the position of sensors, and distributing this information to other nodes can be a complex task, yet workable solutions now exist. And what about the future? Sensor networks for urban applications will form the "next tier of the Internet," lever-

aging the cell phone installed base of acoustic and image sensors to capture readings.

■ *Part of the Problem/Part of the Solution*

Claudia Morrell, University of Maryland; Revi Sterling, University of Colorado; Sophia Huyer, Women and Global Science and Technology

The panel explored the issues affecting women surrounding deployment of Information and Communication Technology (ICT) in developing countries. Nongovernmental organizations (NGOs) are increasingly employing ICT tools as part of their strategy. Whereas ICTs show great promise in alleviating entrenched economic, health, and gender disparities, they may be exacerbating gender gaps. Despite the best intentions, many aid-focused initiatives intended to assist women's unique development goals are challenged with long-term sustainability and the ability to have wide-ranging impact upon the culture. Several women-friendly projects were presented. The Grameen-Phone scheme enables women to buy cell phones and rent usage, SchoolNet Africa provides computers to schools in Africa, and TeleCenters are springing up in Latin America and Africa. In the past year or so, important synergies have emerged. The 1st Women and ICT Symposium was held in 2005. At last it is possible to combine known information on emerging markets, people, and technologies. A taskforce was formed which has been recognized by the UN Global Alliance for ICT for Development as a community of expertise.

■ *The Impact of ICT on Women in Brazil*

Dilma M. Da Silva, T.J. Watson Research Laboratories

In her presentation, Dilma Da Silva reviewed the current status of ICT deployment in Brazil with respect to women. In the 1980s, the government imposed tight controls on the market, encouraging local development of software and hardware where possible. The 1990s were characterized by dramatic improvements in phone, cell, and networking services, but by around 1995 over 35,700 jobs were cut in the IT industry (representing a 48.1% reduction). At present, technology deployment is uneven. The Brazilian business triangle (São Paulo, Rio de Janeiro, and Belo Horizonte) has high-capacity fiber, virtual private networks, and bandwidth on a par with the United States and Europe, whereas most of the countryside has no access at all. But the government is actively promoting the use of electronic voting, and the percentage of tax returns completed online is growing.

So where do women fit into this picture? In 2000, the World Bank reported that Brazil has one of the widest gender gaps in Latin America. Even with the same qualifications as male colleagues, women got only 54% of what men received as wages and comprised only 20% of the IT workforce. Socially, women are not considered suitable for ICT-related jobs and this, in turn, influences women's interests and goals. Dilma also pointed out that despite these disappointing statistics, this is not the whole picture. The situation is reflected in my own experience: In my department there are two very talented Brazilian women studying for their Ph.D.s.