Teaching Statement of Chun-Nam Yu

Machine learning and artificial intelligence are exciting and fun subjects to teach. They have numerous connections to other fields (statistics, mathematics, bioinformatics, linguistics, neuroscience), and attract students from a broad spectrum of backgrounds. The subjects are also developing rapidly with new applications coming up daily, so that even students taking a first introductory course can find many innovative ideas for their class projects. During my PhD studies at Cornell I have worked as a teaching assistant to introductory machine learning and artificial intelligence classes. I was involved with various aspects of teaching, including designing and grading problem sets, teaching guest lectures, and supervising student projects.

The guiding principle in my teaching is to make students learn through both theory and practice. In my guest lectures I helped students understand the key ideas of an algorithm or model through simple running examples. After teaching these basic concepts in lectures I made the students go deeper through homework problems. I was responsible for designing all the biweekly problem sets used in a machine learning course in 2007 and 2008. In designing the homework problems I always tried to strike a good balance between theoretical problems and programming-based practicals. On one hand I wanted the students to develop a good intuition for the properties of data and algorithms, so that they learned to choose which method to use or to modify through simple calculations and reasoning. On the other hand I wanted them to realize that although sometimes the experimental results confirm our theoretical intuitions, there are far more often cases when the assumptions we make are not satisfied by the real data and the performance differ from what we expect. Therefore I always included some reflective questions on the experimental results in the homeworks to see if they could be explained by the theory. I also periodically asked the students questions during lectures why a particular idea would not work without having to implement it. In practice I found a lot of students struggle with these questions at first, but I was glad that over time most students in the class can develop good understanding and intuition about data and algorithms.

Apart from regular teaching duties, I also participated in teaching outreach courses to students outside computer science and engineering. In the summer of 2011, I was invited to give a three-day workshop entitled 'Introduction to machine learning and support vector machines' to a group of undergraduate summer research students and faculty members at Hamilton College, NY. These students and faculty members mainly came from the bioinformatics and chemistry departments. They wanted to know what machine learning is, and how it might apply to scientific data analysis in their research. To accommodate for the science background of this group of students, I mainly employed examples from bioinformatics to illustrate the main concepts in machine learning during the lectures. In the practicals I also substituted the usual programming exercises with data analysis questions using existing software packages, so that they could observe how machine learning can be used to model data and make predictions without having to worry about coding. As a result of the course most students began to appreciate the presence of machine learning and pattern recognition algorithms in the daily technologies they use. Some also started to consider going deeper with the subject in order to apply it to their summer research projects. Overall it was a very interesting teaching experience because it was the first time I tailored such a course to liberal arts college students, and I was also very much encouraged by the enthusiasm of the students towards understanding machine learning and their own summer projects.

During my postdoctoral training at the Unversity of Alberta, I was not directly involved in teaching classes, but I had the chance to interact with most of the graduate students and research programmers in my postdoc supervisor's group. Since many of us, including our collaborators, worked on different but related projects spanning areas like machine learning, bioinformatics, and biomedical sciences, it is not easy for each of us to keep up with the most updated literature in all these areas. Recognizing a need to improve the exchange of ideas and keeping updated with the literature, I started an ongoing joint machine learning and computational biology reading group in the summer of 2011. I encouraged each of the members in the group to take turns to present one recent paper related to his/her current

research project in our weekly meeting. As a result of this we are better able to keep up with the latest development in the fields, and also maintain an active discussion on each other's research. In addition to organizing these discussion groups, I also actively mentored several graduate students through close collaboration on research projects.

Apart from machine learning and artificial intelligence, I can teach classes like introductory programming, data structures and algorithms, and certain graduate topics on statistics and optimization that I used regularly in my research. Teaching has been a very rewarding experience in my academic career, as I got to re-visit and re-think many of the fundamental issues of the subject. I also treasure the time spent interacting with students, who range from the usual undergraduates to graduate students with significant work experience, to even spouses of retired faculty members. I must say I have learned more from talking to them than they from the materials I taught.