

Teaching Statement

The whole point of teaching is not just to impart knowledge to others. It is to enable others to do what they cannot, and knowledge is not knowledge if one cannot act upon it when the time comes. Thus the teacher needs to reduce every single piece of knowledge and every step of derivation into the most basic form: Given the scenario or problem, what you can expect or how you should go forward solving the problem. These all boil down to the clarity of presentation of the things well understood as well as honesty about the things that are less well understood.

Teaching Experiences

I summarize my experiences in teaching classes and peer advising.

While being Teaching Assistant of the graduate level Information Retrieval class, the course instructor Professor Jamie Callan asked me to give a lecture on probabilistic retrieval models. I inherited Jamie's slides which had gone through years of teaching. Jamie is very good at presenting lots of ideas quickly and clearly. Thus, the first difficulty for me was that there was too much content. My solution was to emphasize the more important techniques and ideas, and only briefly introduce the rest. The most important development in probabilistic retrieval models is the Binary Independence Model, which became my second difficulty, because I had to go through the formal derivation of the model in front of a group of students with diverse backgrounds. My solution was to use the whiteboard, so that each step in the derivation was carried out slow enough for every student to get the gist of it, and I could see the students' appreciation for that. Preparing and giving this lecture also benefited me personally. A deeper understanding of the $P(t|R)$ probability was one of the key elements that led to my thesis research.

The best way to learn is to do it yourself. In addition to teaching a MapReduce tutorial, I've created MapReduce (Hadoop) programming homework for the IR class. Two sets of homework spiraled into the core of MapReduce. The first homework was a step-by-step guide to solve a small MapReduce coding problem, serving to familiarize the students with the Hadoop environment. The second homework required a bit of independent problem solving, fitting together the pieces they learned in class while understanding Hadoop more deeply. The homework was setup so that students could also practice iterative programming in the process. All these prepare the students for the real world.

I've also worked with other graduate students. Earlier in my PhD, I guided younger students to take on projects that I had worked on. Later on, I started providing my IR expertise to related research projects. One of such collaborations was with my friend Wei Xu who works on information extraction at New York University. For her PhD thesis, she had the idea of building an information extraction system on a search engine. I would advise on retrieval techniques, experimental setup and evaluation, while Wei would carry out the experiments and report results and observations through emails or phone calls. Being one step away from the data was new to me. In each back and forth interaction, I would carefully suggest given the current circumstances where to look, what to do and what to expect. What was rewarding for me was to see Wei starting to pick up the right approach and doing more and more of the right thing. I was glad that I kept Wei motivated and busy. We published, and enjoyed this process of learning together.

Teaching Interests

I'll be happy to teach introductory or advanced courses in information retrieval, Web data mining, machine learning, and large scale computing in search applications or on the Web. I also appreciate opportunities to teach basic computer science courses like computer languages or data structures, which would allow me to interact with more students from more diverse backgrounds, make CS their favorite tool, and make more impact on a student's career.