Teaching Statement

Chengxi Zang
Department of Population Health Sciences
Weill Cornell Medicine, Cornell University
September, 2020

I believe teaching is one of the most enjoyable ways to a lifelong learning experience and to interact with budding talents. Teaching gives me immense satisfaction and joy for having a job as a professor and at the same time offers me the responsibility for stepping in the initial stages of the learning journey of the students. It also provides me a platform for improving my skills both in terms of expressing myself as well as be a great source of research ideas. As Will Durant rightly said, "Education is the progressive discovery of our own ignorance". In the following, I will first state my teaching philosophy, and then summarize my teaching experience, and finally outline my teaching plans for the future.

1. Teaching Philosophy

Graph machine learning is an interdisciplinary field that lies in the intersection between computer science, computational social science and biomedicine. I am specifically interested in teaching machine learning, data mining in big structured and dynamic data type of courses. I believe the following points are crucial to a successful teaching in these courses.

- Combining algorithms and real-world problems
- Providing hands-on exercise opportunities
- Encouraging novel thoughts and ideas

First, graph machine learning is a shared topic among sub-fields of computer science, including data mining, machine learning, and NLP and CV, and beyond (e.g. graph theory, network science, computational social science, biomedicine, and cheminformatics). We cannot just introduce the latest machine learning algorithms on graph without their development history, their motivation, and their applications in the real world. I think a good way is to put algorithms in social and biomedical contexts, i.e., motivating the algorithms with real-world social and biomedical problems, and introducing the algorithms step by step as solving the motivating problem. In this way, the students can understand better the algorithms and how to use them to solve real problems. We try to cover motivating problems in different fields and encouraging interdisciplinary ideas.

Second, graph machine learning is a practical area. Therefore, in order to make the students equipped with the right skillset and become competitive in their future career, providing hands-on exercise opportunities in class and let them practice what they learned is crucial. In this way, the students can understand what they would face in real data science and machine learning jobs and prepare better.

Third, graph machine learning is a young research field that is developing very fast. Lots of problems are still not solved or even well-defined. Therefore, making the class interactive, throwing more questions to the students to stimulate their creativity, is very important. I think this can also stimulate the curiosity of the students in this field and make them interested and love this field. This will be helpful for the health growth of this field.

2. Teaching Experience

When I was a Ph.D. candidate at the Department of Computer Science, Tsinghua University from Aug. 2013 to Jan 2019, I was a senior teaching assistant of both undergraduate- and graduate-level course "Big Data Analytics" for three semesters. It was one of the most popular courses that CS offered and it attracted 40-50 students from different departments including Electrical Engineering, Statistics, Management Science, and even Physics. In the course, I intensively took part in teaching popular (both classic and advanced) big data analytics methodologies with an emphasis on structured and dynamic data and provided the students with concrete real-world examples where such methods can be used. There were no standard/required textbooks but important literature references for the student to read. The students were also asked to finish several homework assignments, as well as a team project which they need to identify themselves. These experiences gave the students hands-on practice opportunities to gain experience that will be super helpful to their future jobs. Besides, I also involve students from other programs (e.g., tri-institute PhD program and MD programs), other campuses (Cornell Tech and Cornell

Ithaca), and other universities. For example, I have mentored two graduate students from Cornell Tech in our drug discovery project and collaborated with many students/postdocs in the past. I really enjoyed the process of interacting with students.

In addition, I also actively participated in several popular tutorials in top data mining, machine learning and AI conferences. In AAAI 2020, I gave a tutorial "Differential Deep Learning on Graphs and its Applications" which investigates the recent advancements in introducing differential equation theory to the graph neural networks. We show that differential deep learning on graphs are powerful tools for modeling the structures and dynamics of complex systems and generating molecular graphs in drug discovery. This tutorial attracted nearly 80-100 audiences. A more challenging one is our KDD 2020 tutorial "Recent Advances on Graph Analytics and Its Applications in Healthcare" which is online in this pandemic era. We offered both live streaming and video on demand type of tutorials on how modern graph analytics technologies (such as graph embedding, knowledge graph, graph generative models, etc.) can be applied in different healthcare scenarios. Surprisingly, our virtual tutorial have attracted 242 audiences in the Whova system, more than 1300 unique IP visits of our tutorial webpage, and more than 800 views of our 5 video clips in YouTube by Sept. 2020. We also got real-time feedbacks in Twitter. For example, one of the audiences wrote the following in his/her tutorial feedback: "@calvin zcx Thanks for organizing the #KDD2020 tutorial this morning. A very informative and motivational morning for sure!"

I believe my experience of delivering both online and offline courses and tutorials to different audiences including undergraduate, graduate and ML/AI/Data-Mining researchers and let them know more about graph machine learning analytics would be helpful for the healthy growth of this field.

3. Future Teaching Plan

In the future, I can teach 1) Graph Machine Learning, 2) ML/AI for Healthcare, and 3) Fundamental Big Data Analytics courses for graduate students, and 1) Data Structures, or 2) Programming with Python for undergraduate students, which covers both the essential basic skills and applications of the modern AI and big data technologies with applications to healthcare and drug discovery.