

Enabling Students with Cutting-edge Machine Learning Techniques

Yue Zhao

CARNEGIE MELLON UNIVERSITY

Teaching and mentoring have always been my passion. I have rich experience in serving as a lab instructor for programming sessions, teaching and assisting more than **ten graduate-level artificial intelligence and statistic courses** with hundreds of students, and advising junior students for AI and ML research with fruitful outcomes. Also, I have been frequently invited for **guest lecturing on machine learning at top schools** including Carnegie Mellon University and Columbia University. In addition, with a deep understanding of industry needs, I have been **presenting and providing consultation at major firms**, e.g., Morgan Stanley, Wells Fargo Bank, and E&Y Canada. I appreciate the opportunities to exchange my experience, knowledge, passion, and ideas with junior students and experienced industry practitioners, and consider this a unique privilege to keep teaching and learning from people.

Experience and Philosophy of Teaching and Mentoring

My teaching experience widely spans *computer science, statistics, and business analytics* as shown in Fig. 1. I grasped all the chances to teach since I was an undergraduate back in 2012—I applied to and became a lab instructor for *Introduction to Programming* in my junior year. I started forming my hands-on and knowing-why teaching philosophy from this experience, and firmly believed that helping students learn how to debug and identify refactoring opportunities is as important as learning algorithms. Throughout the years, I have led numerous lab sessions like C++ programming at the University of Cincinnati, Embedded Systems with C at the University of Toronto, and Statistics at Carnegie Mellon University. For these labs, I have encouraged students to think about the reasons for using specific algorithms and data structures and to identify optimization opportunities in the implementation.

I encourage students to think about “why” before deciding “how” to implement an algorithm or strategy. This philosophy also applies to all my courses beyond hands-on labs—many business analytics courses also require deep thinking other than directly using analytics tools to solve a problem. Taking my experience as the lead TA of *Digital Transformation* and *Managing Digital Business* as an example, I enjoy spending time discussing with students why to use a specific technology in a business setting, other than simply showing them how to achieve a business goal. When we are presenting the real-time inventory supply-demand, although Visual Basic macros with excel can provide reasonable demonstration, it fails to promptly update with regard to the inventory change. Based on the application needs, I help students to understand why it is more efficient to build a real-time dashboard using tools like PowerBI by connecting it with the inventory database for real-time updates.

I emphasize the importance of helping students unpack complex questions into small actionable items. For instance, I have been the lead TA for *Introduction to Artificial Intelligence (IAI)* at Carnegie Mellon since 2020 (for five consecutive semesters with more than 200 students). This project-based course requires students to design and finish group projects by turning learned AI knowledge into demo products. For students who have no or limited AI experience, designing such a project is daunting—they either propose something too ambitious or something too conservative. To help them scope the project well, I identify their high-level interest through discussion and guide them to craft a minimal viable product. By finishing some easy prototyping, students feel more confident with the content and can further build on top of it. This philosophy also reflects in designing homework for *IAI*, where I motivate students

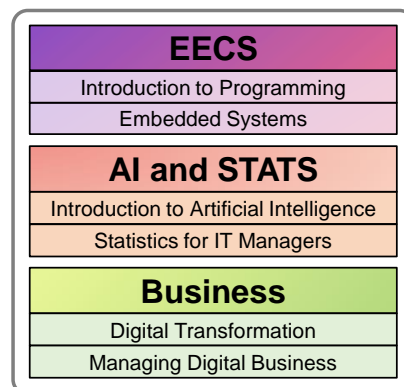


Figure 1: A wide range of courses that I have been a TA or instructor.

with real-world applications and decompose the hard problem into small manageable pieces with skeleton code to follow. In this way, students can make gradual progress and feel empowered with AI techniques.

I have had opportunities to mentor and advise more than 10 master’s and junior Ph.D. students on multiple research projects. My philosophy towards advising is to acknowledge their progress and guide them to explore more and deeper. My mentorship has resulted in multiple high-impact research papers and systems. For example, I advised two junior Ph.D. students to publish their first NeurIPS papers in 2022, and many other papers at top venues including ICDM and AAAI.

Invited Talks, Community Building, and Knowledge Sharing

Besides rich teaching experience at universities, I am frequently invited to talk at both academic institutes and firms. These idea exchanges further nourish my teaching at school. Recently, I have been invited to present my data mining works at the University of Nottingham, Rutgers University Business School, Rice University, Columbia University, etc. Through these talks, we foster collaboration, leading to two NeurIPS papers with the researchers at Rice University. Also, my talks at Tesla, Morgan Stanley, Microsoft Research, and Wells Fargo all received fantastic feedback—some of them deployed my work and achieved substantial profit gain. These opportunities also inspired me to design more realistic and pressing research questions. Meanwhile, I am running the largest anomaly detection research community with more than 750 students, practitioners, and researchers to discuss emerging topics in anomaly detection. I am also a dedicated social media writer, with more than 300 ML articles and 200,000 followers on Zhihu¹—Chinese Quora (200 million+ registered users). Sharing knowledge with others is one of my favorite hobbies, and I feel it is my responsibility to make ML more accessible.

Teaching Interests and Course Proposal

Given my interdisciplinary background in computer science, machine learning, and data science I am well suited for a wide range of existing courses and am happy to design new courses at different levels.

Existing courses: (i) *programming courses* at both introductory levels and advanced levels with data structure and software engineering for life cycle management. (ii) *machine learning and data mining* at undergraduate and graduate levels with wide coverage from popular models to the latest applications and new frontiers (e.g., self-supervised learning, graph neural networks, and diffusion models). (iii) *applied data analytics* including but not limited to predictive modeling, visualization, and cloud computing. I have rich experience in these topics as a business consultant and TA.

Proposed Courses:

1. *Applied Machine Learning*. With the advancement of AI, mastering its concepts as well as knowing when and how to use it in the real world becomes a new challenge for data analysts and industry leaders. In view of this, CMU School of Computer Science has launched an executive education program on AI. With this in mind, I will provide first-hand experiences for data practitioners to understand the usages of AI and how to interpret its results in this project-based course. More emphasis will be given to scalability and interpretability. This course can provide a more applied perspective of ML as a complementary course to the general ML courses.
2. Seminars on *Machine Learning Systems* and *Machine Learning for Risk Modeling*. These graduate-level seminars will help students to be on the frontier of machine learning research, including how to use ML to facilitate system research and vice versa. The MLSys seminar has been successfully established at Washington University and Carnegie Mellon University. Also, ML has been widely used in risk modeling in finance, security, and healthcare. As an expert on using ML to detect anomalies and risky items, I will guide students to see interesting ML applications on this topic as well as the latest research advancement, where my personal research interests lie on.

¹See my ML-related blogs (in Chinese) at Zhihu: <https://www.zhihu.com/people/breaknever>