

As a computer science senior in college, I have always loved getting my hands dirty developing real-world tools and products, to explore how computers can make our lives easier and better. In particular, with a strong interest in big data and machine learning, I put my passion into bridging the gap between cutting-edge AI models/algorithms and real-world, data-intensive products, for example, constructing a machine learning software application with good usability, high scalability, and strong reliability.

In September 2016, I started my college education at the Beijing University of Posts and Telecommunications (BUPT) with a strong emphasis on big data theory and engineering. At BUPT, for example, I took *Probability and Stochastic Processes* to learn how to simulate real-world data using mathematical distribution models. Meanwhile, online education enabled me to take advanced courses besides those offered by BUPT. Andrew Ng's *Machine learning* enlightened me on how to extract information from big data with efficient algorithms, while *Cloud Computing* and *MIT 6.324: Distributed Systems* introduced to me the storage, process, and management of big data. Through these courses, I successfully equipped myself with both theoretical and practical skills to play with big data.

The keyword for my next phase of adventure is "*research*". From May 2018 to August 2019, I enjoyed two experiences as a research assistant at the *Institute of Automation, Chinese Academy of Sciences* (CASIA) and *Penn State University*, respectively.

At CASIA, I researched on graph mining, which focuses on mining information from graph-structured data, such as social networks, in the real world. During my time there, I led a project where my colleagues and I proposed a new graph convolutional network (GCN) model which successfully introduces nonlinear neighborhood interactions into GCNs to capture complex features. The model outperformed the state-of-the-art methods significantly on several tasks. We later summarized this work into a paper submitted to TKDE, one of the top journals in data mining, and the preprint version is available on [arxiv](https://arxiv.org/). Working on this project, I was exposed to all aspects of research: from ideation to implementation and from the organization of information to tweaking the language of our papers.

I continued working on graph mining when I was a summer research intern at Penn State University. During my two-month stay, I worked on graph adversarial learning, a new topic of graph mining, which deals with polluted or maliciously altered graph data. After referencing many papers and open source projects, I implemented several models and designed some variants to analyze the results. This experience was valuable because of the noisy nature of real-world data and helped me understand data from a new perspective.

In September 2019, I applied to and joined MOMO Inc., a top social networking platform in China, as an R&D intern in the 3D reconstruction group, in the hope of gaining experience in the industry. After joining the company, I was assigned to implement a new algorithm to resolve the instability problem in 3D face tracking. Through this task, I became familiar with the development process in the industry and ended up being highly proficient in manipulating image data. After that, I took part in a new project to build a new 3D human face mesh dataset and was in charge of extracting the topological structure from human faces. During this project, I adopted many techniques such as parallel computing to achieve high-fidelity and high-speed processing of big data, and enjoyed this opportunity to be involved in "producing data", a part of "playing with data" that I have never experienced before.

Through these series of rich and varied experiences, I am determined to say that my ultimate goal is to fully tap the predictive capabilities of data to do good for human beings. In the short term, my goal is to become a big data engineer who builds and deploys practical applications with data. And I am particularly interested in solving engineering problems for machine learning algorithms to apply them to real-world applications. To achieve these goals, I believe the postgraduate education is necessary. I look forward to taking advanced courses to further improve myself, as well as networking with professors and students who have different backgrounds and talents. Moreover, I want to seek opportunities to participate in research to explore deeper in big data.

Among all the master's programs in computer science, the MCS provided by Texas A&M is a big draw to me. First, since it is a professional and terminal degree, this program offers students many well-designed courses to train them to be sought-after programmers and engineers, which aligns well with my career goals. Second, I am also attracted by the length of the program. Texas A&M give students large flexibility in deciding when they will

graduate. This is a huge advantage for students with a clear career plan as they can utilize such flexibility to get into the workplace quickly or spend more time taking useful courses and participating into internships. Finally, the university itself is also an important factor. The academic reputation, reasonable cost, Aggie networks, altogether make Texas A&M my dream school.

For all these reasons above, I firmly believe Texas A&M University and specifically, the MCS program is the best path for me to achieve my goals towards a data mastermind.