Statement of Intent

I graduated from the statistical computing master program, Data Mining track at the University of Central Florida(UCF) in August 2021, and I have actively worked in Dr. Larry Tang's research group at the National Center for Forensic Science, UCF from January 2020 to August 2021. The research and study experience at UCF has opened my eyes to a broad range of data science research and motivated me to undertake further study through a Ph.D. program.

Before I started my master's program at UCF, I worked as an auditor for four years in PricewaterhouseCooper, a large multinational auditing firm. The company was great and let me know about many other industries, such as medical, chemical, and automobile. While working there, I have supervised the implementation of over ten projects and gained thorough auditing knowledge, as well as better multitasking abilities, communication, and project management skills. However, the longer I worked there, the more limitations I found in my job. From my perspective, the biggest step backward is that we were still using excel to analyze the financial data, which was time-consuming and inefficient, and would be incompetent shortly since the statistical method is and will be more widely used in many fields.

After careful consideration, I decided to resign and pursue a graduate program related to Data Science. In 2019, I enrolled in the Data Mining MS program at UCF and benefited greatly. The program provided me with a larger and more fascinating world of Data Science, both theoretically and practically, which triggered my intention to undertake further study in this field. I joined Dr. Larry Tang's research group, as a Graduate Research Assistant, in January 2020 at the National Center for Forensic Science at UCF. I enjoyed my research work here. My main work included literature review and R programming for the receiver operating characteristic (ROC) curve. I coded R packages for estimating ROC curves based on forensic comparison scores from computer algorithms and forensic examiners. I also developed a Shiny application (https://menglingheshiny.shinyapps.io/orderROCweb/) that estimates ROC curves and AUC using both conventional and order-constrained methods. Furthermore, under the recommendation of Dr. Tang, I helped Dr. Emanuela Marasco with her papers, especially in the data analysis section. Our two papers got published successfully, and it was a great honor for me to be listed as the second author and to represent both papers at the conferences. The study and research experience greatly enhanced my confidence, encouraging me to continue studying through a Ph.D. program.

What is the purpose of a Doctoral degree and what am I hoping to do with or get out of the degree?

My career goal is to become a researcher who can innovatively solve cross-domain research questions using solid and in-depth knowledge of statistics and data science. My past education and work experiences made me realize the importance of data science. I hope I can have the chance to get in-depth knowledge of statistics and data science. Data science fascinates me, making me want to explore this domain more broadly and profoundly. I hope to get more interdisciplinary research opportunities using advanced statistical methods and data science techniques.

What makes me want to apply to the Bredesen Center, as opposed to other programs at UT or elsewhere?

I got admitted to the Big Data Analytics Ph.D. program in our department in 2021, but then I relocated from Orlando to Knoxville in 2021 with my family because of my husband's job change. So I tried to find a similar program where I can continue my research in Knoxville. It was a very occasional opportunity for me to learn about the Data Science and Engineering Ph.D. program at Bredesen Center. I believe this program is a good fit for me, considering my educational background. I'm impressed by the interdisciplinary opportunities that this program could offer. The strong collaboration with ORNL also attracts me. I heard a lot of excellent research is happening in the lab. I also wish I could have the opportunity to be part of the research and contribute my knowledge and experience to enable better research discovery.

How do I see myself as a leader, and how might I apply that within this program or my research?

Leadership is essential especially in projects. During my third and fourth year working in PWC, I have been team leader for over ten projects. From my perspective, being a leader is different from being a team member. As a team member, it is always important to focus on your work. However it is not enough when you are a leader. Leadership means you should have a general framework and timeline of the project even before it starts, and you should always keep monitoring the project's progress and communicating with your team members. You don't have to know exact detail about other members' work, but you should know what everybody is working on.

What are areas of growth that I'm hoping to address during my program? (i.e. strengths and weaknesses)

- I would like to enhance my knowledge and techniques in statistics and data science through more advanced graduate courses from a different department, e.g., math, statistics, and computer science.
- I wish I could broaden my eyes with the various interdisciplinary research opportunity provided by this program. At the end of the day, I hope I can have not only solid knowledge and techniques in statistics and data science but also a specialized area of application, e.g., health science and social science.
- I've done several research projects with various data analytics and modeling types. However, my experiences were with relatively small sample sizes. Through this program, I hope I can have the opportunity to also work with large-scale datasets and involve more efficient and powerful data analytic and modeling tools.
- I was heavily involved in two research publications, which were published at conference proceedings. I wish to work on more comprehensive and in-depth research, which could end up in several journal publications.