Academic Statement of Purpose

Shuo Liu for Data Science M.S. at UMich

I. Why Data Science at UMich

During the summer vacation in 2017, I travelled to the U.S. and visited the University of Michigan (UMich) for the first time. I was drawn to the gorgeous architecture and beautiful scenery on campus. The thought of studying at UMich came to me after the visit. After conducting some research afterwards, I found that Data Science M.S. at UMich is a great fit for me, as I can make full use of my experiences to better understand the theories taught in this program. In return, UMich's prestige faculty and wide platform will provide me with the opportunities to study in-depth specialization about this subject, and the alumni network of UMich will also be a great asset to my career development. This is why I am going to pursue the Data Science Master's at UMich.

II. Academic & Professional Accomplishments

I accumulated a fair amount of knowledge during my undergraduate studies and developed many interdisciplinary skills from my experiences in different areas. I believe I am qualified for the Data Science Master's program at UMich because of my strong will, hands-on skills, and teamwork spirits shown from the following experiences.

Learning & Competitions
My knowledge of Machine Learning mainly derives from my undergraduate courses and academic competitions. In 2019 Mathematical Contest in Modeling, I led my teammates to complete a Data Mining project about Opioid Crisis. We analyzed data from National Forensic Laboratory Information System, and found the growth pattern of drug reported quantity and important demographic features to make a breakouts prediction, *i.e.*, when and where a drug epidemic will occur. As a result, we recommended the committee to strengthen control for specific demographic groups in some areas. Besides, in my final projects in *Introduction to Data Mining* course (scored 99, ranked 1st), I made a prediction for NBA players' salaries by analyzing the correlation between their performance and earnings from Basketball Reference records. From these experiences, I become familiar with the process and methods of Data Mining, *e.g.*, data preprocessing, correlation analysis, classification algorithms, *etc.*, which could be of great value to the concentration of Prof. Hansen in criminology and medicine analysis, and Prof. Feinberg in demographics analysis.

My interest in Machine Learning involves other fields, too. I started to learn about Computer Vision from my *Artificial Neural Networks* course (scored 99, ranked 1st), where I practiced many kinds of neural networks. To further study, I led a team in a Kaggle competition relevant to cactus identification and achieved a 99.97% accuracy in the test set. Afterwards, I extended my studies to dynamic objects capture through my research in China University of Mining & Technology. I employed Back Propagation Neural Network and Partial Swarm Optimization to track targets and published a paper *Optimal Analysis of Target Dynamic Tracking Strategy Based on Computer Vision*.

Research & Publications
Since sophomore year, I have conducted research in Inplus Lab and focused on the contract and application layers of Blockchain technology. Referring to theories in *A Primer in Game Theory* and *Convex Optimization*, I proposed a two-layer Stackelberg Game data trading mechanism in Blockchain-based Internet of Vehicles and evaluated the robustness and efficiency of my algorithms by implementing several smart contracts on Rinkeby testnet of Ethereum. I completed a paper *Blockchain-Based Digital Goods Trading Mechanism in Internet of Vehicles: A Stackelberg Game Approach* with my colleagues and submitted it to 2020 IEEE Cloud. During this process, I realized that there is a big gap between idea origination and perfect implementation. Specifically, we need to consider more factors to ensure the stability and efficiency of operation in the system, such as security and cost of execution and storage, *i.e.*, gas cost. As this work uses Game Theory methods to facilitate trading, I believe it is aligned with the studies of Prof. Feinberg in socioeconomics. During my research in Inplus, I also participated

in *Perishable Digital Goods Trading Mechanism for Blockchain-based Vehicular Network* and *BCShare: A Decentralized Data Storage and Sharing on Blockchain*, and published a survey *Application of Blockchain in IoT Data Trust and Information Available Technology*. The research in Inplus Lab deepens my understanding of the application of Game Theory in Social Science, which prepares me well for my study at UMich.

Internship
Based on my knowledge and experiences in Machine Learning, I put into practice knowledge of Reinforcement Learning while interning in the Institute of Automation, Chinese Academy of Sciences. Because of my interests in Game AI, I participated in StarCraft team to build StarCraft II Learning Environment with Tensorflow. Initially, I trained the soldiers with Advantage-Actor-Critic and Deep Deterministic Policy Gradient (DDPG) algorithms, but neither brought satisfactory outcomes. After communicating with my advisor and colleagues, I enhanced my DDPG-based work by allowing the soldiers to cooperate. By referring to *Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments*, I applied the novel Multi-Agents Deep Deterministic Policy Gradient (MADDPG) algorithm to this scenario and took more factors into consideration in the rewards. Derived from DDPG, the MADDPG critic not only inputs its own state and action, but also includes others' information to make a global optimization. Due to this revision, the average winning rate of soldiers was improved from 26% to 43% with only 120 training epochs. From this internship, I picked up more knowledge in reinforcement learning, including traditional algorithms, training methods, evaluating criteria, *etc*. Furthermore, I learned more about business cases where Reinforcement Learning methods are being applied, which inspires me to integrate them into Data Mining topics in the future.

III. Interests & Goals

If admitted with honor, I am inclined to work on criminology, medical, demographics, and socioeconomics analysis in Social Mining to make full use of my experiences. Besides, I would like to explore biology and astronomy analysis of Prof. Chen as new fields of study.

In the short term, upon obtaining my Master degree, I expect to complete my program with excellent performance to consolidate my knowledge and to become a reliable employee in a reputable company. In the long run, I would like to set up my own company to promote promising great Data Mining tools or services to make our lives better. I firmly believe I can lay a solid foundation and learn useful skills at UMich to achieve my dream.