Statement of Purpose

From my various volunteering experiences, I have found myself very intrigued by the social bonds formed in the highly connected, yet complicated modern society. I believe that innovations of computational intelligence will lead to a deeper understanding of people's opinions, encouraging the interactions not only between human and technology but also between human through technology. Specifically speaking, my research interest focuses on the social and semantic aspects of recommender systems. I am writing to apply to the M.S. in Computer Science and Engineering (CSE) at the University of California, San Diego to acquire advanced training and professional networking to fulfill my goal of becoming a leading data scientist.

My interest in recommender system was inspired by the personalized music recommendation service of Spotify. When I was a major in computer science at National Chiao Tung University, I built a music streaming service featuring personal recommendation on Raspberry Pi with my teammates for the course Cloud Computing Systems and Applications. Beginning with my initial ideas, we developed the concept into a research proposal. Based on factual data and artificial intelligence APIs, we developed a voice-controlled streaming jukebox based on IBM Bluemix PaaS Service, and I was responsible for integrating IBM Watson API to achieve speech recognition and music recommendation. Our team successfully delivered solid connectivity and creativity with our device demonstrated in front of several IBM managers.

During my college days, I invested a significant amount of time in studying the feasibility of automatic theorem proving by establishing the link between evolutionary algorithms (EA) and proof assistants. Under the guidance of Professor Ying-Ping Chen, I have automatically proven ten theorems in different branches of mathematics. Our idea of applying EA to the realm of theorem proving is unparalleled, and our preliminary results clearly indicate a promising direction of finding formal proofs automatically. In 2016, we published our research results at the IEEE Congress on Evolutionary Computation (for details, see CV).

In the summer of my junior year, I interned as an R&D engineer at Trend Micro Inc.—a leading multinational security software corporations based in Taiwan—and accumulated data-related work experience. Aiming to enhance the performance of the anti-spam mechanism, I developed clustering algorithms on email metadata to filter out known legitimate entries and thus greatly reduced the volume for processing. I also performed frequency analysis on grouped data to simulate the patterns and characteristics of malicious mail attacks that customers had received. My research was highly regarded by my supervisor in the Coretech Department because the

framework I designed was able to identify patterns in the metadata of malicious emails out of originally unstructured data.

To pursue my interest in social computing, I have joined the Computer System and Communication Lab at Academia Sinica since March 2017. My work analyzes keyword trends to understand user intents on Taiwan Open Platform for Educational Resources from user log data. Advised by Dr. Jan-Ming Ho, I pinpointed the seasonality of search keywords and automated a keyword generation pipeline to recommend the most relevant resource to our users at the right timing. In addition to resource recommendation, I also tackle the problems in the field of bioinformatics by proposing an iterative process of *De Novo* Genome Assembly with XGBoost for subset selection of sequencing read. Featuring post-assembly analytics, my approach successfully improved the assembly results from 25% to 35% on N50 statistic.

For my Master's study, I aim to study the underlying theories and methods of building a recommender system with an emphasis on computational social science and linguistics. First, I will examine the foundation of collaborative filtering recommender systems and focus on techniques that address unintended filter bubble due to selection bias. Next, as the first step of extracting personalized features, I plan to widely investigate the word embedding techniques of learning word vectors, e.g., word2vec, from huge in-domain data sets. Lastly, continuing my research experience of developing a streaming jukebox, I will apply my design to the music industry and work on music recommendation services.

I believe that the M.S. CSE at UCSD will best guide me to fulfill my goals of understanding social behaviors and connect people through technology. I am interested in conducting research with Professor Julian McAuley because his works in preference modeling apply thought-provoking research ideas in various social communities, ranging from predicting popular submissions to reddit to bundle recommendations on a game distribution platform. I am confident in succeeding on the diverse campus of UCSD after accumulating international experience by studying as an exchange student at the University of Technology of Compiègne in France. I am eager to join the UCSD community.

My career goal is to tailor personalized recommendation services to promote the sharing of ideas. Studying at UCSD will be vital in reaching that goal. Thank you for your kind consideration.