Statement of Purpose

My first encounter with data science started with Spotify. In 2012, when Spotify was launched in Taiwan, I became fascinated with this music recommendation service and marveled at the apparent omniscience of artificial intelligence, specifically data mining and knowledge discovery. The joy of discovering and utilizing this incredible technology inspired me to investigate the process of using it for solving problems and extracting knowledge. I wondered how corporations utilize data to help people interact not only with technology but also more importantly with each other through technology. Airbnb, for instance, interprets data to make actionable decisions that help increase consumer trust and overcome the stranger-danger bias. Its data science team amplifies the voices of customers by teasing out their preferences from reviews and logs of interaction, which parallels how Spotify works with data. I am truly excited about the potential innovations that data science could bring to the world, and have decided to convert my interest into becoming a data scientist with the goal of bringing people together through technology. Achieving an M.S. in the Computer Science and Engineering program at UCSD will be vital in reaching that goal.

In my undergraduate project at National Chiao Tung University in Taiwan for the *Cloud Computing Systems and Applications* course, I gave tribute to Spotify by building a music streaming service on Raspberry Pi. Our team successfully delivered solid connectivity and creativity with our device demonstrated in front of several IBM managers. I also wrote a thesis-like project proposal to illustrate my core idea of this intellectual device. Featuring personal music recommendations based on factual data and artificial intelligence APIs, the service is a voice-controlled streaming jukebox based on IBM Bluemix PaaS Service.

Moreover, I also spent a significant amount of time researching intelligent agents. Instructed by Professor Ying-Ping Chen, I led a group to examine the feasibility of automatic theorem proving by establishing the link between evolutionary algorithms (EA) and proof assistants. The idea of applying EA to the realm of theorem proving is unparalleled, and our preliminary results clearly indicate the promising direction of research in finding formal proofs automatically. All result were presented under the title of "*Automatically Proving Mathematical Theorems with Evolutionary Algorithms and Proof Assistants*" in the IEEE World Congress on Evolutionary Computation (CEC).

In my senior year, I participated in the *Right Whale Recognition* competition on Kaggle with two other CS undergraduates, and we improved the evaluated score by thirty percent in limited time. This was a huge accomplishment in the competition. We dealt with the identification of

endangered right whales in the North Atlantic via data mining and computer vision techniques. I contributed substantially to the image classification part by instilling deep learning frameworks from Caffe, developed by Berkeley Vision and Learning Center, to a neural network model in MATLAB into our trial. Another contribution I made was to extract distinctive image features into a high dimensional histogram vector trainable through classifiers by adopting a bag-of-features model.

I was delighted and fortunate to gain more data-related work experience in a multi-national technology corporation when I served as an R&D Intern at Trend Micro Inc. My research was highly regarded by the Coretech Department because my designed framework was able to identify new paradigms or behavior out of originally unstructured data. Aiming to enhance the performance of the anti-spam mechanism, I implemented clustering on raw data to filter out known legitimate entries and thus reduce 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.

During my exchange experience at the University of Technology of Compiègne in France, I seized the opportunity to step out of my comfort zone. I have spent three years learning French and traveled as much as possible through backpacking, couch-surfing, volunteering, and hitchhiking. I relish those moments in which I can free my mind to experience a new environment and am confident of integrating these precious life experiences into my ongoing education.

From Spotify to my publication, every one of my endeavors sprang from my ambition to pursue a career in computer science. The philosophy of education in M.S. CS at UCSD will help me fulfill my aspiration to bring people closer through technology, while the innovative curriculum will further enrich my critical understanding of machine learning, and help launch my career as a data scientist. I am particularly drawn to two solid courses that lay the crucial foundation for machine learning: *Machine Learning Theory with Raef Bassily (CSE 250C)* and *Data Mining and Predictive Analytics with Prof. Julian McAuley (CSE 255)*. I look forward to learning from and contributing to the UCSD campus community both as a student and as a future alumnus.