

Dear Admission Committee,

I am writing to express my interest in the Master's program in Computer Science (CS) at the University of Texas at Austin (UT Austin). Having pursued a Bachelor of Computing in CS degree at the National University of Singapore (NUS) with specialisations in Database Systems and Artificial Intelligence, I have had the chance to build up a strong theoretical foundation and then utilise it extensively across a few internships and my first job as a Data Engineer at Sea Group, a global consumer internet company established in Singapore. From all the years in school and across different working environments, I have grown to appreciate the importance of having a good mix of theoretical and practical expertise to truly bring values and quality to the work that I produce as a Data Professional and a Software Engineer. Having this belief deeply rooted in my mind, I thus seek to further broaden and deepen my knowledge with a Master's program in CS at UT Austin with the goal to continue applying CS knowledge to solve problems in the real world. I strongly believe that given my background, experience and strong passion for technology, I would be a great candidate for the program and could benefit tremendously from all the exciting Applications and Theory Courses that the University has kindly offered.

At the end of my first two years of undergraduate study, I was at a crossroad in choosing what to specialise and subsequently pursue as a career in. As a naturally curious person, I hoped to eventually pick up from school the necessary knowledge to better make sense of a world being increasingly influenced and shaped by technological advances. Many of the specialisations which NUS offered were exciting, but data stood out to me particularly as an intersection between various areas of study which could truly make me a more informed individual. Having previously researched about the growing usage of data in decision making, I decided to take up Database Systems and Artificial Intelligence as my focus areas. This marked the beginning of a journey where I dived deep into how a database system organises data for efficient querying and supports concurrent access; learnt to analyse a query plan for tuning and optimisation purposes; horizontally scaled my thinking to understand distributed computing in big data systems; programmed computers to imitate human cognitive abilities by learning hidden patterns in data. In retrospect, even though I might not have fully grasped the essence of these concepts at that point in time, I believe they have laid a strong foundation to prepare me well for my subsequent working experiences.

The aforementioned background landed me my first internship at PSA International, a global port operator and transshipment company based in Singapore. There, I was tasked with researching and prototyping a model based on Reinforcement Learning (RL) to automate the process of reordering the existing stacks of containers in a yard to help optimise the way they can subsequently be loaded onto incoming vessels. This process had been traditionally done manually by the port operators based on their experience and intuition which could be error-prone and time-consuming. Despite having no previous exposure to RL, I began to teach myself from the ground up starting with the impressive *RL Course by David Silver* offered by DeepMind, followed by reading notable research papers and trying out different RL frameworks. Despite many challenges encountered due to the inherent complexity of the

field as well as the difficulty in formalising and modelling the operators' intuition with code, this experience was eye-opening for me as a student back then seeing how theory complemented practice to bring out promising solutions to a very real problem. Thankfully, after six months of iterative work, I finished up a working prototype utilising Proximal Policy Optimisation to hand over to the team as a baseline for future enhancement and usage.

The first internship helped me tremendously in building up the confidence and appreciation of what I had been learning in school. Shortly after, I applied for another internship at my current company, Sea Group, to work in the Data team and continue developing my professional skills. This time, I was introduced to working with massive amounts of data generated from the company's regional e-commerce and gaming businesses, which were used in a big data storage and compute ecosystem known as Hadoop and Spark to build data models for financial reporting. My role was to suggest potential optimisations to improve runtime and resource usage for a few critical yet inefficient data models. Once again, when confronted with a new and challenging technical problem, I first turned to theory before taking actions. After building up a good grasp of the Distributed Computing techniques that Spark leverages, I proposed to the team two key directions for optimisation. The first involved setting the right configuration for each Spark application by relying on the Spark metrics API to resolve inadequate resource allocation issues. Afterwards, I suggested analysing the Spark SQL plan programmatically to identify any serious query logic or data problem. When such optimisation rules showed promising results, I was tasked with building an auto-tuning program to encapsulate the optimisation process so that all members of the team could apply it on their own data models. After the development process, I was rewarded with a full-time offer as a Data Engineer for the team.

As a Data Engineer, I now focus on building reliable pipelines to ingest data timely, efficiently and securely into our big data warehouse for downstream users. Having previous exposure to building data products for both operation and business, I have grown fully aware of the importance of data quality as well as its accessibility to users. At the same time, given my previous tuning experiences, I was entrusted to continue diving deep into the inner workings of big data tools to act as a consultant who provides helpful insights and good practices to other members of the team. My job scope thus spans from maintaining existing Extract-Load-Transform frameworks to building new automated data ingestion services for other internal teams and optimising critical data processing models. One of my most significant projects was to build the engine for a web service allowing users to schedule automated data ingestion pipelines from various data sources such as relational databases, Google Sheets, CSV and Kafka into our Apache Hive data warehouse. These experiences continued to be challenging yet rewarding and reaffirmed my appreciation for CS knowledge.

Despite the varied tasks that I have worked on in the past, what kept consistent to me was the way theory complemented practice to bring out inspiring solutions to difficult problems. For this very reason, I can see great values in furthering my theoretical foundation with a Master's program in CS at UT Austin and am committed to making the best use of my time learning and applying all the knowledge that I will be taught in solving real-world challenges.