Since childhood, I have had a strong interest in numbers. When someone tells me his phone number, I can remember it quickly, and to exercise my calculation ability, I habitually sum up this string of numbers and decompose it into prime numbers. As I grow older, I have been exposed to various subjects both inside and outside of school, such as robotics, programming, statistics, and engineering, and all of these fields require data analysis. In recent years, with the development of computing power, artificial intelligence based on data science has impacted every aspect of our lives. As a result, data science has become a crucial aspect of many industries and could be seen as the foundation of the fourth industrial revolution. As an interdisciplinary field that combines statistics, computer science and engineering, the prospects of its development fascinate me.

During my spare time, I took an introductory course in data science and collaborated with my classmates on a case study involving the identification and categorization of coins using computer visualization techniques. Through this experience, I gained a comprehensive understanding of the application of data science, from data collection and classification to building convolutional neural network(CNN) models and performing data analysis using Python programming. We published a paper “Brazilian Coin Counter Research Report” on this topic and in which present a coin identification and validation model based on the AlexNet convolutional model.

In this process, I realized that my previous understanding of data science was limited to algorithms and programming. However, data collection and labeling form the cornerstone of this process. Having a clear and logical dataset is crucial for accelerating learning by computers.

On the other hand, I find that CNNs heavily rely on matrix operations, such as convolutional filters, which extract features from complex datasets. Related mathematical principles like linear transformations and matrix factorizations underpin the efficacy of CNNs. I first studied linear algebra during my ACT course but did not fully understand its significance in applied scenarios. But during the big data project, I am amazed by the powerfulness of it and would like to learn more about linear algebra systematically in my future university study.

The research of data requires a solid mathematical foundation, so I have actively participated in various mathematical competitions. From number theory to functions, from geometry to probability, all kinds of mathematical problems constantly challenge me. I remember once when preparing for the AMC, I encountered a dice-rolling problem that asked for the probability of getting an odd number before all even numbers appeared once. I was consumed by the problem and ended up skipping dinner , but the feeling of satisfaction after finally arriving at the correct answer before going to bed was incomparable to anything else. From permutation and combination in elementary school math Olympiad to conditional probability in high school, such problems have sparked my great interest in data and statistics.

After mastering a solid mathematical foundation, I could not wait to apply the theories to practical research, so I was honored to participate in the Space City Competition, a global event organized by NASA. In the Asian regional finals, we were tasked with using the Lagrange method to calculate the position of the perigee. The principle behind it is based on the condition of force equilibrium for a mass in a gravitational field potential field. Our team utilized mathematical modeling and equation-solving techniques to determine the location of the Lagrange points during the competition. Concepts such as potential functions and integration methods in mathematics played a crucial role in practical applications. Meanwhile, data also play a key role in this research, and the combination of them reinforces the range of problems they can solve.

During my leisure time, I often read books about mathematics and its applications. Among them, "The Beauty of Mathematics" is one that I frequently browse through. What impressed me the most in the book was the example of using Hidden Markov Models for speech recognition. Speech recognition encounters so many languages and contexts, and I have always been curious about how computers achieve semantic understanding. By learning from a large amount of speech data through such an intuitive and concise probability model, computers can gradually engage in conversations with humans. It is truly fascinating. At the same time, I also realized that the accuracy of data is important for the training of the model, and the cleaning of these underlying data requires enough patience and care.

I also enjoy building various models ranging from robots to paper models. It's fascinating to take a pile of scattered materials and assemble them into a cohesive whole or transform a stack of paper into an airplane or a battleship, which fosters my attention to detail. I often find myself engrossed in this work from morning till night, sitting in one place until I am satisfied with the final outcome.

In the future, I aspire to study Data Science and Engineering and apply it to industries such as finance upon graduation. Hong Kong, as an international financial center and a convergence of Eastern and Western civilizations, has always been a city I yearn for. The University of Hong Kong, being the top university in Asia, is my dream school. When I came across the webpage of the Musketeers Foundation Institute of Data Science and learned that the university considers it first of the ten research institutes in its technological development blueprint for the next decade, I was determined to apply for the Data Science and Engineering program. I hope to have the opportunity to study at the University of Hong Kong. Ultimately, my goal is to continue enjoying the pleasure that data brings to my life.

HKUST

The Hong Kong University of Science and Technology (HKUST), a relatively new but highly reputable university known for its emphasis on science and engineering, holds a leading position in mathematics, statistics, computer science, engineering, and other fields in Asia. I am enthusiastic about the opportunity to study Data Science and Technology at HKUST.

CUHK

The Chinese University of Hong Kong (CUHK) is a leading institution in artificial intelligence and data science in Asia. The Multimedia Laboratory founded by Professor Xiaoou Tang has been one of the world's leading laboratories in this field. It has nurtured industry-leading artificial intelligence companies like SenseTime and has produced top AI scholars, including Dr. Kaiming He. I am highly interested in the opportunity to study Computational Data Science at CUHK.