1. Intro

In 2016, an A.I go program of Google won over a Korean professional go player, Se-Dol Lee. Because of this big event, Koreans could feel that the future has finally come to our days. As a student in college of statistics, I can also feel a bright possibility of machine learning. Although I felt some frustration when I noticed that Korean companies disregard statisticians, I can overcome this when studying more and feeling that statisticians can handle more essential parts of machine learning and data science. More specifically, there are many future technologies that have stalled in development such as genome or nuclear fusion. I expect that a PhD can do a lot of things to realize them. It is most exciting for me that the imaginations of normal people remain as dream but the imaginations of doctor can become reality.

In a few years, a new game will appear; Quantum computer. I expect that IBM will make 1000-qubits quantum computers in 2~3 year. With a 1000-qubits, quantum computer algorithms could come out of paper and become a reality. With the quantum computer, we can implement super complicated models and can deal with the eigenvalue problem by exponentially improving algorithms. Ideally speaking, I believe that with the emergence of quantum computers, statistics itself also can experience groundbreaking developments. If we inject the concepts of superposition and entanglement to former probability theory, we can not only make statistical models fitting with quantum computers but also might be able to redefine the probability, the information and the entire philosophy under the statistics. Maybe we could develop a third branch of statistics; Frequentist, Bayesian and Quantumist. Therefore, I can conclude that I'm mostly interested in improving data science methodology using quantum computers;It might be called as quantum machine learning or quantum information theory. The best scenario for me is getting PhD degree, finding an academic job at university and researching quantum inference for data science.

2. Experience at Yonsei University

During a master of data science at Yonsei university, my main research field were dimensionality reduction and nonparametric model. Because the fundamentals of the field are highly complicated and difficult, I have to spend a lot of time to study real analysis and theory of Hilbert space. However, courses about them in Yonsei university did not target for understanding statistical structure and contained a lot of redundant subjects. Following the advice from my instructor, I studied by myself using "W.Rudin, Real and Complex Analysis" , "L. Debnath, Introduction to Hilbert Spaces" and "B.Li, Sufficient Dimension Reduction". Even though this was really hard time for me, I can understand the structure and organize the them on my github blog. Likewise, I studied quantum computer using the books of "Sutor, Dancing with qubits" and "Schuld, Supervised Learning with Quantum Computers". Fortunately, my advisor, Hakbea. Lee was also deeply intrigued by the quantum computer and managed the Yonsei-IBM cooperation foundation. Thus I can research the field more efficiently. Finally, I can get a "certificate of quantum excellence" at the 2021 Qiskit Global Summer School held by IBM and taught students as TA at the course of "Quantum Computer for Big Data" at Yonsei university. All of these thing are done by myself with just few help of others. Therefore, I expected that such self constructing experience could help me a lot for research of PhD.

I also wrote a paper named "Estimating Deep Learning Prediction Distribution". It is a methodology for interpretable Deep Learning which can be done by a combination of Deep Learning and Sufficient Dimension Reduction. By applying dimension reduction techniques at the last layer of well-designed Deep learning structure, it can make it possible to derive approximated predictive distribution. Even though the idea looks simple, with some beautiful statistical

properties of SDR, the result of them can be interpreted as predictive distribution which is a main goal of the research field of Deep Learning Uncertainty. Moreover, the methodology also can be used to visualize deep learning algorithms. Although it failed to give an exact explanation about structure, by depicting each data point in a 2-dimensional plot, it can give some hints about the result. Also, it can be used to detect some outliers which hinder the model's accuracy. While writing the paper, the main goal of the paper kept varied and I felt some frustration. However, I could learn various data science techniques and understand the mathematical structure of the non-parametric model. I believe that the experience would be greatly helpful to deal with quantum machine learning structure.