Statement of Purpose

Simply with relevant domain data, complicated problems demanding intricately crafted heuristics can be addressed with black-box machine learning (ML) models. The seemingly "magical" capabilities have always captivated my interest, driving me to explore the potential of ML at Shanghai Jiao Tong University. As a Computer Science student, I was initially interested in designing algorithms and solvers to enable computers to solve real-world problems. However, it wasn't until I discovered the world of ML and the potential for learning from large data to revolutionize the computer science that I truly found my calling. At the first class of Machine Learning in the junior year, my professor Shikui Tu exhibited how *transformers* were capable of processing both images and textual questions and generating answers for the questions based on the image accurately. I was immediately drawn by the power of transformer to capture the visual and textual features and the potential “wisdom” to analyze the problems. In the following few weeks, I read over 20 papers on the frontier of AI and audited Andrew Ng’s Stanford CS229 ML course. I became convinced that ML was where I belonged to and excited to witness the future developments in ML. I therefore switched my focus to ML and made it my dream to become a distinguished ML researcher to explore the potential for real-world scenarios.

My journey in ML started when I joined the SJTU-ReThinkLab led by Prof. Junchi Yan. Under his guidance, I began researching graph matching (GM) which aimed at finding point pairs on the given images. With substantial literature review on GM, I found a gap in works on front-end backbones (e.g. CNNs) as most of the existing works focused on the designing back-end modules to better capture the underlying graph structure. Therefore, I proposed to design a GM-specific backbone to make improvements. Inspired by the success of attention mechanism, I decided to build my work on ViT. However, the obstacle was that original ViT extracts global information, rather than the local key-point information. To address this limitation, I devoted several days studying the paper and code of ViT, discussing with the professor to understand ViT's capabilities. After reflection, I proposed modifying ViT by introducing patches centered around the key-point and adopting the cross-attention mechanism to leverage the global information as auxiliary for the local information. I also harnessed the attention mechanism to enhance the back-end module's ability to capture the graph's underlying structure. Consequently, my proposed method achieved the state-of-the-art (SOTA) performance on extensively used GM benchmarks and the proposed backbone improved most existing frameworks by over 5%. The experience for my first research imparted to me the knowledge of how ML research was conducted and further sparked my interest in exploring their potential. I concluded my work in the paper submitted to ICASSP as the first author.

Witnessing the success of my method was truly satisfying and I became eager to explore the potential of ML in other combinatorial optimization (CO) problems. I participated in another research in our lab aiming at designing the guided diffusion model for CO, such as TSP and MIS, and rewrite the generation process to avoid local minima. One of the most challenging problems is the design of guidance. Diffusion model mimics the Brownian motion to denoise a high-quality output from a noised input. Injecting guidance to the generation required reformulating the Bayesian probability distribution, demanding a strong grasp of mathematical skills. I spent substantial efforts learning the structure of diffusion models and conducting mathematical derivation meticulously. After contemplation, I proposed approximating the cost of solutions as guidance using Taylor’s First Order Expansion and Energy Function to represent to posterior probability. Nevertheless, I found guided diffusion model showed marginal improvement on TSP and MIS (two NP-hard CO problems) compared with existing works. I discussed the phenomenon with my group members and proposed to use Bernoulli distribution instead of Gaussian distribution to model the generation process so as to better represent the discrete distribution of 0 and 1, which was more associated with our tasks. This solution addressed the issue of continuous probabilistic modeling and improved the performance by 10%. Eventually, our proposed guided-diffusion model achieved the SOTA performance on TSP and MIS benchmarks compared with learning-based models and reduced the solving time by one order of magnitude compared with exact solvers (e.g. from 1h to 4 min). Our paper was accepted by NeurIPS and I was the second author.

The research in CO honed my programming and mathematical skills and boosted my curiosity to explore ML in more domain-specific problems. Therefore, I joined the ML-PL group led by Prof. Xujie Si affiliated to the University of Toronto and Mila - Quebec AI Institute. I researched on logical puzzles that can be formulated as

SAT expressions, such as Sudoku. My work was built on SATNet, a differentiable MaxSAT solver. Nevertheless, while analyzing the paper of SATNet, I found the learned parameters cannot even represent extremely simple rules such as XOR, contradicting paper's claim. After discussing with group members, we identified a theoretical error in the interpretation. Armed with mathematical knowledge, I intuitively proposed rewriting the SAT conjunctive normal form directly as a parametric matrix. The main challenge in this project was the unfamiliar programming language CUDA, as all my previous projects were written in PyTorch. To implement my idea, I delved into the NVIDIA CUDA tutorial. Although CUDA is a much more obscure programming language compared with PyTorch, I grasped the basic principle and syntax Within three days, and implemented my idea based on SATNet with group members. The proposed idea substantially reduced the number of trainable parameters while remedying theoretical mistakes in the original design. Notably, once the rules are learned, the model can definitely output correct answers, and this representation made the learned rules entirely interpretable, marking a significant breakthrough in the logic reasoning field. Our paper was accepted by NeurIPS and I was the second author.

The positive impact of my proposed ML method on the specific field and the exhilaration once my paper was accepted including at NeurIPS strengthened my determination to be a future prominent researcher in the ML field. Now I dream of designing ML frameworks that surpass human experts, well-designed algorithms and solvers in terms of both accuracy and efficiency for tasks involving logical reasoning. Therefore, I aspire to acquire advanced ML knowledge and pursue further research on ML such as *Symbolic Reasoning*. The MSCS program at Stanford, with its top-notch faculty and impeccably crafted curriculum, is the perfect incubator for this dream. With the elaborately designed AI Specialization, I can not only gain valuable insights into Graph Modeling via *Machine Learning with Graphs* but also advance my knowledge of ML for reasoning with *Introduction to Automated Reasoning*. I am particularly excited to work with the pioneering researcher Prof. Thomas Icard on *Causal Inference.* I am curious to witness how ML exhibits the potential of reasoning and conducts inference step by step, drawing on my previous research experience in ML, particularly in reasoning to help push the boundaries of this exciting field.

With a robust quantitative skill-set, extensive experience, and a deep passion for expanding my knowledge of AI, I firmly believe that I am a well-suited candidate for the MSCS program. I am confident that my Master's journey at Stanford will equip me with the expertise demanded to make insightful contributions to the AI field, pursue further research as a PhD and finally contribute back to the community at Stanford.