Statement of Purpose | **Xinru Wang**

For as long as I can remember, I was the only girl in my Go class. Slightly introverted, I only had as my best friends my two boxes of elegant, black and white Go pieces made of agate and purple jade. Six years of training and tournaments made me a 4-dan at ten and No.1 in the women's Go game in Hangzhou city, one step away from entering the 5-dan rank and becoming a professional chess player. Representing the Youth Group of Jiangxi Province Go Team with my aggressive style and formidable calculation skills, I enjoyed every game and contest. But everything changed when AlphaGo easily defeated Lee Sedol and Ke Jie, both former world No. 1s, in 2016 and 2017, respectively. That was why I chose to study computer science in college, which seemed a far more exciting game. Since graduation, my professional experience and, particularly, daily exposure to astronomical data and immense operational complexity in Meituan, one of the leading e-commerce platforms in China, further convinced me of the necessity of graduate-level training to better harvest AI power.

I am confident in my preparedness for graduate training, given my extensive research and professional experience. For example, I secured a research opportunity at Columbia University's Cathaypath Institute of Science in May 2022. My team aimed to conduct a sentiment analysis of Amazon Kindle Store's classified product reviews. It was a typical natural language processing task, involving tokenization, lemmatization, features vectorization with TF-IDF, and emotion scoring based on fuzzy string matching and NRCLexicon dictionary. To unlock NLP's full potential, we incorporated a time series analysis to understand how long it took for reviews' sentiment to affect sales results. Then, we employed grid search for optimal lag length while training a Decision Tree, a Support Vector Machine, and a Naïve Bayes classifier to predict the ordinal sale changes. With our models achieving prediction accuracies of 72%, 68%, and 69%, respectively, Dr. Patrick Houlihan, our supervisor, encouraged us to submit our work. It was then published with me as the first author[1].

Craving for more professional experience, I took the job offer from Meituan's Daojia Business Group (DBG), which responded to more than 50 million online orders (of food, pharmaceuticals, and other retail goods) and managed 10 million delivery riders daily. Thousands of teams and divisions heavily depend on back-office applications, such as data management, business process management, and operational analysis. To ease the pain of both front-end and back-end developers in developing and deploying such applications, we created a low-code, cloud-based development platform named WOLF. My main contribution was to transform the DSL-based configuration method into a graphic one, with which any business idea could materialize into workflows well-defined by a UI element tree. To further refine its composition, I created the event configurator for rapid event attribute selection and the event link for upstream and downstream events to compose or branch events. Other than that, I devised an expression configurator, a style configurator, and a manager for adapters and interceptors; I invented the overall interaction logic of the continuous integration and delivery pipeline. By the end of 2022, WOLF had helped 1319 developers in 200+ teams and hosted 1064 projects. Because of my performance, I was promoted to L6 within one year, compared to an average of three years for others.

While working on WOLF, I found that all business operators and managers fed on data. To address the concern of excessive time-consumption of gathering and summarizing data for operation supervision, DBG launched a front-end development platform, Diting (諦聽), named after Kshitigarbha's divine beast capable of listening to people's hearts. Based on WOLF, the

[1] **Wang, Xinru** & Jin, Xinran & Wu, Yongqiang & Liao, Yajing & Lei, Dongchen. (2023). Predicting product sales based on sentiment analysis using reviews: The case of Amazon's kindle. *Applied and Computational Engineering*. 6. 1641-1651. 10.54254/2755-2721/6/20230784.

first version of Diting needed a major update due to the expansion of the business, more data sources, more complicated business logic, and more refined operational granularities. Accordingly, I employed a new multi-tenant architecture and enforced the "Leave No Trace Principle" for incremental refactorization, progressively decoupling business logic from data and hiding complex data structures. Other improvements included route separation for different business lines, singleton patterns and adaptors for business logic, a state manager to centralize data source storage, and encapsulation of data access into functions for data independence of different business lines. Immediately, Diting decreased the average cost of developing a front-end application from 7 persons per day (P/D) to 0.5 P/D and reduced the cost of a data-report application from 5 P/D to practically 0 P/D. As of July 2023, we saw a remarkable 186% increase in business developers actively using Diting.

With the advent of ChatGPT and Copilot, I am impressed by their ambitions, which are way bigger than my WOLF low-code platform. While I aimed at low-code, they aspire for "no-code." Motivated by the same vision, I must receive graduate-level training. To that end, the Master of Science in Computer Science program (MSCS) at the Stanford Computer Science Department appeals to me for many reasons. Among the well-defined specializations, Artificial Intelligence aligns with my passion perfectly. I am first attracted by its rigorous curriculum, which reflects the requirements of foundations, breadth, and depth. For instance, Artificial Intelligence: Principles and Techniques (CS211), Deep Reinforcement Learning (CS244R), and Deep Generative Models (CS236) will lay a theoretical groundwork for my quest into code generation. Additionally, I admire the department's first-rate faculty, conducting research on exciting topics, such as formal methods to neural networks by Professor Clark Barret, investigation of bugs of and intervention with genome from the perspective of "code" by Professor Gill Bejerano, and long-tail behavior of a machine learning system by Assistant Professor Tatsunori Hashimoto. Besides the abundant research opportunities from numerous centers and institutes, I look forward to the enlightenment from the academic giants hosting frequent seminars and symposiums, including Stanford Graph Learning Workshop, HAI Seminar, etc.

Upon completing my training at Stanford, I will seek full-time employment at a leading-edge AI-focused company like OpenAI or Google DeepMind. Committed to the vision of a world where programmers don't need to code, I am open to any challenge.