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 publish our work 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 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 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 Domain Specific Language (DSL)-based configuration method into a visualized one, with which business ideas could materialize into workflow schemas described by UI element trees. To advance its functionalities, I implemented an event selector that streamlines the composition and branching of events, along with an expression editor capable of displaying context variables. Other than that, I devised an intuitive style configurator that injects CSS into components, thereby enhancing customization; I also invented the overall interaction logic of the continuous integration and delivery (CI/CD) pipeline. By the end of 2023, WOLF had helped 1898 internal developers in 200+ teams and hosted 1421 projects. Because of my performance, I was promoted to L6 within one year.

While working on WOLF, I found that all business operators and managers fed on data. Due to excessive time consumption of gathering and summarizing data from scattered Excel spreadsheets for operational supervision, DBG launched a CRM platform, Diting (諦聽), named after Kshitigarbha's divine beast capable of listening to people's hearts. Based on

---

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

WOLF, Diting's Minimum Viable Product (MVP) required a framework upgrade to accommodate the surge in data volume, expand the granularity of analysis, and adapt to the ever-evolving requirements. In response, I introduced the multi-tenant architecture for seamless scalability and data isolation while decoupling UI components from data for enhanced flexibility and dynamism. In essence, the new Diting offered broad pre-defined functionalities so that users could just "rent" them after registration and authentication. Then, I leveraged React Redux to maintain consistent and stable global data states, rendering all state changes transparent and trackable. Furthermore, to cope with the ever-growing codebase and ensure better code sustainability and readability, I utilized TypeScript to extend the states and reducers from outside of Redux, thereby achieving maximum elegance in my code. Diting significantly reduced the average construction cost of onboarding new users (tenants) from 7 persons per day (P/D) to 0.5 P/D and broadened the scope of analysis from POI to SPU, SKU, brands, etc. As of July 2023, we saw a remarkable 186% increase in business operators and managers actively using Diting.

With the advent of large language models (LLMs) like ChatGPT and Copilot, I am impressed by their ambitions. While my low-code platform still relies on developers' understanding of the codes, they generate codes from their understanding of developers. Motivated by the same vision, I must receive graduate-level training. To that end, the Master of Science in Computer Science program (MSCS) at Duke's Department of Computer Science appeals to me for many reasons. I am first attracted by the Artificial Intelligence/Machine Learning optional concentration and its course offerings. For instance, Theory and Algorithms for ML (CompSci 671), Reinforcement Learning (CompSci 590.02), Deep Learning Fundamentals (CompSci 547), and Generative Models (CompSci 590.01) will lay an all-around groundwork for understanding contemporary AI techniques and applying them in many real-world scenarios. Additionally, I admire the department's first-rate faculty, conducting transformative research on exciting topics, such as the analysis of medical images for diagnostic purposes, generalist robots, computational journalism, algorithms for structural molecular biology, etc. Besides the insightful guidance from the faculty and abundant research opportunities from numerous labs and groups, I anticipate gaining enlightenment from frequent academic events, including the Triangle Computer Science Distinguished Lecture Series, Identity & Computing Lecture Series, Duke Computer Science Colloquium, and CS-ECE Distinguished Lecture Series, where I can connect with academic and industrial leaders and passionate peers.

Upon completing training at Duke, 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.