

My research interests primarily focus on Natural Language Processing (NLP) and Graph Neural Networks (GNN), especially in sentiment analysis within social media contexts. I am captivated by the human-centric properties of NLP algorithms, while the interpretability of current end-to-end language learning approaches remains a significant challenge. I am enthusiastic about applying for a Ph.D. in Computer Science, aiming to dedicate my passion and endeavor to unravel the interpretability of sentiment analysis, exploring social media interactivity, and enhancing information dissemination. These pursuits are fundamental in comprehending language models and strengthening their credibility. As an old Chinese saying goes, "It is better to understand why it is so than just to know what is so."

## 1. Novel Insight on Interpretability in NLP

My interest began in my freshman year when I came across Dr. Fenghuan Li's research on utilizing GNNs to analyze depressive tendencies among Twitter users. This discovery ignited a profound interest in leveraging language models for real-world challenges, particularly in the intersection of NLP and GNN. Recognizing the critical need for interpretability in high-stakes applications, I embarked on a journey to enhance the interpretability of language models. My experience at the Graph Neural Networks Lab at the Guangdong University of Technology, where I focused on evaluating depressive tendencies through social media analysis, provided insights into improving interpretability.

Our research focused not only on achieving superior performance but also on exploring interpretability. We proposed employing a depression scale and objectively mapping social users' responses through prompt learning, shedding light on implicit psychological states in a novel manner. Additionally, we investigated diverse features reflecting users' depressive states by constructing a heterogeneous information network and utilizing heterogeneous graph and subgraph attention to capture interaction. Through a self-supervised comparative learning approach, we generated more robust discriminative representations. Emphasizing interpretability at the psychological level presented a novel research perspective, with our state-of-the-art approach underscoring its efficacy.

Sharing these findings as the first author with Knowledge-Based Systems was a proud moment, with my manuscript conditionally accepted. Under Dr. Li's insightful guidance, my focus shifted toward journal submissions, emphasizing robust experiments, comprehensive analysis, and data visualization. Publishing patents and software copyrights further refined my capability to extract precise research questions from literature reviews, translate concepts into practical applications, and validate them through meticulously crafted experiments. This research journey earned me the Bachelor's Outstanding Thesis Award, and I took great pride in achieving such milestones, especially given my background and the limited opportunities available at my university.

Proficient programming skills lie at the core of my passion. I translated my ideas into action by participating in Kaggle's Natural Language Processing of Disaster Tweets competition, achieving a top 5% individual ranking. Additionally, as a research assistant alongside professors, I shared my Python expertise with undergraduate students, introducing them to the realm of NLP. These experiences not only bolstered my programming and algorithmic foundations but also honed my teaching abilities, fostering potential

collaborations between faculty and students. I hope to continue to shape minds and inspire curiosity through my teaching in the future, as well as continue to hone my programming skills.

## **2. Curiosity of NLP Performance in Practice**

Driven by curiosity about the practical performance of language models, I participated in the Redbird Challenge Camp at Hong Kong University of Science and Technology, where our team focused on implementing motion-assisted robots. Addressing inefficiencies and low interpretability in voice input methods for space-constrained mobile devices, I structured graphs based on syntax and semantics, utilizing a comparative learning approach to derive scores. I am currently in the process of documenting this work for publication, aiming to contribute to the advancement of practical applications in NLP. This collaborative research experience enhanced my proficiency in task delegation, cooperative teamwork, and applying theories to address cutting-edge challenges.

While working at China Mobile Information Technology Co., I successfully deployed ChatGLM2 in a live production environment to communicate. By optimizing prompt words, I limited the GPU memory usage of a single model to below 2GB. Additionally, I introduced ChatSQL to enable non-specialized database users to query and interact with databases using natural language, enhancing its applicability across diverse data analysis scenarios and significantly boosting retrieval accuracy to over 85%. This hands-on experience deploying language models in a practical setting has substantially enhanced my proficiency in utilizing algorithms to comprehend natural language.

## **3. Pursuing a Ph.D. at UCSC**

The decision to pursue a Ph.D. is oft-debated. My decision solidified after my initial research experience, driven by a passion for exploration, skill refinement, and technological advancement. I am excited to apply to the Ph.D. program in Computer Science at UCSC, envisioning it as an ideal environment for my research growth. Collaborating with Prof. Jeffrey M. Flanigan, whose work in NLP and computational linguistics has inspired me, particularly from a semantic perspective, is a prospect I eagerly anticipate. With confidence in UCSC's abundant resources and collaborative atmosphere, I believe it can provide optimal guidance for my academic journey. I am eager to contribute my positive spirit and unwavering efforts to UCSC.

I firmly believe that language forms the bedrock of communication among individuals and with the world, bridging the gap between machines, individuals, and the broader world. Delving into the nuances of language, the precise connotations of words across diverse contexts, and the interpretability within each language model will offer valuable insights for language models to tackle complex tasks effectively. I am unwavering in my commitment to researching this goal, striving to unravel the complexities of language models and contribute to their progress in addressing challenging tasks.