

## Personal Statement

Every day we are encountering and absorbing a large amount of knowledge. The idea of “knowledge” first struck me in a movie recommendation project, I realized daily events can be transferred to knowledge graphs and contribute to the real problem, which is so surprising and exciting to me. Since that time, I always thinking about how to efficiently link ‘knowledge’ to AI in better service for our daily lives. I was exposed to AI during my graduate studies, and involved in research areas including computer vision, recommendation systems, data mining, and knowledge graph, which strengthened my knowledge of machine learning and shaped my interests in **natural language processing(NLP)**, especially **natural language understanding(NLU)**, **information extraction(IE)**, **information representation(IR)**, and their **application** to real life, I hope to make machines understand nature.

Responsible for NLP as a speech algorithm engineer in the past three years, I came across several excellent researchers and PhD students. I wish I could have a chance to heavily collaborate with salient researchers in the world and contribute to my fields with fulfillment. Getting research results really inspires me and makes me no longer satisfied with my current easy life, and guides me to pursue a Ph.D. I have prepared myself well for the Ph.D. research activities with hardcore programming and application development experience. I have mentored five interns, where I developed teamwork ability and enhanced my problem-solving ability. Motivated by strong research interests in machine learning and making breakthrough contributions to my interested field, I was actively engaging in research projects during my work and study afterward.

I am a motivated NLP researcher with strong learning ability and rich experience with fruitful publications. The following describe my research journey.

We, as human beings, are sometimes difficult to keep track of knowledge, and I was wondering how we can build a model to tackle this issue. My internship in TAL AI Lab offered me research experience in **Knowledge Tracking**. I built a deep knowledge tracking(DKT) model to monitor student knowledge states or skill acquisition levels. We considered the **inherent side relations** of the question sequences, and build questions’ graph based on semantic and feature-based knowledge. We use a **graph embedding** method to obtain the node representations as supervision information to DKT. This method demonstrates the effectiveness of side information and proposes an idea to extract sequence memory based on graph embedding(MovieLens e.g. ). The above research was completed by a Ph.D. and me and led to a co-author conference submission to *AIED 2019*. In this research, I did a lot of data mining work, which have provided me with powerful help to build a knowledge graph. At the same time, I innovatively extend my research to **adaptive learning**, which is critical for tutoring systems. I started to grow up my interest in information extraction and representation.

Apart from knowledge acquisition, I was interested in language understanding regarding semantics and emotions. I gained related research experience when I joined an AI company (Mobvoi) in 2019 as a speech algorithm engineer, during which I divided my work into two sides: semantic NLP and emotion NLP.

So far, I have mainly been involved in **text information extraction(Semantic NLP)**, in the guidance of accurate, natural, and emotional text-to-speech(TTS). The most fundamental work is converting non-standard text into representations with accurate pronunciation, prosody, stress, style and emotion. Modules include Tokenization, text normalization(TN), polyphonic disambiguation(Polyphone), Prosody, etc., based on cross-language (Mandarin, English, etc.).

Earlier, I used feature-based and simple neural networks(LSTM e.g. ) to build these units separately, which had limited performance. Later, with the progress of text representation, I proposed a model architecture called **Unified Frontend**, which aims to transform complex joint models into end-to-end(E2E) models. I built an effective semantic extraction model (pre-trained), as the backbone of all units, combined with multi-task training and knowledge distillation to get the final model. I also verified the maximum capability of each model to ensure the final model's performance. My solution mitigated the problems of out-of-vocabulary(OOV), ambiguity, and performance of the online system. In **multi-level prosody** task, I uses **prompt learning** to solve training difficulties due to data imbalance and inconsistency. Prompt learning effectively isolates the feature space, making model building easier. This work led to a co-author conference submission to *PRML 2022* and five patents. From this project, I deepen my understanding of text semantics. However, the challenges of fine-grained tasks relying on more knowledge, semantic ambiguity, lack of context and robustness drive my motivation to further my study in PhD program.

The **Emotional NLP** project, on the other hand, focuses on the **emotional aspect of language understanding**. Language has the ability to convey information effectively in social networks. For example, "I love you?" and "I love you!" express totally different meanings. Such tasks are often highly subjective, without clear definitions and benchmarks. First, I established the label criterion for style, emotion, and stress by reading related literature and referring to **linguistic knowledge**(definition of data is important). With a small scale of data, combining with **contrastive learning** and **semi-supervised learning**, a good style extraction model is obtained and used as a guidance for TTS. With better expressive TTS, our research has attracted the attention of customers and other companies. In addition, for stress prediction, I propose a semi-supervised progressively **granular** approach. Specifically, we use a two-stage pipeline to construct the coarse-to-fine stress model with auxiliary granularity supervision, which decreases the model deviation and ensures global diversity. This work led to a first-author conference submission to *ICASSP 2023*(reviewing) along with two patents.

In my work experience, I realized that texts contain rich semantics and emotions that are critical to help us understand our society, such as fighting against toxic online comments. I hope to contribute to more precisely understanding and reconstruction of information during my PhD study.

My background may seem diverse, but I won't consider it as a disadvantage. On the one hand, a broader sense in different fields will provide me a strong support to become a "Deep Researcher". On the other hand, pursuing a PhD degree is a lasting commitment, and I have to determine what my interests really are before I take the next move. Although I have research experiences in NLP, I still need more training and more professional theories to possess a better understanding of this field. I believe that USC is a perfect fit for me because of its strong computer science program and its distinguished faculty members.

I'm particularly interested in the following directions. **(a)natural language understanding:** The precise understanding and transmission of semantics are critical for some fine tasks, which contain the ambiguity problems, in addition to extracting contextual information, they also rely on structure analysis and knowledge side information. **(b)information extraction and representation:** Converting information into robust representations acceptable to machines is a fundamental ability. The quality of the representation will determine the maximum capability of the downstream task. I have done knowledge-based representation research which considers the construction of information networks, which is an enlightening experience for me. Introducing knowledge to NLP is still at an exciting field with blanks to be filled. **(c)low resource & robust natural language processing:** A majority of current methods rely heavily on large-scale annotated data. However, some tasks have no well-defined definitions or require domain-specific knowledge, nor do we have the time and energy

to label large-scale corpora manually. Advances in unsupervised and semi-supervised methods can help utilize unlabeled data (granular stress e.g.). At the same time, I am interested to translate research into applications to solve real-world problems, then cost and robustness will be critical considerations.

There're several professors at University of Southern California whose research interests align with mine well: **Professor Muhao Chen** (knowledge acquisition and language understanding) and **Professor Sean (Xiang) Ren** (knowledge acquisition and robust NLP and). I am also interested in **Professor Jieyu Zhao**'s work on efficiently learn knowledge and robust NLP. As I have explored relevant fields, I have the confidence to work with professors and partners to solve practical problems and make further contributions in the future. I believe with their guidance and support, as well as the lively cultural and academic atmosphere at USC, I will be given the great opportunity to work on meaningful research projects, and enjoy the journey along the way. Thank you!