Yang Xu

yangxu579@ucsb.edu · (805)637-1449 · void-b583x2-NULL yang-xu-786147328

Personal Website: https://void-b583x2-null.github.io/

Education

University of California, Santa Barbara Sept. 2024 – June 2026 (Expected)
M. S. in ***Computer Science***

Shanghai Jiao Tong University (SJTU) Sept. 2020 – June 2024
B. Eng. in ***Artificial Intelligence (Honor Class)*** **Score:** 90.8/100 **GPA:** 3.92/4.3

Internship Experiences

Applied Scientist Intern | Amazon Web Services AI Shanghai Lablet July 2023 – August 2024
NLP Team

- Co-author of **RefChecker** (Title changed to **Knowledge-Centric Hallucination Detection**, EMNLP 2024)
  - **Cultivated strong ability of applying Large Language Models (LLMs) to specific tasks** *during* development & experiment on claim extractor in the project using various kinds of LLMs to check the factuality of claims from LLM responses
  - **Managed data flow in utilizing information retrieved and stored in a database along with the construction and testing of the final module** *during* development of module for online information retrieval in released demo of this project
  - **Demonstrated strong practical ability of frontend UI design & implementation including backend integration** *during* construction of data annotation platform to annotate LLM responses
- Research on Long-context Language Modeling (internal exploration)
  - **Dived deeper and got further understanding into the inner structures and pre-training pipeline of LLMs** *during* participation in the project modeled after AutoCompressor, in which I explored more aggressive compression ratio by making multiple designs and validating their effects
  - **Practically implemented ideas and checked the results, conducted deliberated experiments and made detailed analysis & explanations** *to find that* the structure worked well on short sequence reconstruction but encountered trouble when using compressed context for prediction
- Efforts in internal development
  - **Assembled full aspects of information and made a comprehensive survey** *in* reporting the comparison between prevalent LLM frameworks
  - **Grasped usage of more kinds of LLMs** *in* prompt engineering to make precise source attribution in claim extraction
  - **Individually carried out and executed research plans** *in* exploring retrieval-based methods of hallucination detection

Selected Research Projects

Multimodal Large Language Model for Music Generation & Tasks

Collaborate with Music Department at Stanford University*, with advice from* ***Google DeepMind***

- **Implemented the whole versatile pipeline** of LLM structure modification, task prompts, training, evaluation, and inference
- **Implemented a pipeline for preprocessing huge hours of audio data**, including slicing, voice filtering, quality filtering and audio tokenization, based on the open-sourced *demucs*. *audiobox-aesthetics* and *Encodec*, to select 4,000 hours for use
- **Validated the proposed music-related tasks**, including generation, understanding, audio editing, source separation & infill tasks

MULTI: Multimodal Understanding Leaderboard with Text and Images

Multimodal Group*, X-LANCE Lab, SJTU. Advised by* ***Tenure-track Prof. Lu Chen***

- **Successfully handled big data and grasped the techniques to pre-process them into database** *in* preparation for the unique multimodal problem dataset of cognitive questions for LLMs with a total of 18k, including various types of problems with pictures from Chinese school examination questions
- **Enhanced skills of using API and local deployment of using LLMs** *during* evaluating LLMs' performance by their answers

Project for Singing Voice Conversion Challenge (SVCC) 2023

Text-to-Speech Group*, X-LANCE Lab, SJTU. Advised by* ***Prof. Kai Yu***

- **Obtained experience of audio processing and voice conversion system construction** *during* the project that adopted Source-Filter HifiGAN as vocoder with PPG features extracted by Whisper (OpenAI) for baseline
- **Adapted to new strategies of solving problems quickly and implement them into practice in time** *in* training models adapted to target speakers based on open-source repository lora-svc and upsample networks, which finally achieved a better similarity MOS score given by English listeners than all baselines

Skills

- **Programming Languages & Tools:** Python, C++, C#, SQL, Linux, Git, Shell, json, pytorch, huggingface, mongodb, LaTeX
- **Abilities:** Deep Learning, Natural Language Processing, Software development, Database, Audio processing