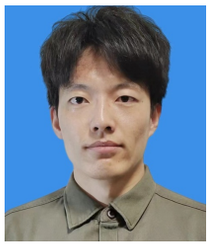
# Qiang Gao



24 years old | Male | personal web: https://cooper12121.github.io |
https://scholar.google.com/citations?user=eoUnS60AAAAJ&hl=en&authuser=1
gaoqiang.nlp@gmail.com

## Education

**Sep 2018 - Jun 2022**

**East China University of Science and Technology (ECUST)** **Energy and Power Engineering (Bachelor)**

Computer Science and Technology(second major).

**Sep 2022 - Jun 2025**

**WuHan University (WHU)** **Computer Science and Technology (Master)**

I am affiliated with the Language and Cognition Computing Laboratory at Wuhan University, under the guidance of Professor Fei Li. The laboratory focuses on natural language processing, primarily engaging in research areas such as information extraction, multimodal content recognition, and analysis of large models. My current work mainly involves research in information extraction, fine-tuning of large models, and multimodal content recognition.

## Publications

- Enhancing Cross-Document Event Coreference Resolution by Discourse Structure and Semantic Information COLING2024, First author

Existing cross-document event coreference resolution models either compute the similarity between event mentions directly or enhance mention representation by extracting event arguments, lacking the capability to utilize document-level information. We propose to enhance document-level representation by using discourse information. By constructing a document-level Rhetorical Structure Theory (RST) tree and cross-document lexical chains, we model the structural and semantic information of documents, respectively.

- Harvesting Events from Multiple Sources: Towards a Cross-Document Event Extraction Paradigm ACL2024 (Finding), First author

In this paper, we introduce a pioneering approach to cross-document event extraction (CDEE), significantly enhancing the extraction and integration of event information from multiple documents. We address the limitations of traditional document-level event extraction by utilizing a novel dataset and a multifaceted pipeline.

- MMLSCU: A Dataset for Multimodal Multi-domain Live Streaming Comment Understanding WWW2024 (Oral), Second author

Interactive audience participation in live-streaming scenarios provides constructive feedback for both streamers and platforms. Analyzing these live comments to uncover underlying intentions is crucial for enhancing the quality of broadcasts . We have introduced a multimodal dataset specific to the live streaming domain, which includes video, audio, and comment text from live sessions. We propose four tasks: audience comment intent detection, intent cause mining, audience comment explanation, and streamer strategy recommendation.

## Internship Experience

**Feb 2024 - Present** **Tencent AI Lab**

since February 2024, I have been engaged as a research intern at Tencent AI Lab, supervised by researcher Jian Li, where I have delved into large language models, including multimodal variations. My work has primarily revolved around enhancing the performance of Mixture of Experts (MoE) models on business data. I have contributed to initiatives aimed at improving the training efficiency and stability of warm-starting MoE models and have conducted detailed analyses of expert distribution strategies.

- **Project 1 - Warm-starting MoE**

Here, we explore the performance of warm-started MoE models on business data. We constructed warm-started Yi-8x6b and Yi-4x6b models based on the Yi-6b model, including both base and Instruct versions (by copying the MLP parameters as experts, and creating randomly initialized routing and load balancing losses). Our goals are:

1. For the Instruct version, fine-tune the constructed MoE model to leverage the strong performance of the original model and achieve better performance on gaming business data with minimal data,
2. For the base version, perform post-pretraining to evaluate the routing distribution strategy and training stability.
3. To determine under what conditions stable routing can be trained.

- **Project 2 - More anthropomorphic NPCs**

Our goal is to make the responses of NPCs in games more anthropomorphic. This means that NPCs should not only answer players' questions but also be able to resonate emotionally with players. This is primarily manifested in the following ways:

1. NPCs should be able to recall content shared by players. This requires NPCs to proactively mention information

previously brought up by players, such as if a player says, "I've been wanting to eat hotpot recently," the NPC's response should not merely mention hotpot, but could proactively mention, "Didn't you say you have a sensitive stomach? Eating hotpot might not be healthy for you," making the player feel like the NPC cares and listens like a human.

2. NPCs should provide emotional support to players.
3. NPCs should have the ability to refuse answers outside of their knowledge domain.

## Research Interest

My research over the past two years has primarily covered traditional natural language processing tasks, Mixture of Experts (MoE) models, and multimodal large models. Recently, I have started to focus on the field of model integration and plan to continue my research on multimodal LLM-MoE models in the future. I hope to further expand my knowledge base in the final year of my master's program, deeply exploring various aspects of LLMs to enrich my options for PhD research.

- Develop tools that facilitate the LLM community
- More interesting AI applications。
- Large Model Safety.
- End to end RAG system.
- Reasoning ability of LLM.

## Self-evaluation

- Throughout my academic journey, I have maintained a profound interest in research and have always been eager to explore new technologies and directions. This passion has driven me to venture into uncharted territories
- I approach my research with persistence and seriousness, and I possess a high enthusiasm for coding and experimentation. This attitude not only aids in my progress but also allows me to remain resolute and focused in the face of challenges.
- I have strong programming skills and have open-sourced several projects on GitHub.
- I have extensive experience with LLM code and am familiar with various common frameworks currently in use.
- I love academia and maintain an optimistic life attitude, skilled at balancing work and life.