Pengfei Zhang Applying for Summer 2024 Internship (starting Jun 2024) with F1 OPT
Email : **pengfz5@uci.edu** Mobile: **9496868833** Github LinkedIn Personal Website

EDUCATION

Year	Degree	Institute	GPA
Fall 2022 -	PHD student in Computer Science	University of California, Irvine	3.98/4.0
2017 - 2021	B.Eng in Computer Science	University of Science and Technology of China	3.45/4.3

TECHNICAL SKILLS

- Languages. Python, C++, Java, R, Web Frontend Languages, Verilog.
- Frameworks. Pytorch, Kafka, Spring Boot, Vue (Ant Design), Docker, Langchain, Flask, Django, MySQL, Redis.
- Research Fields: Generative artificial intelligence, Large Language Models, Diffusion models, Multimodal Generation, AI for Healthcare, Pose Estimation

PUBLICATIONS

- **First Author Publications**
  3. **Agent-based Hirachical Food Recommendation System Over a Large Knowledge Graph** *Submitted to CHASE 2025*
     - Introduced a food recommendation system that can satisfy constraints at different levels. Designed an LLM-based query system that takes user query, healthcare guidelines, and personal preference. Our proposed system has provided a good paradigm on querying Knowledge Graph with LLM and outperforms all previous food recommendation systems.
  2. **Handformer2T: A Lightweight Regression-based model for Interacting Hands Pose Estimation from a single RGB Image** *WACV 2024*
     - Designed a lightweight but high performance model which proposed hand-level tokenization in the transformer based model for interacting hand pose estimation, where only one token was used for each hand. This decreased the model size by 11M parameters and three times increased the FPS.
     - Designed a pose query enhancer module, which can refine the pose prediction iteratively utilizing feature sampling and Residual Log-likelihood Loss, which improved the MPJPE by 5.5mm
  1. **CLMB: deep contrastive learning for robust metagenomic binning.** *RECOMB 2022 (oral)*
     - Designed a contrastive learning framework for training Variational Autoencoder (VAE) in metagenome binning (CLMB), which reconstructed 8-22 more high-quality genomes and 15-32 more middle-quality genomes.
- **Participated Publications**
  2. **DEMENTIA-PLAN: An Agent-Based Framework for Multi-Knowledge Graph Retrieval-Augmented Generation in Dementia Care** *Submitted to AAAI 2025*
     - Introduced an innovative retrieval-augmented generation framework that enhances conversations with mild-stage dementia patients by intelligently integrating multiple knowledge graphs with Large Language Models.
  1. **Knowledge-Infused LLM-Powered Conversational Health Agent: A Case Study for Diabetes Patients.** *EMBC*
     - Developed a nutrition assessment chatbot powered by a large language model (LLM), incorporating external nutrition knowledge sources. This chatbot can analyze user dietary inputs and provide a risk assessment.

SOFTWARE AND PROGRAMMING PROJECTS

- **Distributed Chatroom with LLaMa-Powered Summarization** *Jan. 2024 - Mar. 2024*
  - Implemented a Multi-topic Web Chatroom which can provide summaries and backup on previous conversations
  - Developed the frontend using React, the backend using Spring Boot (Java) and Kafka, and the database using Redis.
  - Implemented Cache Feature on each summarization, which largely decreases latency and improves fault tolerance
- **OpenCHA - an Automatic Conversational Health Agent** *Oct. 2023 - Present*
  - Implemented a Conversational Health Agent framework leveraging LLMs-based agents (ReAct) as problem solvers, which can address health tasks like stress estimation. Built the frontend using Vue and backend using flask and Docker.

Working Experience

- Research Intern in the Chinese University of Hong Kong *May. 2022 - Aug. 2022*
  - Led the project on AI for metagenomic binning and presented the results on Conference 1

AWARDS

- Dean's award from UCI *2022-2023*
- National Encouragement Scholarship (top 20%) from USTC *2020*
- National Encouragement Scholarship (top 20%) from USTC *2018*