Hanyu Zeng

Ph.D. Student, Pittsburgh, USA
Haz207@pitt.edu — +1-412-390-4396
Personal Website — Google Scholar Profile

Education

- **University of Pittsburgh**, Pittsburgh, USA *2024/01 – Present*
  Ph.D. student in Information Science
- **National University of Singapore**, Singapore *2021/07 – 2023/02*
  Master of Intelligent Systems
- **University of Electronic Science and Technology of China**, Chengdu, China *2017/09 – 2021/06*
  Bachelor of Communication Engineering
- **University of Glasgow**, Glasgow, UK *2017/09 – 2021/06*
  Bachelor of Electronic Engineering

Research Experience

- **Self-supervised Learning Anomaly Detection** *2022/02 – 2022/07*
  Advanced Digital Science Center, Singapore
  Proposed a self-supervised learning model for **anomaly detection** trained with minimal labeled data using the BERT model. Achieved 93% accuracy, surpassing other SOTA models by over 15%. "Detecting Cyber Attacks in Smart Grids with Massive Unlabeled Sensing Data" is published on SmartGridComm, 2022.
- **Solution for Imbalanced Dataset in Anomaly Detection** *2022/12 – 2023/07*
  Chengdu, China
  Developed a novel loss function to address imbalanced dataset issues in industrial anomaly detection. Authored paper titled "Unleashing the Power of Unlabeled Data" (under revision).
- **Diabetic Insulin Management System** *2024/04 – 2024/09*
  University of Pittsburgh, USA
  Designed an insulin management system providing injection recommendations based on diet descriptions and historical glucose/injection data. Paper "DIETS" is submitted to MOBICOM 2025.
- **Video-Based Dietary Analysis System** *2024/10 - now*
  University of Pittsburgh, USA
  Designed a dietary analysis system to estimate the nutrient intake of users based on the videos recorded by smart glasses. **Yolo-world**, **llava** and **SAM** are used for multimodal interaction.

Work Experience

- **Machine Learning Engineer**, Meituan, Beijing, China *2023/07 – 2023/12*
  - Developed a multimodal AI assistant system using Java and Python for hotel recommendations, customer Q&A, and chat with **Large Language Model; Llama2** and **Qwen V1.5**.
  - Designed and trained a user intent classification system based on **BERT_Large**, achieving **90%** accuracy.
  - Optimized GIS alignment via NER model, reducing redundancy by **15%** on SQL.

Technical Skills

- **Programming Languages:** Python, Java, C, C++, MATLAB, SQL, R, HTML, JavaScript
- **Frameworks and Libraries:** TensorFlow, PyTorch, Scikit-learn, Pandas, NumPy, FastAPI, HuggingFace
- **Cloud and Platforms:** AWS (EC2, S3, Lambda), Google Cloud, Spark, Hadoop
- **Databases:** MySQL, PostgreSQL, MongoDB, Redis
- **Other Skills:** Git, Docker, Linux/Unix, Tableau, Object Detection/Tracking, Anomaly Detection, CPS, IoT system, Mobile Network, NLP, CV, Signal Processing, Machine Learning