

Zhili Yang

12112711@mail.sustech.edu.cn | +86 13979300723
SUSTech, 1088 Xueyuan Ave., Shenzhen, Guangdong, China, 518055

RESEARCH INTERESTS

- Data Selection and LLM4Science
- LLM Efficient Instruction Fine-tuning
- Embodied AI and Human-Computer Interaction

EDUCATION

Southern University of Science and Technology

Shen Zhen, China

GPA: 3.56/4; Weighted Average Score: 85.92/100

Bachelor of Science in Data Science; Department of Statistics and Data Science

Jan 2023 - Present

- Relevant Coursework: Advanced Natural Language Processing, Deep Learning, Statistical Learning, Data Structure and Algorithm Analysis, Distributed Storage and Parallel Computing, Operational Research and Optimazation, Mathematical Analysis, Probability Theroy, Mathematical Statistics, etc.

Marjor in Industry Design; Department of System design and intelligent manufacturing

Sep 2021 - Dec 2022

- Learned some basic courses of industry, such as fundamentals of materials, design science, rapid technology prototyping, analog circuits. Then switched to my current major: Data science.

University of California, Berkeley

Berkeley, CA

Berkeley International Student Program

Jan 2024 - May 2024

- Relevant Coursework: Introduction to Time Series(STAT 153), Modern Statistical Prediction and Machine Learning(STAT 154/254), Introduction to Artificial Intelligence(CS 188), Audited: Deep Unsupervised Learning(CS 294-158)

RESEARCH EXPERIENCE

Data filtering during insructional fine tuning of LLM

Peking University, SUSTech

Under the guidance of Assistant Professor Guanhua Chen, Research Assistant He Zhu

Sep 2024 - Present

- Reproduced some classical data filtering work under the guidance of RA.
- To be continue.

National University of Singapore, School of Computing Summer Workshop

Singapore, SG

Student in Deep Learning cluster

May 2024 - Jul 2024

- The final grade of the workshop is A(Excellent).
- Trained a specific object detection algorithm based on YOLO and used open source speech2text to design a language module, which can enable a single-chip computer-powered car to capture specific objects and follow basic commands.

3D panoramic simulation for metaverse

SUSTech

Project leader (Guangdong Climbing Program), under the guidance of Professor Zhaoyuan Ma

Oct 2022 - Dec 2022

- A final patent application in progress.
- Led project scheduling and contributed to the fish camera splicing algorithm for a 3D panorama acquisition and transmission product in the metaverse. Designed and modeled the product's appearance as well.

PROJECT EXPERIENCE

Github Homepage: <https://github.com/radicalyyyahaha>

Independent Research Project

- **Poem_Generator:** An attempt to recreate nanoGPT from scratch, then use for poem generate after a huge modification.
- **Protein morphology prediction:** Did some research on how to use DL models for protein sequence and structure prediction.
- **Implementation of classic data selection works:** IFD, Superfiltering, LESS, etc.

Course Project

- **Chess AI:** Built a chess game for a freshman Java course, where a AI bot made random moves based on the player's actions.
- **Integrated Design:** Designed a face recognition system for smart trash can, marking my first experience with AI algorithms.
- **Heuristic Search:** Expand ARA* algorithm for dynamic path planning in a maze, introducing me to reinforcement learning.
- **Sound Data Classification:** Designed a CNN model with residual networks for sound classification using ESC-50 data in a statistical learning project.
- **Improved YOLO v8:** Enhanced tiny object detection by adding an attention module to YOLO v8n, achieving a mAP50-95 of 0.35 in an AI course project.
- **Berkeley CS-188:** <https://github.com/radicalyyyahaha/Berkeley-CS188>

SKILLS

- | | | |
|-------------------------------------|-------------------------------------|---------------------------|
| ■ IELTS 7.0 | ■ Highly self-motivated | ■ Curiosity and Execution |
| ■ Familiar with Github, Huggingface | ■ Leadership and cooperation skills | ■ Good state of mind |