TIANYU LI

Undergraduate student at Dept. Statistics & Data Science, Southern University of Science and Technology

@ jacktianyuli@gmail.com; 12212824@mail.sustech.edu.cn   📞 (+86) 18280416745
📍 Shenzhen, China   ⍾ https://github.com/JackTianyuLi

EDUCATION

B.Sc. in Data Science and Big Data Technology

Current CGPA: 3.87/4.0   rank: 5/50

Core Courses: Probability Theory, Mathematical Statistics, Data Structures and Algorithm Analysis, Statistical Linear Models, Distributed Storage and Parallel Computing, Artificial Intelligence, Statistical Learning, Big Data Analytics Software and Applications

Dept. Statistics & Data Science, Southern University of Science and Technology

📅 August 2022 – June 2026 (expected graduation)   📍 Shenzhen, China

EXPERIENCE

Tencent Music Entertainment Group

Project intern, Smart Platform Group

📅 June 2025 – July 2025   📍 Shenzhen, China

- **Optimized an end-to-end text classification model**, managing the full lifecycle from requirement analysis and large-scale imbalanced data handling to deep fine-tuning of pre-trained models (e.g., StructBERT). Achieved a validation F1-score consistently above 82%, demonstrating proficiency in model performance enhancement.
- Gained practical **Model Engineering & Deployment experience** by configuring the production environment using Linux, Docker, Flask, and Gunicorn, establishing a robust foundation for production launch.

Undergraduate research

Supervisor: Dr. Chao Wang

📅 September 2024 – June 2025   📍 Shenzhen, China

- **Developed a novel Gaussian splatting-based compression model for hyperspectral images**, achieving a significant 4-6 dB PSNR improvement over implicit neural representation methods at equivalent compression rates. This also resulted in reduced encoding time and GPU memory usage.
- Awarded a **15,000 CNY research grant** from the Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation (2025 "Climbing Program" Special Funds), recognizing innovative research potential.
- Authored and submitted a paper manuscript, **LoR-SGS: Hyperspectral Image Compression via Low-rank Spectral Gaussian Splatting**, to IEEE Transactions on Geoscience and Remote Sensing.

Participated in Kaggle competition: Google - Unlock Global Communication with Gemma

📅 December 2024 – January 2025

- Fine-tuned Gemma-2 using the COIG-CQIA collection to enhance its reasoning capabilities in Chinese tasks, showcasing practical experience with large language models (LLMs) and fine-tuning techniques.
- Details available at: `https://www.kaggle.com/code/jacktianyuli/finetuning-gemma2-for-chinese-reasoning`.

ACADEMIC PROJECTS

Course Project for STA321: Big data computing utilizing MapReduce

📅 November 2024 – December 2024

- An HDFS-MapReduce pipeline was designed to compute the main net inflow from the order and trade records of the Shenzhen Stock Exchange, which succeeded in processing the example data with the size of 4.7 GB in 50 seconds.
- See `https://github.com/JackTianyuLi/STA321-project` for details.

Participated in Mathematical Contest in Modeling, 2024

📅 Feburary 2024

- Duty part: Process data from a tennis match to calculate 19 technical indicators across different sections (serving, receiving, scoring), followed by PCA to extract key indicators for flow visualization, and the momentum of the match was visualized using a weighted TOPSIS algorithm.

AWARDS

- **Academic awards**:
  - **Sep. 2024   Second Prize of the National Mathematical Modeling Contest, Guangdong Division**
  - **Feb. 2024   Honorable Mention of the Mathematical Contest in Modeling, 2024**
  - **Dec. 2023   Second Prize of the National Mathematics Contest, Guangdong Division**
- **Other:**
  - Jul. 2024   Best-16 in men's single in the 27th Chinese University Tennis Championships
  - Jan. 2024   "Outstanding student" of SUSTech in year 2023
  - Dec. 2023   4-th place in men's volleyball in the 13th University Volleyball Tournament of Guangdong Province, 2023 representing SUSTech

LANGUAGES

English
Proficiency tests taken:
- CET-6: **615** (taken December 2023)
- GRE: **326** (taken March 8 2025, **V156 Q170 AW3.5**)
- TOEFL: **108** (taken June 28 2025, **R30 L29 S23 W26**)

Mandarin Chinese

SKILLS

Programming: Python (Numpy, Pandas, Matplotlib), Java, R, SQL
Software/API: PyTorch, LaTeX, MS Office, GIT