

Ray (Lisirui) Tang

Cell: +86 17385596205

Email: raytang50@gmail.com

EDUCATION

University of Liverpool

2020.09-2024.06

Bachelor of Engineering in Data Science and Big Data Technology (School of AI and Advanced Computing)

Xi'an Jiaotong-Liverpool University

2020.09-2024.06

Bachelor of Engineering in Data Science and Big Data Technology (School of AI and Advanced Computing)

Core Modules: Calculus, Introduction to Probability and Statistics, Database Development and Design, High Performance Computing, Cloud Computing, Machine Learning(Top 10% of a class of 200)

WORK EXPERIENCE

Microsoft Search Technology Center Asia (STCA)

Beijing, China

DKI Group R&D Team Member

2023.9-2023.12

- Under my mentor's guidance, a Senior Principal Researcher at Microsoft, I have actively contributed to pioneering the application of Large Language Models (LLMs) such as ChatGPT across diverse domains. Notably, we harnessed ChatGPT to produce over 200,000 PowerPoint outlines for OfficePlus in more than 20 different fields.
- We researched the application of large language models in the field of climate change. Through techniques like fine-tuning, we trained AI for Science to distinguish it from generic large language models.

TIKTOK's parent company : ByteDance

Beijing, China

Prompt Engineering Team

2023.6-2023.8

- Engaged in the highly confidential AIGC project at ByteDance, responsible for analyzing over 10,000 training data entries daily for LLM to ensure data quality.
- Utilized Python and SQL for daily data scanning, homogeneity detection, and word count analysis, which significantly improved company efficiency by approximately 50%.
- Contributed to the production and rule formulation of LLM data, including prompt data generation, personalization rule development for SFT data, and product workflow.

Xi'an Jiaotong Liverpool University

Suzhou, China

AI/LLM RA(R&D)

2024.2-2024.6

- Engineered over 1000 prompts and responses to fine-tune 70 academic papers via Prompt engineering, employing statistical and quantitative analysis. Utilized varying levels of fine-tuning data to analyze and visualize experimental Llama and ChatGLM models, using BLEU-4, ROUGE-1, ROUGE-2, and ROUGE-L. Higher levels of data led to improved evaluation scores, showing a clear correlation through rigorous statistical analysis. Llama achieved a maximum ROUGE-1 score of 41.28%, indicating significant statistical performance.

Neusoft

Guiyang, China

Data Analysis Intern

2022.6 - 2022.8

- Participated in the completion of the construction program for upgrading the population management system in Guizhou Province, and modified the problems of incorrect logic and repeated variables in the population system
- Helped the public security system to build and query multiple tables for the required population information

Suzhou Xian Zhiyi E-commerce Co.

Suzhou, China

Co-founder, Executive Director and Shareholder

2020.10-present

- Evitalizing the electric motorcycle industry through secondhand e-commerce to boost sustainability and cut carbon emissions. In the first month of inception, we achieved profitability, and in the second quarter, I led a 20-member team to expand our services to major universities in Suzhou.

RESEARCH EXPERIENCE & PROJECTS

MIT Blockchain and AI in Financial Data Science

- Utilized AI to gain insights into the financial markets, using python, matlab to build metrics and models for market forecasting based on "off-chain data" primarily from Galaxy Digital and other exchanges, and "on-chain data" from blockchain and network analytics.
- Responsible for analyzing exchange rates and Bitcoin price data for major countries from 2015 to 2022, investigating the correlation between them. Utilized LSTM models for time series analysis and successfully predicted Bitcoin price trends.
- The model underwent 200 epochs of training, showcasing a remarkable reduction in loss from 0.1050 to 0.0081 for training data and from 0.1455 to 0.0136 for validation data, demonstrating significant improvement and convergence over time.

Baidu Paddle Paddle

- Participated in practical training, mastering image segmentation through Paddle Paddle, and AI projects such as vehicle and pedestrian traffic light detection and mobile robot detection in smart transportation.
- Harnessed deep learning techniques to curate an extensive dataset comprising over 10 million weather images. Through rigorous supervised learning, the machine learned to discern various weather conditions, resulting in an impressive performance where the model achieved an accuracy exceeding 0.9 in eight out of ten test runs.

Final Year Project: Prediction of Greenhouse Gas Emissions

- Determining the research methodology of the project by reading relevant articles, collected relevant AI and past greenhouse gas emission data.
- Conducted a study to predict CO2 emissions resulting from China's energy consumption between 1965 and 2022. Compared two models: Support Vector Regression (SVR) with different kernel functions and LSTM. The LSTM model showed high accuracy and robustness, with an R^2 value reaching up to 0.9955.

PUBLICATIONS, CONFERENCE TALKS, AND AWARDS

- **Tang, L.**, Zhao, P., & Abdul Majeed, A. P. P. (2023). "Predicting Carbon Dioxide Emissions from Energy Consumption in China with Long Short-Term Memory and Support Vector Regression." The 11th International Conference on Robot Intelligence Technology and Applications (ICRITA 2023).
- **Tang, L.**, Wang, C. Y., & Li, G. G. (2024). "A Large Language Model for HeXie Management Theory." has been submitted to 2024 6th International Conference on Pattern Recognition and Intelligent Systems (PRIS 2024).
- The 28th United Nations Climate Conference (COP28): Youth representative Speaker /AI in climate change
- The Third China Greenhouse Gas Detection Conference, hosted by Fudan University, the research on greenhouse gases using AI.
- The third prize of "China University Student Energy Finance Innovation Challenge" by the UK association of applied mathematics
- Suzhou Huisi Cup Entrepreneurship Competition Best Style Award
- Suzhou Dandelion Entrepreneurship Program

ADDITIONAL SKILLS

LEADERSHIP: Monitor or the Youth League branch secretary(Elementary school-High school)

Position in the Student Union(College)

IT SKILLS: Python, R, Matlab, MS Office, Logic pro, CUDA, C/C++

LANGUAGES: English (Fluent), Chinese (Native)

INTERESTS AND ACHIEVEMENTS: Basketball (Varsity member), Music Producer (released 7 albums, 15 songs, listened to over 6,000 plays), Surfing