

YUE WU

wuy2021@mail.sustech.edu.cn · Personal Website

EDUCATION

Southern University of Science and Technology, SUSTech, Shenzhen 07/2023-05/2024

Bachelor of Statistics, GPA: 3.80/4.00, Rank: 4th

- Main Courses: Statistical Learning(98), Time Series Analysis(92), Operations Research and Optimization(90), Probability Theory and Mathematical Statistics(97), Statistical Computation and Software(91)

RESEARCH EXPERIENCE

Personalized Knowledge Graph Summarization with LLM Research Assistant

Supervisor: Prof. Yujun Yan, Dartmouth 07/2024-present

- Goal: Designed a personalized knowledge graph summarization algorithm using LLM based on historical data.
- Contribution: Conducted research on Knowledge Graph Reasoning algorithms and Question-Answering databases. Utilized all-MiniLM-L6-v2 to construct a text-based question similarity graph and historical data. Reproduced GreaseLM and designed baseline experiments based on this model.

Employee Performance Evaluation Using NLP and LLM Research Assistant

Supervisor: Prof. Wenqi Fan, PolyU 07/2024-present

- Data Augmentation: Enhanced employee interview data using traditional NLP (random swapping, deletion, insertion, etc.) and a custom evolution prompt-based LLM method, achieving 1000x data augmentation and iteratively screening by complexity and diversity.
- Quantitative Scoring: Developed a few-shot and pair-comparing scoring pipeline with LLMs to analyze and quantify employee interview texts for performance and promotion potential evaluation.

Application of Deep Learning in Multi-Agent Financial Simulation Research Assistant

Supervisor: Prof. Peng Yang, SUSTech 07/2023-05/2024

- Model Building: Developed a transformer-based autoencoder with a novel calibration method, improving multi-agent LOB financial simulator's performance by over 50%, and conducted comprehensive experiments with visualizations and ablation studies for model analysis.
- Co-authored a paper and submitted it to NeurIPS. Preprint Link Here.

INTERNSHIP EXPERIENCE

LLM Training Data Quality Enhancement and Prompt Engineer Intern Team Leader

Supervisor: Yipeng Cheng, xDAN AI 03/2024-06/2024

- Data Augmentation: Expanded over 10 renowned datasets through data augmentation and diversification, guided by WizardLM and DEITA and other papers, utilizing advanced tools such as Distilable and DataDreamer. Quantitatively evaluated LLM data based on complexity and quality using UltraFeedBack, embedding similarity, and perplexity.
- Prompt Engineer: Developed diverse prompts using CoT, Few-shot In-Context, and other methods to enhance question variety and complete high-quality specific tasks with several LLM models.

HONORS & AWARD

- | | |
|--|------|
| • First Class (First Place) of the Merit Student Scholarship(1%, SUSTech) | 2023 |
| • iGEM International Competition Gold Medal(10%, Global) | 2023 |
| • First Prize in the China Undergraduate Mathematical Contest in Modelling(10%, Guangdong) | 2023 |
| • Student Cadre Model of Excellence(1%, SUSTech) | 2023 |
| • Second Prize in The Chinese Mathematics Competitions(15%, Guangdong) | 2021 |

SKILLS & LANGUAGES

Computer Skills: Python, Java, R, SAS, Hadoop, Spark, Pytorch, Linux, Latex

Language: Chinese(Native), English(TOEFL 103).