



JIAN GAO

+86 18810800385 | [Github jackgao-thu](#) | jack.gao.working@gmail.com

EDUCATION

Course	University	Year	GPA
All Courses	Tsinghua University	2024	3.73/4.00
Data structures and algorithms etc.	Tsinghua University	2024	4.00/4.00 for 21 courses

Research Experiences

- **Research on space gamma burst time delay positioning method based on nanosatellite detection data | GRID(Gamma Ray Integrated Detectors) | Tsinghua University** November 2022- July 2023
 - As the lead for triangulation methods in gamma-ray burst(GRB) positioning at GRID, I utilized light curves to significantly reduce the uncertainty in angular measurements for GRB locations.
 - I improved the positioning calculation algorithm, substantially reducing the number of attempts needed to calculate time bins, which resulted in an average speed increase of 80%.
- **Basic research on new electrolysis hydrogen production | Institute of Thermal Energy | Tsinghua University** March 2023 - February 2024
 - Investigated the mechanisms and performance of cationic mass transfer electrolytes. Completed microelectrode electrolysis experiments at various concentrations, pH values, and gradient temperatures for solutions including piperazine, imidazole, and potassium dihydrogen phosphate.
 - Compared the performance and mechanisms between microelectrode and planar electrode electrolysis in the same solutions, concluding the mass transfer limitation mechanisms during piperazine solution electrolysis.
- **NLP research on attention in transformers | Department of Electronic Engineering | Tsinghua University** July 2023 - Today
 - I studied the attention characteristics and common features of models such as llama2-7B, llama2-70B, and gpt2 when trained on long sentences with six pieces of information. The findings indicate a strong positive correlation between the size of the model parameters, the attention weight given to key information such as names, and the accuracy of the model in answering questions.
 - By manually controlling the attention allocation within the llama2-7B model to enhance the attention on key information like names or to reduce the attention on irrelevant information like punctuation, I successfully demonstrated that adjusting attention can significantly improve the training outcomes for information positioned later in the six pieces of information, with the best results mirroring the effects of training on short sentences with single pieces of information. Now we are working on attention based self-reflection training, aimed at making LLMs work well on knowledge memory functions.

Skills

- Program languages: C/C++, Python, MATLAB
- Tools and Frameworks: Git, LaTeX, PyTorch, Markdown, Linux servers, [Huggingface](#)
- Techniques: Pretrain and Finetune in NLP, Attention visualization and Attention mask aditing

Awards and Social Work

- National Level: 14th National College Student Mathematics Competition, Non-mathematics Category (Second Prize)
- Leader of the volunteer group of the Youth League Committee of Weiyang Academy: received two 5-star volunteer's projects a year, which was the largest number among all departments in Tsinghua University
- English scores: CET-4: 628; CET-6: 604; [TOEFL: 101](#)