Kean Shi

shikean@tju.edu.cn | +86 156 2095 6849 | Shi's blog | Shi's github

Education

Peking University, M.Eng in Computer Science (provisional admission) Sept 2025 – Jun 2028

Tianjin University, B.E in Computer Science Sept 2021 – Jun 2025

- Advisor: Prof. Wenjun Wang
- GPA - 3.84/4.0; Weighted Score - 91.8/100; Ranking: 4/41
- Other Abilities: CET-4 - 550; CET-6 - 500; CSP - 380

Interests

Natural Language Processing (NLP), Multimodal Large Language Model, Large Vision-Language Model

Projects

Efficient Training for MLLM Oct 2024 - Present

Institute of Computational Linguistics, Peking University

- Training MLLMs like LLaVA-1.5 is computationally expensive (>24h/8 A100s), limiting efficiency.
- As the project lead, I was tasked with improving training efficiency by over 30% while maintaining model performance and reducing hardware costs.
- I proposed a multi-stage training approach combining token pruning and parameter compression. Using LLaVA-1.5 as the baseline, I evaluated the algorithm's performance across multiple benchmarks.
- The method achieved a reduction in training time to 70% of the original (tested on 4×A40 GPUs) without sacrificing performance.

Intelligent preoperative detection system for meningioma Apr 2024 - May 2025

Visual Intelligence Lab, Tianjin University

- Current MRI meningioma research faces challenges like insufficient rare-grade samples and underutilized multimodal data.
- As the algorithm developer, I enhanced the original segmentation model by adding a classification branch and leveraging multimodal fusion to boost performance. I also designed tailored training strategies to address class imbalance.
- For class imbalance: redesigned dataloader to dynamically oversample rare classes. For multimodal fusion: developed a dedicated module to integrate heterogeneous data.
- Built a NestedFormer-based segmentation+classification multitask network enabling efficient MRI meningioma detection (trainable/deployable on a single 3090 GPU).

Industry Knowledge Graph June 2022 - Dec 2022

Smart City and Big Data Intelligent Laboratory, Tianjin University

- Current industry knowledge graphs suffer from limited data volume and incomplete structures.
- As project lead, I orchestrated the end-to-end solution to deliver a production-ready KG system.
- Established phased milestones, coordinated cross-functional teams via weekly syncs, designed schema for entity resolution, and implemented NLP pipelines for text cleansing/alignment.
- Delivered a MediaWiki-based KG with millions of entries, now integrated into our lab's "World Insight Platform".

Algorithm Competition Dec 2021 - Oct 2023

ACM school team member, Tianjin Unversity

- Previously participated in 2022 ICPC Xi'an Site, 2023 ICPC Xi'an Site, 2023 CCPC Qinhuangdao Site, etc.
- 2022, 2023 Tianti Competition Team Tianjin First Prize, and 2024 Blue Bridge Cup Tianjin First Prize.

Awards & Services

- National Scholarship of Undergraduate Student, Merit Student of Tianjin University for two years
- Student reviewer of AAAI