📞 +86 18021009207   ✉ sulkstation8@outlook.com

📍 Shenzhen, China   ✉ 12112906@mail.sustech.edu.cn

LIKAI SU

EDUCATION

09/2021 - Present
Graduation: 06/2025

Southern University of Science and Technology (SUSTech), China   **Undergraduate**
- Major: Big Data Management and Application
- Courses: Data Mining, Deep Learning, Data Statistics and Econometrics, Database, Distributed Storage and Parallel Computing, Data Structure and Algorithm, etc.
- GPA: 3.51/4.0

RESEARCH EXPERIENCE

06/2024 - Present

Diagnosis of HIV-Associated ANI based on sMRI and fMRI (at Nanjing University)   **Research Intern**
- Utilizing structural-MRI and functional-MRI to identify Asymptomatic Neurocognitive Impairment (ANI) and help doctors gain understandings of potential biomarkers to improve diagnosis in clinical practice.
- Conducting experiments on several multimodal fusion method to integrate sMRI and fMRI features, including cross attention and multiway transformer.
- Testing several attribution methods on deep neural networks to explain the diagnosis result, including integrated gradient and its variants and iterative search attribution, etc.
`Multimodal Fusion` / `Model Interpretability` / `Asymptomatic Neurocognitive Impairment`

09/2024 - Present

Alibaba Tianchi Imaging Algorithm Competition of Thyroid B-mode Ultrasound   **Individual Participant**
- Develop algorithms based on the thyroid grayscale B-mode ultrasound image data to realize the automatic segmentation and classification (salignant and malignant) of thyroid nodules in static and dynamic images.
- Organizer: Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai Artificial Intelligence Laboratory, Alibaba Cloud Tianchi Platform
- *Currently ranked 4th out of 306 teams.*
`Medical Image Segmentation` / `Nodules Classification`

PROJECT EXPERIENCE

10/2023-11/2023

DEFCON31 @ AI Village Capture the Flag   **Individual**
- Collected flags by evading, deceiving, or exploiting vulnerabilities in AI/ML models.
- Identified weaknesses in deep learning models by using methods such as prompt injection, adversarial attack, machine learning, and genetic algorithms.
- *Ranked in the top 15%* of the competition with 18 flags collected.
`Machine Learning` / `Prompt Injection` / `FGSM` / `AI Security`

01/2024-04/2024

Passenger Flow Prediction Model for Tianjin Urban Traffic System   **Team Member**
- Used multiple LightGBM models to predict future passenger flow based on historical passenger flow and weather forecast SMS provided by the Tianjin Municipal Government.
- Introduced event data and detailed meteorological data, leading to better robustness to fluctuations caused by temporary events and use Prophet model (time-series model) to construct several features.
- *Our team ranked 6th out of 1060 teams*, with MAPE 3.04% on the test data.
`Machine Learning` / `Ensemble Methods` / `Prophet` / `BERT` / `Web Crawler`

02/2024

Tennis Match Prediction Model Based on Momentum   **Team Member**
- Incorporated "match momentum" into the prediction of next round result, using the number of consecutive wins after 10 rounds as a concrete proxy.
- Utilized LSTM to predict the outcome of the next round based on the sequence of historical 10 matches using sliding windows, by comparing the predicted consecutive wins.
- The model reached MSE 0.402 on consecutive wins and accuracy 84% on next round result.
`LSTM` / `Sliding Window`

05/2024

Knowledge Q&A System Based on LLaMA3-8B + RAG   **Individual**
- Based on the LLaMA3-8B model, used FastGPT as the front-end and RAG framework, and deployed the service in Docker form.
- Using specific datasets, fine-tuned the model with LoRA and reward modeling.
- Used the vLLM inference engine for the LLaMA model, achieving concurrent handling of 100 user requests on a single 3090 GPU without significantly affecting response speed.
`Fine-tune` / `LLaMA` / `RAG` / `Docker` / `Reward Modeling`

INTERNSHIPS

06/2023-09/2023 **Tencent Technology (Shenzhen) Co., Ltd.**   Data Analyst Intern
- Designed a quality inspection algorithm used to screen the structured content parsed by OCR. The algorithm automatically filtered, corrected, and supplemented each piece of data using consistency checks and text similarity methods.
- *The workload of manual inspection was reduced by about 32.22%, the filling rate of missing data was 48.7% with accuracy 81%.*
- Built a data dashboard for this business using BI tools together with colleagues.
  `Data Quality Inspection Algorithm  /  BI`

09/2023-06/2024 **Shuge Technology (Shenzhen) Co., Ltd.**   Data Product Intern
- Provided data products or consulting services to clients based on massive real-time recruitment data.
- **Highlight Works:**
  1. **College Students' Resume Analysis and Career Insight System (Team Leader)**: Utilized NLP and LLMs to analyze user resumes, and developed a full-stack system that recommends job types based on skill-job requirement matching and provides data-driven improvement suggestions. The system also provides users with big data insights into the current job market and recruitment analysis of interested companies.
  2. **Big Data Insights into the Recruitment Market (Member)**:Built a dashboard that leveraged multi-dimensional data analytics and knowledge graphs to provide decision-making support. The project is showcased at the 25th China International Hi-Tech Fair.
  `Data Visualization  /  NLP  /  Power BI  /  Full-Stack Development`

CERTIFICATIONS

11/2023 **Tencent Cloud Big Data Developer (TCP Senior Engineer Certification)**   Tencent Cloud

ACHIEVEMENTS

09/2021 **Freshman Scholarship**
by SUSTech

08/2023 **Third Prize in "Huashu Cup" College Students' Mathematical Modeling Contest**
by Big Data and Mathematical Modeling Professional Committee of China Future Research Association

10/2023 **Outstanding Student Scholarship**
by SUSTech

10/2023 **Third Prize in China Undergraduate Mathematical Contest in Modeling**
by China Society for Industrial and Applied Mathematics

ENGLISH PROFICIENCY

TOEFL - 103, **IELTS** - 7.0, **CET6** - 638

SKILLS

Languages: Python, Java.
Technical Skills: Pytorch, Scikit-learn, Docker, Hadoop, Web Scraping, Pandas, Linux.

OTHER INFORMATION

- Volunteer service duration of 90 hours+.
- Capable of quickly learning and applying new technologies, most of the skills currently mastered were self-taught outside of coursework. Proficient at project-oriented learning of new knowledge and skills, and adept at turning ideas into reality.
- Honest and reliable, meticulous in work, with good moral quality and excellent communication and coordination skills.