Linan Wu

Mobile: +86 181 3805 5392 | Email: linan_huilaioffer@163.com

EDUCATION

Beijing Normal – Hong Kong Baptist University (BNBU)

*Bachelor of Science (Honors) in **Data Science** (GPA 3.7/4.0)* *09/2022-06/2026*

*Minor in **Economics** (GPA 3.8/4.0)*

- Guangdong Provincial Third Prize in 2024 China Undergraduate Mathematical Contest in Modeling
- BNBU Second-tier Scholarship (2022-23 & 2023-24)

COMPUTER TECHNIQUES

- Programming languages: Python, R, C, Java
- Deep learning frameworks: PyTorch, TensorFlow, Hugging Face Transforers, JAX
- Machine learning & NLP: Transformer (GPT, BERT, RoBERTa, LLaMA), CNN, RNN, LSTM
- LLM applications and methods: LangChain, Prompt Engineering
- Data processing & visualization: Pandas, NumPy, SQL, Matplotlib, Seaborn, Plotly
- Big Data and Distributed Computing: Hadoop, Spark, Docker, Kubernetes

INTERNSHIP

SFMAP Technology (Shenzhen) Co., Ltd. (subsidiary to Shunfeng Holding Co., Ltd)

Algorithm Engineer Intern *07-08/2025*

Participated in the development of a Flask-based platform for driver safety analytics, mainly modules such as automated anomaly detection, AI-driven insight generation, and Text-to-SQL, to achieve end-to-end intelligence in data analytics

- Designed and implemented five APIs (overview, analysis list, operational analysis, AI-driven insights, and automatic insights), achieved the interfacing of the data service layer and the unifying of JSON serialization, to support frontend consumption
- Develop an automated insight engine: fulfilled the AnomalyInsight data structure, integrating Z-Score (threshold 3.0/2.5) + IQR detection methods and 4 business rules (safety score decreased by 10, accident rate month-over-month increased by 50%+, regional disparity decreased by 30%, and proportion of high-risk driver decreased by 20%); output the top 30 abnormal insights sorted by confidence level and deviation, achieving a coverage rate of 90%+, and identify high-risk drivers 1–2 weeks in advance
- Built a Text-to-SQL Self-Service Query Function: combined SentenceTransformer + Chroma Vector Retrieval with Qwen2.5-72B LLM to implement the conversion of natural language to SQL (limited to SELECT, first 20 rows), ensuring data security, and covering 95%+ common query scenarios
- Introduced accessing LLMs via a Bastion Host: integrated LLM services in an offline environment to automatically generate Chinese security reports and management summaries, supporting insight archiving and cross-month queries through a Flask interface to enhance managerial decision-making efficiency
- Optimized database performance: configured SQLAlchemy connection pool parameters (pool_pre_ping, pool_recycle=3600, pool_size=10), and, via combination with the logging module, monitored SQL execution and automated insight processes, significantly improving system stability and concurrency capability

Guangzhou Yunke Information Technology Co., Ltd.

Intern of Backend Development *06-08/2024*

1. *Participated in the development and optimization of APIs for the logistics management module of the company's SaaS ERP system*

- Assisted in developing core APIs such as driver QR code check-in and vehicle transport statistics; conducted self-testing via Postman and collaborated with QA teams to carry out functional validation
- Learned and utilized the Spring Boot framework to implement API logic, and, according to the team coding standards, optimized data processing logic via Java Collections Utility Classes; assisted in developing fundamental APIs for logistics tracking by studying using thread pools to handle concurrent data writes, enabling scheduled reception and storage of vehicle GPS data
- Assisted in troubleshooting the problem of 'slow generation of weighing data report' by identifying unindexed join queries through analysis of slow query logs; shortened the report generation time from 15 minutes to 3 minutes

2. *Participated in developing the data interaction layer of the company's LED vehicle queuing system*

- Learned the use of Redis string and hash structures, and assisted in building a Redis cache basic model for storing temporary data of vehicle queuing status
- Wrote scheduled task scripts to realize the automatic synchronization of queuing information to the Redis cache every 5 minutes, reducing query pressure of database
- Studied the use of distributed locks to solve the problem of dirty data caused by multiple threads modifying queuing numbers at the same time, to deal with the high-concurrency scenario of real-time update of vehicle queuing information

PROJECTS

Image Matching Challenge 2025 (Kaggle) *04-06/2025*

Aimed to construct a complete system of image matching and 3D reconstruction inference that can predict camera's extrinsic parameters and scene affiliation in different scenarios

- Utilized FAISS cosine similarity retrieval, based on DINOv2 and CLIP global features, to generate candidate image pairs, ensuring coverage of matched pairs under cross-view and cross-illumination conditions
- Employed ALIKED to extract local key points and descriptors, combined with LightGlue for efficient matching; implemented a fallback to LoFTR when matched points are insufficient, boosting recall by over 15% in scenes with weak textures or repetitive structures
- Adopted PyCOLMAP for incremental mapping to estimate camera rotation matrices and translation vectors, while automatically performing scene clustering to ensure mapping purity and robustness
- Designed a mechanism for parameter tuning and caching (including SIM threshold, number of feature points, and fallback strategy) to significantly reduce inference time while maintaining accuracy; persist intermediate results in HDF5 supported resumable execution
- Implemented an automated submission module that writes the cluster, R, and t into a submission.csv file compliant with competition specifications; the coverage and localization accuracy on the local validation set remained stable and met the requirements for submission

BYU - Locating Bacterial Flagellar Motors 2025 (Kaggle) *04-06/2025*

Aimed to precisely localize bacterial flagellar motors in three-dimensional electron tomographic images reconstructed from a series of two-dimensional projection images, which are characterized by significant noise and substantial variations

- Constructed a two-stage inference pipeline: the first stage utilizes YOLOv8/YOLO11 combined with SAHI sliced inference and multi-resolution TTA to generate candidate points; the second stage employs rotation + zoomed cropping, Midpoint Reasoning (merging closely located detection points), and Bypass Logic (skipping the second stage for high-confidence detections) for refined screening and acceleration
- Trained YOLOv8l/YOLO11l by integrating official and externally corrected datasets; additionally trained a localized model utilizing random cropping to enhance robustness in detecting small-scale regions
- Performed model ensemble learning and constructed multiple submission configurations to explore various combinations of SAHI activation, different scaling factors (z1.5/z2/z3), and parallel pipelines, ensuring the model's generalization capability

Construction of Large Language Model Inference System *02-04/2025*

Aimed to build a high-accuracy large language model inference system to answer complex natural science multiple-choice questions generated by GPT-3.5

- Built a multi-model architecture based on open-source models such as Mistral-7B, Yi-34B, and LLaMA 2, using Zero-shot, Few-shot, and SFT fine-tuning strategies, combined with H2O LLM Studio, to optimize the model, achieving an accuracy improvement of over 10% in specific fields such as physics/biology
- Used LoRA fine-tuning and cached past_key_values, increasing inference speed by 60%
- Dynamically enhanced context by integrating RAG and LangChain, building a corpus containing 60million paragraphs based on multi-source Wikipedia data, developing a FAISS vector retrieval system, and entering pre-designed multi-round Prompt templates
- Implemented confidence integration based on results of multiple models, combined with methods such as TF-IDF re-ranking and Embedding similarity, to increase the Top3 hit rate to 93%; and the integrated model was 15% more accurate than the single model

Identification of Offensive Language in Chinese Social Media Environment *10-12/2024*

Aimed to build an automated language detection & recognition system, based on deep learning, to help platforms manage implicit and explicit offensive remarks in the Chinese social media environment

- Conducted the collection, cleaning, and annotation of Chinese social media data; built and optimized multiple dedicated datasets including COLD, ToxiCN, and ToxiCloakCN
- Built and trained LSTM neural networks and pre-trained language models (BERT and RoBERTa) to achieve accurate identification and classification of aggressive language
- Evaluated and optimized the model through performance evaluation indicators (Accuracy, ROC-AUC, F1-score); eventually decided on the RoBERTa model, which achieved an accuracy of 78.54% and performed outstandingly in complex language recognition

ADDITIONAL

BNBU Dream Weavers Club *10/2022-Now*

- Have served in the club, one of the most renowned student organizations at BNBU, from initially an ordinary member to currently a vice president who manages a team of 50+ members
- Help plan, organize, and coordinate various student activities including fund-raising charity campaigns, guest speeches, volunteer programs, and so forth; independently led my team to launch a new activity of monthly guest speech from scratch
- Contact and negotiate with external organizations to acquire sponsorship of RMB 10,000+