Shang-Hung Shih (Albert)

Software Engineer · Full-Stack Web Development and Machine Learning

(+886) 910-024730 | shshih8497@gmail.com | shanghungshih.github.io | shanghungshih | shanghungshih

Education

National Taiwan University and Academia Sinica — **Taipei, Taiwan**

M.S. in Genome and Systems Biology Degree Program — *Sep. 2018 - Jan. 2021*

- Worked with Prof. Chien-Yu Chen who focus on Genomic AI applications.
- Related courses: Data Structures and Algorithms, Database Management System, Biomedical Data Mining, NGS, Bioinformatics and Genomic Medicine, etc.

Chung Shan Medical University — **Taichung, Taiwan**

B.S. in Biomedical Sciences — *Sep. 2014 - Jul. 2018*

- Joined a Bioinformatic research team of Prof. Yu-Fan Liu which focus on NGS big data analysis.
- Related courses: Linear Algebra, Java Programming, Bioinformatics, etc.

Professional Experience

Software Engineer, Taiwan AILabs — **Taipei, Taiwan**

TypeScript, ReactJS, Node.js, NestJS, FastAPI, Python, Redis, Neo4j, MongoDB, PostgreSQL, Qdrant, pgvecto.rs, Prefect, Docker, Kubernetes — *Jan. 2020 - Present*

- Spearheaded the design and implementation of flow-hub infrastructure for LLM chain and workflow development, enabling cross-environment end-to-end testing and accelerating development cycles.
- Co-designed and implemented the AutoKB service for FedGPT, developing unstructured data indexing with hybrid search and RAG retrieval, significantly improving embedding search context recall.
- Orchestrated the on-premise deployment of GenDiseak gene analysis platform using k3s, including variant detail pages and a redesigned pagination variant table viewer that improved loading speed by 76.9%.
- Developed key features for TAIGenomics web app, including signup system, DNA sequencing viewer, and pedigree editor, alongside rule-based ACMG and ML-based v-score interpretation systems for variant prioritization.
- Created PubmedKB web app for NLP model result visualization and proteomics data visualization for QCheck quality control.
- Optimized the algorithm team's processor build and deployment pipeline, reducing time by over 90%.

Software Engineer Intern, Taiwan AILabs — **Taipei, Taiwan**

Python, MySQL, SQLite, Docker — *Mar. 2019 - Dec. 2019*

- Engineered a high-performance variant annotator using multi-processing and MySQL, optimizing annotation of 4 million variants against 9 billion data points. Reduced annotation time by 58%, cutting whole genome sample processing from 2 hours to 50 minutes.

M.S. student and Server Administrator, NTU — **Taipei, Taiwan**

Python, MySQL — *Sep. 2018 - Dec. 2020*

- Developed advanced deep learning models to construct transcription factor binding profiles, enhancing genomic data interpretation and prediction accuracy.
- Conducted comprehensive analysis of 1,498 whole genome sequencing datasets from Taiwan Biobank using Taiwania supercomputer, contributing to large-scale genomic studies and personalized medicine initiatives.

Side Projects and Honors

1st place in PIXNET 6th Hackathon Final: Travel Tech (Travel Charger) — **Taipei, Taiwan**

Python, Neo4j, Docker — *Aug. 2019 - Sep. 2019*

- Developed a personalized attraction recommendation system using Pixnet articles for correlation analysis and Instagram data for sentiment analysis, with primary focus on constructing relation graphs for attractions.

Skills

Programming Language

TypeScript(ReactJS, Node.js, NestJS), Python

Others

SQL(MySQL, PostgreSQL, SQLite), NoSQL(MongoDB, Redis, Qdrant), GQL(Neo4j), Git, Linux, Docker, Kubernetes, k3s, Data-Science