Yuan Hai Shu

Mobile: +65 85427025   Email:haishuyuan@yahoo.com.sg

EDUCATION

National University of Singapore (NUS)   **Aug 2019 – Dec 2023**

- Bachelor of Engineering (Beng) Chemical Engineering, Minor in Data Engineering

WORK EXPERIENCE

Patsnap — *Data Engineer*   **Jun 2025 – Present**

- Redesigned and refactored legacy database for Patsnap Eureka for faster searches, and lower storage
- Redesigned a monolithic database into a layered data architecture to improve scalability and maintainability
- Implemented incremental processing, along with traceable fingerprints for the 100% of the pipeline
- Implemented chemical and materials engineering domain based rules processing across the full dataset for data validation

Silent Eight — *Data Scientist*   **Jan 2024– Jun 2025**

- Fine-tuned ML models using various ML algorithms in Tensorflow to improve overall efficiency and effectiveness of solve rate by 5% of alerts raised by clients
- Wrote in-house python scripts to improve quality of policy steps used in reasoning branch reports
- Improved in-house machine learning agent by incorporating LLMs, increasing the efficiency by 10%, and decreasing the false negative rate by close to 12%
- Implemented ETL pipeline in python and spark, developing API payload to in-house services for client bank
- Used NLP algorithms to extract key entities, relationships from client data, to improve solve rates

GlobalFoundries — *DTE Modelling Intern*   **Jun 2023 – Dec 2023**

- Automated the verification of results collected by the characterisation of the wafer machines, and identification of bad dies, reduced time taken from a week to 2 days
- Trained and developed the model which identifies which test has been run, and for which device type, and how huge the discrepancy is from the standard result, using past data collected which are queried
- Linked it to a Microsoft BI dashboard where the severity of the discrepancy is colour coded, to indicate retesting

Panasonic — *Technology Analyst Intern*   **May 2022 - Aug 2022**

- Spearheaded the change from Excel to Tableau for data visualisation
- Built the Tableau dashboards which shows the trends, and other analysis from litigation data provided.
- Significantly reducing the man-hour requirements for this entire project by 30% and ultimately improving our client deliverables efficiency

STEngineering — *Data Analyst Intern*   **May 2021 - Aug 2021**

- Embarked on project with focus on complete solution for predictive maintenance
- Maintained and updated various Qlikview dashboards for visualisation.
- Trained machine learning models to predict breakdowns and delays in services.

COMPETITIONS

Champion in CDL GCNS Young SDG Leaders award 2022

Came up with practical solutions for Vena Energy to solve sustainability issue

Did data cleaning and analysis on the trends based on the data, presented in the final report and presentation submitted.

Champion for Samsung Solve for Tomorrow 2021

Built encapsulated perovskite solar cell with guidance from SERIS, along with an asset management system

SKILLS

Programming & Query	Big Data & Streaming	Databases & Query Engines
Python, Java, SQL	Spark, Flink, Kafka	MySQL, PostgreSQL, Trino, TiDB
AI / ML (Applied)	**Cloud Platforms**	**Data Storage & Compute**
LLM, NLP, TensorFlow, PyTorch	AWS, Tencent Cloud	Cloud Object and Distributed Storage