WU Yanjing

Email : Wu.Yanjing@outlook.com | Mobile : +65 84038231 LinkedIn : <u>https://www.linkedin.com/in/wuyanjing</u> | GitHub : <u>https://github.com/sycamore-st</u>

PERSONAL SUMMARY

Data Scientist with over 2 years of experience in developing predictive models and analyzing complex datasets. I hold a Master's degree in Applied Economics from the National University of Singapore (NUS), and I specialize in statistical modeling and machine learning. My background enables me to transform raw data into actionable insights, supporting strategic decision-making and driving business growth.

EDUCATION

National University of Singapore (NUS)

- Master of Economics (Applied Economics)
- Sun Yat-sen University (one of China's top 10 universities)
 - Bachelor of Economics in Finance
 - UBS Campus Ambassador: Recognized by UBS for outstanding contribution to campus marketing efforts, demonstrating initiative, communication skills, and alignment with UBS values.

TECHNICAL SKILLS

- Programming: Python (Scikit-learn, TensorFlow, PyTorch, Hugging Face Transformers for LLMs), SQL, Stata, R.
- Core Competencies: Natural Language Processing (NLP), Large Language Models, Statistical Modeling, Machine Learning
- Data Engineering & ETL: Building and optimizing scalable data pipelines, feature engineering.
- Environments & Tools: Git, CI/CD, Apache Airflow, Docker, AWS (EC2, S3, Lambda), Cloud-based deployment and automation.
- Financial Market Knowledge: Real estate market analytics, capital markets, financial risk modeling.
- Language: English (PTE: 84, GMAT: 710), Mandarin (Native), Cantonese (Native)

WORKING EXPERIENCE

Real Estate Analytics. PTE. LTD	Singapore
Data Scientist	Jan 2023 – Present
Data Analyst Intern – <i>Data Science Team</i>	Sep 2022 – Dec 2022

- Machine Learning & Predictive Modelling:
 - Developed and optimized predictive models for housing demand forecasting, profitability analysis, and price index formulation using algorithms such as XGBoost, CatBoost, and Random Forest.
 - Utilized NLP and large language models for advanced feature extraction, intent categorization, and other data processing tasks.
- Big Data Engineering and Scalable Data Pipelines:
 - Designed and maintained ETL pipelines using SQL, Python, and Apache Airflow, ensuring seamless data ingestion, transformation, and storage for large-scale real estate data.
 - \circ $\;$ Automated data pipelines with CI/CD integration using Git.
- Data Visualization & Business Insights:
 - Created interactive dashboards and automated reports using Power BI and Tableau to present model outputs, performance trends, and strategic insights to technical and non-technical stakeholders
- Key Projects:
 - Automated Valuation Model Tool: AI-driven automatic valuation model for real estate properties.
 - Real Estate Forecasting Models: Housing Index Formulation, National Housing Index Forecast, National Housing Demand Forecast, Project Demand Forecast, and developer building and pricing strategies recommendations achieving an error rate below 3% (*Patent Filing in progress*).
 - Real Estate Listing Chatbot: Developed a chatbot-powered real estate assistant that leverages GPT-4 for intent detection, automatically generating SQL queries for property listings and market analytics.
 (GitHub: <u>https://github.com/sycamore-st/real-estate-chatbot.git</u>)

Sep 2017 – Jun 2021

Aug 2021 – Jan 2023