WU HANYU

whanyu47@gmail.com　　https://github.com/whanyu1212　　+65 81274309

Profile

I am a seasoned data scientist with a strong commitment to delivering results through data-driven solutions. With a passion for statistical analysis and model building, I take pride in translating data points into insights that illuminate the trajectory towards business goals. Keeping up with the latest advancements in the field, I possess a well-rounded skill set in machine learning, data visualization, and programming, making me a valuable asset to any team.

Work Experience

Data Scientist　　　　　　　　　　　　　　　　　　　　　　　　　　　　**Jan 2022 – Present**
Micron Technology Singapore

- Designed and implemented an end-to-end interpretable machine learning pipeline for an international assignment, successfully identifying root causes of deviation in run time and achieving an approximate 10% reduction, thus significantly enhancing manufacturing capabilities
- Collaborated closely with subject matter experts to develop an energy efficiency prediction model, significantly contributing to cost reduction initiatives
- Built interactive web interfaces with Angular to create real-time monitoring of manufacturing processes and conduct periodic performance evaluations, with a primary focus on detecting bottlenecks and inefficiencies across different workstations and tools
- Conducted an in-depth project centered around the analysis of expendables within the manufacturing process. Leveraged a combination of causal inference techniques and time series studies to derive the optimal parameter ranges for these expendable materials. Additionally, the project entailed a comprehensive examination of their degradation trends, enabling the identification of precise intervention points for timely replacements, thereby enhancing operational efficiency and cost-effectiveness

Data Science Consultant　　　　　　　　　　　　　　　　　　　　　　　　**Feb 2021 - Dec 2021**
KPMG Singapore

- Initiated and executed data processing and cleaning, establishing the groundwork for robust BI tools. Concurrently, advocated for data schema standardization and data quality improvements, resulting in enhanced operational efficiency, and the ability for clients in the tourism sector to derive actionable insights from key tourism figures
- Collaborated with NUS School of Medicine to develop a comprehensive, state-of-the-art dashboard tailored for educators. This dashboard empowered educators to proactively predict and monitor students at risk of facing academic challenges, fostering a supportive learning environment.
- As a part of a consulting team, provided expert guidance to Sentosa Development Corporation on their CRM system, data governance, and the seamless flow of data across systems. This established a robust foundation for future data-driven solutions and expanded use cases.

Associate Data Scientist　　　　　　　　　　　　　　　　　　　　　　　　**Jul 2020 – Feb 2021**
Singapore Toursim Board

- Conducted sentiment analysis using NLP on web-scraped reviews to gain insights into tourist preferences, specifically evaluating aspects such as dining, sightseeing, and local cuisine, to understand the perceived strengths and weaknesses of Singapore in these areas.
- Spearheaded a differential privacy project aimed at enhancing the handling and analysis of sensitive data, with a primary focus on promoting secure data collection and sharing not only within the organization but also with vendors and the public
- Implemented a genetic algorithm-based solution to the survey design optimization problem, resulting in the ability to conduct shorter micro surveys. This led to higher response rates and increased willingness among tourists to participate

in surveys, providing more valuable insights

Education

National University of Singapore | *Singapore* **Aug 2023 – Present**
- Master of Science - Digital Financial Technology (Part time)

National University of Singapore | *Singapore* **Aug 2016 – Jun 2020**
- Bachelor of Science(Hon) - Statistics Major
- Performed well in data science and statistics related coursework
- Ex-president and training coordinatorof the NUS Chinese Debate Team
- Represented school in various debate tournaments, securing several victories and accolades

Personal Projects:

- Designed and implemented a real-time FX trading algorithm for OANDA, utilizing Q-Learning to optimize decision-making in a dynamic market environment

- Developed a comprehensive stock market volatility forecasting model utilizing LSTM-GARCH, SVR-GARCH, and Stacking methodologies to predict both realized and implied volatility. Incorporated diverse data sources, including pricing data, news, and press releases, to enhance forecasting accuracy

- Developed a comprehensive basketball scouting report analysis tool leveraging Langchain, OpenAI's API, and the Streamlit framework. The solution is designed to process uploaded PDFs, extract and aggregate player biographical data, and distill insights into strengths and weaknesses via an intuitive radar chart visualization. Additionally, it integrates a question-and-answer feature to address further user inquiries, enhancing the interactive experience and providing deeper analytical value

- Developed an asset allocation tool leveraging Streamlit, which provides portfolio weight recommendations based on the concepts of Sharpe ratio, Global Minimum Variance (GMV), and the Efficient Frontier. This project enables users to dynamically adjust their investment portfolios, utilizing Scipy optimizers to optimize the balance between risk and return on the fly based on their risk appetite.

- Experiemented with text to SQL translation using Langchain and OPENAI API, making databases more accessible to non-technical users, enhancing data analytics workflows, and simplifying the interaction with complex datasets. Pending the testing offunctionality into a suitable user interface

Skills

PYTHON | JULIA | GOOGLECLOUD | GITHUB | MARKDOWN | TYPESCRIPT

PYTORCH | SCIKIT-LEARN | TENSORFLOW | MLFLOW | LINUX | JIRA

MYSQL | POSTGRES | APACHE AIRFLOW | GIT | GITHUB ACTIONS

DOCKER | LEETCODE | ANGULAR | PREFECT | LATEX