TONY WANG

0403 725 801 tonywang205@yahoo.com.au
LinkedIn GitHub @ Blog

EDUCATION

University of New South Wales
Master of Mathematics. *WAM: 93.2*
Thesis: Theoretical Results for the Lasso and a New Root-Log Regulariser in High-Dimensional Linear Regression.
Relevant coursework: Machine Learning, Neural Networks, Bayesian Inference, Statistical Inference, Time Series.

University of Sydney
Bachelor of Science (Advanced), Honours in Pure Mathematics. *First Class Honours: 92.0*
Thesis: Existence of Solutions to Anisotropic Elliptic Partial Differential Equations.

University of California, Berkeley
Exchange student. *GPA: 3.9/4*

CAREER

Quantium *February 2023 - Present*
Graduate Data Analyst in the Product Analytics vertical.

- Member of the winning team in the 2023 1-day Quantium GenAI Hackathon, with a submission based off using Large Language Models for transaction data labelling in Quantium's Q.Refinery offering.

Work Projects
Snowflake Cost Optimisation
- Conducted exploratory data analysis to identify beneficial Snowflake configuration changes which maintain performance and client satisfaction while also reducing spend. This included suggesting and evaluating rules for a query re-routing strategy.
- Analysed cost savings and performance impact after configuration changes were pushed to production.
- Achieved $500,000 in projected annual cost savings in 6 months.

ETL Migration
- Worked in a 2-man team to scope out and migrate an ETL process from our decommissioned on-prem cluster to Snowflake. This reduced a 10-hour, manual and unstable Spark-based ETL to a one-click Snowflake SQL script, with outputs feeding into a dashboard used by Woolworth's media advertising arm.

Q.Refinery Analytics
- Created a dashboard in Dash (Python) displaying coverage and accuracy metrics for each of Q.Refinery's clients.
- The dashboard was made to be configurable, with new clients and charts automatically generated based on a YAML input file.

Research Projects
Anomaly Detection in Structural Sensing Data
- Used matrix profiling techniques to identify regime changes in noisy sensor data, with real world applications in detecting major structural damage automatically.
- Created a streaming data solution for regime change detection, removing the need for slow, batch processing. (Online): See our website.

SKILLS AND TECHNOLOGIES

Tech: Python, Spark (Scala, PySpark), Snowflake, SQL, Excel.
Languages: English, Mandarin