

WORK EXPERIENCE

PwC

Associate, Digital, Cloud and Data

July 2022 – Present

- Integrated HR and Finance data to analyze the financial costs of HR for a multinational bank.
- Developed ETL pipelines for cleaning and transforming data, and staging of data from a Data Lake to a Synapse pool for dashboard development for a local infrastructure and urban development consultancy.

Energy Market Authority

Intern, Research & Statistics Unit

May 2021 – October 2021

- Created over 100 automated Machine Learning models for predicting forced outages using DataRobot, achieving a F1 score of 0.7.
- Merged and cleansed over 2 years of electricity market time series data using R and engineered over 20 new domain-related features for machine learning.

Housing and Development Board

Intern, Information Services Group

Dec 2020 – Jan 2021

- Applied anomaly detection techniques to detect unusual requests for access in HDB's internal systems.
- Reduced significant the memory and processing time of existing methods using hash environments.

EDUCATION & CERTIFICATIONS

Singapore Management University

BSc (Economics), Second Major in Data Science and Analytics

2018 – 2022

- Cumulative GPA: 3.84/4.00 with Summa Cum Laude (High Distinction)
- Placed on Dean's List in AY 18/19, 19/20 and recipient of Wong Kwok Leong Scholarship in AY 20/21
- President of SMU Athletics in 2020

Microsoft Azure

- Power BI Data Analyst Associate 2022
- Azure Data Engineer Associate 2022

PERSONAL PROJECTS

EY BWB Data Challenge

Semi-finalist

May 2022 – Jun 2022

- Created a ADABOOST model for predicting and understanding the distribution of frog species across Australia using ecological variables and achieved a F1 score of 0.76.

Data Preparation and Auto-ML App Development

Jul 2021- Aug 2021

- Developed 2 easy-to-use no-code applications using Shiny; 1 for data preparation and 1 for creating and evaluating machine learning models.
- Implemented over 10 machine learning modelling and advanced data preparation techniques that can be applied to user-uploaded data.

Shopee Code League

Participant

Jun 2020 – Aug 2020

- Created a neural network for sentiment analysis using BERT to analyse reviews on Shopee products.
- Applied Principal Component Analysis to reduce the dimensionality of large data sets and identify key components explaining engagement rates of Shopee's emails.

SKILLS

- **Languages:** English, Mandarin
- **Programming & Visualisation:** Python, R, SQL, Tableau, Power BI, SAS, Alteryx
- **Frameworks & Libraries:** Pandas, Tidyverse, Plotly, Ggplot, Imblearn, Scikit-learn, Selenium, Tensorflow, Keras, Flask, Streamlit, Shiny
- **Data Science & Machine Learning:** NLP, Statistical Modelling, Spatial Data Analysis, Time Series Analysis, Deep Learning, ARM, DataRobot
- **Big Data & Cloud:** Pyspark, Sparklyr, Databricks, Azure Data Factory, Azure Data Storage