

Working Experience

Machine Learning Engineer / Huawei Ltd.

Dec 2022 – Present

- Built a pipeline to automatically visualize data tables using LSTM network trained on ChatGPT-generated data with pairwise loss method, achieving 80% recall@5 on 100+ internal test cases.
- Designed and implemented a novel SISR method that enhanced WIFI-signal simulations for office buildings by achieving 10x speedup compared to physics-based simulation with negligible loss in accuracy (1% MAE) on over 80 large-scale office layouts.

Machine Learning Research Engineer / Dyson Ltd.

Sept 2021 – Dec 2022

- Implemented an object localization model in a few-shot context by semi-supervised training. The model achieved comparable results to professional software with improved adaptability and robustness.
- Designed and implemented an air quality estimation model, using LGBM, Bayesian Regression, etc., with geographical and meteorological features. Demonstrated its advantages over spatial interpolated methods and deployed the pipeline with Metaflow framework on AWS services.

ML Research Assistant / NUS-Singtel Cyber Security Lab

Sept 2020 – July 2021

- Identified anomalies from system logs leveraging DBSCAN and hierarchical clustering for model training.
- Developed an information retrieval method for web-attack strategy identification from system and firewall logs. The recall@3 rate achieved 80% on 100+ hand-labelled samples.

Data Analyst Intern / GIC Pte. Ltd.

Dec 2018 – July 2019

- Deployed an R application that forecasts the mid-term returns of portfolio with visualization using R shiny.
- Optimized the coefficients of a mean reversion forecasting model using the Genetic Algorithm.

Data Analyst Intern / PropertyGuru

May 2018 – Aug 2018

- Developed dashboards in Tableau to analyze the user behaviors and listings' performance to better match user demand to agents' recommendations.
- Implemented a POC to calculate and geographically visualize the liveability score for properties.

Education

MComp in Artificial Intelligence / NUS

Aug 2020 – Sept 2021

- **School of Computing:** CAP 4.42/5.0
- **Teaching Assistant:** Advanced Analytics and Machine Learning 2021

Jan 2021 – May

BSc. (Hons) in Business Analytics / NUS

Aug 2016 – June 2020

- **School of Computing:** CAP 4.15/5.0, Dean's List in Semester 3 AY 2018/2019
- **Distinction:** [Analytics Techniques Knowledge Area](#)
- **Teaching Assistant:** Programming Methodology in python

Dec 2020

Aug 2017 – June 2018

Relevant Projects

Distilling ChatGPT for finetuning image captioning models

Jan 2023 – Present

- Employed Chain-of-Thought with verification prompting technique on ChatGPT to create 10k+ accurate captions from the xView annotations. Fine-tuned a GIT image captioning [model](#) and significantly improved the CIDEr score from 11.59 to 85.93 over 2k RSICD samples.

Dialogue Response Generation (Master Thesis) / NUS NExT++ Lab

Nov 2020 – Aug 2021

- Built an enriched task-oriented response generation by implementing copy-mechanism on GPT-2 using Pytorch. The proposed model is capable of naturally incorporating external tips/user reviews about venues into responses. The generated response outperforms many state-of-the-art models on user satisfaction.

Property Resale Price Prediction

Jan 2021 – May 2021

- Fitted CatBoost, LGBM, XGBoost on 43k pieces of property sales data. Selected features by correlation and information gain. Engineered new features describing properties' livability. Reduced data dimensionality with WOE encoding. The final ensemble methods' accuracy achieved 5th/64 place.

Skills

- **Python (Pytorch, Tensorflow), R:** Machine Learning, Deep Learning, Data processing
- **SQL, Spark:** Data query and big data
- **Tableau, PowerBI:** Visualization development
- **Java, Git, Scala, JavaScript, HTML, CSS:** Software Development