

ZHANG Simian

Mobile: (+65) 84560290 | Email: zhang.simian@u.nus.edu | GitHub: <https://github.com/simonCodeZzz>

EDUCATION

National University of Singapore (NUS), BSc (Hons) **Aug 2019 – May 2023(Expected)**

- Major in Data Science and Analytics, Second Major in Economics, Minor in Computer Science
- Cumulative Average Points: 4.77/5 (Expected First Class Honors)
- Awards: Dean's List AY2021/22 Semester 1 & 2, Undergraduate Science & Technology Scholarship

INTERNSHIP & RESEARCH EXPERIENCES

Student Research Assistant, NUS Institute of Operations Research (IORA), Singapore **Jan 2022 – Present**

Academic institute with research focus on the optimization, analysis and management of service systems.

- Supervised by Prof Jussi Keppo from NUS Business School
- Scraped and processed data of Kaggle competition submissions, performed Bayesian estimation for crucial model parameters and assisted in model diagnostics in mechanism design project
- Researched on herd detection in financial forecasts using large-scale Bayesian structural estimation methods

Data Engineer Intern, ByteDance, Singapore **May 2021 – July 2021**

Chinese multinational internet technology company with products including Tik Tok and Douyin.

- Constructed stable data pipelines using Hive, Spark, Kafka to account for costs and revenues of video services such as storage, transcoding, playing and content delivery network (CDN)
- Collaborated with product managers and engineers to optimize existing data pipelines and provide accurate and timely dashboard reporting
- Built forecasting model in Python to predict usage of bandwidth by video services with mean absolute error less than 3% (improved from > 10%) to help budget planning

Data Analytics Intern, International SOS, Singapore **Dec 2020 – Jan 2021**

World's leading health and security risk services company.

- Assisted data scientists and data analysts with timely data cleaning, modelling and reporting
- Wrote R scripts to automate data quality checking of more than 1,000,000 entries in data migration project
- Analyzed medical text data, performed feature engineering and applied clustering model to help management gain operational insights

Data Analytics Intern, Microsoft, China **Jun 2020 – Jul 2020**

Leading global tech company with services including software, mobile and cloud services.

- Analyzed spending data of churned customer to extract actionable insights
- Explored feature engineering and clustering in Python to segment and label customers based on 7-day spending
- Performed time series analysis on 30-day customer spending data and developed regression models to help clients establish operation strategy and optimize management

Managing Director, Brighster Education, China **Jun 2019 – Jul 2019**

Education startup providing coaching and pre-overseas-education training services.

- Assisted 12 students to get full scholarship and study opportunities in Singapore, overall saved more than \$100K of tuition fees for those in need
- Ran and managed a team of 13 with promoting plans including road show and campus branding to enroll candidates and met targets

PROJECTS

Large-Scale Optimization for Classification Problem **Jan 2022 – Apr 2022**

https://github.com/simonCodeZzz/DSA4212_Project

- Led a team of 4 to explore impacts of different image processing procedures and optimization algorithms on large-scale image classification with best test accuracy of 0.93 and AUC of 0.97
- Implemented framework of multi-class logistic regression and SGD variance reduction algorithms from scratch in Jax and investigated their empirical performance on large-scale data sets

- Achieved faster convergence rate and lower computation cost compared to baseline algorithm

Singapore Mixed Veg Rice (Cai Fan) Recognition

Jan 2022 – Apr 2022

https://github.com/willpyt98/CS4243_Project

- Scraped and cleaned image data of popular Singapore Mixed Vegetable Rice dishes from Google image in Python
- Implemented deep learning models including MLP and CNN variants for classification in PyTorch (test error < 9%)
- Researched and applied weakly supervised learning technique Class Activation Map (CAM) for dish detection

Enterprise Database System

Sep 2021 – Nov 2021

<https://github.com/simonCodeZzz/cs2102-assignment>

- Developed enterprise database system for employee information storage, room booking, contact tracing and report viewing in PostgreSQL
- Translated business requirements into ER diagram, relational schema and PL/pgSQL procedures/functions
- Fortified data consistency by creating triggers and extensive test cases

E-Commerce Transaction Prediction System

Aug 2021 – Sep 2021

<https://github.com/simonCodeZzz/dsa3101>

- Led a team of 7 to predict e-commerce transaction helping companies to increase purchase conversion rates based on budget constraints and sales targets
- Directed project flow from literature review, data preprocessing, exploratory data analysis to modelling, metrics design and report writing
- Achieved AUC-ROC score of 0.87 on training data and 0.82 on unseen test data

Using Respiratory Sound for Respiratory Disease Detection

Feb 2021 – Apr 2021

<https://drive.google.com/file/d/1oPO05w8G-ZdavPDpp5c-x0jD4aEOEb90/view?usp=sharing>

- Predicted respiratory health condition with machine learning models on respiratory sound data, achieved near 0.90 precision and recall
- Performed different models including XGBoost Decision Tree and Logistic Regression on annotated data to compare results with raw sound data

LEADERSHIP

President, NUS 90 Degree Chinese Drama Club

Apr 2020 – Apr 2021

- Led a team of 30 to come out 2 drama productions in Singapore Drama Center that was featured in local media Lianhe Zaobao (联合早报)
- Produced and co-directed a school-themed drama series of 6 episodes
- Administered drama training sessions, recruitments and marketing with an achievement of 100 audiences and \$5000 sponsorship

EXCO, NUS Active Community Engagement Migrant Committee

Sep 2019 – Apr 2020

- Collaborated with a team of 5 focusing on Singapore migrant workers' rights, equality and well-beings, engaged to bring close them with students to eliminate misunderstandings
- Planned and executed engagement sessions for NUS students in College of Alice & Peter Tan and foreign domestic workers in Centre for Domestic Employees
- Over 20 foreign domestic workers attended sessions with feedback of appreciation

ADDITIONAL INFORMATION

- Programming Tools: Python, R, SQL, Java, Tableau, Hive, Spark, Kafka, Flink, MATLAB, STATA
- Languages: English(fluent), Chinese(native), Japanese(elementary)