

Harry Chang

Portfolio: harrychangjr.streamlit.app

Github: github.com/harrychangjr

Email: harrychang.work@gmail.com

LinkedIn: linkedin.com/in/harrychangjr

Mobile: +65-8428-5869

EDUCATION

- National University of Singapore** Singapore
Bachelor of Science - Data Science and Analytics Aug 2020 - May 2024 (Expected)
Coursework: Convex Optimization, Data Science in Practice, Data Structures and Algorithms, Data Visualization, Database Technology and Management, Linear Algebra, Multivariable Calculus, Optimization for Large-Scale Data-Driven Inference, Probability, Regression Analysis, Statistical Learning

SKILLS

- Languages:** R, Python, SQL, Java, Stata, MATLAB
- Machine Learning:** Scikit, NLTK, TensorFlow, Keras, JAX
- Cloud and Database Systems:** MySQL, PostgreSQL, SQLite, BigQuery, GCP, AWS, Git, Docker
- Data Visualization:** Numpy, Pandas, Matplotlib, Seaborn, ggplot2, Folium, Tableau, Power BI, Domo, Gephi
- Design and Front-End:** Figma, Canva, Streamlit, HTML, CSS
- Miscellaneous:** Microsoft Office, Retool, Google Ads, Google Analytics

EXPERIENCE

- Sephora** Singapore
Product Data Analyst Intern Jun 2023 - Dec 2023 (Expected)
 - Database Queries:** Built, documented and hosted SQL queries and processes to enable reproducible and effective pipelines, analysis and dashboards using BigQuery
 - Data Visualization:** Utilised Domo to create dataflows and visualizations that provide quick insights into product health and performance of e-commerce features (web and mobile)
 - A/B Testing:** Implemented A/B testing to measure potency of new e-commerce features before reporting results
- SCOR** Singapore
Actuarial Intern May 2022 - Aug 2022
 - Actuarial Analysis:** Performed actuarial analysis of reinsurance treaties in various APAC markets, including entry of client portfolio and loss data into xAct (treaty pricing system)
 - Insurance Market Databases:** Regularly updated and analysed risk profiles and claims databases for insurance markets in Pakistan, Thailand and Vietnam
 - Credit Product Pricing:** Trained machine learning models (logistic regression, random forest) to predict insurance claims, with an average accuracy of 80% for each model
- Quest** Singapore
Data Analytics Intern Feb 2022 - May 2022
 - Product Analytics:** Conducted cohort analysis to optimise user acquisition and retention rates
 - Growth Analytics:** Collected, analysed and interpreted trends within user data to improve company's growth and marketing strategies
 - Dashboard Processes:** Built visualizations and dashboards using RStudio and Tableau to report monthly key metrics of company's mobile application
- Immigration@SG LLP** Singapore
Data Migration Intern Oct 2020 - Jan 2021
 - Data Cleaning:** Performed data cleaning of over 30,000 records using Excel and Pandas in Python to facilitate smooth data migration into new customer relationship management (CRM) systems
 - Customer Segmentation:** Derived customer segmentation models using machine learning techniques (regression, decision trees, market basket analysis) to improve company's marketing strategies
 - Time Series Analysis:** Completed time series analysis using past sales data to forecast future monthly revenue

PROJECTS

- PassionPassport:** Built and integrated web application with ChatGPT that utilises hobbies as inputs to recommend itineraries for overseas travel. Awarded Top 15 Finalist for NUS LifeHack 2023. Tech: Python, HTML, CSS, Firebase, Streamlit, OpenAI API, Folium, BeautifulSoup4. (Jun '23)
- Enhanced TikTok Analytics Dashboard:** Open-source dashboard for TikTok users to analyse performance (views, shares, likes & comments) of videos using hashtag analysis. Tech: Python, Streamlit, NLTK, Gensim, Plotly, Matplotlib & Seaborn. (May '23)
- Optimising Article Quality with ChatGPT and Natural Language Processing:** Evaluated performance of ChatGPT in improving original write-ups of science news articles using NLP methods. Tech: Python, Streamlit, NLTK, Textstat & Matplotlib. (May '23)
- Statistical Learning - Analysis on Video Game Sales:** Utilised various regression methods (random forest, XGBoost) and linear mixed-effects model to investigate effects of predictors in predicting global sales. Tech: R (Apr '23)

- **Statistical Learning - Nourish Your Body with Data:** Built matrix-factorisation based recommendation system to suggest healthier alternatives for grocery shoppers with varying demographics Tech: R, Python, Streamlit (Apr '23)
- **Data Science Project on Biopics Dataset from Kaggle:** Built content-based recommendation system using cosine similarity to recommend similar movies based on input title. Tech: Python, R. (Apr '23)
- **Optimisation for Large-Scale Data-Driven Inference - Word Embedding:** Trained Word2Vec model on 20 Newsgroups dataset from scikit-learn package in Python, which provides a number of similar words based on input word. Also evaluated usefulness of model by applying sentiment analysis (86.4% accuracy). Tech: Python, NLTK, Word2Vec, Scikit-Learn. (Mar '23)
- **Optimisation for Large-Scale Data-Driven Inference - Anime Recommendation System:** Built content-based and matrix factorisation-based recommendation systems to recommend similar anime series based on input title while minimising MSE. Tech: Python. (Mar '23)

COMPETITIONS AND AWARDS

- Top 15 Finalist at NUS LifeHack 2023 - June, 2023
- Theme Best (Safety) and First Runner Up (Overall) out of 117 team submissions at NUS LifeHack 2022 - July, 2022
- First Runner Up at NUS Fintech Month Hackathon 2021 - January, 2021
- SMU-LIT (Legal-in-Tech) Hackathon 2021 - July, 2021
- NUS Economics Case Competition 2021 - June, 2021
- Shopee Product and Design Challenge 2021 - December, 2020
- Singapore Business Case Competition 2020 - October, 2020

ACTIVITIES

- **Co-founder and President, NUS Product Club** Singapore
Spearheaded logo design and oversaw operations, partnerships and workshops of new club Apr 2023 - Apr 2024
- **President, NUS Statistics and Data Science Society** Singapore
Oversaw efforts of 56 club members in organising events and technical workshops May 2022 - Nov 2022
- **Track and Field (Science) Captain, NUS Inter-Faculty Games** Singapore
Led and trained 24 runners for track relay events (4x100m, 4x200m, 4x400m) May 2022 - Aug 2022
- **Publicity Executive, NUS Inter-Faculty Games** Singapore
Increased publicity of event through extended outreach to over 5,000 students Feb 2022 - May 2022
- **Deputy Head of Finance, Google Developer Student Clubs NUS** Singapore
Compiled team budgets to ensure sufficient funding for workshops and hackathon Sep 2021 - Apr 2022
- **Marketing Director, NUS Statistics and Data Science Society** Singapore
Secured \$19,000 worth of sponsorships for 850 students in Data Analytics Competition May 2021 - Apr 2022
- **Volunteer Management Co-head, Project Safe Space** Singapore
Managed welfare of volunteers while planning recreational activities at Anglican Care Centre Mar 2021 - Jul 2021
- **Public Relations Executive, NUS Open Day Student Village** Singapore
Liaised with participating bodies to increase awareness of event to prospective students Jan 2021 - Mar 2021