Yudi Su

Github LinkedIn yudi.su@mail.mcgill.ca

Education

McGill Universit, Desautels Faculty of Management - Bachelor of Commerce 2022 - E2025
General Management Major (Business Analytics & Operation Management), Minor in Statistics *GPA: 3.72/4.0*

Coursework

Courses: Operation Management, Operation Research, Decision Analytics, Machine Learning, Python Programming, Linear Algebra, Calculus, Probability & Statistics

Awards

- McGill Entrance Scholarship (25,000 CAD)
- IMSF Impact Award (3,000 CAD for Project Funding)

Skills

Languages: Python, SQL, JavaScript, HTML/CSS, LaTeX
Tools: Git/GitHub, VS Code, PyCharm, Anaconda
Libraries: Pandas, NumPy, Matplotlib/Seaborn, Pytorch, Re, Selenium, Sklearn, Gurobi

Research Experience

Investigating Attributes in COVID Testing Groups with Positive Results | *Supervisor: Rob Glew* In Progress

- Analyzed a dataset with **25,000** students with **50+** features. Developed a random forest model to select most influential group attributes based on the feature importance.
- Utilized a parametric cumulative normal function to construct the conditional probability of test groups obtaining positive results using.
- Checked model robustness by using a linear propensity model and a logit model.

McGill Centre for the Convergence of Health and Economics | *Researh Assistant* Sept. 2023 - April 2024

- Scraped data of traditional food information and conducted text analytics to evaluate traditional and grocery store food options, involving text cleaning, English-French translation and similarity scoring calculation using regular expressions, NLP techniques, and libraries like fuzzywuzzy and unidecode.
- Analyzed data of **36,000+** transactions from local grocery chains. Applied a Tobit regression model to explore what influence indigenous people to buy traditional foods.
- Visualized the data by using Python and developed an interactive dashboard through PowerBI, and presented the research findings through research poster.

Project Experience

Case Study - American Office Systems, Inc. | *MGSC 404 - Decision Analytics* Feb. 2024

- Developed a Linear Programming model to optimize American Office Systems' fund distribution across production and advertising, considering a range of financial and operational factors to maximize profits.
- Formulated a Linear Programming model with **80** decision variables and **94** constraints, and coded in Python and used Gurobi to solve the model.
- Performed sensitivity analysis to investigate the optimal allocation strategy under **10** different scenarios.

Yelp Star Rating Initiative | *INSY 446 - Data Mining* March 2024

- Analyzed big datasets with **170,000+** businesses and **125** attributes, aimed at examining the influence of attributes of businesses and customer reviews on star ratings.
- Performed feature engineering and dimensionality reduction techniques, like random forests and PCA to select most influential factors, and developed a logistic regression model (for better interpretabilities) and obtained the **10-fold CV accuracy of 80.96%**.
- Utilized NLP techniques to perform keyword and sentiment (polarity & subjectivity) analysis for business in Yelp using the library TextBlob.

Case Study - Optimizing Battery Capacity for Solar Energy | *MGSC 404 - Decision Analytics* April 2024

- Developed Python simulation programs to determine the battery size with the lowest cost under different stochastic scenarios, such as weather, battery aging, etc.
- Optimized teammates' code by using vectorization instead of for loops, increasing the speed of code by **80%**.

Prediction on the Success Rate of Kickstarter Projects | *INSY 446 - Data Mining* April 2024

- Analyzed a dataset with **15,000+** projects and **45** features. Performed EDA and prepared the data for predictive model through data cleaning, one-hot encoding, and feature engineering.
- Developed a Gradient Boosting Classifier model, using the parameters based on the highest accuracy score from 5-fold GridSearchCV. The final model had an accuracy of **75.76%** and f1 score of **65.4%**.

Working Experience

IDEMIA | *Shenzhen, China (On-site, Full-time) - Data Analyst Intern* May 2023 - August 2023

- Designed new feature in the SQL database that allows for more efficient calculation of inventory level forecasting.
- Automatized the calculation of inventory forecasting using Python, visualized through PowerBI, increasing calculation efficiency by **50%** and enhancing accuracy by **80%**.
- Developed project tracking systems using Microsoft Power Automate visualized the progress through PowerBI, reducing the operational cost by **5%**.

New England Educational Consulting Group | *China (Remote, Part-time) - Math Teacher* April 2023 – Present

- Taught high school and advanced math subjects, including High School Maths, AP Calculus, and AP Statistics.
- Taught **10** students, resulting in an average grade increase of **25%** and a **100%** success rate of scoring 5 on AP exams for attendees.

Leadership and Extracurricula

Integrated Management Student Fellowship (IMSF), McGill University | *LINK* Fall 2023 - Winter 2024

- Conducted empirical research on the topic of senior homelessness and social isolation, obtaining knowledge by interviewing with **20+** stakeholders (seniors, local organizations, community workers).
- Designed a tech workshop strategy aimed at equipping seniors with the skills of engaging with the local community to prevent them falling into homelessness due to isolation.

International Buddy Program, The New Student Mentorship Program | *Mentor* Fall 2023 - Winter 2024

- Paired with 4 new students (both international and local) to offer linguistic support, cultural guidance, academic assistance, and information about life at McGill University.