Justin Shum

jshum2117@gmail.com linkedin.com/in/jshum17 github.com/jshum17

Technical Skills

Languages: Python, R, SQL, Java, HTML, CSS, C++, C, JavaScript, TypeScript
Technologies: MySQL, Flask, Vue.js, Jupyter Notebook, PowerBI, Git
Libraries: pandas, NumPy, scikit-learn, PyTorch, Matplotlib, seaborn, statsmodels, Chart.js, spaCy, NLTK

Projects

UBC Lost & Found Analysis and Forecasting | GitHub **January 2025 – Present**

Data Analysis | *Python, Jupyter, scikit-learn, PyTorch*

- Conducted a comprehensive data analysis to uncover trends and identify actionable recommendations to improve the operational efficiency and communication channels of the UBC Lost and Found System
- Performed topic modeling using spaCy and Latent Dirichlet Allocation (LDA) to categorize items based on descriptions
- Created clear, data-driven visualizations using seaborn to showcase patterns in temporal data and high-traffic areas
- Implemented an ARIMA model using statsmodels to forecast the volume of items lost over time
- Designed and trained a LSTM model in PyTorch, improving forecast accuracy by 92% over ARIMA

Web Dashboard | *Flask, Vue.js, Chart.js, MySQL, Python*

- Developed an interactive web dashboard using Flask, Vue.js and Chart.js, enabling users to visualize real-time insights and forecast lost item volumes
- Utilized Flask to handle API calls and manage data flow, integrating with MySQL for structured storage and optimized SQL queries for efficient retrieval of 200+ lost item records

Airbnb Listings Analysis | *Python, Jupyter* | GitHub **June 2024**

- Analyzed 22,000+ Bangkok Airbnb listings with pandas, aggregating data on pricing and booking trends
- Visualized the distribution of nightly rates and booking patterns using matplotlib through simple and effective visuals

NBA Statistics Dashboard | *PowerBI, Excel* | GitHub **May 2024**

- Utilized Power BI to create an interactive dashboard that showcases various NBA teams and players' performance metrics
- Designed intuitive visualizations of player statistics including points, assists, rebounds, and shooting percentages
- Implemented interactive filters that allow users to filter data by player, team, and position to customize their analysis

Heart Disease Data Analysis | *R, Jupyter* | GitHub **March 2024 – April 2024**

- Built a K-NN classification model in R to predict heart disease using patient data from the Cleveland Clinic
- Applied data preprocessing and hyperparameter tuning techniques to improve model performance
- Achieved an accuracy rate of 74% based on over 300 patient records

IngredientIQ (nwPlus HackCamp 2023) | *HTML, CSS, JavaScript, JSON, Figma* | GitHub | Demo **Nov 2023**

- Developed a website that serves as a food substitution finder for people with dietary restrictions
- Utilized ChatGPT to construct a JSON database with parameters for extracting alternative food options
- Designed an intuitive user interface using Figma, enabling users to find alternatives with a comprehensive analysis of pros, cons, and calorie content for each alternative

Experience & Extracurricular Activities

Marketing Coordinator **October 2024 – Present**

UBC Data Science Club *Vancouver, BC*

- Developed and executed marketing strategies to promote club events, leading to a 25% increase in event attendance
- Designed visually engaging marketing materials and social media posts using Canva

Junior Instructor **July 2024 – August 2024**

UBC Geering Up Engineering Outreach *Richmond, BC*

- Assisted in delivering STEM lessons and guiding elementary students through coding-based activities in summer camp

Data Science Tutor **May 2024 – June 2024**

UBC Centre for Accessibility *Vancouver, BC*

- Provided one-on-one DSCI 100 tutoring sessions for a student with a disability
- Guided student through data analysis assignments and facilitated their understanding of data science concepts

Education

University of British Columbia **September 2022 – Present**

3rd Year Bachelor of Science, Computer Science Major (Expected Graduation May 2027) *Vancouver, BC*