Ben Lau

(+852) 6234-0301 | benyplau@gmail.com | github.com/benlau6 | benlau6.github.io

Work Experience

Bull Butler FinTech Services Limited – Data Scientist   05/2022 – Present
- Conducted an event study on the impacts of listed company announcements
- Built a B2B matchmaking web application hosted on Azure cloud platform using Rust, Python, and JavaScript
- Designed an end-to-end recommender system
- Conducted qualitative and quantitative researches in legal and finance service industries
- Meet with end users and domain experts regularly to gather requirements and knowledge for agile development

Hong Kong Public Opinion Research Institute – Executive (Data Science)   08/2021 – 05/2022
- Performed statistical analysis routinely on survey results
- Implemented Bayesian data imputation model to tackle missing data problem
- Built a directed graph data model for survey question flow
- Automated part of the data pipeline for survey analysis

MTR Corporation Limited (via eCloudvalley) – Consultant   03/2021 – 08/2021
- Implemented Kalman filter for estimating actual value from noisy sensor data
- Built a Gaussian process model for forecasting
- Built a change point detection model for detecting behavioral change of malformed railway tracks

eCloudvalley Digital Technology Co., Ltd. – Data Engineer   08/2020 – 08/2021
- Built several serverless APIs in AWS cloud platform
- Built automated demand forecasting pipeline with AWS products
- Migrated data workflows from on-premise to AWS
- Led regular meetings with clients to meet their business needs

Research Experience

The University of Hong Kong – Research Assistant (Operations Research)   07/2024 – Present
Vehicle Routing Problem Solver with Visualization – benlau6.github.io/projects/routing-visualization
Demo – routing-visualization.vercel.app
- Build a workflow in Python to solve daily scheduling problem with hundreds of daily vehicle requests
- Conducted extensive literature reviews on mathematical programming and vehicle routing problem (VRP)
- Applied constraint programming to solve VRP with time window and capacity constraints in Python using OR-Tools
- Visualized geospatial data on map of Hong Kong interactively in JavaScript

Gaussian Process for Time-To-Event Analysis – Master's Thesis   01/2022 – 05/2022
benlau6.github.io/documents/master-thesis.pdf
- Built a hierarchical Bayesian model with Gaussian process for survival analysis
- Conducted extensive literature reviews on survival analysis and Gaussian process

Education

The University of Hong Kong – Master of Statistics   09/2020 – 07/2022
- Equipped enough statistical foundation to conduct research and implement novel algorithms independently
- Deepened knowledge in advanced statistics such as Bayesian modeling and time series modeling
- Conducted research on Gaussian process for survival analysis

The University of Hong Kong – BEng in Computing and Data Analytics   09/2015 – 05/2020
- Spent 90% course credits on Computer Science and Statistics
- Studied algorithms, data structures, programming, statistical analysis and machine learning

Certificates

AWS Certified Machine Learning – Specialty   11/2020

AWS Certified Cloud Practitioner   11/2020

Skills

Quantitative Skills: Bayesian Modeling, Causal Inference, Mathematical optimization, Visualization, Statistical Inference, Machine Learning, Probabilistic Programming, Constraint Programming
Programming Skills: Data processing, Optimization, Full Stack development, Design Patterns
Programming Tools: Pandas, PyMC, OR-Tools, Svelte, HTML, CSS, Docker, Git, FastAPI, Markdown, LaTeX, Vim, AWS
Programming Languages: Python, SQL, Rust, JavaScript, R
Languages: Cantonese (native), English (fluent), Mandarin (fluent)

Projects

COVID-19 Bayesian Analysis – benlau6.github.io/projects/covid · 09/2024 – 09/2024
- Performed Bayesian analysis to estimate effective production rate and daily new cases with PyMC in Python

Full Stack Website in Rust and HTMX – benlau6.github.io/projects/fullstack-rust · 09/2024 – 09/2024
Demo – pokemon-auth.shuttleapp.rs
- Built an API in Rust and a website in HTMX and tailwindcss
- Used traits in Rust to define shared behavior as contracts to be fulfilled by services, handlers, and ETL procedures
- Implemented user authentication and authorization from scratch
- Built a scrapping program to collect data and stored it in Postgres database

Bayesian Change Point Detection – benlau6.github.io/projects/change-point · 08/2021 – 08/2021
- Implemented several hierarchical Bayesian models to detect change points in time series data
- Tried linear regression, Gaussian processes, Kalman filter, Gaussian mixture, and Student's t mixture models

FastAPI backend in Python – github.com/benlau6/fastapi-fullstack · 07/2021 – 08/2021
- Built a FastAPI backend in Python
- Examined relational and non-relational databases
- Dockerized entire application for the ease of deployment
- Explored static typing and testing in Python

Sentiment Analysis on Amazon Reviews – Bachelor's Final Year Project · 09/2019 – 05/2020
- Applied sentiment analysis on millions of product reviews
- Examined different neural network models, including CNN, RNN, Transformer

Topic classification on Cantonese Forum Posts · 02/2019 – 05/2020
- Implemented text classification on a Cantonese forum to categorize the contents of 100,000 posts into topics

Volunteer Experience

github.com/probml/pyprobml – Contributor · 03/2022 – 04/2022
Python code for "Probabilistic Machine learning" book by Kevin Murphy
- Implemented algorithms from machine learning research papers
- Migrated existing machine learning program into other framework

Simon K. Y. Lee Hall – Floor Representative · 10/2016 – 10/2017
- Led 24 residents to organize activities routinely
- Cooperated with student association
- Maintained relationships between 24 residents

Simon K. Y. Lee Hall – Orientation Camp Organizing Committee · 03/2016 – 10/2016
- Led 120 student helpers to organize and run a week-long camp for 80 freshmen
- Held three week-long discussion forums for 100 residents