Kevin Yang

MSc. Bioinformatics

(604) 710-7454
Vancouver, BC
kevinyang10@gmail.com

keviny2.github.io/
github.com/keviny2
linkedin.com/in/keyang2

SKILLS

Tools and Languages	Python, R, SQL, Linux, Git, Azure, ImageJ, Jupyter Notebook, SLURM
Packages	PyTorch, NumPy, Pandas, Scikit-learn, Scipy, pysam, matplotlib, seaborn, Pyro, NumPyro
Machine Learning	Deep Learning, Bayesian Networks, Supervised/Unsupervised Learning, Clustering and Classification

TECHNICAL EXPERIENCE

3D Image Reconstruction of Cancer Tumours   01/2021 — Present
BC Cancer Research Centre   *Vancouver, BC*

- Decrease validation error by 60% for a computer vision problem through rapid prototyping with deep learning libraries
- Exploit image analysis tools such as ImageJ for preprocessing and decrease memory usage by over 75%
- Accelerate neural network training over 100x using GPUs with CUDA programming in PyTorch

Bayesian Probabilistic Graphical Model for Early Cancer Detection   01/2021 — Present
BC Cancer Research Centre   *Vancouver, BC*

- Design reliable machine learning models to infer cancer population structure and reduce error by 75%
- Exploit parallel programming with HPC clusters to gain 10x speed up on pipeline executions
- Coordinate with a team of over 10 researches from two countries for deploying machine learning models to the cloud
- Apply noise reduction techniques to enhance data analysis by over 20%

Teaching Assistant   09/2019 — 08/2020
The University of British Columbia   *Vancouver, BC*

- Communicated key concepts related to data structures and basic algorithms, receiving 100% positive feedback from students
- Coached students through coding assignments and exam preparation for up to 50 students during weekly office hours

Application Developer / Property Value Website   01/2019 — 08/2019
Statistics Canada   *Ottawa, ON*

- Stored, managed and visualized Property Value data by writing 10+ SQL stored procedures in relational databases
- Implement test-driven code for a program to process web scraping data and achieved over 80% function coverage
- Collaborated with 4–5 senior programmers to develop innovative working solutions, optimizing an $O(n^2)$ algorithm to $O(n)$

PROJECTS (SEE GITHUB)

Data Science Regression Application

- Created a GUI Application to perform 6 different types of statistical regressions on user inputted data sets
- Applied Object Oriented Design Factory Pattern to maintain code clarity and cleanliness

Clustering for Pan-Cancer Analysis

- Extended a machine learning clustering algorithm to perform multivariate analysis
- Cut memory usage by 75% by creating an R package to process and wrangle high-dimensional datasets

Vancouver Temperature Forecasting

- Analyzed time series data using correlograms, partial correlograms, and differencing
- Forecasted the next 12 observations in sequence Vancouver's temperature in 2020 using Holt-Winter Exponential Smoothing

Effects of Heat Treatment on Grapevine Performance and Botrytis Disease

- Conducted A/B testing using the Kruskal-Wallis test and a one-way ANOVA, uncovering 2 statistically significant variables

EDUCATION

MSc. Bioinformatics - *GPA: 4.0*, *The University of British Columbia*   2021-2022
- *Canada Graduate Scholarships - Master's (CIHR) ($17 500 over 1 year)*
BSc. Computer Science & Statistics - *GPA: 3.9*, *The University of British Columbia*   2015-2020
- *Dean's Honour List*