Li, Feifei

github.com/ff98li
ff.li@mail.utoronto.ca

RESEARCH INTERESTS

Computational biology algorithms, Computational genomics, Machine learning in bioinformatics

EDUCATION

University of Toronto, Toronto, Ontario Canada

B.Sc., Bioinformatics and Computational Biology Specialist **September, 2018 - present**

COMPUTING SKILLS

- Data Analysis: Hypothesis Testing, Statistical Estimation, Sampling Methods, Data Visualization
- Machine Learning: k-NN, k-Means, Linear/Logistic Regression, SVM, PCA
- Deep Learning: CNN, RNNs, Transformers, NLP Models, VAE, GAN
- Languages: R, Python, Java, C, SQL, LaTeX, Linux/UNIX shell scripting.
- Applications: Version control systems for team projects (SVN and Git), Docker, Conda, Common spreadsheet and presentation software
- Mathematics: Multivariable Calculus, Linear Algebra, Computational theories, Data Structures and algorithm designs.
- Operating Systems: UNIX/Linux, Windows.

LABORATORY SKILLS

- General Laboratory: Pipetting, Filtration, Titration, Media Preparation, Glassware cleaning.
- Biochemistry: PCR, Electrophoresis, Blotting, theoretical knowledge of chromatography, column, and florescent antibody techniques
- Instruments: Optic microscopes, Vortex mixers, Centrifuges, Spectrophotometers.

EXPERIENCE

University of Toronto, Toronto, Ontario, Canada *Undergraduate*

Software Development **October - December, 2018**

- As a team of 4 together developed an Android application.
- Gained experience in managing a development team as a leader by planing a project schedule, teaching the team members how to work remotely using version control system, assigning tasks to the team members, holding multiple in-person meetings and organising a team project presentation.

Analysis of Toronto Break and Enter Crime Data in 2019 **March, 2020**

- A collaborative project jointly conducted by the University of Toronto and the Toronto Police Service to analyse real data on Break and Enters in Toronto during 2019.
- As an accidental leader in charge of a team of 4, responsible for selecting appropriate statistical methods for analysing large datasets, integrating team members' ideas and render them into actual code.

STEM Fellowship Big Data Challenge 2020 **May, 2020**

- Statistical analysis employing supervised and unsupervised machine learning methods to explain severity of COVID-19 transmission in 144 countries using variables that reflect their economic development and population heath status.
- Assembled a team of 4 as a leader who signed up for the contest, planed a project schedule, assigning roles and responsibilities, provided guidance to the team members and organised multiple online meetings while also responsible for performing all the data analyses and revising the project report.

Research Student at the BHK Lab **May-August, 2022**

- Researched synergy scoring models for quantifying interactive effects of drug pairs.
- Implemented synergy scoring models for PharmacoGx.

Course Projects

Game Centre - *An Android application containing three classic games available.*

- Developed with the basic software design principle (SOLID) in mind, and employed various design patterns such as adapter and abstract factory to enhance code reusability and extendibility.
- Implemented three mobile games, a user profile management system, a score-based and a time-based player ranking system.

Analysis of Toronto Break and Enter Crime Data in 2019

- A statistical study to discover factors behind B&E crime occurrences in Toronto

- Applied various hypothesis testings and supervised machine learning methods and translated the result into readable visual representation.

scRGNet - *An R implementation of the scGNN's feature autoencoder using torch for R*

- Encapsulates the feature autoencoder module of the single cell graph neural network (scGNN) framework into an R package.

FeelsAmazingMan

- Combining static and contextual representations to improve Twitch sentiment analysis

- Performed emote-based augmentation on the labeled training data.
- Explored the generalisation performance of fine-tuned BERT for downstream classification tasks under various settings of freezing Transformer encoders and re-initialising pretrained weights.
- Developed setup for integrating BERT pooler representations and Word2Vec embeddings using a GRU cell in a sentiment classifier for fine-tuning BERT.