

Mehdi FOROOZANDEH SHAHRAKI

PhD Student | Research Assistant

+1 (778) 228 2027 @ mehdiroozandehsh@gmail.com @ mfa76@sfu.ca 8888 University Dr., Burnaby, B.C. V5A 1S6
i GitHub i Google Scholar i LinkedIn i ResearchGate Personal Website

I am a PhD student in the Computing Science department at SFU. My research interests are machine learning, probabilistic modelling, and genomics. I am a graduate research assistant in the Computational Biology lab supervised by Dr. Maxwell Libbrecht.

EDUCATION

- 2022 - 2026 **Ph.D. Computing Science**, Simon Fraser University
2020 - 2022 **MSc. Computing Science**, Simon Fraser University, *CGPA : 4.0 / 4.33*,
Courses : Machine Learning, Statistical Machine Learning, Design and Analysis of Algorithms, Problem-based Learning in Bioinformatics | *Thesis Title : Evaluating reproducibility of segmentation and genome annotation (SAGA) algorithms*
2015 - 2019 **BSc. Cell and Molecular Biology- Microbiology**, University of Tehran, *GPA (Final Two Years) : 3.68 / 4*,
Selected Courses : Biostatistics, Molecular Genetics, Epigenetics, Virology, Evolutionary Biology, Developmental Biology, Animal Cell Differentiation

EXPERIENCE

- September 2020
Present **Research Assistant, SIMON FRASER UNIVERSITY, Department of Computing Sciences**
➤ At Libbrecht Computational Biology Lab
➤ Research on Human Genome Annotation using epigenomic data working on the evaluation of SAGA methods such as Segway and ChromHMM and the reproducibility of their annotations. In this project, we are aiming to get robust annotations of the human genome.
➤ Research on Drug Resistance Prediction in Bacteria Using Machine Learning. Particularly, I helped to create and optimize deep learning models capable of predicting bacterial antibiotic resistance from genome data.
Machine Learning Probabilistic Modeling Genomics Epigenomics Deep Learning Drug Resistance
- January 2019
September 2020 **Research Assistant, UNIVERSITY OF TEHRAN, IBB- Department of Bioinformatics**
➤ At Complex Biological Systems and Bioinformatics (CBB) lab, led by Dr. Kavousi.
➤ Research on developing machine learning and deep learning models for predicting enzyme properties.
➤ Research on the discovery of novel enzymes with particular applications from metagenomic sources.
➤ Collaboration in writing and editing academic publications.
Bioinformatics Machine Learning Metagenomics Enzymes

PUBLICATIONS

- Foroozandeh Shahraki, Mehdi, *et al.* "A computational learning paradigm to targeted discovery of biocatalysts from metagenomic data : a case study of lipase identification" *Biotechnology and Bioengineering* :21-130.R1 (2022).
Foroozandeh Shahraki, Mehdi, *et al.* "A generalized machine-learning aided method for targeted identification of industrial enzymes from metagenome : A xylanase temperature dependence case study." *Biotechnology and Bioengineering* (2020).
Foroozandeh Shahraki, Mehdi, *et al.* "MCIC : Automated Identification of Cellulases From Metagenomic Data and Characterization Based on Temperature and pH Dependence." *Frontiers in microbiology* 11 (2020).
Maleki, Morteza, *et al.* "A novel thermostable cellulase cocktail enhances lignocellulosic bioconversion and biorefining in a broad range of pH." *International journal of biological macromolecules* 154 (2020) : 349-360.
Ariaeenejad, Shohreh, *et al.* "A novel high performance in-silico screened metagenome-derived alkali-thermostable endo- β -1, 4-glucanase for lignocellulosic biomass hydrolysis in the harsh conditions." *BMC biotechnology* 20.1 (2020) : 1-13.
Foroozandeh Shahraki, Mehdi, *et al.* "Evaluating the reproducibility of segmentation and genome annotation (SAGA) algorithms" Poster at Machine Learning in Computational Biology 2021 (full submission in progress).

PROJECTS

REPRODUCIBILITY OF SAGA ALGORITHMS

2021 - 2022

 github.com/mehdiforoozandeh/SAGAconf  [Video Summary](#)

Research on evaluating the reproducibility of predictions from a well-known class of probabilistic graphical models known as SAGA algorithms. These models, that are usually based on Hidden Markov Models or Dynamic Bayesian Networks, are essentially clustering algorithms for sequence data (genomic and epigenomic data). Similar to most unsupervised learning algorithms, evaluation of their prediction performance is challenging. In this research project, I use statistical principles of reproducibility analysis to evaluate these probabilistic machine learning algorithms.

probabilistic modeling machine learning reproducibility epigenomics genomics

DRUG RESISTANCE PREDICTION USING MACHINE LEARNING

2020 - 2021

 github.com/mehdiforoozandeh/DRML

In a series of related projects, I designed and optimized several machine learning and deep learning frameworks to understand and predict anti-microbial drug resistance from genomic data. In this project, I explored various avenues such as customized feature extraction from high dimensional genomic data, comparing several machine learning algorithms for the same prediction problem, and most importantly, implementing Bayesian hyper-parameter optimization to optimize our models more effectively. Particularly, I applied Bayesian hyper-parameter optimization to optimize a massive deep learning model that could not possibly be optimized using conventional grid-searching method.

Deep Learning Bayesian Optimization Drug Resistance

PREDICTION OF ENZYMATIC PROPERTIES USING MACHINE LEARNING AND ENZYME MINING FROM METAGENOME

2019-2020

 github.com/mehdiforoozandeh/MeTarEnz  github.com/mehdiforoozandeh/MCIC

Research on developing machine learning and deep learning models for predicting enzyme properties heavily relying on amino acid sequence data for model training and feature extraction. I also designed and developed automated enzyme mining software to facilitate discovery of biocatalysts from metagenome data. Finally, using the developed computational infrastructures, we discovered and mined various novel enzymes with particular applications from metagenomic sources.

Enzyme mining Metagenomics Machine Learning Automation

SKILLS

Programming Python, MATLAB

Tools and Software TensorFlow, Keras, Scikit-learn, Scikit-optimize, Numpy, Pandas, Biopython, NCBI Toolkit, Segway, ChromHMM

Other Unix Shell Scripting, \LaTeX , OOP, Linux, Web Scraping

Laboratory Skills Microbial Culture, PCR, DNA extraction, Gene Cloning, Gel Electrophoresis

Languages English, Persian (Native)

CERTIFICATIONS

2019 Machine Learning by Stanford University on Coursera (Grade : 100/100)

2019 Deep Learning with TensorFlow at Sharif University of Technology

2019 Molecular Cloning Practical Workshop, University of Tehran

2018 International Summer School of Brain and Cognitive Sciences at Royan Institute

2017 PCR, Electrophoresis and Drawing of Phylogenetic Tree Practical Workshop, University of Tehran

REFERENCES

Maxwell Libbrecht
Assistant Professor, SFU
@ maxwl@sfu.ca
☎ (+1) 778-782-3045

Kaveh Kavousi
Assistant Professor, UT
@ kkavousi@ut.ac.ir
☎ (+98)-21-8899 4062

Shohreh Ariaeenejad
Assistant Professor, ABRII
@ Sh.ariaee@abrii.ac.ir
☎ (+98)-263 270 3536