

Tian Bai, APPLICANT FOR A PHD IN STATISTICS

Undergraduate student at McGill University Montreal, Canada
tian.bai3@mail.mcgill.ca — +1 (514) 441-2550 — LinkedIn — Github — Personal Website

EDUCATION

McGill University, Montreal, Canada 09.2021 — 12.2024 (exp.)
Bachelor of Science in Honours Mathematics and Computer Science Cumulative GPA: 4.00/4.00
Minor in Statistics

RESEARCH EXPERIENCE

Remote Research Collaboration Remote
Research Assistant 08.2024 — Present

- Developed OptCS, a broad extension of conformal selection that enabled simultaneous model selection and candidate selection by addressing the selection bias, achieving enhanced statistical power
- Proved general statistical theory for distribution-free FDR control under data-driven score optimization
- Conducted extensive simulations to demonstrate the superior performance of the purpose methods
- Applied these innovations to challenges in drug discovery and large language model (LLM) alignment
- Supervisor: Dr. Ying Jin (*Incoming Professor of Statistics at the University of Pennsylvania*)

McGill University Montreal, Canada
Research Assistant 05.2024 — Present

- Applied model-free selection technique based on conformal inference to drug research and development
- Collaborated with statisticians and scientists at Merck to innovate old methods lacking statistical assurance
- Purposed and proved generalizations of the conformal selection method to multivariate targets
- Published a preprint summarizing our findings and presented at McGill Undergraduate Research Conference
- Earned the Science Undergraduate Research Award (SURA)
- Supervisor: Prof. Archer Yi Yang

McGill University & MILA Montreal, Canada
Research Assistant 01.2024 — 08.2024

- Implemented several graph-based machine learning models (GAT, KCN) for single-cell velocity analysis
- Enhanced expertise in model implementation and gained experience in biomedical data analysis
- Supervisor: Prof. Jun Ding, Prof. Archer Yi Yang

Lady Davis Institute for Medical Research & McGill University Montreal, Canada
Research Assistant 09.2023 — 05.2024

- Developed multiple machine learning models for triage acuity prediction in emergency departments
- Achieved superior prediction accuracy versus the traditional method with Canadian Triage Acuity Scale (CTAS)
- Improved model interpretability by evaluating predictor significance using the Shapley score
- Co-authored a research paper published in the Canadian Journal of Emergency Medicine
- Supervisor: Prof. Lars Grant

McGill University Montreal, Canada
Research Assistant 11.2022 — 04.2023

- Designed and developed a software in C for low-latency distance approximation and presented a live demo
- Acquired hand-on experience working with microcontrollers (ESP32), Raspberry Pi and low-level C programming
- Integrated accelerometers, magnetometers, and gyroscopes with ESP32 for future applications
- Supervisor: Prof. Muthucumar Maheswaran

WORK EXPERIENCE

Voffice, Inc. Longueuil, Canada
Software Developer Intern 05.2023 — 08.2023

- Developed digital business management platforms using Microsoft Business Central and Microsoft Azure
- Achieved expertise in Microsoft AL and C# through hands-on development and implementation
- Partnered with functional consultants to deliver custom solutions with seamless integration
- Conducted peer code reviews to ensure high-quality standards and maintain best practices
- Managed source code repositories and workflows using Git, Git Bash, and Microsoft Azure DevOps
- Supervisor: Yisheng Cao

PUBLICATIONS

Journal Paper

- Grant L, Diagne M, Aroutiunian R, Hopkins D, Bai T, Kondrup F, Clark G. Machine learning outperforms the Canadian Triage and Acuity Scale (CTAS) in predicting need for early critical care. *Canadian Journal of Emergency Medicine* 2024. <https://doi.org/10.1007/s43678-024-00807-z>

Preprint

- Bai T, Tang P, Xu Y, Svetnik V, Khalili A, Yu X, Yang A. Conformal Selection for Efficient and Accurate Compound Screening in Drug Discovery. *ChemRxiv* 2024. <https://doi.org/10.26434/chemrxiv-2024-pf3ph>
- Bai T, Jin Y. Optimized Conformal Selection: Powerful Selective Inference After Conformity Score Optimization. *ArXiv* 2024. <https://arxiv.org/abs/2411.17983>

In Progress

- Multivariate Conformal Selection. *TBD*

PRESENTATIONS

- Model-Free Selection Inference for Drug Discovery via Conformal Prediction. *McGill 7th Undergraduate Research Conference, August 2024.*

SELECTED ADVANCED COURSES

- MATH 680 Computation Intensive Statistics*
- COMP 551 Applied Machine Learning*
- COMP 579 Reinforcement Learning*
- MATH 523 Generalized Linear Models*
- MATH 423 Applied Regression
- MATH 454 Honours Analysis 3 (Measure Theory)
- MATH 455 Honours Analysis 4 (Functional Analysis)
- MATH 457 Honours Algebra 4 (Galois Theory)
- MATH 447 Introduction to Stochastic Processes
- COMP 421 Database Systems

*Graduate Courses. Click on the courses for more details.

AWARDS

Science Undergraduate Research Award (SURA) <i>15 weeks of full-time research under the supervision of a professor, with financial support</i> Amount: \$8700	Montreal, Canada 03.2024
Tomlinson Engagement Award for Mentoring <i>For the Helpdesk Tutor positions</i> Amount: \$300	Montreal, Canada 12.2023, 03.2024
Sir Edward Beatty Memorial Scholarships in Mathematics <i>For high academic standing</i> Amount: \$2100	Montreal, Canada 09.2024
A. D. Pelletier Scholarship in Mathematics and Statistics <i>For high academic standing</i> Amount: \$4400	Montreal, Canada 09.2023
Robert W Wilson Scholarship <i>For high academic standing</i> Amount: \$2500	Montreal, Canada 09.2022
Dean's Honour List <i>Top 10% of continuing students</i>	Montreal, Canada 09.2022, 09.2023

CERTIFICATES

Microsoft Certified: Azure Data Engineer Associate	06.2023
Microsoft Certified: Azure Developer Associate	08.2023

OTHER EXPERIENCES

McGill University Montreal, Canada
CSUS (Computer Science) Helpdesk Tutor 01.2024 — 04.2024

- Helped students tackle their assessments and understand coursework through effective teaching techniques

McGill University Montreal, Canada
Math Helpdesk Tutor 09.2023 — 12.2023

- Provided intuitive explanations and active assistance on mathematics problems

McGill University Montreal, Canada
Teaching Assistant 09.2023 — 04.2024

- (Fall 2023) MATH 350 Honours Discrete Mathematics taught by Prof. Sergey Norin
- (Winter 2024) MATH 457 Honours Algebra 4 taught by Prof. Henri Darmon

McGill University Montreal, Canada
Note-taker 09.2023 — 04.2024

- Voluntary note taking and sharing for students with disabilities at McGill University
- (Fall 2023) COMP 302 Programming Language and Paradigms
- (Winter 2024) COMP 579 Reinforcement Learning and COMP 421 Database Systems

EXTRACURRICULAR ACTIVITIES

Competitive Programming

- Contestant, Compete McGill (competitive programming club at McGill). Team ranked 5th in the 2023 NAQ contest.
- Presentation on Mo's algorithm.
- Presentation on Advanced Dynamic Programming techniques.

SKILLS

- **Programming:** Python (Pytorch, Scikit-learn), R, C/C++, SQL
- **Software/Tool:** ~~La~~TeX, Git, Powershell, Microsoft Azure, Rstudio

LANGUAGES

- **English** TOEFL 112 (R 30, L 29, S 27, W 26); GRE 331 (V 161, Q 170, AWA 4.5)
- **Chinese (Mandarin)** Native