CANSSI Distinguished Postdoctoral Fellows Projects List

Project Name: *Statistical foundations of invariant representation learning, with applications to data-driven compression and algorithm design*

Supervisor:

- Name: Benjamin Bloem-Reddy
- Affiliation: Assistant Professor, Dept. of Statistics, University of British Columbia
- Email: benbr@stat.ubc.ca
- Website: https://www.stat.ubc.ca/~benbr/

Co-Supervisor:

- Name: Chris J. Maddison
- Affiliation: Assistant Professor, Depts. of Computer Science and, Statistical Sciences, University of Toronto
- Email: hkj@yorku.ca
- Website: http://www.yorku.ca/~hkj/

Mentor/Collaborator

- Name: James V. Zidek
- Affiliation: Professor Emeritus, Department of Statistics, University of British Columbia
- Email: jim@stat.ubc.ca
- Website: https://www.stat.ubc.ca/~jim/

Location/University:

- University of British Columbia
- University of Toronto

Abstract:

Although advances in deep learning methods have been driven largely by computation and engineering, some of the most successful methods incorporate an appropriate notion of invariance and/or sufficiency. This project will establish the statistical theory for such methods, elucidating the respective roles of statistical invariance and sufficiency in model deep learning, and suitable approximate versions learned from data. The DPF will collaborate with the Maddison group to develop novel methods based on these ideas for data compression and data-driven algorithm design (e.g., discrete optimization), and develop specialized statistical theory for those settings.

Interdisciplinary/applied experience:

In addition to the theoretical objectives of the research project, the DPF will collaborate with Professor Maddison and his research group to develop theoretically sound representation learning methods for data compression and data-driven algorithm design. Such methods are at the forefront of research into data driven solutions to core computer science problems and pose challenging methodological and theoretical questions. Successful development of new methods will require designing appropriate deep

neural network models and optimization objectives, implementing them to be trained efficiently on massive data sets, and running reproducible large-scale computer experiments to assess empirical performance. Moreover, collaborating on the release of clean, open-source code will be part of the project. Overall, the methodological component will complement the theoretical component, and generate tangible, salient theoretical questions.

Prof. Maddison's research group develops cutting-edge machine learning methods in a number of areas, including data-driven approaches to compression and algorithm design. Two PhD students will be involved as potential collaborators with the DPF: one who is working on principled representation learning methods and data compression, and another who is working on representation learning for reasoning and combinatorial optimization (data-driven algorithm design). The DPF will benefit from the group's expertise, access to computational resources, and massive data sets.

The collaboration will occur throughout the term of the fellowship, via regular virtual meetings between the DPF, supervisor, co-supervisor, and PhD student collaborators; and with 2 research visits of 1-2 weeks per year by the DPF to Prof. Maddison's group in Toronto. In order to start the collaboration, the first of those will be early in the DPF's first academic term. Prof. Maddison will also visit UBC once per year for collaborative work.

Teaching/training/education:

Teaching - In order to facilitate the DPF's development as a teacher and research mentor, they will have the opportunity to teach a half-term graduate-level topics course in Year 1 and a full-term undergraduate course in Year 2, both in the Department of Statistics at UBC. They will be supported in these endeavors by participating in the UBC Centre for Teaching and Learning's postdoc-specific Instructional Skills Workshop in Year 1,1 and other teaching-related training as the DPF elects. The DPF will work with the supervisor and mentor to design the topics course, with the goal of integrating their teaching and research. Moreover, the full-term undergraduate course will be pair-taught with an experienced instructor. Although the paired teaching program will be informal (conducted entirely within the Statistics Department), it will be modeled after the successful Teaching Startup2 program for new faculty in UBC Faculty of Science.
This program includes targeted feedback and regular reflection on teaching practices, in addition to learning about educational technology and research-based teaching techniques.

Mentoring of research students - The DPF will work as part of the supervisor's research group, collaborating with 1-2 PhD students with related interests, and training to supervise student research. This will include jointly supervising a student research project with Prof. Bloem-Reddy, taking on more responsibility over time. The DPF will also regularly observe meetings between Prof. Bloem-Reddy and research students and reflect with Prof. Bloem-Reddy on the meeting afterward. This will help the DPF develop a sense of how to give effective guidance while offering opportunities for student growth and prepare the DPF for their future career.

CANSSI Distinguished Postdoctoral Fellows Projects List

Mentoring of the postdoctoral fellow:
Regular meetings between the DPF, supervisors, and mentor will support the DPF's research progress. After the initial few meetings, the DPF will set the agenda, sharing progress and identifying specific areas needing attention. Monthly "big-picture" lunches with the supervisor and mentor will ensure that the DPF's progress towards their fellowship and post-fellowship goals is on track.

Communication - The DPF will be encouraged to present their work at workshops and conferences, with practice and feedback sessions ahead of time. The supervisor's group also has regular meetings at which members present informal research updates, providing additional opportunities for practice and feedback. The DPF will collaborate on all aspects of paper-writing and, if it aligns with their career goals, they may have the opportunity to participate in the writing of a grant proposal.

Professional development - Each term the DPF will be encouraged and supported to participate in 1-2 professional development workshops/seminars offered by the UBC Postdoctoral Fellows Office (PDFO). This will include, in the Fall Term of Year 1, participating in an EDI training. The UBC PDFO also has access to programs through Mitacs and the National Center for Faculty Development and Diversity, which host a wide range of programs on topics such as stress management, project management, networking, and career planning.

Enrichment and opportunities - UBC and UoT have world-class environments for machine learning research. Through each university's AI institute (Vector Institute at Toronto and CAIDA at UBC), the DPF will have access to seminars, training, networking opportunities, collaborations, and a peer support network. Moreover, the DPF will be supported to attend conferences and workshops, with the goal of presenting their work starting in summer of Year 1.

Proposed schedule:

	Research	Teaching	Development
Year 1			
Sept.-Dec.	Begin research project Begin method. collab. Visit Maddison group Attend conference/workshop	Instructional Skills Workshop	EDI training
Jan.-Apr.	Maddison visits UBC	Teach grad topics course	Elective PD training
May-Aug.	Visit Maddison group Present work at conference		Career planning training
Year 2			
Sept.-Dec.	Visit Maddison group Present work at conference	Elective teaching training	
Jan.-Apr.	Maddison visits UBC	Pair-teach UG course	Elective PD training
May-Aug.	Visit Maddison group Present work at conference		

CANSSI Distinguished Postdoctoral Fellows Projects List

Qualifications:

- PhD or equivalent in Statistics, Computer Science, or a related field such as Mathematics or Physics.
- An excellent publication record along with effective teaching and supervision skills.
- Expertise in a core area of theoretical statistics, such as statistical learning theory, functional analysis, or information theory. Experience in algorithm analysis and design will be considered valuable.
- Strong programming skills.
- Willingness to learn.