



Danica J. Sutherland

Version as of May 8, 2025.

1. ACADEMIC POSITIONS AND EDUCATION

Assistant Professor, Computer Science	University of British Columbia	2020 –
Canada CIFAR AI Chair	Alberta Machine Intelligence Institute	2021 –
Research Assistant Professor	Toyota Technological Institute at Chicago	2019 – 2020
Research Associate	University College London	2016 – 2019
Ph.D., Computer Science	Carnegie Mellon University	2011 – 2016
B.A., Computer Science	Swarthmore College	2007 – 2011

2. TEACHING

(a) *Courses Taught at UBC*

CPSC 532D: Modern Statistical Learning Theory is a graduate course giving students the broad background in learning theory required for mathematical understanding of machine learning theory.

CPSC 440/550: Advanced Machine Learning is a second course in machine learning, covering a variety of topics with a probabilistic bent, intended for advanced undergraduates or beginning graduate students.

CPSC 340: Machine Learning and Data Science is an introductory course in machine learning.

Session	Course Number	Sched. Hours	Class Size	Hours Taught	
				Lectures	Other
24-25 W2	CPSC 440+550	3	67+22	3/week	1h/week office hours
24-25 W1	CPSC 532D	3	9	3/week	2h/week office hours
23-24 W2	CPSC 440+550	3	96+16	3/week	1h/week office hours
23-24 W1	CPSC 532D	3	15	3/week	2h/week office hours
22-23 W2	CPSC 440+540	3	69+30	3/week	1h/week office hours
22-23 W1	CPSC 532D	3	22	3/week	1h/week office hours
21-22 W2	CPSC 532S	3	23	3/week	2h/week office hours
21-22 W1	CPSC 340	3	89+139	3/week	1h/week office hours

CPSC 440 and 550 (previously 540) are undergraduate and graduate cross-listings of one course; sizes show the enrolment in each listing.

CPSC 532D was formerly known as 532S.

CPSC 340 in 2021-22 W1 was run jointly with another section taught by Mike Gelbart with an enrolment of 139, for a total of 228 students across the two sections.

(b) *Graduate Research Supervision*

Postdoc Supervision

Student	Program	Start	Finish	Supervisory Role
Bingshan Hu	Postdoc, UBC CS	2023/11	ongoing	Primary supervisor

Bingshan is co-supervised with Mathias Lécuyer and Trevor Campbell; I provide the majority of her funding.

PhD Supervision

Student	Program	Start	Finish	Supervisory Role
Nathaniel Xu	PhD, UBC CS	2023/09	ongoing	Supervisor
Zheng He	PhD, UBC CS	2023/09	ongoing	Supervisor
Hamed Shirzad	PhD, UBC CS	2021/09	ongoing	Supervisor
Wonho Bae	PhD, UBC CS	2020/09	ongoing	Supervisor
Yi Joshua Ren	PhD, UBC CS	2020/09	ongoing	Supervisor

Wonho is on track to graduate in July, and has an accepted a job offer at Apple.

MSc Thesis Supervision

Student	Program	Start	Finish	Supervisory Role
Mohamad Amin Mohamadi	MSc, UBC CS	2021/05 ^S	2023/04	Supervisor
Namrata Deka	MSc, UBC CS	2020/09	2023/04	Supervisor
Milad Jalali	MSc, UBC CS	2020/09	2023/08	Supervisor

Mohamad Amin and Namrata, after graduating, continued for some time in non-student research positions (Mohamad Amin for eight months, Namrata for three); they then began PhDs at TTI-Chicago and Carnegie Mellon, respectively. Milad moved on to a job at Vancouver biomedical imaging company Prenuvo.

PhD Supervisory Committee Memberships

Student	Program	Start	Finish	Committee Type
Chen Fan	PhD, UBC CS	2025/03	ongoing	PhD committee member
Vasileios Lioutas	PhD, UBC CS	2024/09	2025/04	PhD committee member
Tina Behnia	PhD, UBC ECE	2024/08	ongoing	PhD committee member
Victor Portella	PhD, UBC CS	2023/01	2024/07	PhD committee member

Research Proficiency Exam Committee Memberships

Student	Program	Start	Finish	Committee Type
Seungyeon Baek	PhD, UBC CS	2024/05	2024/07	RPE committee member
Matthew Niedoba	PhD, UBC CS	2024/01	2024/02	RPE committee member
Chen Fan	PhD, UBC CS	2023/05	2023/07	RPE committee member
Saiyue Lyu	PhD, UBC CS	2023/05	2023/07	RPE committee member
Qiaoyue Tang	PhD, UBC CS	2023/04	2023/06	RPE committee member
Amrutha Ramesh	PhD, UBC CS	2022/09	2022/11	RPE committee member

The RPE is a step in the degree for PhD students at UBC CS roughly corresponding to a qualifying exam.

MSc Committee Memberships

Student	Program	Start	Finish	Committee Type
Shadab Shaikh	MSc, UBC CS	2024/03	2024/09	MSc committee member
Yilin (Justin) Yang	MSc, UBC CS	2024/02	2024/04	MSc committee member
Melika Farahani	MSc, UBC CS	2022/06	2022/08	MSc committee member

(c) Graduate Program Supervision

Student	Program	Start	Finish	Supervisory Role
Anubhav Garg	MSc (essay), UBC CS	2022/10 ^S	2024/04	Supervisor
Arsh Jhah	MSc (essay), UBC CS	2022/06 ^S	2023/10	Supervisor

Arsh and Anubhav both came to me as essay-option students, switching from other advisors, rather than switching from a research program under my supervision.

(d) *Continuing Education Activities*

2023 Instructor, Lifting Inference with Kernel Embeddings (LIKE23) (at the University of Bern). *(Deep) Kernel Mean Embeddings for Representing and Learning on Distributions*. Seventy-five minute lecture, about 40 participants.

Instructor, Winter School on Deep Learning (at the Indian Statistical Institute).^ν *Neural Tangent Kernels, Finite and Infinite*. Two hour course, about 300 participants (not all present live).

2022 Instructor, Research School on Uncertainty in Scientific Computing, Corsica (ETICS).^{\$} *Modern Kernel Methods in Machine Learning*. Six hour course, about 55 students.

2021 Instructor, Data Science Summer School (DS3) (at École Polytechnique, Paris).^{\$ν} *Kernel Methods: From Basics to Modern Applications*. Three hour course, repeated twice, about 12 students each.

2019 Tutorial, Neural Information Processing Systems (NeurIPS).^{\$C} *Interpretable Comparison of Distributions and Models*. With Arthur Gretton and Wittawat Jitkrittum. Two hour tutorial; thousands of attendees.

Instructor, Data Science Summer School (DS3) (at École Polytechnique, Paris).^{\$} *Learning with Positive Definite Kernels: Theory, Algorithms and Applications*. With Bharath Sriperumbudur. Six hour course, repeated twice, about 25 students each.

Instructor, Machine Learning Crash Course (MLCC) (at the University of Genoa).^{\$} *Introduction to Generative Adversarial Networks*. 1.5 hour course, about 100 students.

(e) *Other*

Undergraduate Research Supervision

Student	Program	Start	Finish	Supervisory Role
Aaron Wei	CPSC 449	2024/09	2025/04	Honours thesis supervisor
Achinth Bharadwaj	Science Undergraduate Research Experience	2023/06	2023/08	Summer research supervisor

Achinth worked on a research project titled “Demographic representativeness via kernel distribution distances”; he went on to a job at Palantir and then a graduate research program at Virginia Tech. Aaron’s thesis focused on the theoretical foundations of statistical measures of conditional independence; he will continue this work with me over this summer.

3. SCHOLARLY AND PROFESSIONAL ACTIVITIES

(a) *Research or equivalent grants*

Agency, Program	Title	\$/year	Year(s)	PI (Co-PIs in italics)
Google, unrestricted gift	Conditional independence measures for fairness and representation learning ^C	\$41,000	2024	D. J. Sutherland
DRAC, Resources for Research Groups	Empirically Understanding Deep Learning with Theory ^C	\$24,449 [≈]	2024–25	D. J. Sutherland

(continued...)

^ν Presented virtually. ^{\$} Expenses paid. ^C Competitively selected.

Agency, Program	Title	\$/year	Year(s)	PI (Co-PIs in <i>italics</i>)
NSERC, CREATE	Advanced Machine Learning Training Net-work ^C	\$1,075,650*	2023–29	F. Wood; A. Bouchard-Côté; B. Bloem-Reddy; T. Campbell; K. Leyton-Brown; M. van de Panne; D. Poole; M. Schmidt; L. Sigal; D. J. Sutherland; K. M. Yi
NSERC + FRQNT, NOVA	Towards Improved Understanding and Efficient Utilization of Depthwise Computation in Modern Neural Networks ^C	\$75,000	2023–26	G. Wolf; E. Belilovsky; D. J. Sutherland; J. Tian
UBC Data Science Institute, Postdoctoral Matching Fund	Anytime-Valid PAC-Bayes for Industrial Applications ^C	\$40,000	2023–24	T. Campbell; D. J. Sutherland
DRAC, Resources for Research Groups	Empirically Understanding Deep Learning with Theory ^C	\$28,737 [≈]	2023–24	D. J. Sutherland
CIFAR	Canada CIFAR AI Chair, Amii ^C	\$78,333	2021–26	D. J. Sutherland
NSERC Discovery	Learning Better Representations with Kernels ^C	\$29,000	2021–26	D. J. Sutherland
NSERC Discovery	Launch Supplement	\$12,500	2021	D. J. Sutherland
UBC	Startup Fund	\$200,000	2020	D. J. Sutherland

(b) *Invited Presentations*

Invited tutorials and research school lectures are listed instead in section 2(d).

Feb 2025 Snowflake.^ℒ *Data-efficient learning, in general and in LLM preference tuning.*

Dec 2024 Mathematics of Machine Learning (workshop at the Canadian Mathematical Society Winter Meeting, CMS).^N *Expander Graphs and Low-Distortion Embeddings for Learning on Graphs.*

Nov 2024 Learning on Graphs Vancouver Meetup (LoG-Vancouver).^L *Scaling Graph Transformers with Sparse and Sparsified Attention.*

Jun 2024 Simon Fraser University, AI seminar.^L *Scaling Graph Transformers with Expander Graphs.*

Feb 2024 Statistical Aspects of Trustworthy Machine Learning (workshop at the Banff International Research Station).^N *Conditional independence measures for fairer, more reliable models.*

Jun 2023 Lifting Inference with Kernel Embeddings (LIKE23) (workshop at the University of Bern).^ℒ *Learning conditionally independent representations with kernel regularizers.*

Jun 2023 Gatsby25.^ℒ *Learning conditionally independent representations with kernel regularizers.*

Apr 2023 Pacific Conference on Artificial Intelligence (PCA).^L *Are these datasets different? Two-sample testing for data scientists.*

Mar 2023 AI Seminar, University of Michigan.^ℒ *A Defense of (Empirical) Neural Tangent Kernels.*

Mar 2023 Robotics DEI Seminar, University of Michigan.^{ℒ\$} *Post-Publication Name Change Policies, Why they Matter, and Whether they Work.*

Mar 2023 Microsoft Research Montréal.^{N^ν} *In Defence of (Empirical) Neural Tangent Kernels.*

* Includes contributions from various UBC departments and partners; NSERC amount is on average \$275,000/year. [≈] Estimated equivalent cost for computational resources. ^ℒ International venue. ^N National venue. ^L Local venue. ^ν Presented virtually.

\$ Expenses paid.

- Nov 2022 Queer in AI Workshop (workshop at Neural Information Processing Systems, NeurIPS).^ℳ *Name Change Policies: A Brief (Personal) Tour.*
- Apr 2022 Toronto Womxn in Data Science Conference.^{Nν} *Are These Datasets The Same? Learning Kernels for Efficient and Fair Two-sample Tests.*
- Feb 2022 TrustML Young Scientist Seminars.^{ℳ\$ν} *Are These Datasets The Same? Learning Kernels for Efficient and Fair Two-sample Tests.*
- Jan 2022 Lifting Inference with Kernel Embeddings (LIKE22) (workshop at the University of Bern).^{ℳν} *Better deep learning (sometimes) by learning kernel mean embeddings.*
- Nov 2021 NYU Center for Data Science Lunch Seminar Series.^{Nν} *Can Uniform Convergence Explain Interpolation Learning?*
- Jan 2021 Kickoff Workshop (workshop at the Pacific Interdisciplinary Hub on Optimal Transport).^{Lν} *Deep kernel-based distances between distributions.*
- Oct 2020 Penn State, Statistics colloquium.^{Nν} *Can Uniform Convergence Explain Interpolation Learning?*
- Oct 2019 University of Massachusetts Amherst, College of Information and Computer Sciences.^{N\$} *Better GANs by Using Kernels.*
- Jul 2019 Distance Metrics and Mass Transfer Between High Dimensional Point Clouds (workshop at the International Congress on Industrial and Applied Mathematics, ICIAM).^ℳ *Kernel distances between distributions for generative models.*
- Sep 2018 Advances in Kernel Methods (workshop at the Gaussian Process Summer School, GPSS).^{N\$} *Kernel Distances for Better Deep Generative Models.*
- Jun 2018 Facebook AI Research New York.^ℳ *Better GANs by using the MMD.*
- Jun 2018 Gatsby Tri-Center Meeting.^{ℳ\$} *Efficiently Estimating Densities and Scores with Kernel Exponential Families.*
- Jun 2018 Machine Learning reading group, Google New York.^ℳ *Better GANs by using the MMD.*
- Jun 2018 Machine Learning reading group, Columbia University.^ℳ *Better GANs by using the MMD.*
- May 2018 Machine Learning Seminar, University of Sheffield.^{N\$} *Advances in GANs based on the MMD.*
- Dec 2017 Approximating high dimensional functions (workshop at the Alan Turing Institute).^L *Efficient and principled score estimation with kernel exponential families.*
- Dec 2017 Computational Statistics and Machine Learning seminar, University College London.^L *Efficient and principled score estimation with kernel exponential families.*
- Aug 2017 Implicit Models (workshop at the International Conference on Machine Learning, ICML).^ℳ *Evaluating and Training Implicit Generative Models with Two-Sample Tests.*
- Apr 2017 Theory of Generative Adversarial Networks (workshop at Data Analysis, Learning, and Inference, DALI).^ℳ *Two-Sample Tests, Integral Probability Metrics, and GAN Objectives.*
- Feb 2017 Computational Statistics and Machine Learning seminar, Oxford University.^{N\$} *Generative Models and Model Criticism via Optimized Maximum Mean Discrepancy.*

^ℳ International venue. ^{\$} Expenses paid. ^N National venue. ^ν Presented virtually. ^L Local venue.

(c) *Other Presentations*

- 2021 Workshop on Widening NLP (WiNLP) (workshop at the Conference on Empirical Methods in Natural Language Processing, EMNLP). *How to Make Virtual Conferences Queer-Friendly: A Guide*.
- 2019 International Conference on Machine Learning (ICML). *Learning deep kernels for exponential family densities*.
- 2017 Learning on Distributions, Functions, Graphs and Groups (workshop at Neural Information Processing Systems, NeurIPS). *Bayesian Approaches to Distribution Regression*.
- 2018 Artificial Intelligence and Statistics (AISTATS). *Efficient and principled score estimation with Nyström kernel exponential families*.
- 2013 Knowledge Discovery and Data Mining (KDD). *Active learning and search on low-rank matrices*.

(d) *Conference Participation*

Reviewing-based roles are in section 5(c); tutorials are in section 2(d).

- 2020-25 Area Chair, Conference on Artificial Intelligence and Statistics (AISTATS).
- 2024 Panel participant, *Out-of-Distribution Generalization: Shortcuts, Spuriousness, and Stability*, NeurIPS tutorial.
- 2020-22 Area Chair, Conference on Neural Information Processing Systems (NeurIPS).
- 2022 Panel Moderator, *Faculty and Queerness*, Queer in AI NeurIPS workshop.
- 2022 Area Chair, International Conference on Machine Learning (ICML).
- 2022 D&I event organizer (Queer in AI), AAAI Conference on Artificial Intelligence (AAAI).
- 2021-22 Senior Program Committee, AAAI Conference on Artificial Intelligence (AAAI).
- 2018 Panel participant, *Advances in Kernel Methods*, Gaussian Process Summer School workshop.
- 2017 Panel participant, *Implicit Models*, ICML workshop.

4. SERVICE TO THE UNIVERSITY

(a) *Areas of special interest and accomplishments*

Informal mentorship for trans undergrads Since beginning at UBC, I've occasionally met with several trans undergrads in Computer Science and related fields for mentorship, career advice, and community.

Preferred names. Many UBC IT systems ignore preferred names, a significant problem for e.g. trans people who have not or cannot change their legal name in BC. I made a list of over a dozen in 2021-22, some of which I worked directly towards getting fixed, some of which I handed off to the Equity Office or others. Since I changed my legal name in 2022, it's been harder for me to find more instances.

(b) *Memberships on committees*

- 2024–25 Graduate Affairs Committee, Computer Science Department.
- 2023–24 Faculty recruiting committee: Artificial Intelligence Methods for Scientific Impact.
- 2023 Ad-hoc faculty recruiting committee (single candidate).
- 2023–24 AIM-SI Steering Committee (temporary position as a leave replacement).
- 2022–23 Graduate Recruiting and Admissions Committee, Computer Science Department.
- 2021–22 Faculty recruiting committee: CRC Tier 2 in Quantum Computing, Computer Science Department.
- 2021 Graduate Recruiting and Admissions Committee, Computer Science Department.

(c) *Other service*

2023– Recruitment Chair, Advanced Machine Learning Training Network.

2025 Internal review for CFI JELF applications.

2024 Chair of thesis proposal exam: William Harvey, Muchen Li, Wanxin Li.

2023–24 Assistance with the AIM-SI recruiting committee in Statistics, to help work around conflicts of interest.

2023 Chair of thesis proposal exam: Christian Weilbach, Shih-Yang Su.

Chair of thesis defense: Andreas Munk.

UBC Women in Science mentorship program, occasionally meeting first-year undergraduates.

2022 Chair of thesis proposal exam: Vaden Masrani.

5. SERVICE TO THE COMMUNITY

(a) *Memberships in other societies*

2022– Queer in AI, core organizer. (Also various capacities from 2018 to 2021.) My work with this international nonprofit organization has included trans-inclusive advocacy to conferences, particularly making enormous improvements to the way in which NeurIPS, ICML, ICLR, AISTATS, and MLSys handle attendees with professional names different from their legal names; working to improve post-publication name change policies at several venues; co-organizing social, mentoring, and workshop events at conferences; serving on a hiring committee for paid administrator positions; initiating and maintaining an informal program of feedback and mentoring for faculty candidates; and more.

2021– Name Change Policy Working Group, member. This loosely-organized group (no formal positions) advocates with scientific publishers for post-publication name change policies. I gave two invited talks on these issues in 2022-23.

(b) *Memberships on other committees*

2025 Canadian Artificial Intelligence Association Best Thesis Award Committee, member.

2022-23 CIFAR Pan-Canadian AI Strategy's National Program Committee, member.

(c) *Reviewer*

Area chair / meta-reviewer roles are listed under section 3(d).

2024 Conference reviewing: ICML, NeurIPS.

2024 Conference reviewing: ICML, ICLR.

Grant reviewing: CIFAR AI Catalyst grant.

2023 Journal reviewing: JRSS-A.

Conference reviewing: ICLR, NeurIPS.

Grant reviewing: NSERC Discovery Grant.

2022 Journal reviewing: JMLR.

Conference reviewing: UAI, ICML.

Grant reviewing: NSERC Discovery Grant.

2021 Journal reviewing: JMLR, Bernoulli.

Conference reviewing: ICML.

2014–20 Conference reviewing: NeurIPS, ICML, ICLR, AISTATS, AAAI, COLT, SoCG, IJCAI, ECML-PKDD.

Journal reviewing: JMLR, IEEE TSP, IEEE T-PAMI, Bernoulli, MLJ, Comptes rendus.

(d) *External examiner*

Year	Student Name	Institution	Degree
2024	Ankit Vani	Université de Montréal / Mila	Doctoral
2024	Kiarash Zahirnia	Simon Fraser University	Doctoral
2020	Iryna Korshunova	Ghent University	Doctoral
2020	Tong Che	University of Cambridge	Masters

6. AWARDS AND DISTINCTIONS

(a) *Awards for Teaching*

Letters from the Dean of the Faculty of Science for “some of the highest student evaluations in the Faculty of Science” for: CPSC 340 in 2021W1, CPSC 532S in 2021W2, CPSC 532D in 2024W1.

(b) *Awards for Scholarship*

2021 Canada CIFAR AI Chair, Amii. From CIFAR: “The goal of the Chairs Program is to recruit and retain in Canada some of the world’s leading researchers in AI and provide them with long-term, dedicated research funding to support their research programs and help them train the next generation of AI leaders.” Accompanied by funding, listed under grants.

2014–16 Sandia Campus Executive Program. A PhD fellowship, renewed in 2015.

2013 National Science Foundation Graduate Research Fellowship: Honorable mention.

2011 Undergraduate awards: Ivy Award, one of two awards to the graduating class for being “outstanding in leadership, scholarship, and contributions to the college community.” Elected Phi Beta Kappa.

Additionally, the following papers received best paper awards, or competitive presentations selected by area and program chairs at large conferences which might be thought of as comparable to best paper awards at a smaller venue:

2025 [C37]: ICLR outstanding paper: 3 of 11,672 submissions (0.03%)

2023 [C28]: FAccT best paper: 6 of 608 submissions (1%)

[C27]: ICLR notable (top 5%): 91 of 4,955 submissions (2%)

2021 [C18]: NeurIPS oral: 55 of 9,122 submissions (0.6%)

[C15]: AISTATS oral: 48 of 1,527 submissions (3%)

2020 [C14]: NeurIPS spotlight: 386 of 9,467 submissions (4%)

2018 [C9]: AISTATS oral: 31 of 645 submissions (5%)

2013 [C3]: KDD oral: 66 of 726 submissions (9%)

(c) *Awards for Service*

Reviewing awards: NeurIPS 2018 (top 7%) and 2019 (“Top Reviewer”); ICML 2018 (“Outstanding Reviewer”), 2019 (“Best Reviewer”), 2020 (“Top Reviewer”), 2021 (“Best Reviewer”).



Student authors. Authors marked below in bold were my formal supervisees for the work that resulted in the paper in question.

Author order. Machine learning publications typically first list those primarily performing the bulk of the work (usually students and postdocs) in decreasing order of contribution, followed by supervisors in increasing order of contribution and seniority. Shared authorship (usually co-first, but on occasion co-last) is denoted below with *. In my few papers with a large number of co-authors – [C28, J3, N6] – the full author list is suppressed for brevity, but author order (outside of the first) is arbitrary, and none of these papers include my students as co-authors.

Publication venues. The primary venues in machine learning work are conferences, which conduct rigorous review processes with low acceptance rates (which are listed below where known). NeurIPS and ICML are traditionally considered the most prestigious; ICLR, AISTATS, and COLT are also primary venues in my subfield. Although a deeply flawed metric, Google Scholar’s h5-index (as of May 2025) lists NeurIPS as #7, ICLR as #10, and ICML as #17 among all publication venues in *any* field.

Citation metrics. Citation metrics are from Google Scholar as of May 2025. My work has been cited over 5,400 times (last five years: over 4,800), with an *h*-index of 27 (last five years: 26), and 40 papers with at least ten citations (last five years: 37). Highly-cited papers are noted below (thresholds: 10, 50, 100, 500, 1000).

7. REFEREED PUBLICATIONS

(a) *Conference Proceedings*

- [C38] **Wonho Bae**, Danica J. Sutherland*, and Gabriel Oliveira*. “Uncertainty Herding: One Active Learning Method for All Label Budgets.” *International Conference on Learning Representations (ICLR)*. 2025.
AR 3,704/11,672 = 32%
- [C37] **Yi Ren** and Danica J. Sutherland. “Learning Dynamics of LLM Finetuning.” *International Conference on Learning Representations (ICLR)*. 2025. **Outstanding Paper Award**.
AR outstanding paper: 3/11,672 = 0.03%; overall: 3,704/11,672 = 32%
- [C36] **Hamed Shirzad**, Honghao Lin, Balaji Venkatachalam, Ameya Velingker, David P. Woodruff, and Danica J. Sutherland. “Even Sparser Graph Transformers.” *Neural Information Processing Systems (NeurIPS)*. 2024.
AR 4,035/15,671 = 26%
- [C35] **Yi Ren**, Shangmin Guo, Linlu Qiu, Bailin Wang, and Danica J. Sutherland. “Bias Amplification in Language Model Evolution: An Iterated Learning Perspective.” *Neural Information Processing Systems (NeurIPS)*. 2024.
AR 4,035/15,671 = 26%
- [C34] **Wonho Bae**, Jyunhug Noh, and Danica J. Sutherland. “Generalized Coverage for More Robust Low-Budget Active Learning.” *European Conference on Computer Vision (ECCV)*. 2024.
AR 2,395/8,585 = 28%
- [C33] **Wonho Bae**, Jing Wang, and Danica J. Sutherland. “Exploring Active Learning in Meta-Learning: Enhancing Context Set Labeling.” *European Conference on Computer Vision (ECCV)*. 2024.
AR 2,395/8,585 = 28%
- [C32] **Mohamad Amin Mohamadi**, Zhiyuan Li, Lei Wu, and Danica J. Sutherland. “Why Do You Grok? A Theoretical Analysis on Grokking Modular Addition.” *International Conference on Machine Learning (ICML)*. 2024.
AR 2,944/9,653 = 30%
- [C31] **Yi Ren**, Samuel Lavoie, Mikhail Galkin, Danica J. Sutherland, and Aaron Courville. “Improving Compositional Generalization using Iterated Learning and Simplicial Embeddings.” *Neural Information Processing Systems (NeurIPS)*. 2023.
AR 3,218/12,345 = 26% ≥ 10 cites

- [C30] **Hamed Shirzad***, Ameya Velingker*, Balaji Venkatachalam*, Danica J. Sutherland, and Ali Kemal Sinop. “Exphormer: Scaling Graph Transformers with Expander Graphs.” *International Conference on Machine Learning (ICML)*. 2023. AR 1,828/6,538 = 28% ≥ 100 cites
- [C29] **Mohamad Amin Mohamadi, Wonho Bae**, and Danica J. Sutherland. “A Fast, Well-Founded Approximation to the Empirical Neural Tangent Kernel.” *International Conference on Machine Learning (ICML)*. 2023. AR 1,828/6,538 = 28% ≥ 10 cites
- [C28] Organizers of QueerInAI et al. “Queer in AI: A Case Study in Community-Led Participatory AI.” *ACM Conference on Fairness, Accountability, and Transparency (FAccT)*. 2023. **Best Paper award**. AR best paper: 6/608 = 1%; overall: 153/608 = 25% ≥ 10 cites
- [C27] Roman Pogodin*, **Namrata Deka***, Yazhe Li*, Danica J. Sutherland, Victor Veitch, and Arthur Gretton. “Efficient Conditionally Invariant Representation Learning.” *International Conference on Learning Representations (ICLR)*. 2023. **Selected as notable (top 5%), i.e. as an oral**. AR notable (top 5%): 91/4,955 = 2%; overall: 1,575/4,955 = 32% ≥ 10 cites
- [C26] **Yi Ren**, Shangmin Guo, **Wonho Bae**, and Danica J. Sutherland. “How to prepare your task head for finetuning.” *International Conference on Learning Representations (ICLR)*. 2023. AR 1,575/4,955 = 32% ≥ 10 cites
- [C25] **Namrata Deka** and Danica J. Sutherland. “MMD-B-Fair: Learning Fair Representations with Statistical Testing.” *Artificial Intelligence and Statistics (AISTATS)*. 2023. AR 490/1,689 = 29% ≥ 10 cites
- [C24] Lijia Zhou*, Frederic Koehler*, Pragya Sur, Danica J. Sutherland, and Nathan Srebro. “A Non-Asymptotic Moreau Envelope Theory for High-Dimensional Generalized Linear Models.” *Neural Information Processing Systems (NeurIPS)*. 2022. AR 2,671/10,411 = 26% ≥ 10 cites
- [C23] **Mohamad Amin Mohamadi***, **Wonho Bae***, and Danica J. Sutherland. “Making Look-Ahead Active Learning Strategies Feasible with Neural Tangent Kernels.” *Neural Information Processing Systems (NeurIPS)*. 2022. AR 2,671/10,411 = 26% ≥ 10 cites
- [C22] **Hamed Shirzad**, Kaveh Hassani, and Danica J. Sutherland. “Evaluating Graph Generative Models with Contrastively Learned Features.” *Neural Information Processing Systems (NeurIPS)*. 2022. AR 2,671/10,411 = 26% ≥ 10 cites
- [C21] Jinhwan Seo, **Wonho Bae**, Danica J. Sutherland, Jyunhug Noh, and Daijin Kim. “Object Discovery via Contrastive Learning for Weakly Supervised Object Detection.” *European Conference on Computer Vision (ECCV)*. 2022. AR 1,650/5,803 = 28% ≥ 10 cites
- [C20] **Wonho Bae**, Jyunhug Noh, **Milad Jalali Asadabadi**, and Danica J. Sutherland. “One Weird Trick to Improve Your Semi-Weakly Supervised Semantic Segmentation Model.” *International Joint Conference on Artificial Intelligence (IJCAI)*. 2022. AR 680/4,535 = 15%
- [C19] **Yi Ren**, Shangmin Guo, and Danica J. Sutherland. “Better Supervisory Signals by Observing Learning Paths.” *International Conference on Learning Representations (ICLR)*. 2022. AR 1,095/3,328 = 33% ≥ 10 cites
- [C18] Frederic Koehler*, Lijia Zhou*, Danica J. Sutherland, and Nathan Srebro. “Uniform Convergence of Interpolators: Gaussian Width, Norm Bounds and Benign Overfitting.” *Neural Information Processing Systems (NeurIPS)*. 2021. **Selected for oral presentation**. AR oral: 55/9,122 = 0.6%; overall: 2,334/9,122 = 26% ≥ 50 cites
- [C17] Yazhe Li*, Roman Pogodin*, Danica J. Sutherland, and Arthur Gretton. “Self-Supervised Learning with Kernel Dependence Maximization.” *Neural Information Processing Systems (NeurIPS)*. 2021. AR 2,334/9,122 = 26% ≥ 50 cites
- [C16] Feng Liu*, Wenkai Xu*, Jie Lu, and Danica J. Sutherland. “Meta Two-Sample Testing: Learning Kernels for Testing with Limited Data.” *Neural Information Processing Systems (NeurIPS)*. 2021. AR 2,334/9,122 = 26% ≥ 10 cites
- [C15] Pritish Kamath, Akilesh Tangella, Danica J. Sutherland, and Nathan Srebro. “Does Invariant Risk Minimization Capture Invariance?” *Artificial Intelligence and Statistics (AISTATS)*. 2021. **Selected for oral presentation**. AR oral: 48/1,527 = 3%; overall: 455/1,527 = 30% ≥ 100 cites
- [C14] Lijia Zhou, Danica J. Sutherland, and Nathan Srebro. “On Uniform Convergence and Low-Norm Interpolation Learning.” *Neural Information Processing Systems (NeurIPS)*. 2020. **Selected for spotlight presentation**. AR spotlight: 386/9,467 = 4%; overall: 1,899/9,467 = 20% ≥ 10 cites

- [C13] Feng Liu*, Wenkai Xu*, Jie Lu, Guangquan Zhang, Arthur Gretton, and Danica J. Sutherland. "Learning Deep Kernels for Non-Parametric Two-Sample Tests." *International Conference on Machine Learning (ICML)*. 2020. AR 1,084/4,990 = 22% ≥ 100 cites
- [C12] Li Wenliang*, Danica J. Sutherland*, Heiko Strathmann, and Arthur Gretton. "Learning deep kernels for exponential family densities." *International Conference on Machine Learning (ICML)*. 2019. AR 773/3,424 = 23% ≥ 50 cites
- [C11] Michael Arbel*, Danica J. Sutherland*, Mikołaj Bińkowski, and Arthur Gretton. "On gradient regularizers for MMD GANs." *Neural Information Processing Systems (NeurIPS)*. 2018. AR 1,009/4,856 = 21% ≥ 100 cites
- [C10] Mikołaj Bińkowski*, Danica J. Sutherland*, Michael Arbel, and Arthur Gretton. "Demystifying MMD GANs." *International Conference on Learning Representations (ICLR)*. 2018. AR 314/981 = 32% $\geq 1,000$ cites
- [C9] Danica J. Sutherland*, Heiko Strathmann*, Michael Arbel, and Arthur Gretton. "Efficient and principled score estimation with Nyström kernel exponential families." *Artificial Intelligence and Statistics (AISTATS)*. 2018. **Selected for oral presentation.** AR oral: 31/645 = 5%; overall: 214/645 = 33% ≥ 10 cites
- [C8] Ho Chung Leon Law*, Danica J. Sutherland*, Dino Sejdinovic, and Seth Flaxman. "Bayesian Approaches to Distribution Regression." *Artificial Intelligence and Statistics (AISTATS)*. 2018. AR 214/645 = 33% ≥ 10 cites
- [C7] Danica J. Sutherland, Hsiao-Yu Tung, Heiko Strathmann, Soumyajit De, Aaditya Ramdas, Alex Smola, and Arthur Gretton. "Generative Models and Model Criticism via Optimized Maximum Mean Discrepancy." *International Conference on Learning Representations (ICLR)*. 2017. AR 198/507 = 39% ≥ 100 cites
- [C6] Danica J. Sutherland*, Junier B. Oliva*, Barnabás Póczos, and Jeff Schneider. "Linear-time Learning on Distributions with Approximate Kernel Embeddings." *AAAI Conference on Artificial Intelligence (AAAI)*. 2016. AR 549/2,132 = 26% ≥ 10 cites
- [C5] Danica J. Sutherland and Jeff Schneider. "On the Error of Random Fourier Features." *Uncertainty in Artificial Intelligence (UAI)*. 2015. AR 99/291 = 34% ≥ 100 cites
- [C4] Yifei Ma*, Danica J. Sutherland*, Roman Garnett, and Jeff Schneider. "Active Pointillistic Pattern Search." *Artificial Intelligence and Statistics (AISTATS)*. 2015. AR 127/442 = 29% ≥ 10 cites
- [C3] Danica J. Sutherland, Barnabás Póczos, and Jeff Schneider. "Active learning and search on low-rank matrices." *Knowledge Discovery and Data Mining (KDD)*. 2013. **Selected for oral presentation.** AR oral: 66/726 = 9%; overall: 125/726 = 17% ≥ 10 cites
- [C2] Barnabás Póczos, Liang Xiong, Danica J. Sutherland, and Jeff Schneider. "Nonparametric kernel estimators for image classification." *Computer Vision and Pattern Recognition (CVPR)*. 2012. AR 465/1,776 = 26% ≥ 50 cites
- [C1] Andrew Stromme, Danica J. Sutherland, Alexander Burka, Benjamin Lipton, Nicholas Felt, Rebecca Roelofs, Daniel-Elia Feist-Alexandrov, Steve Dini, and Allen Welkie. "Managing User Requests with the Grand Unified Task System (GUTS)." *Large Installation System Administration (LISA)*. 2012.

(b) *Journals*

- [J7] Yilin Yang, Kamil Adamczewski, Xiaoxiao Li, Danica J. Sutherland, and Mijung Park. "Differentially Private Neural Tangent Kernels (DP-NTK) for Privacy-Preserving Data Generation." *Journal of Artificial Intelligence Research* (2024). ≥ 10 cites
- [J6] **Wonho Bae, Yi Ren**, Mohamad Osama Ahmed, Frederick Tung, Danica J. Sutherland, and Gabriel Oliveira. "AdaFlood: Adaptive Flood Regularization." *Transactions on Machine Learning Research* (2024).
- [J5] Lijia Zhou*, Frederic Koehler*, Danica J. Sutherland, and Nathan Srebro. "Optimistic Rates: A Unifying Theory for Interpolation Learning and Regularization in Linear Regression." *ACM/IMS Journal of Data Science* (2023). ≥ 10 cites
- [J4] Frederik Harder, **Milad Jalali Asadabadi**, Danica J. Sutherland, and Mijung Park. "Pre-trained Per-

ceptual Features Improve Differentially Private Image Generation.” *Transactions on Machine Learning Research* (2023). ≥ 10 cites

[J3] Rémi Flamary et al. “POT: Python Optimal Transport.” *Journal of Machine Learning Research*. Machine Learning Open Source Software Paper (2021). ≥ 1,000 cites

[J2] Michelle Ntampaka, Hy Trac, Danica J. Sutherland, Sebastian Fromenteau, Barnabás Póczos, and Jeff Schneider. “Dynamical Mass Measurements of Contaminated Galaxy Clusters Using Machine Learning.” *The Astrophysical Journal* 831.2 (2016), p. 135. ≥ 50 cites

[J1] Michelle Ntampaka, Hy Trac, Danica J. Sutherland, Nicholas Battaglia, Barnabás Póczos, and Jeff Schneider. “A Machine Learning Approach for Dynamical Mass Measurements of Galaxy Clusters.” *The Astrophysical Journal* 803.2 (2015), p. 50. ≥ 100 cites

(c) Workshop Proceedings

Workshops in machine learning play a similar role to conferences in most fields; submissions are peer-reviewed but acceptance rates are high (and typically not published). Most of these papers have (or will have) a corresponding conference version above.

[W12] Alan Milligan, Frederik Kunstner, **Hamed Shirzad**, Mark Schmidt, and Danica J. Sutherland. “Normalization Matters for Optimization Performance on Graph Neural Networks.” *Optimization for Machine Learning (NeurIPS workshop)*. 2024.

[W11] **Hamed Shirzad**, Honghao Lin, Ameya Velingker, Balaji Venkatachalam, David P. Woodruff, and Danica J. Sutherland. “A Theory for Compressibility of Graph Transformers for Transductive Learning.” *Machine Learning and Compression (NeurIPS workshop)*. 2024.

[W10] **Yi Ren** and Danica J. Sutherland. “Understanding Simplicity Bias towards Compositional Mappings via Learning Dynamics.” *Compositional Learning: Perspectives, Methods, and Paths Forward (NeurIPS workshop)*. 2024.

[W9] Yilin Yang, Kamil Adamczewski, Danica J. Sutherland, Xiaoxiao Li, and Mijung Park. “Differentially Private Neural Tangent Kernels for Privacy-Preserving Data Generation.” *Privacy-Preserving Artificial Intelligence (AAAI workshop)*. 2024.

[W8] **Hamed Shirzad**, Balaji Venkatachalam, Ameya Velingker, Danica J. Sutherland, and David P. Woodruff. “Low-Width Approximations and Sparsification for Scaling Graph Transformers.” *New Frontiers in Graph Learning (NeurIPS workshop)*. 2023.

[W7] **Mohamad Amin Mohamadi**, Zhiyuan Li, Lei Wu, and Danica J. Sutherland. “Grokking modular arithmetic can be explained by margin maximization.” *Mathematics of Modern Machine Learning (NeurIPS workshop)*. 2023.

[W6] **Namrata Deka** and Danica J. Sutherland. “Learning Privacy-Preserving Deep Kernels with Known Demographics.” *Privacy-Preserving Artificial Intelligence (AAAI workshop)*. 2022.

[W5] Organizers of QueerInAI, A Pranav, MaryLena Bleile, Arjun Subramonian, Luca Soldaini, Danica J. Sutherland, Sabine Weber, and Pan Xu. “How to Make Virtual Conferences Queer-Friendly: A Guide.” *Workshop on Widening NLP (EMNLP workshop)*. 2021.

[W4] Ho Chung Leon Law*, Danica J. Sutherland*, Dino Sejdinovic, and Seth Flaxman. “Bayesian Approaches to Distribution Regression.” *Learning on Distributions, Functions, Graphs and Groups (NeurIPS workshop)*. 2017. **Selected for oral presentation.**

[W3] Jay Jin, Kyle Miller, Danica J. Sutherland, Simon Labov, Karl Nelson, and Artur Dubrawski. “List Mode Regression for Low Count Detection.” *IEEE Nuclear Science Symposium (IEEE NSS/MIC)*. 2016.

[W2] Danica J. Sutherland*, Junier B. Oliva*, Barnabás Póczos, and Jeff Schneider. “Linear-time Learning on Distributions with Approximate Kernel Embeddings.” *Feature Extraction: Modern Questions and Challenges (NeurIPS workshop)*. 2015.

[W1] Yifei Ma*, Danica J. Sutherland*, Roman Garnett, and Jeff Schneider. “Active Pointillistic Pattern Search.” *Bayesian Optimization (NeurIPS workshop)*. 2014.

8. NON-REFEREED PUBLICATIONS

(a) *Other*

- [N7] Danica J. Sutherland and **Namrata Deka**. *Unbiased estimators for the variance of MMD estimators*. 2019. arXiv: 1906.02104. ≥ 10 cites
- [N6] Michelle Ntampaka et al. *The Role of Machine Learning in the Next Decade of Cosmology*. 2019. arXiv: 1902.10159. ≥ 50 cites
- [N5] Danica J. Sutherland. *Fixing an error in Caponnetto and de Vito (2007)*. 2017. arXiv: 1702.02982.
- [N4] Seth Flaxman, Danica J. Sutherland, Yu-Xiang Wang, and Yee Whye Teh. *Understanding the 2016 US Presidential Election using ecological inference and distribution regression with census microdata*. 2016. arXiv: 1611.03787. ≥ 10 cites
- [N3] Junier B. Oliva*, Danica J. Sutherland*, Barnabás Póczos, and Jeff Schneider. *Deep Mean Maps*. 2015. arXiv: 1511.04150.
- [N2] Danica J. Sutherland, Liang Xiong, Barnabás Póczos, and Jeff Schneider. *Kernels on Sample Sets via Nonparametric Divergence Estimates*. 2012. arXiv: 1202.0302. ≥ 10 cites
- [N1] Matthew Bodenhamer, Thomas Palmer, Danica J. Sutherland, and Andrew H. Fagg. *Grounding Conceptual Knowledge with Spatio-Temporal Multi-Dimensional Relational Framework Trees*. 2012.