
ACCOMPLISHMENTS

Highly experienced artificial intelligence (AI) researcher, with over 7 years of expertise leading and publishing novel research in reinforcement learning, machine learning, deep learning and optimization. Proven track record of publishing consistently in top-tier AI conferences including NeurIPS, ICLR, and AAAI. Excellent knowledge and experience with the Python ecosystem, including TensorFlow, PyTorch, JAX, Gurobi, scikit-learn. Expertise in developing complex AI applications for research and industry, and providing end-user documentation, troubleshooting and support. Excellent knowledge of generative models (e.g., diffusion) and Bayesian machine learning. Ability to communicate complex concepts verbally and in writing, and proven track record of working in collaborative and team-focused environments.

AREA OF EXPERTISE

Reinforcement Learning/Robotics — Machine Learning — Deep Learning — Optimization — Statistics — AI Research — Transfer Learning — Data Analysis — Software Development

PROGRAMMING SKILLS

• **Languages:** Python, C++, Java **Technologies:** TensorFlow, PyTorch, JAX, Gurobi, scikit-learn, Docker

EXPERIENCE

- **Robot Vision and Learning Lab, Computer Science, University of Toronto** Toronto, Canada
Postdoctoral Research Fellow Jan. 2024 - Present
 - **Offline Reinforcement Learning:** Aims to evaluate or optimize a reinforcement learning (RL) policy using available data collected from another policy. Offline RL is challenging in high-dimensional long-horizon problems.
 - **Generative Models:** Exploring deep generative models (score-matching, flow-matching, and denoising diffusion) for scaling offline policy evaluation to high-dimensional long-horizon problems.
 - **Active Learning:** Exploring budgeted active (i.e. online) data collection to make offline RL more robust.
- **Data-Driven Decision Making Lab, Engineering, University of Toronto** Toronto, Canada
Postdoctoral Research Fellow Jan. 2023 - Sep. 2023
 - **Gradient-Based Planning:** Developed research in gradient-based planning for high-dimensional problems. Introduced model relaxations based on fuzzy logic to facilitate gradient propagation in discrete problems.
 - **pyRDDL Gym:** Led the development of pyRDDL Gym, an ecosystem of Python packages and general framework for modelling and solving sequential decision making problems.
 - **Planning Competition:** Co-hosted the 2023 International Planning Competition (probabilistic track), designed benchmark problems and evaluation metrics, wrote a Docker image for evaluating submitted codebases and a summary paper.
- **Google DeepMind, Reinforcement Learning Team** London, UK (Remote)
Research Scientist Summer 2022
 - **Lifelong Reinforcement Learning:** Involves the design of optimal policies in long-lived environments where the goal changes over time, and it can be seen as a step towards artificial general intelligence.
 - **Transfer Learning:** Proposed a novel solution to the lifelong RL problem, in which a limited number of policies corresponding to previously observed goals are cached. Transfer learning was proposed to use cached policies to improve performance across future goals. A white paper and software package were presented internally.
- **Russell Investments** Toronto, Canada
Research Analyst Oct. 2014 - May. 2015
 - **Data Analysis:** Performed data collection and analysis in excel and VBA to summarize trends in clients' institutional asset allocations and competing mutual funds.
 - **Natural Language Processing:** Developed a VBA application from scratch to automate data analysis of semi-structured client data, using fuzzy string search to reduce manual work by as much as 90%.

SELECTED PROJECTS

- **pyRDDLgym**: Open source library for automated planning in Python. Includes a large set of benchmark problem descriptions and growing ecosystem of planning tools such as gradient-based planning with JAX.
- **Constructive Neural Networks**: Package for training cascade correlation neural networks (CCNN) on classification and regression data sets using Tensorflow and scikit-learn. CCNNs add neurons iteratively during training, thus avoid both the moving target problem and the need for architecture tuning of deep neural networks.
- **Deep Successor Features**: Transfer learning in RL using deep successor features in Tensorflow and Keras.
- **DeepAlpha Inc.**: Co-founder of a startup dedicated to the development of cutting-edge financial investing and trading tools using artificial intelligence, reinforcement learning and quantum computing.

EDUCATION

- **University of Toronto** Toronto, ON
PhD in Industrial Engineering; GPA: 3.93/4 Sep. 2017 – Dec. 2022
- **University of Toronto** Toronto, ON
Master in Applied Science; GPA: 3.93/4 Sep. 2015 – Sep. 2017
- **Schulich School of Business, York University** Toronto, ON
Bachelor in Business Administration (Finance); GPA: 8.2/9 (with Distinction) Sep. 2010 – Jun. 2014

SELECTED PUBLICATIONS

Gimelfarb, Michael, Ayal Taitler, and Scott Sanner. Bounded-Error Policy Optimization for Mixed Discrete-Continuous MDPs via Constraint Generation in Nonlinear Programming. *International Conference on the Integration of Constraint Programming, Artificial Intelligence, and Operations Research*. 2025

Xiaotian Liu, Jihwan Jeong, Ayal Taitler, Michael Gimelfarb, and Scott Sanner. ModelDiff: Symbolic Dynamic Programming for Model-aware Policy Transfer in Deep Q-Learning. *Association for the Advancement of Artificial Intelligence*. Acceptance rate: 23.4% 2025

Gimelfarb, Michael, Ayal Taitler, and Scott Sanner. JaxPlan and GurobiPlan: Optimization Baselines for Replanning in Discrete and Mixed Discrete-Continuous Probabilistic Domains. *International Conference on Automated Planning and Scheduling*. Acceptance rate: 21.6% 2024

Ayal Taitler, Ron Alford, Joan Espasa, Gregor Behnke, Daniel Fišer, Michael Gimelfarb, Florian Pommerening, Scott Sanner, Enrico Scala, Dominik Schreiber, Javier Segovia-Aguas, and Jendrik Seipp. The 2023 International Planning Competition. *AI Magazine*. 2024

Jeong, Jihwan, Xiaoyu Wang, Michael Gimelfarb, Hyunwoo Kim, Baher Abdulhai, and Scott Sanner. Conservative Bayesian Model-Based Value Expansion for Offline Policy Optimization. *International Conference on Learning Representations*. Acceptance rate: 31.8% 2023

Patton, Noah*, Jihwan Jeong*, Michael Gimelfarb*, and Scott Sanner. A Distributional Framework for Risk-Sensitive End-to-End Planning in Continuous MDPs. *Association for the Advancement of Artificial Intelligence*. Acceptance rate: 15.0% (* equal contribution) 2022

Gimelfarb, Michael, André Barreto, Scott Sanner, and Chi-Guhn Lee. Risk-Aware Transfer in Reinforcement Learning using Successor Features. *Advances in Neural Information Processing Systems*. Acceptance rate: 25.7% 2021

Gimelfarb, Michael, Scott Sanner, and Chi-Guhn Lee. Contextual Policy Transfer in Reinforcement Learning Domains via Deep Mixtures-of-Experts. *Uncertainty in Artificial Intelligence*. Acceptance rate: 26.3% 2021

Gimelfarb, Michael, Scott Sanner, and Chi-Guhn Lee. Bayesian Experience Reuse for Learning from Multiple Demonstrators. *International Joint Conference on Artificial Intelligence*. Acceptance rate: 13.9% 2021

Gimelfarb, Michael, Scott Sanner, and Chi-Guhn Lee. Epsilon-BMC: A Bayesian Ensemble Approach to Epsilon-Greedy Exploration in Model-Free Reinforcement Learning. *Uncertainty in Artificial Intelligence*. Acceptance rate: 26.0% 2020

Gimelfarb, Michael, Scott Sanner, and Chi-Guhn Lee. Reinforcement learning with multiple experts: A bayesian model combination approach. *Advances in Neural Information Processing Systems*. Acceptance rate: 20.8% 2018