

Daniel McNeela

daniel.mcneela@gmail.com

(760) 685-7926

SKILLS SUMMARY

- **High-Level:** Machine Learning, Deep Learning, HPC, Large Language Models, NLP, Graph Neural Networks
- **Languages:** Python, C, C++, CUDA, Spark, Javascript
- **Frameworks:** PyTorch, PyTorch Lightning, JAX, HuggingFace, Weights and Biases, Django, React.js
- **Tools:** Docker, Git, Bash, Slurm, HTCondor
- **Platforms:** Linux, AWS (S3, EC2, Sagemaker)
- **Communication:** Technical Writing, Content Writing

EXPERIENCE

- **Valence Labs** Montreal, QC
Machine Learning Research Intern *January 2024 - July 2024*
 - **Machine Learning:** In 4 months, built and trained a large-scale deep learning model using PyTorch and PyTorch Lightning for modeling of non-covalent interactions in atomistic systems. Trained model with millions of SAPT0 and DFT simulation data samples to predict atomic energies. Achieved 30% greater GPU utilization and 50% increase in model parameter size over previous models.
 - **Distributed Training:** Used DDP and multi-GPU training to scale to extremely large model sizes and datasets. Managed jobs on an HPC Slurm cluster.
 - **Data Engineering:** Helped architect a scalable and performant backend system for processing and storage of over 100 million data points of atomic simulation data.
 - **Tools:** Worked extensively in Python, PyTorch, and PyTorch Lightning
- **University of Wisconsin, Madison** Madison, WI
Researcher *June 2021 - July 2024*
 - **Machine Learning:** Performed research in computational biology, equivariant neural networks, graph representation learning, and graph neural networks under the guidance of Professors Anthony Gitter and Fred Sala. Published novel research in top-tier venues such as ICML and NeurIPS.
 - **Tools:** Worked extensively in Python, PyTorch, and PyTorch Lightning
- **Eli Lilly** San Diego, CA
Software Engineer *June 2020 - May 2021*
 - **Backend Architecture:** Served as backend engineering lead for drug discovery platform. Developed and maintained pipeline for in-silico drug discovery of monoclonal antibodies and immunotherapeutics using Cromwell, Python, R, and Django. Scaled pipeline to handle 100x the number of parallel runs (thousands of concurrent jobs) relative to the previous system using cloud tools such as AWS and Docker. Applied rigorous test driven development using pytest and pytest-wdl.
 - **Frontend:** Contributed to the frontend for a drug discovery platform used by scientists across the company. Developed extensively in React.js.
- **Human Longevity, Inc.** San Diego, CA
Machine Learning Scientist *February 2019 - October 2019*
 - **Bioinformatics:** Built bioinformatics pipelines for variant calling incorporating GATK tools. Architected with Python and workflow management with Luigi. Migrated existing pipelines from local environment to the AWS cloud.
- **machineVantage** Berkeley, CA
Machine Learning Engineer *August 2017 - February 2019*
 - **Deep Learning and NLP:** Implemented deep learning models from current NLP research papers and performed novel research in machine learning. Focused on word and sentence embedding models such as word2vec, GloVe, and FastText.
 - **Frontend and Backend Engineering:** Built an in-house web app which allowed marketers to rapidly search through related cultural/brand concepts. Mined page links in Wikipedia, wrote code in C++ to collate links, generated embeddings, and served data using ElasticSearch. Built a microservice using Flask as a backend with Javascript frontend.

EDUCATION

- **University of Wisconsin, Madison** Madison, WI
Master of Science - Biomedical Data Science *May 2023*
- **University of California, Berkeley** Berkeley, CA
Bachelor of Arts - Applied Mathematics *May 2017*

PUBLICATIONS

- **GROG: Reducing LLM Hallucinations for Improved Legal Reasoning** D McNeela
ICML 2024 Workshop on Generative AI and the Law July 2024
- **Mixed-Curvature Representation Learning for Biological Pathway Graphs** D McNeela, F Sala, A Gitter
ICML 2023 Workshop on Computational Biology July 2023
- **Almost Equivariance via Lie Algebra Convolutions** D McNeela
Under Review for NeurIPS 2024 October 2023
- **Almost Equivariance via Lie Algebra Convolutions – Extended Abstract** D McNeela
NeurIPS 2023 Workshop on Symmetry and Geometry in Neural Representations December 2023

PROJECTS

- **C++ & CUDA Recommender System** Madison, WI
 - : Wrote a recommender system in C++ and CUDA implementing classic matrix-factorization algorithms. Created efficient parallel implementation across many GPU cores.
- **Distributed Data and Model Parallel Transfer Learning System** Madison, WI
 - : Helped develop a distributed system using PyTorch DDP that allows users to transfer learn many models simultaneously on multiple datasets.

ACADEMIC SERVICE

- **ML for Life and Material Science ICML Workshop** Vienna, Austria
Program Committee Member July, 2024
- **ICML Topology and Geometry in Machine Learning Workshop** Honolulu, HI
Reviewer July, 2023
- **NeurIPS Symmetry and Geometry in Neural Representations Workshop** New Orleans, LA
Program Committee Member, Reviewer December, 2023

TECHNICAL AND CONTENT WRITING

- **Various Companies** Remote
Freelance Writer 2019 - Present
 - **Blog Posts, Whitepapers, Documentation:** Wrote blog posts for company websites on machine learning topics and architectures such as transformers and large language models. Created whitepapers to market company technologies. Built out documentation for software products. Worked for a number of companies including Snorkel AI, Scale AI, MosaicML, Neptune.ai, MonaLabs, Kaskada, Divio, Cortical.io, New Relic, iMerit, Rev.ai, Hyper-Space.io, Algorithmia, and Clarifai.