Michael Munn, PhD

mikemunn@gmail.com | +1 646 409 1398
munnm.github.io linkedin.com/in/munnm | Google scholar

Accomplished Research Scientist with a proven track record in foundational ML research, learning theory and mathematics; demonstrated ability to lead diverse cross-functional research teams.
Experienced Software Engineer with success leading, scoping, developing, and productionzing large scale machine learning solutions across multiple verticals and product areas. Published O'Reilly author of ML/AI books, research articles, and blog posts

Skills

Programming: Python, SQL, bash, R, C++ (beginner)
Tools, Skills and Areas of Interest: JAX/Flax/Pax, Tensorflow/Keras, PyTorch, Scikit-Learn, Text Generation (GPT, Transformers), Image Generation (GANs, diffusion models), fine-tuning and transfer learning, explainable/interpretable AI and XAI tools (Captum, SHAP, LIME, LIT), statistical learning theory and generalization, implicit regularization, differential & Riemannian geometry, optimal transportation, Google Cloud Platform, AWS, Kubeflow/Kubernetes, BigQuery/Dremel, Dataflow/Flume, Colab/Jupyter, Flask

Experience

Google Research, Senior Software Engineer **Feb/2022 – present**
My research is focused on foundational machine learning theory and its application to improving efficiencies in training and fine-tuning large-scale models at Google. I have strong contributions to research communities within and outside of Google and my research has appeared in top tier conferences
• **20% projects** AI Coach, Google.org GenAI Accelerator
  • built a bio-materials discovery lab assistant using Gemini Pro working with UK-based NGO Materiom
  • engineered a tool using Google's Gemini 1.5 Pro to monitor and analyze online videos to surface misinformation and harmful health claims working with UK-based FullFact

Google Cloud, ML Solutions Engineer **Apr/2018 – Feb/2022**
Within Google's AI Services, I built and productionized AI solutions for Cloud customers across all verticals. As lead of the Advanced Solutions Lab, we provided dedicated ML training covering best practices for developing end-to-end machine learning solutions.
• **Tech Lead**, Google.org Fellowship with The Trevor Project
  • developed and deployed a custom conversation simulator model using PyTorch and GPT2/GPT3
  • led a team of six ML SWEs across Google and Trevor
  • determined scope for the 6 month project from exploratory ML development to a deployed model currently in production and being used by Trevor trainees
  • trainees used our conversation simulator for 4,919 hrs of training, removing 35-50 role-play shifts for trainers and the potential to graduate an additional 200-300 new counselors per year
  • regularly coordinated with business stakeholders with Trevor and Google to align on delivery and key objectives to identify and resolve blockers
  • external press: MIT Tech Review, Time Magazine's Best Inventions of 2021, Google.org blog post, Today show segment, and Google I/O '21
• **Engineering and technical Lead** for multiple ML engagements with Google Cloud customers across various verticals delivering end-to-end ML solutions from development to productionisation
  • Led scoping and ML/AI Feasibility engagements with Cloud customers & AI Accelerator grantees to assess use cases and shape delivery
• **Delivered dedicated ML Instruction** for Google's Advanced Solutions Lab, an intensive, immersive ML training course combining Advanced ML curriculum with Tensorflow/GCP and sprint-based Open Project work.
  • consistently received 95+% Overall CSAT and NPS, and 100% Trainer Communications and 100% Trainer Technical skill on customer feedback
  • developed and maintained the code base consisting of hands-on labs used during training delivery

Accenture, Data Science Consultant **Jan/2017 – Apr/2018**

• Developed Tensorflow model to predict future server failure, implemented LIME and other explainability methods to provide insight to root cause analysis
• Created a multi-class boosted tree model in sk-learn and bi-directional RNN in Tensorflow to classify aircrafts in flight, applied SMOTE to handle class imbalance, implemented end-to-end pipeline handling data ingestion through BigQuery, preprocessing via Dataflow, modeling in Tensorflow and accuracy analysis
• Applied data driven analytics and ML modeling to recognize $8MM in reducible cost for one of largest utility companies in the Northeast US
• Built robust modeling pipelines in python for supervised (regularized regression, boosting, random forest) and unsupervised (k-means) learning, cutting down analysis time from days to hours

Insight Health Data Science Fellow **Sept/2016 – Dec/2016**

• Analyzed user FitBit sleep and activity data using python/pandas and mixed effect models in R
• Fine-tuned multiple ML models to provide a personalized user sleep recommendation based on their personal data and other crowdsourced data
• Deployed an interactive webapp on AWS using Flask and Bootstrap in python

(prior academic work experience)

New York University, Courant Institute, Clinical Assistant Professor **2014 – 2016**
University of Missouri, Assistant Professor **2011 – 2014**
University of Warwick, National Science Foundation Postdoctoral Fellow **2009 – 2011**
CUNY, NYCCT, Assistant Professor **2008 – 2011**

Publications

Books

Explainable AI for Practitioners, (with D. Pitman) O'Reilly, November 2022
Machine Learning Design Patterns, (with V. Lakshmanan and S. Robinson) O'Reilly, November 2020

Recent Research Papers

Leveraging free energy in pretraining model selection for improved fine-tuning (with S. Wei) *in review*

Training in reverse: How iteration order influences convergence and stability in deep learning (with B. Dherin, B. Avelin, A. Karlsson, H. Mazzawi, J. Gonzalvo) *in review*

The Impact of Geometric Complexity on Neural Collapse in Transfer Learning (with B. Dherin, X. Gonzalvo), NeurIPS 2024
A margin-based multiclass generalization bound via geometric complexity (with B. Dherin, X. Gonzalvo), ICML 2023, Topology, Algebra and Geometry Workshop (with B. Dherin, X. Gonzalvo)
Unified Functional Hashing in Automated Machine Learning (with R. Gillard, S. Jonany, Y, Miao, C. Souza, J. Dungay, E. Real, Q. V. Le)
Why neural networks find simple solutions: the many regularizers of geometric complexity (with B. Dherin, M. Rosca, D. Barrett) NeurIPs 2022
The Geometric Occam's Razor Implicit in Deep Learning (with B. Dherin, D. Barrett) NeurIPS 2021, Workshop on Optimization
COT-GAN: Generating Sequential Data via Causal Optimal Transport (with T. Xu, L. Wenliang, B. Acciaio) NeurIPS 2020

Pre-Google

On the Size of a Ricci Flow Neckpinch via Optimal Transport (w/ S. Lakzian) Analysis and Geometry of Metric Measure Spaces, 2021
Three-Dimensional Alexandrov spaces with positive or nonnegative Ricci curvature (w/ Q. Deng, F. Galaz-Garcia, L. Guijarro) Potential Analysis, 2017
Geometric singularities and a flow tangent to the Ricci flow (w/ L. Bandara, S. Lakzian) Annali S.N.S di Pisa, 2015

Alexandrov spaces with large volume growth, Journal of Mathematical Analysis and Applications, 2015
Super Ricci flow for disjoint unions (w/ S. Lakzian) Analysis and Geometry of Metric Measure Spaces, 2012
Volume growth and the topology of pointed Gromov-Hausdorff limits, Diff Geo. and Its Applications, 2010
Volume growth and the topology of manifolds with nonnegative Ricci curvature, Journal of Geometric Analysis, 2010
On the appearance of Eisenstein series through degeneration (w/ D. Garbin, J. Jorgenson) Commentarii Mathematici Helvetici, 2008

Blog Posts

How to deploy interpretable models on Google Cloud Platform, Toward Data Science, 2020
Building a document understanding pipeline with Google Cloud, Google Cloud AI blog, 2019

Education

City University of New York, NY
Ph.D in Mathematics (focus area: geometric analysis and topology)
M.Phil in Mathematics
University of Notre Dame, South Bend, IN
B.S. in Honors Mathematics