

Mohammad Norouzi

Staff Research Scientist and Manager,
Google Inc., Toronto, ON, Canada

Email: mnorouzi@google.com
Homepage: norouzi.github.io

Employment

Staff Research Scientist and Manager, Google Inc.
Google Brain team, Toronto, Oct 2019 – present.

Senior Research Scientist, Google Inc.
Google Brain team, Toronto, Mar 2017 – Oct 2019.

Research Scientist, Google Inc.
Google Brain team, Mountain View, Jan 2016 – Mar 2017.

Education

Ph.D. in Computer Science
University of Toronto, ON Canada, Sep 2010 – Dec 2015.

Thesis: Learning Compact Discrete Representations for Scalable Similarity Search
Advisor: [David Fleet](#)

M.Sc. in Computer Science
Simon Fraser University, BC Canada, Jan 2008 – Dec 2009.

Thesis: Convolutional Restricted Boltzmann Machines for Unsupervised Feature Discovery
Advisor: [Greg Mori](#)

B.Sc. in Computer Engineering
Sharif University of Technology, Tehran, Iran, Sep 2003 – Aug 2007.

Research Interests

Machine Learning, Neural Networks, Reinforcement Learning, Generative Models, Computer Vision, Natural Language Processing

Publications

Shekoofeh Azizi, Basil Mustafa, Fiona Ryan, Zachary Beaver, Jan Freyberg, Jonathan Deaton, Aaron Loh, Alan Karthikesalingam, Simon Kornblith, Ting Chen, Vivek Natarajan, Mohammad Norouzi, “Big Self-Supervised Models Advance Medical Image Classification”, [ArXiv](#).

Simon Kornblith, Honglak Lee, Ting Chen, Mohammad Norouzi, “What’s in a Loss Function for Image Classification?”, [ArXiv](#).

Michael R Zhang, Thomas Paine, Ofir Nachum, Cosmin Paduraru, George Tucker, Ziyu Wang, Mohammad Norouzi, "Autoregressive Dynamics Models for Offline Policy Evaluation and Optimization", [ICLR 2021](#).

Justin Fu, Mohammad Norouzi, Ofir Nachum, George Tucker, Ziyu Wang, Alexander Novikov, Mengjiao Yang, Michael R Zhang, Yutian Chen, Aviral Kumar, Cosmin Paduraru, Sergey Levine, Thomas Paine, "Benchmarks for Deep Off-Policy Evaluation", [ICLR 2021](#).

Nanxin Chen, Yu Zhang, Heiga Zen, Ron J Weiss, Mohammad Norouzi, William Chan, "WaveGrad: Estimating gradients for waveform generation", [ICLR 2021](#).

Danijar Hafner, Timothy Lillicrap, Mohammad Norouzi, Jimmy Ba, "Mastering Atari with discrete world models", [ICLR 2021](#).

Will Grathwohl, Jacob Kelly, Milad Hashemi, Mohammad Norouzi, Kevin Swersky, David Duvenaud, "No MCMC for me: Amortized sampling for fast and stable training of energy-based models", [ICLR 2021](#).

Ting Chen, Simon Kornblith, Kevin Swersky, Mohammad Norouzi, Geoffrey Hinton, "Big self-supervised models are strong semi-supervised learners", [NeurIPS 2020](#).

Sajad Norouzi, David J Fleet, Mohammad Norouzi, "Exemplar VAE: Linking Generative Models, Nearest Neighbor Retrieval, and Data Augmentation", [NeurIPS 2020](#).

Yijie Guo, Jongwook Choi, Marcin Moczulski, Shengyu Feng, Samy Bengio, Mohammad Norouzi, Honglak Lee, "Memory Based Trajectory-conditioned Policies for Learning from Sparse Rewards", [NeurIPS 2020](#).

Caglar Gulcehre, Ziyu Wang, Alexander Novikov, Thomas Paine, Sergio Gómez, Konrad Zolna, Rishabh Agarwal, Josh S. Merel, Daniel J. Mankowitz, Cosmin Paduraru, Gabriel Dulac-Arnold, Jerry Li, Mohammad Norouzi, Matthew Hoffman, Nicolas Heess, Nando de Freitas, "RL Unplugged: A Collection of Benchmarks for Offline Reinforcement Learning", [NeurIPS 2020](#).

Chitwan Saharia, William Chan, Saurabh Saxena, Mohammad Norouzi, "Non-autoregressive machine translation with latent alignments", [EMNLP 2020](#).

Xuanli He, Gholamreza Haffari, Mohammad Norouzi, "Dynamic Programming Encoding for Subword Segmentation in Neural Machine Translation", [ACL 2020](#).

Timothy Jeruzalski, Boyang Deng, Mohammad Norouzi, John P Lewis, Geoffrey Hinton, Andrea Tagliasacchi, "NASA: Neural articulated shape approximation", [ECCV 2020](#).

William Chan, Chitwan Saharia, Geoffrey Hinton, Mohammad Norouzi, Navdeep Jaitly, "Imputer: Sequence modelling via imputation and dynamic programming", [ICML 2020](#).

Ting Chen, Simon Kornblith, Mohammad Norouzi, Geoffrey Hinton, "A Simple Framework for Contrastive Learning of Visual Representations", [ICML 2020](#).

Yucen Luo, Alex Beatson, Mohammad Norouzi, Jun Zhu, David Duvenaud, Ryan Adams, Ricky Chen, "SUMO: Unbiased Estimation of Log Marginal Probability for Latent Variable Models", [ICLR 2020](#).

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, Mohammad Norouzi, "Dream to control: Learning behaviors by latent imagination", [ICLR 2020](#).

- James Lucas, George Tucker, Roger B Grosse, Mohammad Norouzi,
"Don't Blame the ELBO! A Linear VAE Perspective on Posterior Collapse", [NeurIPS 2010](#).
- Rishabh Agarwal, Chen Liang, Dale Schuurmans, Mohammad Norouzi,
"Learning to Generalize from Sparse and Underspecified Rewards", [ICML 2019](#).
- Simon Kornblith, Mohammad Norouzi, Honglak Lee, Geoffrey Hinton,
"Similarity of Neural Network Representations Revisited", [ICML 2019](#).
- Zafarali Ahmed, Nicolas Le Roux, Mohammad Norouzi, Dale Schuurmans,
"Understanding the Impact of Entropy on Policy Optimization", [ICML 2019](#).
- Sara Sabour, William Chan, Mohammad Norouzi,
"Optimal Completion Distillation for Sequence Learning", [ICLR 2019](#).
- Jongwook Choi, Yijie Guo, Marcin Moczulski, Junhyuk Oh, Neal Wu, Mohammad Norouzi, Honglak Lee,
"Contingency-Aware Exploration in Reinforcement Learning", [ICLR 2019](#).
- Supasorn Suwajanakorn, Noah Snively, Jonathan Tompson, Mohammad Norouzi,
"Discovery of Latent 3D Keypoints via End-to-end Geometric Reasoning", [NeurIPS 2018 \(Oral\)](#)
- Chen Liang, Mohammad Norouzi, Jonathan Berant, Quoc Le, Ni Lao,
"Memory Augmented Policy Optimization for Program Synthesis with Generalization", [NeurIPS 2018 \(Spotlight\)](#).
- Xuanli He, Gholamreza Haffari, Mohammad Norouzi,
"Sequence to Sequence Mixture Model for Diverse Machine Translation", [CoNLL 2018](#).
- Nicolas Ford, Daniel Duckworth, Mohammad Norouzi, George E Dahl,
"The Importance of Generation Order in Language Modeling", [EMNLP 2018](#).
- Bhuwan Dhingra, Christopher J Shallue, Mohammad Norouzi, Andrew M Dai, George E Dahl,
"Embedding Text in Hyperbolic Spaces", [TextGraphs Workshop 2018](#).
- Ofir Nachum, Mohammad Norouzi, George Tucker, Dale Schuurmans,
"Smoothed Action Value Functions for Learning Gaussian Policies", [ICML 2018](#).
- Yun Liu, Timo Kohlberger, Mohammad Norouzi, George E Dahl, Jenny L Smith, Arash Mohtashamian, Niels Olson, Lily H Peng, Jason D Hipp, Martin C Stumpe,
"Artificial Intelligence-Based Breast Cancer Nodal Metastasis Detection: Insights Into the Black Box for Pathologists", [Archives of Pathology & Laboratory Medicine 2018](#).
- Adams Wei Yu, David Dohan, Minh-Thang Luong, Rui Zhao, Kai Chen, Mohammad Norouzi, Quoc V Le,
"QANet: Combining Local Convolution with Global Self-Attention for Reading Comprehension", [ICLR 2018](#).
- Ofir Nachum, Mohammad Norouzi, Kelvin Xu, Dale Schuurmans,
"Trust-PCL: An Off-Policy Trust Region Method for Continuous Control", [ICLR 2018](#).
- Chris J Maddison, Dieterich Lawson, George Tucker, Nicolas Heess, Mohammad Norouzi, Andriy Mnih, Arnaud Doucet, Yee Whye Teh,
"Filtering Variational Objectives", [NIPS 2017](#).
- Ofir Nachum, Mohammad Norouzi, Kelvin Xu, Dale Schuurmans,
"Bridging the Gap Between Value and Policy Based Reinforcement Learning", [NIPS 2017](#).

Ryan Dahl, Mohammad Norouzi, Jonathon Shlens,
"Pixel Recursive Super Resolution", [ICCV 2017](#).

Sergio Guadarrama, Ryan Dahl, David Bieber, Mohammad Norouzi, Jonathon Shlens, Kevin Murphy,
"PixColor: Pixel Recursive Colorization", [BMVC 2017](#).

Michael Gygli, Mohammad Norouzi, Anelia Angelova,
"Deep Value Networks Learn to Evaluate and Iteratively Refine Structured Outputs", [ICML 2017](#).

Azalia Mirhoseini, Hieu Pham, Quoc V Le, Benoit Steiner, Rasmus Larsen, Yuefeng Zhou, Naveen Kumar, Mohammad Norouzi, Samy Bengio, Jeff Dean,
"Device Placement Optimization with Reinforcement Learning", [ICML 2017](#).

Jesse Engel, Cinjon Resnick, Adam Roberts, Sander Dieleman, Douglas Eck, Karen Simonyan, Mohammad Norouzi,
"Neural Audio Synthesis of Musical Notes with WaveNet Autoencoders", [ICML 2017](#).

Irwan Bello, Hieu Pham, Quoc V. Le, Mohammad Norouzi, Samy Bengio,
"Neural Combinatorial Optimization with Reinforcement Learning", [ICLR Workshop 2017](#).

Ofir Nachum, Mohammad Norouzi, Dale Schuurmans,
"Improving Policy Gradient by Exploring Under-appreciated Rewards", [ICLR 2017](#).

Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Lukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes, Jeffrey Dean,
"Google's neural machine translation system: Bridging the Gap between Human and Machine Translation", [Technical Report 2016](#).

Mohammad Norouzi, Samy Bengio, Zhifeng Chen, Navdeep Jaitly, Mike Schuster, Yonghui Wu, Dale Schuurmans,
"Reward Augmented Maximum Likelihood for Neural Structured Prediction", [NIPS 2016](#).

Mohammad Norouzi,
"Compact Discrete Representations for Scalable Similarity Search", [PhD thesis 2016](#).

Mohammad Norouzi, Maxwell D. Collins, David J. Fleet, Pushmeet Kohli,
"CO₂ Forest: Improved Random Forest by Continuous Optimization of Oblique Splits" [Technical Report 2015](#).

Mohammad Norouzi, Maxwell D. Collins, Matthew Johnson, David J. Fleet, Pushmeet Kohl,
"Efficient Non-greedy Optimization of Decision Trees", [NIPS 2015](#).

Mohammad Norouzi, Tomas Mikolov, Samy Bengio, Yoram Singer, Jonathon Shlens, Andrea Frome, Greg S. Corrado, Jeffrey Dean,
"Zero-Shot Learning by Convex Combination of Semantic Embeddings", [ICLR 2014 \(Oral\)](#).

Mohammad Norouzi, Ali Punjani, David J. Fleet,
"Fast Exact Search in Hamming Space with Multi-Index Hashing", [TPAMI 2014](#), vol. 36, no. 6.

Mohammad Norouzi, David J. Fleet,
"Cartesian k-means", [CVPR 2013 \(Oral\)](#).

Mohammad Norouzi, David J. Fleet, Ruslan Salakhutdinov,
"Hamming Distance Metric Learning", [NIPS 2012](#).

Mohammad Norouzi, Ali Punjani, David J. Fleet,
“Fast Search in Hamming Space with Multi-index Hashing”, [CVPR 2012](#).

Mohammad Norouzi, David J. Fleet,
“Minimal Loss Hashing for Learning Compact Binary Codes”, [ICML 2011](#).

Mohammad Norouzi, Mani Ranjbar, and Greg Mori,
“Restricted Boltzmann Machines for Learning Shift-Invariant Features”, [CVPR 2009](#).

Academic Service

Area Chair for ICLR’18, NIPS’18, ICLR’19, ICML’19.

Reviewer for NIPS, ICML, ICLR, TPAMI, IJCV, CVPR, ICCV, ECCV, BMVC.

Awards & Honors

2014 Google US/Canada PhD Fellowship in Machine Learning.

Ontario Graduate Scholarship, 2013-14, as an international student.

Ontario Graduate Scholarship, 2012-13, as an international student.

SFU Graduate Fellowship, 2009.

2nd Team, ACM ICPC Regional Contest, Tehran, 2005. 1st and 3rd teams advanced to ACM Final.

5th place (silver medal) in [Central European Olympiad in Informatics \(CEOI 2003\)](#), Munster, Germany.

Team member, International Olympiad in Informatics, 2003. *Iranian team could not attend IOI’03 in US.*

Gold medal in Iranian National Olympiad in Informatics, 2003. *Exempt from the national university entrance exam.*

Silver medal in Iranian National Olympiad in Informatics, 2002.

Previous Work Experience

Research Intern, Google Inc., Mountain View, CA, USA. Mentors: Samy Bengio and Yoram Singer,
June 2013 – Sep 2013.

Research Intern, Microsoft Research Ltd., Cambridge, UK. Mentor: Pushmeet Kohli,
March 2013 – May 2013.

Teaching Assistant – UofT: Machine Learning, Data Structures, Algorithms, Introduction to Theory of Computation

Teaching Assistant – SFU: Discrete Mathematics

Lead Software Developer, ACM training Web application for Sharif University, 2006.

Member of the scientific committee, ACM ICPC Regional Contest, 2006 and 2007.

Member of the scientific committee, Iranian National Olympiad in Informatics, 2003-2007

Teaching in training camps and designing problems for Iranian National Informatics Olympiad.

Teaching Assistant – Sharif University: Data Structure and Algorithms, Advanced Programming in Java, Discrete Structures.

Miscellaneous

Data Structures, Algorithms, and Graph Theory: Excellent theoretical and practical background, as demonstrated in the programming competitions such as CEOI and ACM ICPC.

Programming languages skills: Proficient in C, C++, Python, TensorFlow, Java, Matlab.

Web development skills: Experienced in Apache Struts, Hibernate, Spring, JSF, GWT, PHP, HTML.

Sports: avid cyclist, long round-trips between: Horseshoe bay & Whistler, Vancouver & Harrison hot springs, Toronto & Waterloo, Toronto & Niagara Falls.

Member of the organizing committee for Sharif university hiking club, 2006.

Last updated: March 25, 2021