Maithra (Maithreyi) Raghu

@maithra_raghu

maithrar@gmail.com
http://maithraraghu.com

Education

Cornell University — New York
PhD Computer Science, Advisor: Jon Kleinberg

University of Cambridge (Trinity College) — Cambridge, UK
Masters in Mathematics (MMath)

University of Cambridge (Trinity College) — Cambridge, UK
BA Mathematics, First Class Honours.

Research and Professional Experiences

Google Brain — Senior Research Scientist
with Quoc Le, Samy Bengio and Geoffrey Hinton — August 2018 – present
Insights on AI representations for design and human-AI decision making

Google Brain — Research Scientist
with Quoc Le, Samy Bengio and Jascha Sohl-Dickstein — January 2016 – August 2018
Understanding and Explaining Deep Learning

UC Berkeley, Simons Institute — Visiting Researcher
Foundations of Deep Learning Program — May 2019 – August 2019
Invited participant for summer program on foundations and frontiers of deep learning.

UC Berkeley, Simons Institute — Visiting Researcher
Foundations of Machine Learning Workshop — January 2017 – June 2017
Invited participant for workshop on better understanding fundamentals of machine learning.

Stanford University — Visiting Researcher
with Surya Ganguli — January 2016 – January 2017
Understanding fundamentals of deep neural networks and connections to neuroscience

Google Research, Strategic Technologies — Research Intern
with Andrew Tomkins, Ravi Kumar and Tamas Sarlos — May 2015 – August 2015
Predicting user trails (website visits, song listens, location checkins) with Markov models and LSTMs

Brown University — Visiting Researcher
with Eli Upfal — July 2013 – September 2013
Developed and analysed a random sampling procedure that could greatly reduce the amount of metadata storage required during garbage collection in flash memory.

Tata Institute of Fundamental Research — Visiting Scholar
with C.S. Rajan — June 2012 – August 2012
Scholarship student at summer school aimed at exposing students to advanced mathematics. Studied Galois Theory and Elliptic Curves.

Selected Awards

STAT 2020 Wunderkind — 2020
One of STAT's 2020 top young scientists for our work on human-AI collaboration.

Forbes 30 Under 30 (Science) — 2019
Named on global list of most impactful leaders and scientists for work on studying representations of deep neural networks and applications to medicine.

MIT EECS Rising Stars 2018 2018
Invited participant in workshop at MIT for top female graduate students in EECS

Cornell PhD McMullen Fellowship 2014 — 2015
Awarded to excellent incoming graduate students.

Trinity College Cambridge Senior Scholarship 2013
Awarded for outstanding results in University of Cambridge Part II Examinations.

China Girls Maths Olympiad: Bronze Medal 2010
One of a team of four representing United Kingdom at an international Olympiad.

British Mathematical Olympiad: Gold Medal 2009 — 2010
Top (nationally) twenty and top ten, 2009, 2010

Selected Professional Activities

Advisory Roles

National Security Commission on AI (NSCAI) 2021
Advisory role as AI expert on research advances in AI, technical considerations, and important areas for future focus.

Program Co-Chair

ICCV Workshop: Machine Learning for Medical Imaging 2021
Taking place with the International Conference on Computer Vision, this workshop focuses on studying AI and ML applications in medical imaging

National Academy of Sciences Colloquium: Science of Deep Learning 2019
Invited to be one of five organizers for colloquium at National Academy of Sciences on The Science of Deep Learning. Invited leading researchers, policy makers and representatives from different government bodies to convene and discuss the advances and important challenges for AI and Deep Learning.

ICML Workshop: Deep Phenomena 2019
Co-located with ICML, this workshop focused on identifying and systematically understanding counterintuitive properties exhibited by deep neural networks.

NeurIPS Workshop: Deep Learning: Bridging Theory and Practice 2017
Co-organized NIPS workshop looking at a systematic exploration of phenomena observed with deep neural networks, attended by over one thousand researchers.

Women in Machine Learning (WiML) Organizer 2015
Co-located with NIPS, this workshop brings together hundreds of leading women and men machine learning researchers, and displays top peer-reviewed results in Machine Learning by female researchers. Organizers are vetted by leading women ML researchers.

Publications

Do Vision Transformers See Like Convolutional Neural Networks?
Maithra Raghu, Thomas Unterthiner, Simon Kornblith, Chiyuan Zhang, Alexey Dosovitskiy 2021
Neural Information Processing Systems (NeurIPS) 2021

Pointer-Value Retrieval: Understanding limits of neural network generalization
Chiyuan Zhang, Maithra Raghu, Jon Kleinberg, Samy Bengio 2021
preprint

Do Wide and Deep Neural Networks Learn the Same Things?
Thao Nguyen, Maithra Raghu, Simon Kornblith 2021
International Conference on Learning Representations (ICLR) 2021

Neural Network Teaching with Commentaries
Aniruddh Raghu, Maithra Raghu, Simon Kornblith, David Duvenaud, Geoff Hinton 2021
International Conference on Learning Representations (ICLR) 2021

Anatomy of Catastrophic Forgetting: Hidden Representations and Task Semantics
Vinay Ramasesh, Ethan Dyer, Maithra Raghu 2021
ICLR 2021, also Best Paper at ICML 2020 Workshop on Continual Learning

A Survey of Deep Learning for Scientific Discovery
Maithra Raghu, Eric Schmidt 2020
Preprint

Rapid Learning or Feature Reuse? Understanding the Effectiveness of MAML
Aniruddh Raghu*, Maithra Raghu*, Samy Bengio, Oriol Vinyals 2020
International Conference on Learning Representations (ICLR) 2020

Transfusion: Understanding Transfer Learning for Medical Imaging
Maithra Raghu*, Chiyuan Zhang*, Jon Kleinberg, Samy Bengio 2019
Neural Information Processing Systems (NeurIPS) 2019

The Algorithmic Automation Problem: Prediction, Triage and Human Effort
Maithra Raghu, Katy Blumer, Greg Corrado, Jon Kleinberg, Ziad Obermeyer, Sendhil Mullainathan 2019
Preprint

Direct Uncertainty Prediction for Medical Second Opinions
Maithra Raghu*, Katy Blumer*, Rory Sayres, Ziad Obermeyer, Sendhil Mullainathan, Jon Kleinberg 2019
International Conference on Machine Learning (ICML) 2019

Insights on Representational Similarity in Neural Networks with CCA
Ari Morcos*, Maithra Raghu*, Jascha Sohl-Dickstein, Samy Bengio 2018
Neural Information Processing Systems (NeurIPS) 2018

Adversarial Spheres
J Gilmer, L Metz, F Faghri, S Schoenholz, Maithra Raghu, M Wattenberg, I Goodfellow 2018
International Conference on Learning Representations (ICLR) Workshop 2018

Can Deep Reinforcement Learning Solve Erdos-Selfridge-Spencer Games?
Maithra Raghu, Alexander Irpan, Jacob Andreas, Robert Kleinberg, Quoc V. Le, Jon Kleinberg 2018
International Conference on Machine Learning (ICML) 2019

SVCCA for Deep Learning Dynamics and Interpretability
Maithra Raghu, Justin Gilmer, Jason Yosinski, Jascha Sohl-Dickstein 2017
Neural Information Processing Systems (NeurIPS) 2017

On the expressive power of deep neural networks
Maithra Raghu, Ben Poole, Jon Kleinberg, Surya Ganguli, Jascha Sohl-Dickstein 2017
International Conference on Machine Learning (ICML) 2017. Also appeared in NeurIPS Interpretable Machine Learning Workshop 2016, Women in Machine Learning (WiML) 2016 Oral.

Explaining the Learning Dynamics of Direct Feedback Alignment
Justin Gilmer, Colin Raffel, Sam Schoenholz, Maithra Raghu, Jascha Sohl-Dickstein 2017
International Conference on Learning Representations (ICLR) Workshop 2017

Exponential expressivity in deep neural networks through transient chaos
B Poole, S Lahiri, M Raghu, J Sohl-Dickstein, S Ganguli 2016
Neural Information Processing Systems (NeurIPS) 2016

Linear Additive Markov Processes
Ravi Kumar, Maithra Raghu, Tamas Sarlos, Andrew Tomkins (alphabetical order) 2016
WWW 2017

Team Performance with Test Scores
Jon Kleinberg, Maithra Raghu (alphabetical order) 2015
In Economics and Computation (EC) 2015. Invited for submission to the journal ACM Transactions on Economics and Computation (TEAC).

Random Walks on Graphs Dissertation (Trinity College)
Maithra Raghu 2013
Awarded the Cambridge Rouse Ball Essay prize.

Invited Talks

TEDx New York Date TBD
AI and Interpretability

NeurIPS Workshop on Human-AI Decision Making 2021
Human-AI collaboration

NYU Machine Learning Seminar 2021
On Vision Transformers and Convolutional Networks

AAAI Workshop 2021
Beyond Performance Measures: Representational Insights for ML Design

STAT Summit 2020
Humans and AI in Healthcare

Weights and Biases 2020
Do Wide and Deep Neural Nets Learn the Same Things?

Simons Institute, UC Berkeley 2020
Anatomy of Catastrophic Forgetting

Yale 2020
Insights from Deep Representations for Machine Learning Systems and Human Collaboration

MSR AI Breakthroughs 2020
Insights from Deep Representations for Machine Learning Systems and Human Collaboration

RAAIS 2020
Insights from Deep Representations for Machine Learning Systems

NYU 2020
Insights from Deep Representations for Machine Learning Systems and Human Collaboration

MIT 2020
Insights from Deep Representations for Machine Learning Systems and Human Collaboration

NVDIA GTC (Keynote) 2020
Insights from Deep Representations for Machine Learning Systems and Human Collaboration

Harvard 2020
Insights from Deep Representations for Machine Learning Systems and Human Collaboration

Stanford 2020
Insights from Deep Representations for Machine Learning Systems and Human Collaboration

Workshop on Theory of Deep Learning, IAS Princeton 2019
Understanding Transfer Learning for Medical Imaging

Artifial Intelligence, O'Reilly Media 2019
Artificial And Human Intelligence in Healthcare

DeepMind 2019
Insights on Deep Representations and Applications to Healthcare

OpenAI 2019
Insights on Deep Representations and Applications to Healthcare

Frontiers of Deep Learning, Simons Institute UC Berkeley 2019
Understanding Transfer Learning with Applications to Medicine

HealthAI, Stanford 2019
Artificial And Human Intelligence in Healthcare

REWORK: Deep Learning Summit, San Francisco 2019
Direct Uncertainty Prediction for Medical Second Opinions

TTIC: Young Researchers Seminar 2018
Insights on Deep Representations and Applications to Healthcare

REWORK: Deep Learning in Healthcare Summit 2018
Direct Uncertainty Prediction for Medical Second Opinions

Simons Institute 2018
Insights from Deep Representations

Facebook AI Research 2018
Insights from Deep Representations with Applications to Healthcare

UMass Amherst 2018
Insights from Deep Representations

Institute of Advanced Study, Princeton 2017
Understanding Generalization in Reinforcement Learning

Columbia University 2017
Understanding Generalization in Reinforcement Learning

OpenAI 2017
Analyzing and Interpreting Deep Representations

Harvard University 2017
Analyzing and Interpreting Deep Representations

Massachusetts Institute of Technology 2017
Neural Network Learning Dynamics

REWORK: Deep Learning Summit 2017
Neural Network Learning Dynamics

World Wide Web Conference 2017
Modelling Sequential Data with Linear Additive Markov Processes

DeepMind 2016
Better interpretability for deep networks.

Women in Machine Learning (WiML) 2016
On the Expressive Power of Deep Neural Networks

New York University 2016
On the Expressive Power of Deep Neural Networks

Northstar Science Film Panel for Arrival 2016
Success of modern methods of machine translation, and women in STEM

Janelia Workshop on Machine Learning and Computer Vision 2016
Interpreting results from Deep Neural Architectures

Economics and Computation 2015
Evaluating Team Performance with Tests

Program Committees and Peer Review

Neural Information Processing Systems (NeurIPS) 2019
International Conference on Machine Learning (ICML) 2019
International Conference on Learning Representations Workshop (ICLR) 2019
Neural Information Processing Systems (NeurIPS) 2018
Conference on Learning Theory (COLT) 2018
International Conference on Machine Learning (ICML) 2018
International Conference on Learning Representations Workshop (ICLR) 2018
International Conference on Learning Representations (ICLR) 2018
NeurIPS Workshop Deep Learning: Bridging Theory and Practice 2017
Women in Machine Learning (WiML) 2017
Neural Information Processing Systems (NeurIPS) 2017
International Conference on Machine Learning (ICML) 2017
International Conference on Learning Representations (ICLR) 2017
Neural Information Processing Systems (NeurIPS) 2016
Women in Machine Learning 2015

Press Coverage

STAT: Human-AI in Healthcare
Fortune: Deep Learning for Scientific Discovery
VentureBeat: Transfer Learning for Medical Imaging
Quanta: The Foundations of Neural Networks
Forbes 30 Under 30 (Science)
Podcast on Talking Machines
WIRED: Interpreting Deep Neural Networks
Quartz: Principled tools to study deep learning
Northstar Science Film Festival
Washington Post: Algorithmic Team Selection and Diversity

Misc.

Ask Me Anything Reddit (r/machinelearning) 2016
with the Google Brain Team, answering questions on Machine Learning and Deep Learning, and academia and industry

Cambridge University Mathematics Society 2013 – 2014

United Kingdom Mathematics Trust, Senior Mentor 2010 – 2012
Mentored talented students for the national mathematical olympiads