NeurIPS 2019 Workshop book

Workshop organizers make last-minute changes to their schedule.
Download this document again to get the lastest changes, or use the NeurIPS mobile application.

Schedule Highlights

Dec. 14, 2019

West Ballroom C, “Do the right thing”: machine learning and causal inference for improved decision making Santacatterina, Joachims, Kallus, Swaminathan, Sontag, Zhou

The Optimization Foundations of Reinforcement Learning Dai, He, Le Roux, Li, Schuurmans, White

Document Intelligence Duffy, Akkiraju, Bedrax Weiss, Bennett, Motahari-Nezhad

Science meets Engineering of Deep Learning Sagun, Gulcehre, Romero, Rostamzadeh, de Freitas

Bridging Game Theory and Deep Learning Miliagkas, Gidel, He, Askari Hemmat, Haghitalab, Lacoste-Julien

NeurIPS Workshop on Machine Learning for Creativity and Design 3.0 Elliott, Dieleman, Roberts, Engel, White, Fiebrink, Mital, Payne, Tokui

ML For Systems Hashemi, Mirhoseini, Goldie, Swersky, Xu, Raiman

Joint Workshop on AI for Social Good Fang, Bullock, Dihac, Green, saltiel, Adjodah, Clark, McGregor, Luck, Penn, Sylvain, Boucher, Swaine-Simon, Tadesse, Cité, Bethke, Bengio

Machine Learning with Guarantees London, Dziugaite, Roy, Joachims, Madry, Shawe-Taylor

Learning Transferable Skills Mattar, Juliani, Lange, Crosby, Beyret

Learning with Temporal Point Processes Rodriguez, Song, Valera, Liu, De, Zha

Context and Compositionality in Biological and Artificial Neural Systems Turek, Jain, Huth, Wehbe, Strubell, Yuille, Linzen, Honey, Cho

Deep Reinforcement Learning Abbeel, Finn, Pineau, Silver, Singh, Achiam, Florensa, Grimmer, Tang, Veeriah

Robot Learning: Control and Interaction in the Real World Calandra, Rakelly, Kamarth, Kragic, Schaaf, Wulfmeier

Program Transformations for ML Lamblin, Baydin, Wiltschko, van Merriënboer, Fertig,Pearlmutter, Duvenaud, Hascoet

Machine Learning for Autonomous Driving McCallister, Rhinehart, Yu, Li, Dragan

Medical Imaging meets NeurIPS Lombaert, Glover, Konukoglu, de Bruijne, Feragen, Oguz, Teuwen

Tackling Climate Change with ML Rolnick, Danti, Kaack, Lacoste, Maharaj, Ng, Platt, Chayes, Bengio

Emergent Communication: Towards Natural Language Gupta, Noukhovitch, Resnick, Jaques, Filos, Ossenkopf, Lazaridou, Foerster, Lowe, Kiela, Cho

The third Conversational AI workshop – today’s practice and tomorrow’s potential Geramifard, Williams, Byrne, Celikyilmaz, Gasic, Hakkani-Tur, Henderson, Lastras, Ostendorf

Sets and Partitions Monath, Zaheer, McCallum, Kobren, Oliva, Poczos, Salakhutdinov

Fair ML in Healthcare Joshi, Chen, Obermeyer, Mullainathan

Machine Learning and the Physical Sciences Baydin, Carrasquilla, Ho, Kashinath, Paganini, Thais, Anandkumar, Cranmer, Melko, Prabhat, Wood

Real Neurons & Hidden Units: future directions at the intersection of neuroscience and AI Lajoie, Shizerman, Puelma Touzel, Thompson, Kording

Privacy in Machine Learning (PriML) Balie, Chaudhuri, Honkela, Koskela, Meehan, Park, Smart, Weller

Dec. 13, 2019

Solving inverse problems with deep networks: New architectures, theoretical foundations, and applications Heckel, Hand, Baraniuk, Bruna, Dimakis, Needell

Perception as generative reasoning: structure, causality, probability Rosenbaum, Garcelo, Battaglia, Allen, Yildirim

Workshop on Human-Centric Machine Learning Angelov, Oliver, Wellner, Rodriguez, Valera, Chiappa, Heidari, Kilbertus

Shared Visual Representations in Human and Machine Intelligence Deza, Peterson, Murty, Griffiths


Workshop on Federated Learning for Data Privacy and Confidentiality Fan, Konečny, Liu, Mcmahan, Smith, Yu

Meta-Learning Calandra, Clavera Gilaberte, Hutter, Vanschoren, Wang

Visually Grounded Interaction and Language Strub, Das, Wijmans, de Vries, Lee, Suhr, Arad Hudson

Robust AI in Financial Services: Data, Fairness, Explainability, Trustworthiness, and Privacy Oprea, Gal, Moulinier, Chen, Veloso, Kumar, Faruquie

Optimal Transport for Machine Learning Cuturi, Peyré, Flamary, Suvorikova

MLSys: Workshop on Systems for ML Lakshmiratan, Sen, Gonzalez, Crankshaw, Bird

Minding the Gap: Between Fairness and Ethics Rubinov, Kondor, Poulson, Warmuth, Moss, Hagerty

CiML 2019: Machine Learning Competitions for All Mendrik, Tu, Guyon, Viegas, LI


Graph Representation Learning Hamilton, van den Berg, Bronstein, Jegelka, Kipf, Leskovec, Liao, Sun, Veličković

Information Theory and Machine Learning Zhao, Song, Han, Choi, Kalluri, Poole, Dimakis, Jiao, Weissman, Ermon


Learning with Rich Experience: Integration of Learning Paradigms Hu, Wilson, Finn, Lee, Berg-Kirkpatrick, Salakhutdinov, Xing

Machine Learning for the Developing World (ML4D): Challenges and Risks De-Arteaga, Coston, Afonja

Biological and Artificial Reinforcement Learning Chua, Zannone, Behbahani, Ponte Costa, Clopath, Richards, Precup

Beyond first order methods in machine learning systems Kyrillidis, Berahas, Roosta, Mahoney

Bayesian Deep Learning Gal, Hernández-Lobato, Louizos, Nalisnick, Ghahramani, Murphy, Welling

AI for Humanitarian Assistance and Disaster Response Gupta, Murphy, Darrell, Heim, Wang, Goodman, Billski

Safety and Robustness in Decision-making Ghavamzadeh, Mannor, Yue, Petrik, Chow
“Do the right thing”: machine learning and causal inference for improved decision making

Michele Santacatterina, Thorsten Joachims, Nathan Kallus, Adith Swaminathan, David Sontag, Angela Zhou

West Ballroom C, Sat Dec 14, 08:00 AM

In recent years, machine learning has seen important advances in its theoretical and practical domains, with some of the most significant applications in online marketing and commerce, personalized medicine, and data-driven policy-making. This dramatic success has led to increased expectations for autonomous systems to make the right decision at the right target at the right time. This gives rise to one of the major challenges of machine learning today that is the understanding of the cause-effect connection. Indeed, actions, intervention, and decisions have important consequences, and so, in seeking to make the best decision, one must understand the process of identifying causality. By embracing causal reasoning autonomous systems will be able to answer counterfactual questions, such as “What if I had treated a patient differently?”, and “What if I had ranked a list differently?” thus helping to establish the evidence base for important decision-making processes.

The purpose of this workshop is to bring together experts from different fields to discuss the relationships between machine learning and causal inference and to discuss and highlight the formalization and algorithmization of causality toward achieving human-level machine intelligence.

This purpose will guide the makeup of the invited talks and the topics for the panel discussions. The panel discussions will tackle controversial topics, with the intent of drawing out an engaging intellectual debate and conversation across fields.

This workshop will lead to advance and extend knowledge on how machine learning could be used to conduct causal inference, and how causal inference could support the development of machine learning methods for improved decision-making.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:45 AM</td>
<td>Opening Remarks</td>
<td>Joachims, Kallus, Santacatterina, Swaminathan, Sontag, Zhou</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Susan Athey</td>
<td>Athey</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Andrea Rotnitzky</td>
<td>Rotnitzky</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Poster Spotlights</td>
<td></td>
</tr>
<tr>
<td>10:15 AM</td>
<td>Coffee break, posters, and 1-on-1 discussions</td>
<td>Lu, Chen, Namkoong, Chapignon, Rudolph, Coston, von Kügelgen, Prasad, Dhillon, Xu, Wang, Markham, Rohde, Singh, Zhang, Hassanpour, Sharma, Lee, Pouget-Abadie, Krijthe, Mahajan, Ke, Wirsberger, Semenova, Mykhaylov, Shen, Takatsu, Sun, Yang, Franks</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Susan Murphy</td>
<td>Murphy</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Ying-Qi Zhao</td>
<td>Zhao</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Tentative topic: Reasoning about untestable assumptions in the face of unknowable counterfactuals</td>
<td></td>
</tr>
<tr>
<td>12:45 PM</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>02:30 PM</td>
<td>Susan Shortreed</td>
<td>Shortreed</td>
</tr>
<tr>
<td>03:00 PM</td>
<td>Contributed talk 1</td>
<td></td>
</tr>
<tr>
<td>03:15 PM</td>
<td>Contributed talk 2</td>
<td></td>
</tr>
<tr>
<td>03:30 PM</td>
<td>Poster Spotlights</td>
<td></td>
</tr>
<tr>
<td>03:45 PM</td>
<td>Coffee break, posters, and 1-on-1 discussions</td>
<td>von Kügelgen, Rohde, Schumann, Charles, Veitch, Semenova, Demirer, Syrgkanis, Nair, Puli, Uehara, Gopalan, Ding, Ng, Khosravi, Sherman, Zeng, Wieczorek, Liu, Gan, Hartford, Oprescu, D’Amour, Boehnke, Sailo, Griveau-Billion, Modi, Karimov, Berrevoets, Graham, Josse, Sridhar, Dahabreh, Mishler, Wadsworth, Qureshi, Ladhania, Morishita</td>
</tr>
<tr>
<td>04:30 PM</td>
<td>Tentative topic: How machine learning, and causal inference work together: cross-pollination and new challenges.</td>
<td></td>
</tr>
<tr>
<td>05:15 PM</td>
<td>Closing Remarks</td>
<td></td>
</tr>
</tbody>
</table>

Abstracts (1):

Abstract 8: Tentative topic: Reasoning about untestable assumptions in the face of unknowable counterfactuals in “Do the right thing”: machine learning and causal inference for improved decision making, 12:00 PM

Tentative topic: Reasoning about untestable assumptions in the face of unknowable counterfactuals

The Optimization Foundations of Reinforcement Learning
Interest in reinforcement learning (RL) has boomed with recent improvements in benchmark tasks that suggest the potential for a revolutionary advance in practical applications. Unfortunately, research in RL remains hampered by limited theoretical understanding, making the field overly reliant on empirical exploration with insufficient principles to guide future development. It is imperative to develop a stronger fundamental understanding of the success of recent RL methods, both to expand the useability of the methods and accelerate future deployment. Recently, fundamental concepts from optimization and control theory have provided a fresh perspective that has led to the development of sound RL algorithms with provable efficiency. The goal of this workshop is to catalyze the growing synergy between RL and optimization research, promoting a rational reconsideration of the foundational principles for reinforcement learning, and bridging the gap between theory and practice.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00 AM</td>
<td>Opening Remarks</td>
<td>Dai, He, Le Roux, Li, Schuurmans, White</td>
</tr>
<tr>
<td>08:10 AM</td>
<td>Plenary Talk</td>
<td>Agrawal</td>
</tr>
<tr>
<td>08:50 AM</td>
<td>Contributed Talk</td>
<td></td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Poster and Coffee Break</td>
<td></td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Plenary Talk</td>
<td>Kakade</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Contributed Talk</td>
<td></td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Poster Spotlight</td>
<td></td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Plenary Talk</td>
<td>Van Roy</td>
</tr>
</tbody>
</table>

Abstracts (7):

Abstract 2: Plenary Talk in The Optimization Foundations of Reinforcement Learning. Agrawal 08:10 AM

TBA


TBA

Abstract 9: Plenary Talk in The Optimization Foundations of Reinforcement Learning. Van Roy 02:00 PM

TBA

Abstract 10: Plenary Talk in The Optimization Foundations of Reinforcement Learning. Wang 02:40 PM

TBA
Document Intelligence

Nigel Duffy, Rama Akkiraju, Tania Bedrax Weiss, Paul Bennett, Hamid Reza Motahari-Nezhad

Sat Dec 14, 08:00 AM

Business documents are central to the operation of business. Such documents include sales agreements, vendor contracts, mortgage terms, loan applications, purchase orders, invoices, financial statements, employment agreements and a wide many more. The information in such business documents is presented in natural language, and can be organized in a variety of ways from straight text, multi-column formats, and a wide variety of tables. Understanding these documents is made challenging due to inconsistent formats, poor quality scans and OCR, internal cross references, and complex document structure. Furthermore, these documents often reflect complex legal agreements and reference, explicitly or implicitly, regulations, legislation, case law and standard business practices.

The ability to read, understand and interpret business documents, collectively referred to here as “Document Intelligence”, is a critical and challenging application of artificial intelligence (AI) in business. While a variety of research has advanced the fundamentals of document understanding, the majority have focused on documents found on the web which fail to capture the complexity of analysis and types of understanding needed across business documents. Realizing the vision of document intelligence remains a research challenge that requires a multi-disciplinary perspective spanning not only natural language processing and understanding, but also computer vision, knowledge representation and reasoning, information retrieval, and more -- all of which have been profoundly impacted and advanced by neural network-based approaches and deep learning in the last few years.

We propose to organize a workshop for AI researchers, academics and industry practitioners to discuss the opportunities and challenges for document intelligence.

Schedule

08:00 AM  Opening Remarks
08:30 AM  David Lewis
09:00 AM  Ndapa Nakashole
10:00 AM  Coffee Break
Bridging Game Theory and Deep Learning

Abstracts (1):


Since we are a small workshop, we will hold the poster sessions during the day, including all the breaks as the authors wish.
Nonconvex-Nonconcave Minimax Optimization? in Bridging Game Theory and Deep Learning, Jin 11:30 AM

Minimax optimization has found extensive applications in modern machine learning, in settings such as generative adversarial networks (GANs), adversarial training and multi-agent reinforcement learning. As most of these applications involve nonconvex-nonconcave formulations, a very basic question arises—"what is a proper definition of local optima?"

Most previous work answers this question using classical notions of equilibria from simultaneous games, where the min-player and the max-player act simultaneously. In contrast, most applications in machine learning, including GANs and adversarial training, correspond to sequential games, where the order of which player acts first is crucial (since minimax is in general not equal to maximin due to the nonconvex-nonconcave nature of the problems). The main contribution of this paper is to propose a proper mathematical definition of local optimality for this sequential setting—local minimax, as well as to present its properties and existence results. Finally, we establish a strong connection to a basic local search algorithm—gradient descent ascent (GDA): under mild conditions, all stable limit points of GDA are exactly local minimax points up to some degenerate points.

Contributed talk: Exploiting Uncertain Real-Time Information from Deep Learning in Signaling Games for Security and Sustainability, Bondi 03:00 PM

This paper investigates the convergence of learning dynamics in Stackelberg games on continuous action spaces, a class of games distinguished by the hierarchical order of play between agents. We establish connections between the Nash and Stackelberg equilibrium concepts and characterize conditions under which attractors of simultaneous gradient descent are Stackelberg equilibria in zero-sum games. Moreover, we show that the only stable attractors of the Stackelberg gradient dynamics are Stackelberg equilibria in zero-sum games. Using this insight, we develop two-timescale learning dynamics that converge to Stackelberg equilibria in zero-sum games and the set of stable attractors in general-sum games.

Contributed talk: On Solving Local Minimax Optimization: A Follow-the-Ridge Approach in Bridging Game Theory and Deep Learning, Wang 02:30 PM

Many tasks in modern machine learning can be formulated as finding equilibria in sequential games. In particular, two-player zero-sum sequential games, also known as minimax optimization, have received growing interest. It is tempting to apply gradient descent to solve minimax optimization given its popularity in supervised learning. However, we note that naive application of gradient descent fails to find local minimax -- the analogy of local minima in minimax optimization, since the fixed points of gradient dynamics might not be local minimax. In this paper, we propose a Follow-the-Ridge (FR) algorithm that locally converges to and only converges to local minimax. We show theoretically that the algorithm addresses the limit cycling problem around fixed points, and is compatible with preconditioning and positive momentum. Empirically, FR solves quadratic minimax problems and improves GAN training on simple tasks.

Contributed talk: Exploiting Uncertain Real-Time Information from Deep Learning in Signaling Games for Security and Sustainability in Bridging Game Theory and Deep Learning, Bondi 03:00 PM

Motivated by real-world deployment of drones for conservation, this paper advances the state-of-the-art in security games with signaling. The well-known defender-attacker security games framework can help in planning for such strategic deployments of sensors and human patrollers, and warning signals to ward off adversaries. However, we show that defenders can suffer significant losses when ignoring real-world uncertainties, such as detection uncertainty resulting from imperfect deep learning models, despite carefully planned security game strategies with signaling. In fact, defenders may perform worse than forgoing drones completely in this case. We address this shortcoming by proposing a novel game model that integrates signaling and sensor uncertainty; perhaps surprisingly, we show that defenders can still perform well via a signaling strategy that exploits the uncertain real-time information primarily from deep learning models. For example, even in the presence of uncertainty, the defender still has an informational advantage in knowing that she has or has not actually detected the attacker; and she can design a signaling scheme to "mislead" the attacker who is uncertain as to whether he has been detected. We provide a novel algorithm, scale-up techniques, and experimental results from simulation based on our ongoing deployment of a conservation drone system in South Africa.
Generative machine learning and machine creativity have continued to grow and attract a wider audience to machine learning. Generative models enable new types of media creation across images, music, and text - including recent advances such as StyleGAN, MuseNet and GPT-2. This one-day workshop broadly explores issues in the applications of machine learning to creativity and design. We will look at algorithms for generation and creation of new media, engaging researchers building the next generation of generative models (GANs, RL, etc). We investigate the social and cultural impact of these new models, engaging researchers from HCI/UX communities and those using machine learning to develop new creative tools. In addition to covering the technical advances, we also address the ethical concerns ranging from the use of biased datasets to the use of synthetic media such as “DeepFakes”. Finally, we’ll hear from some of the artists and musicians who are adopting machine learning including deep learning and reinforcement learning as part of their own artistic process. We aim to balance the technical issues and challenges of applying the latest generative models to creativity and design with philosophical and cultural issues that surround this area of research.

### Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:15 AM</td>
<td>Welcome and Introduction</td>
</tr>
<tr>
<td>08:30 AM</td>
<td>Yann LeCun</td>
</tr>
<tr>
<td>08:45 AM</td>
<td>Sougwen Chung</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Al Art Gallery Overview</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Giorgio Patrini</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Neural Painters: A learned differentiable constraint for generating brushstroke paintings</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Transform the Set: Memory Attentive Generation of Guided and Unguided Image Collages</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Deep reinforcement learning for 2D soft body locomotion</td>
</tr>
<tr>
<td>11:40 AM</td>
<td>Towards Sustainable Architecture: 3D Convolutional Neural Networks for Computational Fluid Dynamics Simulation and Reverse Design Workflow</td>
</tr>
<tr>
<td>11:50 AM</td>
<td>Human and GAN collaboration to create haute couture dress</td>
</tr>
<tr>
<td>01:30 PM</td>
<td>Poster Session 1</td>
</tr>
<tr>
<td>02:30 PM</td>
<td>Alec Radford</td>
</tr>
<tr>
<td>03:45 PM</td>
<td>Claire Evans</td>
</tr>
<tr>
<td>04:15 PM</td>
<td>MidiMe: Personalizing a MusicVAE model with user data</td>
</tr>
<tr>
<td>04:25 PM</td>
<td>First Steps Towards Collaborative Poetry Generation</td>
</tr>
<tr>
<td>04:35 PM</td>
<td>Panel Discussion</td>
</tr>
<tr>
<td>05:00 PM</td>
<td>Poster Session 2</td>
</tr>
</tbody>
</table>

---

**ML For Systems**

*Milad Hashemi, Azalia Mirhoseini, Anna Goldie, Kevin Swersky, Xinlei Xu, Jonathan Raiman*

Sat Dec 14, 08:00 AM

Compute requirements are growing at an exponential rate, and optimizing these computer systems often involves complex high-dimensional combinatorial problems. Yet, current methods rely heavily on heuristics. Very recent work has outlined a broad scope where machine learning vastly outperforms these traditional heuristics: including scheduling, data structure design, microarchitecture, compilers, circuit design, and the control of warehouse scale computing systems. In order to continue to scale these computer systems, new learning approaches are needed. The goal of this workshop is to develop novel machine learning methods to optimize and accelerate software and hardware systems.

Machine Learning for Systems is an interdisciplinary workshop that brings together researchers in computer architecture and systems and machine learning. This workshop is meant to serve as a platform to promote discussions between researchers in the workshops target areas.

This workshop is part two of a two-part series with one day focusing on ML for Systems and the other on Systems for ML. Although the two workshops are being led by different organizers, we are coordinating our call for papers to ensure that the workshops complement each other and that submitted papers are routed to the appropriate venue.

### Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:15 AM</td>
<td>Welcome and Introduction</td>
</tr>
<tr>
<td>08:30 AM</td>
<td>Yann LeCun</td>
</tr>
<tr>
<td>08:45 AM</td>
<td>Sougwen Chung</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Al Art Gallery Overview</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Giorgio Patrini</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Neural Painters: A learned differentiable constraint for generating brushstroke paintings</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Transform the Set: Memory Attentive Generation of Guided and Unguided Image Collages</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Deep reinforcement learning for 2D soft body locomotion</td>
</tr>
<tr>
<td>11:40 AM</td>
<td>Towards Sustainable Architecture: 3D Convolutional Neural Networks for Computational Fluid Dynamics Simulation and Reverse Design Workflow</td>
</tr>
<tr>
<td>11:50 AM</td>
<td>Human and GAN collaboration to create haute couture dress</td>
</tr>
<tr>
<td>01:30 PM</td>
<td>Poster Session 1</td>
</tr>
<tr>
<td>02:30 PM</td>
<td>Alec Radford</td>
</tr>
<tr>
<td>03:45 PM</td>
<td>Claire Evans</td>
</tr>
<tr>
<td>04:15 PM</td>
<td>MidiMe: Personalizing a MusicVAE model with user data</td>
</tr>
<tr>
<td>04:25 PM</td>
<td>First Steps Towards Collaborative Poetry Generation</td>
</tr>
<tr>
<td>04:35 PM</td>
<td>Panel Discussion</td>
</tr>
<tr>
<td>05:00 PM</td>
<td>Poster Session 2</td>
</tr>
</tbody>
</table>
Joint Workshop on AI for Social Good

Fei Fang, Joseph Bullock, Marc-Antoine Dilhac, Brian Green, natalie saltiel, Dhaval Adjodah, Jack Clark, Sean McGregor, Margaux Luck, Jonnie Penn, Tristan Sylvain, Geneviève Boucher, Sydney Swaine-Simon, Girmaw Abebe Tadesse, Myriam Côté, Anna Bethke, Yoshua Bengio

Sat Dec 14, 08:00 AM

The accelerating pace of intelligent systems research and real world deployment presents three clear challenges for producing "good" intelligent systems: (1) the research community lacks incentives and venues for results centered on social impact, (2) deployed systems often produce unintended negative consequences, and (3) there is little consensus for public policy that maximizes "good" social impacts, while minimizing the likelihood of harm. As a result, researchers often find themselves without a clear path to positive real world impact.

The Workshop on AI for Social Good addresses these challenges by bringing together machine learning researchers, social impact leaders, ethicists, and public policy leaders to present their ideas and applications for maximizing the social good. This workshop is a collaboration of three formerly separate lines of research (i.e., this is a "joint" workshop), including researchers in applications-driven AI research, applied ethics, and AI policy. Each of these research areas are unified into a 3-track framework promoting the exchange of ideas between the practitioners of each track.

We hope that this gathering of research talent will inspire the creation of new approaches and tools, provide for the development of intelligent systems benefiting all stakeholders, and converge on public policy mechanisms for encouraging these goals.

Schedule

08:00 AM  Opening remarks  Bengio

08:05 AM  Track 1: Producing Good Outcomes  Dietterich, Gomes, Luengo-Oroz, Dilkina, Cornebise

10:30 AM  Break

11:00 AM  Track 1: Producing Good Outcomes  Abebe, Kleinberg, Lucier, Cuthbertson, Mathewson, Schumann, Chi, Babirye, Lim, Rane, Owoeye, Da San Martino, Kimura, Rutkowski, Fruehwirt, Rho, Chraptignon, Konya, Ben Daya, Thomas, Abdulrahim, Ssendiwalala, Namanya, Akera, Manandhar, Greeff, Verma, Nyman, Kermode, Narain, Johnson, Yanagiwara, Sugiyama, Sharma, Dey, Sarbaja, Govindaraj, Cornebise, Dulhanty, Deglint, Blich, Masood, Varga, Gomes, Dietterich, Luengo-Oroz, Dilkina, Mironova

12:00 PM  Lunch - on your own

02:00 PM  Track 2: From Malicious Use to Responsible AI  Dobbe, Shamout, Clifton, Whittlestone, Kinsey, Elhalal, Bajaj, Wall, Tomasev, Green

03:00 PM  Break

03:30 PM  Track 2: From Malicious Use to Responsible AI  Yang, Lin, Tomasev, Raicu, Vincent

04:00 PM  Track 3: Public Policy  Sun, Veeramachaneni, Cuesta-Infante, Etzayn, Gamper, Schim van der Loeff, Green

Machine Learning with Guarantees

Ben London, Gintare Karolina Dziugaite, Dan Roy, Thorsten Joachims, Aleksander Madry, John Shawe-Taylor

Sat Dec 14, 08:00 AM
As adoption of machine learning grows in high-stakes application areas (e.g., industry, government and health care), so does the need for guarantees: how accurate a learned model will be; whether its predictions will be fair; whether it will divulge information about individuals; or whether it is vulnerable to adversarial attacks. Many of these questions involve unknown or intractable quantities (e.g., risk, regret or posterior likelihood) and complex constraints (e.g., differential privacy, fairness, and adversarial robustness). Thus, learning algorithms are often designed to yield (and optimize) bounds on the quantities of interest. Beyond providing guarantees, these bounds also shed light on black-box machine learning systems.

Classical examples include structural risk minimization (Vapnik, 1991) and support vector machines (Cristianini & Shawe-Taylor, 2000), while more recent examples include non-vacuous risk bounds for neural networks (Dziugaite & Roy, 2017, 2018), algorithms that optimize both the weights and structure of a neural network (Cortes, 2017), counterfactual risk minimization for learning from logged bandit feedback (Swaminathan & Joachims, 2015; London & Sandler, 2019), robustness to adversarial attacks (Schmidt et al., 2018; Wong & Kolter, 2018), differentially private learning (Dwork et al., 2006, Chaudhuri et al., 2011), and algorithms that ensure fairness (Dwork et al., 2012).

This one-day workshop will bring together researchers in both theoretical and applied machine learning, across areas such as statistical learning theory, adversarial learning, fairness and privacy, to discuss the problem of obtaining performance guarantees and algorithms to optimize them. The program will include invited and contributed talks, poster sessions and a panel discussion. We particularly welcome contributions describing fundamentally new problems, novel learning principles, creative bound optimization techniques, and empirical studies of theoretical findings.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
<th>Presenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:45 AM</td>
<td>Welcome Address</td>
<td>London</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>TBD</td>
<td>Roth</td>
</tr>
<tr>
<td>09:45 AM</td>
<td>TBD</td>
<td></td>
</tr>
</tbody>
</table>

Abstracts (2):


10:15 AM  **Break / Poster Session 1**

11:00 AM  **Mehryar Mohri**

11:45 AM  TBD

12:15 PM  Lunch Break

02:00 PM  **Po-Ling Loh**

02:45 PM  **TBD**

03:15 PM  **Coffee Break / Poster Session 2**

04:00 PM  **Emma Brunskill**

04:45 PM  **TBD**

05:15 PM  **Discussion Panel**

Presenters without NIPS accounts: jindong.gu@siemens.com

Abstract 10: **Coffee Break / Poster Session 2 in Machine Learning with Guarantees**, 03:15 PM

Same presenters as Poster Session 1
Learning Transferable Skills

Marwan Mattar, Arthur Juliani, Danny Lange, Matthew Crosby, Benjamin Beyret

Sat Dec 14, 08:00 AM

After spending several decades on the margin of AI, reinforcement learning has recently emerged as a powerful framework for developing intelligent systems that can solve complex tasks in real-world environments. This has had a tremendous impact on a wide range of tasks ranging from playing games such as Go and StarCraft to learning dexterity. However, one attribute of intelligence that still eludes modern learning systems is generalizability. Until very recently, the majority of reinforcement learning research involved training and testing algorithms on the same, sometimes deterministic, environment. This has resulted in algorithms that learn policies that typically perform poorly when deployed in environments that differ, even slightly, from those they were trained on. Even more importantly, the paradigm of task-specific training results in learning systems that scale poorly to a large number of (even interrelated) tasks.

Recently there has been an enduring interest in developing learning systems that can learn transferable skills. This could mean robustness to changing environment dynamics, the ability to quickly adapt to environment and task variations or the ability to learn to perform multiple tasks at once (or any combination thereof). This interest has also resulted in a number of new data sets and challenges (e.g. Obstacle Tower Environment, Animal-AI, CoinRun) and an urgency to standardize the metrics and evaluation protocols to better assess the generalization abilities of novel algorithms. We expect this area to continue to increase in popularity and importance, but this can only happen if we manage to build consensus on which approaches are promising, and, equally important, how to test them.

The workshop will include a mix of invited speakers, accepted papers (oral and poster sessions) and a panel discussion. The workshop welcomes both theoretical and applied research, in addition to novel data sets and evaluation protocols.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 AM</td>
<td>Opening Remarks</td>
<td>Mattar, Juliani, Crosby, Beyret, Lange</td>
</tr>
<tr>
<td>09:15 AM</td>
<td>Raia Hadsell (DeepMind)</td>
<td>Hadsell</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Environments and Data Sets</td>
<td>Cobbe, De Fabritiis</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>11:15 AM</td>
<td>Vladlen Koltun (Intel)</td>
<td>Koltun</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>01:30 PM</td>
<td>Oral Presentations</td>
<td>Grau-Moya, Pascual-Diaz, Marinier, Pietquin, Efros, Isola, Darrell, Lu, Pathak, Ferret</td>
</tr>
<tr>
<td>02:15 PM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Learning with Temporal Point Processes

Manuel Rodriguez, Le Song, Isabel Valera, Yan Liu, Abir De, Hongyuan Zha

Sat Dec 14, 08:00 AM

In recent years, there has been an increasing number of machine learning models and algorithms based on the theory of temporal point processes, which is a mathematical framework to model asynchronous event data. These models and algorithm have found a wide range of human-centered applications, from social and information networks and recommender systems to crime prediction and health. Moreover, this emerging line of research has already established connections to deep learning, deep generative models, Bayesian nonparametrics, causal inference, stochastic optimal control and reinforcement learning. However, despite these recent advances, learning with temporal point processes is still a relatively niche topic within the machine learning community—there are only a few research groups across the world with the necessary expertise to make progress. In this workshop, we aim to popularize temporal point processes within the machine learning community at large. In our view, this is the right time to organize such a workshop because, as algorithmic decisions becomes more consequential to individuals and society, temporal point processes will play a major role on the development of human-centered machine learning models and algorithms accounting for the feedback loop between algorithmic and human decisions, which are inherently asynchronous events. Moreover, it will be a natural follow up of a very successful and well-attended ICML 2018 tutorial on learning with temporal point processes, which two of us recently taught.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30 AM</td>
<td>Welcome Address and Introduction</td>
<td></td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Opening Remarks</td>
<td>Mattar, Juliani, Crosby, Beyret, Lange</td>
</tr>
<tr>
<td>09:15 AM</td>
<td>Raia Hadsell (DeepMind)</td>
<td>Hadsell</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Environments and Data Sets</td>
<td>Cobbe, De Fabritiis</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>11:15 AM</td>
<td>Vladlen Koltun (Intel)</td>
<td>Koltun</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>01:30 PM</td>
<td>Oral Presentations</td>
<td>Grau-Moya, Pascual-Diaz, Marinier, Pietquin, Efros, Isola, Darrell, Lu, Pathak, Ferret</td>
</tr>
<tr>
<td>02:15 PM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Context and Compositionality in Biological and Artificial Neural Systems

Javier Turek, Shailee Jain, Alexander Huth, Leila Wehbe, Emma Strubell, Alan Yuille, Tal Linzen, Christopher Honey, Kyunghyun Cho

Sat Dec 14, 08:00 AM

The ability to integrate semantic information across narratives is fundamental to language understanding in both biological and artificial cognitive systems. In recent years, enormous strides have been made in NLP and Machine Learning to develop architectures and techniques that effectively capture these effects. The field has moved away from traditional bag-of-words approaches that ignore temporal ordering, and instead embraced RNNs, Temporal CNNs and Transformers, which incorporate contextual information at varying timescales. While these architectures have lead to state-of-the-art performance on many difficult language understanding tasks, it is unclear what representations these networks learn and how exactly they incorporate context. Interpreting these networks, systematically analyzing the advantages and disadvantages of different elements, such as gating or attention, and reflecting on the capacity of the networks across various timescales are open and important questions.

On the biological side, recent work in neuroscience suggests that areas in the brain are organized into a temporal hierarchy in which different areas are not only sensitive to specific semantic information but also to the composition of information at different timescales. Computational neuroscience has moved in the direction of leveraging deep learning to gain insights about the brain. By answering questions on the underlying mechanisms and representational interpretability of these artificial networks, we can also expand our understanding of temporal hierarchies, memory, and capacity effects in the brain.

In this workshop we aim to bring together researchers from machine learning, NLP, and neuroscience to explore and discuss how computational models should effectively capture the multi-timescale, context-dependent effects that seem essential for processes such as language understanding.

We invite you to submit papers related to the following (non-exhaustive) topics:

* Contextual sequence processing in the human brain
* Compositional representations in the human brain
* Systematic generalization in deep learning
* Compositionality in human intelligence
* Compositionality in natural language
* Understanding composition and temporal processing in neural network models
* New approaches to compositionality and temporal processing in language
* Hierarchical representations of temporal information
* Datasets for contextual sequence processing
* Applications of compositional neural networks to real-world problems

Submissions should be up to 4 pages excluding references, and should be NIPS format and anonymous. The review process is double-blind.
We also welcome published papers that are within the scope of the workshop (without re-formatting). This specific papers do not have to be anonymous. They will only have a very light review process.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00 AM</td>
<td>Opening Remarks</td>
<td>Huth</td>
</tr>
<tr>
<td>08:15 AM</td>
<td>Patricia Churchland</td>
<td>Churchland</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Gina Kuperberg</td>
<td>Kuperberg</td>
</tr>
<tr>
<td>09:45 AM</td>
<td>Poster Session + Break</td>
<td></td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Spotlights - TBA</td>
<td></td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Joshua Bengio</td>
<td>Bengio</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Poster Session + Lunch</td>
<td></td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Tom Mitchell</td>
<td>Mitchell</td>
</tr>
<tr>
<td>02:45 PM</td>
<td>Liina Pylkkanen</td>
<td>Pylkkanen</td>
</tr>
<tr>
<td>03:30 PM</td>
<td>Poster Session + Break</td>
<td></td>
</tr>
<tr>
<td>04:15 PM</td>
<td>Panel</td>
<td>Willke, Fedorenko, Lee, Smolensky, Marcus</td>
</tr>
<tr>
<td>05:55 PM</td>
<td>Closing remarks</td>
<td>Webbe</td>
</tr>
</tbody>
</table>

Abstracts (7):

Abstract 1: **Opening Remarks in Context and Compositionality in Biological and Artificial Neural Systems**, Huth 08:00 AM

Note: schedule not final and may change

Abstract 2: **Patricia Churchland in Context and Compositionality in Biological and Artificial Neural Systems**, Churchland 08:15 AM

Note: schedule not final and may change

Abstract 3: **Gina Kuperberg in Context and Compositionality in Biological and Artificial Neural Systems**, Kuperberg 09:00 AM

Note: schedule not final and may change

Abstract 6: **Joshua Bengio in Context and Compositionality in Biological and Artificial Neural Systems**, Bengio 11:00 AM

Note: schedule not final and may change

Abstract 8: **Tom Mitchell in Context and Compositionality in Biological and Artificial Neural Systems**, Mitchell 02:00 PM

Note: schedule not final and may change

Deep Reinforcement Learning

**Pieter Abbeel, Chelsea Finn, Joelle Pineau, David Silver, Satinder Singh, Joshua Achiam, Carlos Florensa, Christopher Grimm, Haoran Tang, Vivek Veeriah**

Sat Dec 14, 08:00 AM

In recent years, the use of deep neural networks as function approximators has enabled researchers to extend reinforcement learning techniques to solve increasingly complex control tasks. The emerging field of deep reinforcement learning has led to remarkable empirical results in rich and varied domains like robotics, strategy games, and multiagent interaction. This workshop will bring together researchers working at the intersection of deep learning and reinforcement learning, and it will help interested researchers outside of the field gain a high-level view about the current state of the art and potential directions for future contributions.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:45 AM</td>
<td>Welcome Comments</td>
<td></td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Invited Talk</td>
<td>Vinyals</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Contributed Talks</td>
<td></td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Invited Talk</td>
<td>Todorov</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Invited Talk</td>
<td>Brunskill</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Contributed Talks</td>
<td></td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Invited Talk</td>
<td>Fei-Fei</td>
</tr>
<tr>
<td>01:30 PM</td>
<td>Invited Talk</td>
<td>Whiteson</td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Contributed Talks</td>
<td></td>
</tr>
</tbody>
</table>
Abstract 7: Contributed Talks in Deep Reinforcement Learning, 11:30 AM
(Speakers and details forthcoming.)

Abstract 8: Invited Talk in Deep Reinforcement Learning, Fei-Fei Li
12:00 PM
(Speaker and details forthcoming.)

Abstract 9: Invited Talk in Deep Reinforcement Learning, Whiteson
01:30 PM
(Talk title and abstract TBD.)

Abstract 10: Contributed Talks in Deep Reinforcement Learning, 02:00 PM
(Speaker and details forthcoming.)

Abstract 12: NeurIPS RL Competitions Results Presentations in Deep Reinforcement Learning, 04:00 PM

16:00 - 16:15 Learn to Move: Walk Around
16:15 - 16:30 Animal Olympics
16:30 - 16:45 Robot open-Ended Autonomous Learning (REAL)
16:45 - 17:00 MineRL

Abstract 13: Invited Talk in Deep Reinforcement Learning, Littman
05:00 PM
(Talk title and abstract TBD.)

Abstract 14: Panel Discussion in Deep Reinforcement Learning, 05:30 PM
(Topic and panelists TBA.)

Robot Learning: Control and Interaction in the Real World

Roberto Calandra, Kate Rakelly, Sanket Sayaji Kamthe, Danica Kragic, Stefan Schaal, Markus Wulfmeier

Sat Dec 14, 08:00 AM

The growing capabilities of learning-based methods in control and robotics has precipitated a shift in the design of software for autonomous systems. Recent successes fuel the hope that robots will increasingly perform varying tasks working alongside humans in complex, dynamic environments. However, the application of learning approaches to real-world robotic systems has been limited because real-world scenarios introduce challenges that do not arise in simulation. In this workshop, we aim to identify and tackle the main challenges to learning on real robotic systems. First, most machine learning methods rely on large quantities of labeled data. While raw sensor data is available at high rates, the required variety is hard to obtain and the human effort to annotate or design reward functions is an even larger burden. Second, algorithms must guarantee some measure of safety and robustness to be deployed in real systems that interact with property and people. Instantaneous reset mechanisms, as common in simulation to recover from even critical failures, present a great challenge to real robots. Third, the real world is significantly more complex and varied than...
curated datasets and simulations. Successful approaches must scale to this complexity and be able to adapt to novel situations.

Schedule

09:10 AM Invited Talk 1 Deisenroth
09:45 AM Coffee Break
11:30 AM Poster 1 Genc, Clavera Gilaberte, Zimmer, Smith, Xiao, Fu, Ding, Stepputtis, Mallya, Bodapati, Lin
12:00 PM Lunch Break
08:30 AM Opening statements
08:40 AM Jan-Willem van de Meent - TBA
09:30 AM Applications of a disintegration transformation Narayanan
09:50 AM Coffee break
10:30 AM Christine Tasson - TBA Tasson
11:20 AM The Differentiable Curry Vytiniotis
12:00 PM Lunch break & Poster session Considine, Siskind, Innes, Phan, Maclaurin, Zuidberg Dos Martires, Radul, Gowda, Sharma, Sennesh, Kochurov, Piotkin, Wieckl, Kukreja, Shan, Johnson, Belov, Pradhan, Manhaeve, Meert, Kimmig, De Raedt, Patton, Hoffman, A. Saourous, Roy, Bingham, Jankowiak, Carroll, Lao, Paull, Abadi, ROJAS JIMENEZ, Chen
02:00 PM Zachary DeVito - TBA DeVito
02:50 PM Skye Wanderman-Mline - TBA Wanderman-Mline
03:40 PM Coffee break
04:20 PM Generalized Abs-Linear Learning Griewank
04:40 PM Towards Polyhedral Automatic Differentiation Hueckelheim
05:00 PM Taylor-Mode Automatic Differentiation for Higher-Order Derivatives in JAX Bettencourt
05:20 PM Panel and general discussion

Program Transformations for ML

Pascal Lamblin, Atilim Gunes Baydin, Alexander Wiltshcko, Bart van Merriënboer, Emily Fertig, Barak Pearlmutter, David Duvenaud, Laurent Hascoet

Sat Dec 14, 08:00 AM

Machine learning researchers often express complex models as a program, relying on program transformations to add functionality. New languages and transformations (e.g., TorchScript and TensorFlow AutoGraph) are becoming core capabilities of ML libraries. However, existing transformations, such as automatic differentiation (AD), inference in probabilistic programming languages (PPL), and optimizing compilers are often built in isolation, and limited in scope. This workshop aims at viewing program transformations in ML in a unified light, making these capabilities more accessible, and building entirely new ones. Program transformations are an area of active study. AD transforms a program performing numerical computation into one computing the gradient of those computations. In PPL, a program describing a sampling procedure can be modified to perform inference on model parameters given observations. Other examples are vectorizing a program expressed on one data point, and learned transformations where ML models use programs as inputs or outputs. This workshop will bring together researchers in the fields of AD, programming languages, compilers, and ML, with the goal of understanding the commonalities between disparate approaches and views, and sharing ways to make these techniques broadly available. It would enable ML practitioners to iterate faster on novel models and architectures (e.g., those naturally expressed through high-level constructs like recursion).

Topics:
— Abstractions and syntax (beyond meta-programming and operator overloading) to naturally express a program (expression, or procedure) as an object to be manipulated.
— Techniques from AD and PPL the ML community could adopt to enable research on new models
— How to overcome challenges due to the ML’s specific hardware (GPUs, specialized chips) and software (Python) stacks, and the particular demands of practitioners for their tools
— Greater collaboration between ML and programming languages communities
Machine Learning for Autonomous Driving

Rowan McAllister, Nick Rhinehart, Fisher Yu, Li Erran Li, Anca Dragan

Sat Dec 14, 08:00 AM

Autonomous vehicles (AVs) provide a rich source of high-impact research problems for the machine learning (ML) community at NeurIPS in diverse fields including computer vision, probabilistic modeling, gesture recognition, pedestrian and vehicle forecasting, human-machine interaction, and multi-agent planning. The common goal of autonomous driving can catalyze discussion between these subfields, generating a cross-pollination of research ideas. Beyond the benefits to the research community, AV research can improve society by reducing road accidents; giving independence to those unable to drive; and inspiring younger generations towards ML with tangible examples of ML-based technology clearly visible on local streets.

As many NeurIPS attendees are key drivers behind AV-applied ML, the proposed NeurIPS 2019 Workshop on Autonomous Driving intends to bring researchers together from both academia and industries to discuss machine learning applications in autonomous driving. Our proposal includes regular paper presentations, invited speakers, and technical benchmark challenges to present the current state of the art, as well as the limitations and future directions for autonomous driving.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:45 AM</td>
<td>Welcome McAllister, Rhinehart, Dragan</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Invited Talk Urtasun</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Contributed Talks</td>
</tr>
<tr>
<td>09:45 AM</td>
<td>Coffee + Posters</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Invited Talk</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Invited Talk</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Invited Talk</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Lunch + Posters</td>
</tr>
<tr>
<td>01:30 PM</td>
<td>Invited Talk</td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Invited Talk</td>
</tr>
<tr>
<td>02:30 PM</td>
<td>Invited Talk</td>
</tr>
<tr>
<td>03:00 PM</td>
<td>Contributed Talks</td>
</tr>
</tbody>
</table>

Medical Imaging meets NeurIPS

Hervé Lombaert, Ben Glocker, Ender Konukoglu, Marleen de Bruijne, Aasa Feragen, Ipek Oguz, Jonas Teuwen

Sat Dec 14, 08:00 AM

Medical imaging and radiology are facing a major crisis with an ever-increasing complexity and volume of data along an immense economic pressure. The current advances and widespread use of imaging technologies now overload the human capacity of interpreting medical images, dangerously posing a risk of missing critical patterns of diseases. Machine learning has emerged as a key technology for developing novel tools in computer aided diagnosis, therapy and intervention. Still, progress is slow compared to other fields of visual recognition, which is mainly due to the domain complexity and constraints in clinical applications, i.e., robustness, high accuracy and reliability.

“Medical Imaging meets NeurIPS” aims to bring researchers together from the medical imaging and machine learning communities to discuss the major challenges in the field and opportunities for research and novel applications. The proposed event will be the continuation of a successful workshop organized in NeurIPS 2017 and 2018 (https://sites.google.com/view/med-nips-2018). It will feature a series of invited speakers from academia, medical sciences and industry to give an overview of recent technological advances and remaining major challenges.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:15 AM</td>
<td>Opening Remarks Lombaert, Glocker, Konukoglu, de Bruijne, Feragen, Oguz, Teuwen</td>
</tr>
<tr>
<td>08:30 AM</td>
<td>Session 1 (Invited Talk + presentations) Schnabel, Vidal, Sodickson, Grady</td>
</tr>
</tbody>
</table>
and close collaboration with diverse fields and practitioners. This workshop is intended as a forum for those in the machine learning community who wish to help tackle climate change.

Schedule

10:30 AM Coffee Break + Poster Session

11:00 AM Session 2 (Invited Talk + presentations)

12:30 PM Lunch

01:30 PM Session 3 (Invited Talk + presentations)

03:00 PM Coffee Break + Poster Session

03:30 PM Session 4 (Invited Talk + presentations)

05:00 PM fastMRI Challenge Talks

06:00 PM Closing Remarks

Abstracts (1):

Abstract 9: fastMRI Challenge Talks in Medical Imaging meets NeurIPS, Yakubova 05:00 PM

tentative

Tackling Climate Change with ML

David Rolnick, Priya Donti, Lynn Kaack, Alexandre Lacoste, Tegan Maharaj, Andrew Y Ng, John Platt, Jennifer Chayes, Yoshua Bengio

Sat Dec 14, 08:00 AM

Climate change is one of the greatest problems society has ever faced, with increasingly severe consequences for humanity as natural disasters multiply, sea levels rise, and ecosystems falter. Since climate change is a complex issue, action takes many forms, from designing smart electric grids to tracking greenhouse gas emissions through satellite imagery. While no silver bullet, machine learning can be an invaluable tool in fighting climate change via a wide array of applications and techniques. These applications require algorithmic innovations in machine learning

Emergent Communication: Towards Natural Language

Abhinav Gupta, Michael Noukhovitch, Cinjon Resnick, Natasha Jaques, Angelos Filos, Marie Ossenkopf, Angeliki Lazaridou, Jakob Foerster, Ryan Lowe, Douwe Kiela, Kyunghyun Cho

Sat Dec 14, 08:00 AM

Communication is one of the most impressive human abilities but historically it has been studied in machine learning on confined datasets of natural language, and by various other fields in simple low-dimensional spaces. Recently, with the rise of deep RL methods, the questions around the emergence of communication can now be studied in new, complex multi-agent scenarios. Two previous successful workshops (2017, 2018) have gathered the community to discuss how, when, and to what end communication emerges, producing research that was later published at top ML venues such as ICLR, ICML, AAAI. Now,
we wish to extend these ideas and explore a new direction: how emergent communication can become more like natural language, and what natural language understanding can learn from emergent communication.

The push towards emergent natural language is a necessary and important step in all facets of the field. For studying the evolution of human language, emerging a natural language can uncover the requirements that spurred crucial aspects of language (e.g. compositionality). When emerging communication for multi-agent scenarios, protocols may be sufficient for machine-machine interactions, but emerging a natural language is necessary for human-machine interactions. Finally, it may be possible to have truly general natural language understanding if agents learn the language through interaction as humans do. To make this progress, it is necessary to close the gap between artificial and natural language learning.

To tackle this problem, we want to take an interdisciplinary approach by inviting researchers from various fields (machine learning, game theory, evolutionary biology, linguistics, cognitive science, and programming languages) to participate and engaging them to unify the differing perspectives. We believe that the third iteration of this workshop with a novel, unexplored goal and strong commitment to diversity will allow this burgeoning field to flourish.

Schedule

08:00 AM Posters
08:55 AM Intro Remarks
09:00 AM Invited Talk - 1 Gibson
09:45 AM Contributed Talk - 1
10:00 AM Coffee Break / Poster Session
10:30 AM Invited Talk - 2 Zaslavsky
11:15 AM Contributed Talk - 2 Cowen-Rivers
11:30 AM Spotlight presentations x5
02:00 PM Invited Talk - 3 Eisner
02:45 PM Contributed Talk - 3 Brown
03:00 PM Invited Talk - 4 Andreas
03:45 PM Coffee Break / Poster Session
04:15 PM Invited Talk - 5 Lee
05:00 PM Panel Discussion
05:55 PM Closing Remarks

The third Conversational AI workshop – today’s practice and tomorrow’s potential

Alborz Geramifard, Jason Williams, Bill Byrne, Asli Celikyilmaz, Milica Gasic, Dilek Hakkani-Tur, Matt Henderson, Luís Lastras, Mari Ostendorf

Sat Dec 14, 08:00 AM

In the span of only a few years, conversational systems have become commonplace. Every day, millions of people use natural-language interfaces such as Siri, Google Now, Cortana, Alexa and others via in-home devices, phones, or messaging channels such as Messenger, Slack, Skype, among others. At the same time, interest among the research community in conversational systems has blossomed: for supervised and reinforcement learning, conversational systems often serve as both a benchmark task and an inspiration for new ML methods at conferences which don’t focus on speech and language per se, such as NIPS, ICML, IJCAI, and others. Such movement has not been unnoticed by major publications. This year in collaboration with AAAI community, the AI magazine will have a special issue on conversational AI (https://tinyurl.com/y6shq2td). Moreover, research community challenge tasks are proliferating, including the seventh Dialog Systems Technology Challenge (DSTC7), the Amazon Alexa prize, and the Conversational Intelligence Challenge live competitions at NIPS (2017, 2018).

Following the overwhelming participation in our last two NeurIPS workshops:

2017: 9 invited talks, 26 submissions, 3 oral papers, 13 accepted papers, 37 reviewers
2018: 4 invited talks, 42 submission, 6 oral papers, 23 accepted papers, 58 reviewers, we are excited to continue promoting cross-pollination of ideas between academic research centers and industry. The goal of this workshop is to bring together researchers and practitioners in this area, to clarify impactful research problems, understand well-founded methods, share findings from large-scale real-world deployments, and generate new ideas for future lines of research.

This one day workshop will include invited talks and a panel from academia and industry, contributed work, and open discussion.

Schedule

08:30 AM Opening Geramifard, Williams
08:40 AM Invited talk - Gabriel Skantze Skantze
09:10 AM Invited talk - Zhou Yu Yu
09:40 AM Poster lighting round Zheng, Segaard, Saleh, Jang, Gong, Florez, Li, Madotto, Nguyen, Kulkov, einolghozati, Wang, Eric, Hansen, Lubis
09:55 AM Posters + coffee break
10:40 AM Contributed talk 1
10:55 AM Contributed talk 2
11:10 AM Contributed talk 3
11:25 AM Contributed talk 4
Sets and Partitions

Nicholas Monath, Manzil Zaheer, Andrew McCallum, Ari Kobren, Junier Oliva, Barnabas Poczos, Ruslan Salakhutdinov

Sat Dec 14, 08:00 AM

Classic problems for which the input and/or output is set-valued are ubiquitous in machine learning. For example, multi-instance learning, estimating population statistics, and point cloud classification are all problem domains in which the input is set-valued. In multi-label classification the output is a set of labels, and in clustering, the output is a partition. New tasks that take sets as input are also rapidly emerging in a variety of application areas including: high energy physics, cosmology, crystallography, and art. As a natural means of succinctly capturing large collections of items, techniques for learning representations of sets and partitions have significant potential to enhance scalability, capture complex dependencies, and improve interpretability. The importance and potential of improved set processing has led to recent work on permutation invariant and equivariant representations (Ravanbakhsh et al, 2016; Zaheer et al, 2017; Ilse et al, 2018; Hartford et al, 2018; Lee et al, 2019, Cotter et al, 2019, Bloom-Reddy & Teh, 2019, and more) and continuous representations of set-based outputs and partitions (Tai and Lin, 2012; Belanger & McCallum, 2015; Wiseman et al, 2016; Caron et al, 2018; Zhang et al, 2019; Vikram et al, 2019).

The goal of this workshop is to explore:
- Permutation invariant and equivariant representations; empirical performance, limitations, implications, inductive biases of proposed representations of sets and partitions, as well as rich models of interaction among set elements;
- Inference methods for predicting sets or clusterings; approaches based on gradient-descent, continuous representations, amenable to end-to-end optimization with other models;
- New applications of set and partition-based models.

The First Workshop on Sets and Partitions, to be held as a part of the NeurIPS 2019 conference, focuses on models for tasks with set-based inputs/outputs as well as models of partitions and novel clustering methodology. The workshop welcomes both methodological and theoretical contributions, and also new applications. Connections to related problems in optimization, algorithms, theory as well as investigations of learning approaches to set/partition problems are also highly relevant to the workshop. We invite both paper submissions and submissions of open problems. We hope that the workshops will inspire further progress in this important field.

Organizing Committee:
Andrew McCallum, UMass Amherst
Ruslan Salakhutdinov, CMU
Barnabas Poczos, CMU
Junier Oliva, UNC Chapel Hill
Manzil Zaheer, Google Research
Ari Kobren, UMass Amherst
Nicholas Monath, UMass Amherst
with senior advisory support from Alex Smola.

Invited Speakers:
Siamak Ravanbakhsh
Abhishek Khetan
Eunsu Kang
Amr Ahmed
Stefanie Jegelka

Schedule

08:45 AM Opening Remarks
Zaheer, Monath, Kobren, Oliva, Poczos, Salakhutdinov, McCallum
09:00 AM Invited Talk - Siamak Ravanbakhsh
Ravanbakhsh
09:45 AM Coffee Break & Poster Session 1
10:30 AM Invited Talk - Stefanie Jegelka
Jegelka
11:15 AM Contributed Talk - Towards deep amortized clustering
Lee, Lee, Teh
11:30 AM Contributed Talk - Fair Hierarchical Clustering
Ahmadian, Epasto, Knittel, Kumar, Mahdian, Pham
### 11:45 AM
**Invited Talk - Abhishek Khetan**

- Khetan

### 02:00 PM
**Contributed Talk - Limitations of Deep Learning on Point Clouds**

- Bueno

### 02:15 PM
**Contributed Talk - Chirality Nets: Exploiting Structure in Human Pose Regression**

- Yeh, Hu, Schwing

### 02:30 PM
**Invited Talk - Eunsu Kang**

- Kang

### 03:15 PM
**Coffee Break & Poster Session 2**

### 04:15 PM
**Invited Talk - Amr Ahmed**

- Ahmed

---

### Abstracts (3):

**Abstract 3: Coffee Break & Poster Session 1 in Sets and Partitions.**

Zhang, Hare, Prügel-Bennett, Leung, Flaherty, Wiratuchotsatian, Epasto, Lattanzi, Vassilvitskii, Zadimoghaddam, Tulabandhula, Fuchs, Kosiorek, Posser, Hang, Goldie, Ravi, Mirhoseini, Xiong, Ren, Liao, Urtasun, Zhang, Borassi, Luo, Trapp, Dubourg-Felonneau, Kussad, Bender, Zaheer, Oliva, Stypulkowski, Zieba, Dill, Li, Ge, Kang, Parker Jones, Wong, Payne, Li, Nazi, Erdem, Erdem, O'Connor, Garcia, Zamorski, Chorowski, Sinha, Clifford 09:45 AM

**Poster Session 1 Paper Titles & Authors:**

- Deep Set Prediction Networks. Yan Zhang, Jonathon Hare, Adam Prügel-Bennett
- FSPool: Learning Set Representations with Featurewise Sort Pooling. Yan Zhang, Jonathon Hare, Adam Prügel-Bennett
- Deep Learning Features Through Dictionary Learning with Improved Clustering for Image Classification. Shengda Luo, Alex Po Leung, Haici Zhang
- Globally Optimal Model-based Clustering via Mixed Integer Nonlinear Programming. Patrick Flaherty, Pitchaya Wiratuchotsatian, Andrew C. Trapp
- Sliding Window Algorithms for k-Clustering Problems. Michele Borassi, Alessandro Epasto, Silvio Lattanzi, Sergei Vassilvitski, Morteza
- Optimized Recommendations When Customers Select Multiple Products. Prasoon Patidar, Deeksha Sinha, Theja Tulabandhula
- Manipulating Person Videos with Natural Language. Levent Karakan, Mehmet Gunel, Aykut Erdem, Erkut Erdem
- Permutation Invariance and Relational Reasoning in Multi-Object Tracking. Fabian B. Fuchs, Adam R. Kosiorek, Li Sun, Olivi Parker Jones, Ingmar Posner.
- Clustering by Learning to Optimize Normalized Cuts. Azade Nazi, Will Hang, Anna Goldie, Sujith Ravi, Azaïla Mirhoseini
- Deformable Filter Convolution for Point Cloud Reasoning. Yuwen Xiong, Mengye Ren, Renjie Liao, Kelvin Wong, Raquel Urtasun
- Learning Embeddings from Cancer Mutation Sets for Classification Tasks. Geoffrey Dubourg-Felonneau, Yasmeen Kussad, Dominic Kirkham, John Cassidy, Harry W Clifford
- Exchangeable Generative Models with Flow Scans. Christopher M. Bender, Kevin O'Connor, Yang Li, Juan Jose Garcia, Manzil Zaheer, Junier Oliva
- Conditional Invertible Flow for Point Cloud Generation. Stypulkowski Michal, Zamojski Maciej, Zieba Maciej, Chorowski Jan
- Getting Topology and Point Cloud Generation to Mesh. Austin Dill, Chun-Liang Li, Songwei Ge, Eunsu Kang
- Distributed Balanced Partitioning and Applications in Large-scale Load Balancing. Aaron Archer, Kevin Aydin, Mohammadhossein Bateni, Vahab Mirrokni, Aaron Schild, Ray Yang, Richard Zhuang

**Abstract 8: Contributed Talk - Limitations of Deep Learning on Point Clouds in Sets and Partitions.**

- Bueno 02:00 PM

**Poster Session 2 Paper Titles & Authors:**

- Towards deep amortized clustering. Juho Lee, Yoonho Lee, Yee Whye Teh
- Chirality Nets: Exploiting Structure in Human Pose Regression. Raymond Yeh, Yuan-Ting Hu, Alexander Schwing
- Fair Hierarchical Clustering. Sara Ahmadian, Alessandro Epasto, Marina Knittel, Ravi Kumar, Mohammad Mahdian, Philip Pham

- Distributed Balanced Partitioning and Applications in Large-scale Load Balancing. Aaron Archer, Kevin Aydin, Mohammadhossein Bateni, Vahab Mirrokni, Aaron Schild, Ray Yang, Richard Zhuang

**Abstract 11: Coffee Break & Poster Session 2 in Sets and Partitions.**

- Lee, Lee, Teh, Yeh, Hu, Schwing, Ahmadian, Epasto, Knittel, Kumar, Mahdian, Bueno, Sanghi, Jayaraman, Arroyo-Fernández, Hyunikowski, Mathur, Singh, Haddadan, Portilheiro, Zhang, Yukselgonul, Arias Figueroa, Maurya, Ravindran, NIelsen, Pham, Payan, McCallum, Mehta, Sun 03:15 PM

**Poster Session 2 Paper Titles & Authors:**

- Towards deep amortized clustering. Juho Lee, Yoonho Lee, Yee Whye Teh
- Chirality Nets: Exploiting Structure in Human Pose Regression. Raymond Yeh, Yuan-Ting Hu, Alexander Schwing
- Fair Hierarchical Clustering. Sara Ahmadian, Alessandro Epasto, Marina Knittel, Ravi Kumar, Mohammad Mahdian, Philip Pham
Clinical healthcare has been a natural application domain for ML with a few modest success stories of practical deployment. Inequity and healthcare disparity has long been a concern in clinical and public health for decades. However, the challenges of fair and equitable care using ML in health has largely remained unexplored. While a few works have attempted to highlight potential concerns and pitfalls in recent years, there are massive gaps in academic ML literature in this context. The goal of this workshop is to investigate issues around fairness that are specific to ML based healthcare. We hope to investigate a myriad of questions via the workshop.

Schedule

**Fair ML in Healthcare**

**Shalmali Joshi, Irene Y Chen, Ziad Obermeyer, Sendhil Mullainathan**

**Sat Dec 14, 08:00 AM**

Machine learning methods have had great success in learning complex representations that enable them to make predictions about unobserved data. Physical sciences span problems and challenges at all scales in the universe: from finding exoplanets in trillions of sky pixels, to finding machine learning inspired solutions to the quantum many-body problem, to detecting anomalies in event streams from the Large Hadron Collider. Tackling a number of associated data-intensive tasks including, but not limited to, segmentation, 3D computer vision, sequence modeling, causal reasoning, and efficient probabilistic inference are critical for furthering scientific discovery. In addition to using machine learning models for scientific discovery, the ability to interpret what a model has learned is receiving an increasing amount of attention.

In this targeted workshop, we would like to bring together computer
scientists, mathematicians and physical scientists who are interested in applying machine learning to various outstanding physical problems, in particular in inverse problems and approximating physical processes; understanding what the learned model really represents; and connecting tools and insights from physical sciences to the study of machine learning models. In particular, the workshop invites researchers to contribute papers that demonstrate cutting-edge progress in the application of machine learning techniques to real-world problems in physical sciences, and using physical insights to understand what the learned model means.

By bringing together machine learning researchers and physical scientists who apply machine learning, we expect to strengthen the interdisciplinary dialogue, introduce exciting new open problems to the broader community, and stimulate production of new approaches to solving open problems in sciences. Invited talks from leading individuals in both communities will cover the state-of-the-art techniques and set the stage for this workshop.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30 AM</td>
<td>Opening Remarks</td>
<td>Baydin, Carrasquilla, Ho, Kashinath, Paganini, Thais, Anandkumar, Cranmer, Melko, Prabhat, Wood</td>
</tr>
<tr>
<td>08:40 AM</td>
<td>Bernhard Schölkopf</td>
<td>Schölkopf</td>
</tr>
<tr>
<td>09:20 AM</td>
<td>Contributed Talk 1</td>
<td>Schoenholz</td>
</tr>
<tr>
<td>10:40 AM</td>
<td>Lenka Zdeborova</td>
<td>Zdeborová</td>
</tr>
<tr>
<td>11:20 AM</td>
<td>Contributed Talk 2</td>
<td>Metodiev</td>
</tr>
<tr>
<td>11:40 AM</td>
<td>Contributed Talk 3</td>
<td>Jimenez Rezende</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Contributed Talk 4</td>
<td>Yu</td>
</tr>
<tr>
<td>12:20 PM</td>
<td>Lunch Break</td>
<td></td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Katie Bouman</td>
<td>Bouman</td>
</tr>
<tr>
<td>02:40 PM</td>
<td>Maria Schuld</td>
<td>Schuld</td>
</tr>
</tbody>
</table>

**Real Neurons & Hidden Units: future directions at the intersection of neuroscience and AI**

*Guillaume Lajoie, Eli Shlizerman, Maximilian Puelma Touzel, Jessica Thompson, Konrad Kording*

**Sat Dec 14, 08:00 AM**

Recent years have witnessed an explosion of progress in AI. With it, a proliferation of experts and practitioners are pushing the boundaries of the field without regard to the brain. This is in stark contrast with the field's transdisciplinary origins, when interest in designing intelligent algorithms was shared by neuroscientists, psychologists and computer scientists alike. Similar progress has been made in neuroscience where novel experimental techniques now afford unprecedented access to brain activity and function. However, it is unclear how to maximize them to truly advance an end-to-end understanding of biological intelligence. The traditional neuroscience research program, however, lacks frameworks to truly advance an end-to-end understanding of biological intelligence. For the first time, mechanistic discoveries emerging from deep learning, reinforcement learning and other AI fields may be able to steer fundamental neuroscience research in ways beyond standard uses of machine learning for modelling and data analysis. For example, successful training algorithms in artificial networks, developed without biological constraints, can motivate research questions and hypotheses about the brain. Conversely, a deeper understanding of brain computations at the level of large neural populations may help shape future directions in AI. This workshop aims to address this novel situation by building on existing AI-Neuro relationships but, crucially, outline new directions for artificial systems and next-generation neuroscience experiments. We invite contributions concerned with the modern intersection between neuroscience and AI and in particular, addressing questions that can only now be tackled due to recent progress in AI on the role of recurrent dynamics, inductive biases to guide learning, global versus local learning rules, and interpretability of network activity. This workshop will promote discussion and showcase diverse perspectives on these open questions.

Schedule
Privacy in Machine Learning (PriML)

Borja Balle, Kamalika Chaudhuri, Antti Honkela, Antti Koskela, Casey Meehan, Mi Jung Park, Mary Anne Smart, Adrian Weller

Sat Dec 14, 08:00 AM

The goal of our workshop is to bring together privacy experts working in academia and industry to discuss the present and the future of privacy-aware technologies powered by machine learning. The workshop will focus on the technical aspects of privacy research and deployment with invited and contributed talks by distinguished researchers in the area. The programme of the workshop will emphasize the diversity of points of view on the problem of privacy. We will also ensure there is ample time for discussions that encourage networking between researchers, which should result in mutually beneficial new long-term collaborations.

Schedule

08:15 AM  TBA: Brendan McMahan  McMahan

10:30 AM  TBA: Ashwin Machanavajjhala  Machanavajjhala

11:30 AM  Poster Session

02:00 PM  TBA: Lalitha Sankar  Sankar

04:15 PM  TBA: Philip Leclerc  Leclerc

Abstracts (5):

Abstract 1: TBA: Brendan McMahan in Privacy in Machine Learning (PriML), McMahan 08:15 AM

Tentative schedule, details TBA.

Abstract 2: TBA: Ashwin Machanavajjhala in Privacy in Machine Learning (PriML), Machanavajjhala 10:30 AM
Abstract 3: **Poster Session in Privacy in Machine Learning (PriML)**

Canonne, Jun, Neel, Wang, vietri, Song, Lebensold, Zhang, Gondara, Li, Mireshghallah, Dong, Sarwate, Koskela, Jääkö, Kusner, Chen, Park, Machanavajjhala, Kalpathy-Cramer, Feldman, Tomkins, Phan, Esfandiari, Jaiswal, Sharma, Druce, Meehan, Zhao, Hsu, Railsback, Flaxman, Adebayo, Korolova, Xu, Holohan, Basu, Joseph, Thai, Yang, Vitercik, Hutchinson, Wang, Yauney, Tao, Jin, Lee 11:30 AM

Tentative schedule, details TBA.

Schedule is not final. Details TBA.

Abstract 4: **TBA: Lalitha Sankar in Privacy in Machine Learning (PriML)**

Sankar 02:00 PM

Tentative schedule, details TBA.

Abstract 5: **TBA: Philip Leclerc in Privacy in Machine Learning (PriML)**

Leclerc 04:15 PM

Tentative schedule, details TBA.
Solving inverse problems with deep networks: New architectures, theoretical foundations, and applications

Reinhard Heckel, Paul Hand, Richard Baraniuk, Joan Bruna, Alex Dimakis, Deanna Needell

Fri Dec 13, 08:00 AM

There is a long history of algorithmic development for solving inverse problems arising in sensing and imaging systems and beyond. Examples include medical and computational imaging, compressive sensing, as well as community detection in networks. Until recently, most algorithms for solving inverse problems in the imaging and network sciences were based on static signal models derived from physics or intuition, such as wavelets or sparse representations.

Today, the best performing approaches for the aforementioned image reconstruction and sensing problems are based on deep learning, which learn various elements of the method including i) signal representations, ii) stepsizes and parameters of iterative algorithms, iii) regularizers, and iv) entire inverse functions. For example, it has recently been shown that solving a variety of inverse problems by transforming an iterative, physics-based algorithm into a deep network whose parameters can be learned from training data, offers faster convergence and/or a better quality solution. Moreover, even with very little or no learning, deep neural networks enable superior performance for classical linear inverse problems such as denoising and compressive sensing. Motivated by those success stories, researchers are redesigning traditional imaging and sensing systems.

However, the field is mostly wide open with a range of theoretical and practical questions unanswered. In particular, deep-neural network based approaches often lack the guarantees of the traditional physics based methods, and while typically superior can make drastic reconstruction errors, such as fantasizing a tumor in an MRI reconstruction.

This workshop aims at bringing together theoreticians and practitioners in order to chart out recent advances and discuss new directions in deep neural network based approaches for solving inverse problems in the imaging and network sciences.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Speaker(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:10 AM</td>
<td>Robust One-Bit Recovery</td>
<td>Qiu, Wei, Yang</td>
</tr>
<tr>
<td></td>
<td>via ReLU Generative Networks:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Improved Statistical Rate and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Global Landscape Analysis</td>
<td></td>
</tr>
<tr>
<td>09:40 AM</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Laura Waller</td>
<td>Waller</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Michael Unser</td>
<td>Unser</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Neural Reparameterization</td>
<td>Hoyer, Sohl-Dickstein, Greydanus</td>
</tr>
<tr>
<td></td>
<td>Improves Structural</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Optimization</td>
<td></td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Lunch Break</td>
<td></td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Raquel Urtasun</td>
<td>Urtasun</td>
</tr>
<tr>
<td>02:30 PM</td>
<td>Josh Batson</td>
<td>Batson</td>
</tr>
<tr>
<td>03:00 PM</td>
<td>Venkat Chandrasekaran</td>
<td>Chandrasekaran</td>
</tr>
</tbody>
</table>

Abstracts (2):

Abstract 4: Robust One-Bit Recovery via ReLU Generative Networks: Improved Statistical Rate and Global Landscape Analysis in Solving inverse problems with deep networks: New architectures, theoretical foundations, and applications, Qiu, Wei, Yang 09:10 AM

We study the robust one-bit compressed sensing problem whose goal is to design an algorithm that faithfully recovers any sparse target vector $\theta_0$ from $m$ quantized noisy measurements. Under the assumption that the measurements are sub-Gaussian, to recover any $k$-sparse $\theta_0$ (with high probability), the best known computationally tractable algorithm requires...
an algorithm is "computationally tractable" if it has provable convergence guarantees. The notation $\tilde{\log}(\mathcal{O})/\tilde{\mathcal{O}}(\cdotp)$ omits a logarithm factor of $\log^*(\cdotp)$ and $\mathcal{O}(\cdotp)$ in the one-bit sensing problem where the sparsity is implicitly enforced via a low-dimensional representation $x_0$ through a known $\log$-layer ReLU generative network $G: \tilde{\mathcal{O}}(\cdotp)$ such that $G(x_0)$ is via an unconstrained empirical risk minimization (ERM) problem under a much weaker measurement assumption. For such a problem, we establish a joint statistical and computational analysis. In particular, we prove that the ERM estimator in this new framework achieves an improved statistical rate of $\mathcal{O}(\cdotp)$ recovering any $G(x_0)$ uniformly up to an error $\mathcal{O}(\cdotp)$. Moreover, from the lens of computation, despite non-convexity, we prove that the objective of our ERM problem has no spurious stationary point, that is, any stationary point is equally good for recovering the true target up to scaling with a certain accuracy. Our analysis sheds some light on that the objective of our ERM problem has no spurious stationary point, that is, any stationary point is equally good for recovering the true target up to scaling with a certain accuracy. Our analysis sheds some light on the possibility of inverting a deep generative model under partial and quantized measurements, complementing the recent success of using deep generative models for inverse problems.

Abstract 8: Neural Reparameterization Improves Structural Optimization in Solving inverse problems with deep networks: New architectures, theoretical foundations, and applications, Hoyer, Sohl-Dickstein, Greydanus 11:30 AM

Structural optimization is a popular method for designing objects such as bridge trusses, airplane wings, and optical devices. Unfortunately, the quality of solutions depends heavily on how the problem is parameterized. In this paper, we propose using the implicit bias over functions induced by neural networks to improve the parameterization of structural optimization. Rather than directly optimizing densities on a grid, we instead optimize the parameters of a neural network which outputs those densities. This reparameterization leads to different and often better solutions. On a selection of 116 structural optimization tasks, our approach produces an optimal design 50% more often than the best baseline method.

Perception as generative reasoning: structure, causality, probability

Dan Rosenbaum, Marta Garnelo, Peter Battaglia, Kelsey Allen, Ilker Yildirim

Fri Dec 13, 08:00 AM

Many perception tasks can be cast as 'inverse problems' where the input signal is the outcome of a causal process and perception is to invert that process. For example in visual object perception, the image is caused by an object and perception is to infer which object gave rise to that image. Following an analysis-by-synthesis approach, modelling the forward and causal direction of the data generation process is a natural way to capture the underlying scene structure, which typically leads to broader generalisation and better sample efficiency. Such a forward model can be applied to solve the inverse problem (inferring the scene structure from an input image) using Bayes rule, for example. This workflow stands in contrast to common approaches in deep learning, where typically one first defines a task, and then optimises a deep model end-to-end to solve it. In this workshop we propose to revisit ideas from the generative approach and advocate for learning-based analysis-by-synthesis methods for perception and inference. In addition, we pose the question of how ideas from these research areas can be combined with and complement modern deep learning practices.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:50 AM</td>
<td>Opening Remarks</td>
<td>Rosenbaum, Garnelo, Battaglia, Allen, Yildirim</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Sanja Fidler</td>
<td>Fidler</td>
</tr>
<tr>
<td>09:35 AM</td>
<td>Spotlights 1</td>
<td>Wu</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Jiajun Wu</td>
<td>Lopez-Guevara</td>
</tr>
<tr>
<td>11:05 AM</td>
<td>Tatiana Lopez-Guevara</td>
<td>Fidler</td>
</tr>
<tr>
<td>11:40 AM</td>
<td>Spotlights 2</td>
<td>Mitra</td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Fei-Fei Li</td>
<td>Graber, Hu, Fang, Hamrick, Giannone, Co-Reyes, Deng, Crawford, Dittadi, Karkus, Dirks, TRIVEDI, Raj, Felip, Leon, Chan, Chorowski, Orchard, Stan, Kortylewski, Zinberg, Zhou, Sun, Mansinghka, Li, Cusumano-Towner</td>
</tr>
</tbody>
</table>

Workshop on Human-Centric Machine Learning

Plamen P Angelov, Nuria Oliver, Adrian Weller, Manuel Rodriguez, Isabel Valera, Silvia Chiappa, Hoda Heidari, Niki Kilbertus

Fri Dec 13, 08:00 AM

The growing field of Human-centric ML seeks to minimize the potential harms, risks, and burdens of big data technologies on the public, and at the same time, maximize their societal benefits. In this workshop, we address a wide range of challenges from diverse, multi-disciplinary viewpoints. We bring together experts from a diverse set of backgrounds. Our speakers are leading experts in ML, human-computer interaction, ethics, and law. Each of our speakers will focus on one core human-centred challenge (namely, fairness, accountability, interpretability, transparency, security, and privacy) in specific application domains (such as medicine, welfare programs, governance, and regulation). One of the main goals of this workshop is to help the community understand where it stands after a few years of rapid technical development and identify promising research directions to pursue in the years to come.
These directions are further debated in one of our panel discussions.

**Schedule**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:20 AM</td>
<td>Welcome and introduction</td>
</tr>
<tr>
<td>08:30 AM</td>
<td>Invited talks #1,2</td>
</tr>
<tr>
<td>09:20 AM</td>
<td>Contributed talks</td>
</tr>
<tr>
<td>10:05 AM</td>
<td>Panel #1: On the role of industry, academia, and government in developing HCML</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Coffee break</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Invited talk #3</td>
</tr>
<tr>
<td>11:25 AM</td>
<td>Spotlight talks</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Lunch and poster session</td>
</tr>
<tr>
<td>01:30 PM</td>
<td>Invited talks #4,5</td>
</tr>
<tr>
<td>02:20 PM</td>
<td>Contributed talks</td>
</tr>
<tr>
<td>03:05 PM</td>
<td>Coffee break</td>
</tr>
<tr>
<td>03:30 PM</td>
<td>Invited talks #6,7</td>
</tr>
<tr>
<td>04:20 PM</td>
<td>Contributed talks</td>
</tr>
<tr>
<td>04:50 PM</td>
<td>Panel #2: Future research directions and interdisciplinary collaborations in HCML</td>
</tr>
<tr>
<td>05:15 PM</td>
<td>Poster session</td>
</tr>
<tr>
<td>06:15 PM</td>
<td>Closing remarks</td>
</tr>
</tbody>
</table>

**Shared Visual Representations in Human and Machine Intelligence**

*Arturo Deza, Joshua Peterson, Apurva Ratan Murty, Tom Griffiths*

**Fri Dec 13, 08:00 AM**

The goal of the Shared Visual Representations in Human and Machine Intelligence workshop is to disseminate relevant, parallel findings in the fields of computational neuroscience, psychology, and cognitive science that may inform modern machine learning methods. In the past few years, machine learning methods—especially deep neural networks—have widely permeated the vision science, cognitive science, and neuroscience communities. As a result, scientific modeling in these fields has greatly benefited, producing a swath of potentially critical new insights into human learning and intelligence, which remains the gold standard for many tasks. However, the machine learning community has been largely unaware of these cross-disciplinary insights and analytical tools, which may help to solve many of the current problems that ML theorists and engineers face today (e.g., adversarial attacks, compression, continual learning, and unsupervised learning).

Thus we propose to invite leading cognitive scientists with strong computational backgrounds to disseminate their findings to the machine learning community with the hope of closing the loop by nourishing new ideas and creating cross-disciplinary collaborations.

**Schedule**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:50 AM</td>
<td>Opening Remarks</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Olivier Henaff (DeepMind)</td>
</tr>
<tr>
<td>09:25 AM</td>
<td>Irina Higgins (DeepMind)</td>
</tr>
<tr>
<td>09:50 AM</td>
<td>Bill Freeman (MIT)</td>
</tr>
<tr>
<td>10:15 AM</td>
<td>Q&amp;A from the Audience: Ask a Neuro / Cognitive Scientist</td>
</tr>
<tr>
<td>10:45 AM</td>
<td>Coffee Break</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Ruairidh Battleday (Princeton)</td>
</tr>
<tr>
<td>11:15 AM</td>
<td>Will Xiao (Harvard)</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Erin Grant (UC Berkeley)</td>
</tr>
<tr>
<td>11:45 AM</td>
<td>Andrei Barbu (MIT)</td>
</tr>
<tr>
<td>12:10 PM</td>
<td>Mike Tarr (CMU)</td>
</tr>
<tr>
<td>12:35 PM</td>
<td>James DiCarlo (MIT)</td>
</tr>
<tr>
<td>01:00 PM</td>
<td>Lunch on your own</td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Poster Session</td>
</tr>
<tr>
<td>03:00 PM</td>
<td>Q&amp;A from the Audience: Ask the Grad Students</td>
</tr>
<tr>
<td>03:30 PM</td>
<td>Talia Konkle (Harvard)</td>
</tr>
</tbody>
</table>
The last decade has seen both machine learning and biology transformed: the former by the ability to train complex predictors on massive labelled data sets; the latter by the ability to perturb and measure biological systems with staggering throughput, breadth, and resolution. However, fundamentally new ideas in machine learning are needed to translate biomedical data at scale into a mechanistic understanding of biology and disease at a level of abstraction beyond single genes. This challenge has the potential to drive the next decade of creativity in machine learning as the field grapples with how to move beyond prediction to a regime that broadly catalyzes and accelerates scientific discovery.

To seize this opportunity, we will bring together current and future leaders within each field to introduce the next generation of machine learning specialists to the next generation of biological problems. Our full-day workshop will start a deeper dialogue with the goal of Learning Meaningful Representations of Life (LMRL), emphasizing interpretable representation learning of structure and principles. The workshop will address this challenge at five layers of biological abstraction (genome, molecule, cell, system, phenome) through interactive breakout sessions led by a diverse team of experimentalists and computational scientists to facilitate substantive discussion.

We are calling for short abstracts from computer scientists and biological scientists. Submission deadline is Friday, September 20. Significant travel support is also available. Details here:

https://lmrl-bio.github.io/call
https://lmrl-bio.github.io/travel

Schedule

08:30 AM  Daphne Koller  Koller
09:15 AM  Aviv Regev  Regev
09:45 AM  Rediet Abebe  Abebe
10:15 AM  Max Welling  Welling
11:05 AM  Cell  Carpenter, Lopez, Wu, Johnston, Cammarata, Chayes, Sasse, Jones
12:30 PM  Genome  Marks, Peer, Wang, sander, Angelino, Kusner, Dias, Yamins, Blumberg, Wang, Tseng, Siraj, Papaxanthos, Zhang
02:30 PM  Molecules and their Evolution  Huang, Ovchinnikov, Jones, Aspuru-Guzik, Morris, Mostafavi, Clevert, Rush, Lindorff-Larsen
Abstracts (4):

Abstract 5: **Cell in Learning Meaningful Representations of Life**, Carpenter, , Lopez, Wu, Johnston, Cammarata, Chayes, Sasse, Jones

11:05 AM

Anne Carpenter
Orr Ashenberg
Anna Goldenberg
Joshua Weinstein
Mor Nitzan
Romain Lopez
Zachary Wu
Kadina Johnston
Louis Cammarata
Lior Pachter
Peter Karchenko
Jennifer Chayes
Riccardo Zecchina
Regina Barzilay
Yonina Eldar
Alexander Sasse

Geoff Schiebinger
Ray Jones
anne@broadinstitute.org
orr@broadinstitute.org
anna.goldenberg@utoronto.ca
jaweinst@uchicago.edu
mornitz@gmail.com
romain_lopez@berkeley.edu
zwu@caltech.edu
kjohnston@caltech.edu
louis_cammarata@g.harvard.edu
lpachter@caltech.edu
peter.kharchenko@post.harvard.edu
jchayes@microsoft.com
riccardo.zecchina@unibocconi.it
regina@csail.mit.edu
yonina.eldar@weizmann.ac.il
alexander.sasse@mail.utoronto.ca
geoff@broadinstitute.org
thouis@broadinstitute.org


12:30 PM

Pamela Silver
Debora Marks
Dana Pe’er
Samantha Morris
Yixin Wang
Allon Klein
Harris Wang
Chris Sander
Elaine Angelino
Matt Kusner
Mafalda Dias
Dan Yamins
Jeffrey Miller
Andrew J. Blumberg
Jennifer Doudna
Kyogo Kawaguchi
Dennis Wang, University of Sheffield,
Wei-Cheng Tseng*, National Tsing Hua University,
Anika Gupta
Layla Siraj
Laetitia Papaxanthos
Alina Selega
Libby Zhang

psilver@broadinstitute.org
debbie@hms.harvard.edu
peerster@gmail.com
s.morris@wustl.edu
yixin.wang@columbia.edu
allon_klein@hms.harvard.edu
hw2429@columbia.edu
chris@sanderlab.org
elaine@eeecs.berkeley.edu
matt.kusner@gmail.com
mafalda_dias@hms.harvard.edu
Abstract 7: Molecules and their Evolution in Learning Meaningful Representations of Life, Huang, Ovchinnikov, Jones, Aspuru-Guzik, Morris, Mostafavi, Clevert, Rush, Lindorff-Larsen 02:30 PM

Possu Huang
Sergey Ovchinnikov
David Jones
Alan Aspuru-Guzik
Carol Robinson
Itsik Pe’er
Quaid Morris
Sara Mostafavi
Wen Torng
Djork-Amé Clevert
Perry Palemdo
Alexander Sasha Rush


Soumya Raychaudhuri
Nir HaCohen
Yakir Reshef
David Reshef
Matt Johnson
Samantha Morris
Aurel Nagy
Gokcen Eraslan
Meromit Singer
Eli van Allen
Smita Krishnaswamy
Casey Greene

Scott Linderman
Miriam Shiffman
Danilo Bzdok

Machine Learning for Health (ML4H): What makes machine learning in medicine different?

Andrew Beam, Tristan Naumann, Brett Beaulieu-Jones, Irene Y Chen, Sam Finlayson, Emily Alsentzer, Adrian Dalca, Matthew McDermott

Fri Dec 13, 08:00 AM

The goal of the NeurIPS 2019 Machine Learning for Health Workshop (ML4H) is to foster collaborations that meaningfully impact medicine by bringing together clinicians, health data experts, and machine learning researchers. Attendees at this workshop can also expect to broaden their network of collaborators to include clinicians and machine learning researchers who are focused on solving some of the most important problems in medicine and healthcare. The organizers of this proposal have successfully run NeurIPS workshops in the past and are well-equipped to run this year’s workshop should this proposal be accepted.

This year’s theme of “What makes machine learning in medicine different?” aims to elucidate the obstacles that make the development of machine learning models for healthcare uniquely challenging. To speak to this theme, we have received commitments to speak from some of the leading researchers and physicians in this area. Below is a list of confirmed speakers who have agreed to participate.

Luke Oakden-Raynor, MBBS (Adelaide)
Russ Altman, MD/PhD (Stanford)
Lilly Peng, MD/PhD (Google)
Daphne Koller, PhD (in situ)
Jeff Dean, PhD (Google)

Attendees at the workshop will gain an appreciation for problems that are unique to the application of machine learning for healthcare and a better understanding of how machine learning techniques may be leveraged to solve important clinical problems. This year’s workshop builds on the last two NeurIPS ML4H workshops, which were both attended by more than 500 people each year, and helped form the foundations of an emerging research community.
## Overview

Privacy and security have become critical concerns in recent years, particularly as companies and organizations increasingly collect detailed information about their products and users. This information can enable machine learning methods that produce better products. However, it also has the potential to allow for misuse, especially when private data about individuals is involved. Recent research shows that privacy and utility do not necessarily need to be at odds, but can be addressed by careful design and analysis. The need for such research is reinforced by the recent introduction of new legal constraints, led by the European Union’s General Data Protection Regulation (GDPR), which is already inspiring novel legislative approaches around the world such as Cyber-security Law of the People’s Republic of China and The California Consumer Privacy Act of 2018.

An approach that has the potential to address a number of problems in this space is federated learning (FL). FL is an ML setting where many clients (e.g., mobile devices or whole organizations) collaboratively train a model under the orchestration of a central server (e.g., service provider), while keeping the training data decentralized. Organizations and mobile devices have access to increasing amounts of sensitive data, with scrutiny of ML privacy and data handling practices increasing correspondingly. These trends have produced significant interest in FL, since it provides a viable path to state-of-the-art ML without the need for the centralized collection of training data – and the risks and responsibilities that come with such centralization. Nevertheless, significant challenges remain open in the FL setting, the solution of which will require novel techniques from multiple fields, as well as improved open-source tooling for both FL research and real-world deployment.

This workshop aims to bring together academic researchers and industry practitioners with common interests in this domain. For industry participants, we intend to create a forum to communicate what kind of problems are practically relevant. For academic participants, we hope to make it easier to become productive in this area. Overall, the workshop will provide an opportunity to share the most recent and innovative work in FL, and discuss open problems and relevant approaches. The technical issues encouraged to be submitted include general computation based on decentralized data (i.e., not only machine learning), and how such computations can be combined with other research areas, such as differential privacy, secure multi-party computation, computational efficiency, coding theory, etc. Contributions in theory as well as applications are welcome, including proposals for novel system design. Work on fully-decentralized (peer-to-peer) learning will also be considered, as there is significant overlap in both interest and techniques with federated learning.

### Call for Contributions

We welcome high quality submissions in the broad area of federated learning (FL). A few (non-exhaustive) topics of interest include:

- Optimization algorithms for FL, particularly communication-efficient algorithms tolerant of non-IID data
- Approaches that scale FL to larger models, including model and gradient compression techniques
- Novel applications of FL
- Theory for FL
- Approaches to enhancing the security and privacy of FL, including cryptographic techniques and differential privacy
- Bias and fairness in the FL setting
- Attacks on FL including model poisoning, and corresponding defenses

---

## Workshop on Federated Learning for Data Privacy and Confidentiality

**Lixin Fan, Jakub Konečný, Yang Liu, Brendan McMahan, Virginia Smith, Han Yu**

Fri Dec 13, 08:00 AM

### Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:45 AM</td>
<td>Daphne Koller Talk</td>
<td>Koller</td>
</tr>
<tr>
<td>09:15 AM</td>
<td>Emily Fox Talk</td>
<td>Fox</td>
</tr>
<tr>
<td>10:15 AM</td>
<td>Luke Oakden-Rayner Talk</td>
<td>Oakden-Rayner</td>
</tr>
<tr>
<td>10:45 AM</td>
<td>Paper spotlight talks</td>
<td></td>
</tr>
<tr>
<td>11:15 AM</td>
<td>Poster Session I</td>
<td></td>
</tr>
<tr>
<td>02:45 PM</td>
<td>Lily Peng talk</td>
<td>Peng</td>
</tr>
<tr>
<td>03:15 PM</td>
<td>Anna Goldenberg Talk</td>
<td>Goldenberg</td>
</tr>
<tr>
<td>03:45 PM</td>
<td>Poster Session II</td>
<td></td>
</tr>
<tr>
<td>04:45 PM</td>
<td>Deepmind Talk</td>
<td>Tomasev</td>
</tr>
<tr>
<td>05:15 PM</td>
<td>Panel Discussion</td>
<td></td>
</tr>
<tr>
<td>06:15 PM</td>
<td>Message from sponsor</td>
<td></td>
</tr>
</tbody>
</table>

---

*NeurIPS 2019 Workshop book*
We won't be accepting such submissions unless they have been a previous conference with published proceedings?

Can a paper be submitted to the workshop that has already appeared at multiple workshops. Our suggestion is to pick one workshop to submit to. We discourage this, as it leads to more work for reviewers across workshop in parallel?

Can a submission to this workshop be submitted to another NeurIPS addition to the main submission). the paper and be limited to a reasonable amount (max 10 pages in supplementary material should also follow the same NeurIPS format as supplementary material is at the discretion of the reviewers. The ensure that the main paper is self-contained, since looking at the best contribution that provides code to reproduce their results. Submission link: https://easychair.org/conferences/?conf=flneurips2019

Important Dates (2019)
Submission deadline: Sep 9
Author notification: Sep 30
Camera-Ready Papers Due: TBD
Workshop: Dec 13

Organizers:
Lixin Fan, WeBank
Jakub Konečný, Google
Yang Liu, WeBank
Brendan McMahan, Google
Virginia Smith, CMU
Han Yu, NTU

Invited Speakers:
Francoise Beaufays, Principal Researcher, Google
Shahrokh Dajavud, Distinguished Research, IBM
Dawn Song, Professor, University of California, Berkeley
Ameet Talwalkar, Assistant Professor, CMU; Chief Scientist, Determined AI
Max Welling, Professor, University of Amsterdam; VP Technologies, Qualcomm
Qiang Yang, Hong Kong University of Science and Technology, Hong Kong; Chief AI Officer, WeBank

FAQ
Can supplementary material be added beyond the 4-page limit and are there any restrictions on it?
Yes, you may include additional supplementary material, but you should ensure that the main paper is self-contained, since looking at supplementary material is at the discretion of the reviewers. The supplementary material should also follow the same NeurIPS format as the paper and be limited to a reasonable amount (max 10 pages in addition to the main submission).

Can a submission to this workshop be submitted to another NeurIPS workshop in parallel?
We discourage this, as it leads to more work for reviewers across multiple workshops. Our suggestion is to pick one workshop to submit to. Can a paper be submitted to the workshop that has already appeared at a previous conference with published proceedings?

We won't be accepting such submissions unless they have been adapted to contain significantly new results (where novelty is one of the qualities reviewers will be asked to evaluate).

Can a paper be submitted to the workshop that is currently under review or will be under review at a conference during the review phase? It is fine to submit a condensed version (i.e., 4 pages) of a parallel conference submission, if it also fine for the conference in question. Our workshop does not have archival proceedings, and therefore parallel submissions of extended versions to other conferences are acceptable.

==================================================================================================
Accepted papers:
2. Xin Yao, Tianchi Huang, Rui-Xiao Zhang, Ruiyu Li and Lifeng Sun. Federated Learning with Unbiased Gradient Aggregation and Controllable Meta Updating
3. Daniel Peterson, Pallika Kanani and Virendra Marathe. Private Federated Learning with Domain Adaptation
5. Sebastian Caldas, Jakub Konečný, H. Brendan McMahan and Ameet Talwalkar. Mitigating the Impact of Federated Learning on Client Resources
7. Sebastian Caldas, Sai Meher Karthik Duddu, Peter Wu, Tian Li, Jakub Konečný, H. Brendan McMahan, Virginia Smith and Ameet Talwalkar. Leaf: A Benchmark for Federated Settings
13. Yang Liu, Xiong Zhang, Shuqi Qin and Xiaoping Lei. Differentially Private Linear Regression over Fully Decentralized Datasets
14. Florian Hartmann and Sunah Suh. Federated Learning for Ranking Browser History Suggestions
15. Aleksei Triastcyn and Boi Faltings. Federated Learning with Bayesian Information Theory

Of note, the workshop does not have archival proceedings, and therefore parallel submissions of extended versions to other conferences are acceptable.
Differential Privacy


18. Mingshu Cong, Zhongming Ou, Yanxin Zhang, Han Yu, Xi Weng, Jiabao Qu, Siu Ming Yu, Yang Liu and Qiang Yang. Neural Network Optimization for a VCG-based Federated Learning Incentive Mechanism


20. Suyi Li, Yong Cheng, Yang Liu and Wei Wang. Abnormal Client Behavior Detection in Federated Learning


22. Shicong Cen, Huishuai Zhang, Yuejie Chi, Wei Chen and Tie-Yan Liu. Convergence and Regularization of Distributed Stochastic Variance Reduced Methods

23. Zhaorui Li, Zhicong Huang, Chaochao Chen and Cheng Hong. Quantification of the Leakage in Federated Learning


25. Boyue Li, Shicong Cen, Yuxin Chen and Yuejie Chi. Communication-Efficient Distributed Optimization in Networks with Gradient Tracking


27. Felix Sattler, Klaus-Robert Müller and Wojciech Samek. Clustered Federated Learning

28. Ziteng Sun, Peter Kairouz, Ananda Theertha Suresh and Brendan McMah. Backdoor Attacks on Federated Learning and Corresponding Defenses

29. Neta Shoham, Tomer Avidor, Aviv Keren, Nadav Israel, Daniel Bendikis, Liron Mor-Yosef and Itai Zeil. Overcoming Forgetting in Federated Learning on Non-IID Data

30. Ahmed Khaled and Peter Richtárik. Gradient Descent with Compressed Iterates

31. Jiahuan Luo, Xueyang Wu, Yun Luo, Anbu Huang, Yunfeng Huang, Yang Liu and Qiang Yang. Real-World Image Datasets for Federated Learning

32. Ahmed Khaled, Konstantin Mishchenko and Peter Richtárik. First Analysis of Local GD on Heterogeneous Data

33. Dashan Gao, Ce Ju, Xiguang Wei, Yang Liu, Tianjian Chen and Qiang Yang. HHHFL: Hierarchical Heterogeneous Horizontal Federated Learning for Electroencephalography

The workshop schedule (tentative):

<table>
<thead>
<tr>
<th>Schedule</th>
<th>Time</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opening remarks</td>
<td>08:45 AM</td>
<td>Fan</td>
</tr>
<tr>
<td>Contributed talk #0</td>
<td>08:50 AM</td>
<td>Yang</td>
</tr>
<tr>
<td>Qiang Yang Talk</td>
<td>09:00 AM</td>
<td>Talwalkar</td>
</tr>
<tr>
<td>Ameet Talwalkar Talk</td>
<td>09:30 AM</td>
<td>Welling</td>
</tr>
<tr>
<td>Coffee break and poster</td>
<td>10:00 AM</td>
<td></td>
</tr>
<tr>
<td>Contributed talk #1</td>
<td>10:30 AM</td>
<td></td>
</tr>
<tr>
<td>Contributed talk #2</td>
<td>10:40 AM</td>
<td></td>
</tr>
<tr>
<td>Max Welling Talk</td>
<td>10:50 AM</td>
<td></td>
</tr>
<tr>
<td>Contributed talk #3</td>
<td>11:20 AM</td>
<td></td>
</tr>
<tr>
<td>Contributed talk #4</td>
<td>11:30 AM</td>
<td></td>
</tr>
<tr>
<td>Dawn Song Talk</td>
<td>11:40 AM</td>
<td></td>
</tr>
<tr>
<td>Lunch break and poster</td>
<td>12:10 PM</td>
<td></td>
</tr>
<tr>
<td>Dan Ramage Talk</td>
<td>01:30 PM</td>
<td>Ramage</td>
</tr>
<tr>
<td>Contributed talk #5</td>
<td>02:00 PM</td>
<td></td>
</tr>
<tr>
<td>Contributed talk #6</td>
<td>02:10 PM</td>
<td></td>
</tr>
<tr>
<td>Francoise Beaufays Talk</td>
<td>02:20 PM</td>
<td></td>
</tr>
<tr>
<td>Contributed talk #7</td>
<td>02:50 PM</td>
<td></td>
</tr>
<tr>
<td>Contributed talk #8</td>
<td>03:00 PM</td>
<td></td>
</tr>
<tr>
<td>Raluca Popa Talk</td>
<td>03:10 PM</td>
<td></td>
</tr>
<tr>
<td>Contributed talk #9</td>
<td>03:40 PM</td>
<td></td>
</tr>
<tr>
<td>Contributed talk #10</td>
<td>04:15 PM</td>
<td></td>
</tr>
<tr>
<td>Yiqiang Chen Talk</td>
<td>04:25 PM</td>
<td></td>
</tr>
<tr>
<td>Panel discussion</td>
<td>05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Closing Remark</td>
<td>06:00 PM</td>
<td></td>
</tr>
</tbody>
</table>

Meta-Learning
Recent years have seen rapid progress in metalearning methods, which learn (and optimize) the performance of learning methods based on data, generate new learning methods from scratch, and learn to transfer knowledge across tasks and domains. Metalearning can be seen as the logical conclusion of the arc that machine learning has undergone in the last decade, from learning classifiers, to learning representations, and finally to learning algorithms that themselves acquire representations and classifiers. The ability to improve one’s own learning capabilities through experience can also be viewed as a hallmark of intelligent beings, and there are strong connections with work on human learning in neuroscience. The goal of this workshop is to bring together researchers from all the different communities and topics that fall under the umbrella of metalearning. We expect that the presence of these different communities will result in a fruitful exchange of ideas and stimulate an open discussion about the current challenges in metalearning, as well as possible solutions.

**Schedule**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 AM</td>
<td>Invited Talk 1</td>
<td>Abbeel</td>
</tr>
<tr>
<td>09:40 AM</td>
<td>Invited Talk 2</td>
<td>Clune</td>
</tr>
<tr>
<td>10:10 AM</td>
<td>Poster Spotlights 1</td>
<td>Takagi, Javed, Sommer, Sharaf, D'Oro, Wei, Doveh, White, Gonzalez, Nguyen, li, Yu, Ramalho, Nomura, Alvi, ton, Huang, Lee, Flienerhag, Zhang, Friesen, Blomstedt, Dubatovka, Bartunov, Yi, Shcherbatyi, Simon, Shang, MacLeod, Liu, Fowl, Parente Paiva Mesquita</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Coffee/Poster session 1</td>
<td></td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Invited Talk 3</td>
<td>Grant</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Discussion 1</td>
<td></td>
</tr>
<tr>
<td>01:00 PM</td>
<td>Invited Talk 4</td>
<td>Abel</td>
</tr>
<tr>
<td>01:30 PM</td>
<td>Invited Talk 5</td>
<td>Hadsell</td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Poster Spotlights 2</td>
<td>Song, Mangla, Salinas, Zhuang, Feng, Hu, Puri, Maddox, Raghu, Tossou, Yin, Dasgupta, Lee, Alet, Xu, Franke, Harrison, Warrell, Dhillon, Zela, Qiu, Siems, Mendonca, Schliessinger, Li, Manolache, Dutta, Glass</td>
</tr>
<tr>
<td>03:00 PM</td>
<td>Coffee/Poster session 2</td>
<td></td>
</tr>
<tr>
<td>04:00 PM</td>
<td>Contributed Talk 1</td>
<td></td>
</tr>
<tr>
<td>04:45 PM</td>
<td>Contributed Talk 2</td>
<td></td>
</tr>
<tr>
<td>05:00 PM</td>
<td>Invited Talk 6</td>
<td>Lake</td>
</tr>
<tr>
<td>05:30 PM</td>
<td>Discussion 2</td>
<td></td>
</tr>
</tbody>
</table>

Visually Grounded Interaction and Language

**Florian Strub, Abhishek Das, Erik Wijmans, Harm de Vries, Stefan Lee, Alane Suhr, Drew Arad Hudson**

Fri Dec 13, 08:00 AM

The dominant paradigm in modern natural language understanding is learning statistical language models from text-only corpora. This approach is founded on a distributional notion of semantics, i.e. that the "meaning" of a word is based only on its relationship to other words. While effective for many applications, this approach suffers from limited semantic understanding -- symbols learned this way lack any concrete groundings into the multimodal, interactive environment in which communication takes place. The symbol grounding problem first highlighted this limitation, that "meaningless symbols (i.e. words) cannot be grounded in anything but other meaningless symbols".

On the other hand, humans acquire language by communicating about and interacting within a rich, perceptual environment -- providing concrete groundings, e.g. to objects or concepts either physical or psychological. Thus, recent works have aimed to bridge computer vision, interactive learning, and natural language understanding through language learning tasks based on natural images or through embodied agents performing interactive tasks in physically simulated environments, often drawing on the recent successes of deep learning and reinforcement learning. We believe these lines of research pose a promising approach for building models that do grasp the world's underlying complexity.

The goal of this third ViGIL workshop is to bring together scientists from various backgrounds - machine learning, computer vision, natural language processing, neuroscience, cognitive science, psychology, and philosophy - to share their perspectives on grounding, embodiment, and interaction. By providing this opportunity for cross-discipline discussion, we hope to foster new ideas about how to learn and leverage grounding in machines as well as build new bridges between the science of human cognition and machine learning.

**Schedule**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:20 AM</td>
<td>Opening Remarks</td>
<td>Strub, de Vries, Das, Lee, Wijmans, Arad Hudson, Suhr</td>
</tr>
<tr>
<td>08:30 AM</td>
<td>TBA</td>
<td></td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Invited Talk</td>
<td>Jay McClelland</td>
</tr>
<tr>
<td>09:50 AM</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Spotlight</td>
<td>Timothy Lillicrap</td>
</tr>
<tr>
<td>10:50 AM</td>
<td>Invited Talk</td>
<td>Jesse Thomason</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Discussion 6</td>
<td></td>
</tr>
</tbody>
</table>

Page 34 of 55
Robust AI in Financial Services: Data, Fairness, Explainability, Trustworthiness, and Privacy

Abstracts (6):

Abstract 3: **Jay McClelland in Visually Grounded Interaction and Language**, McClelland 09:10 AM
Note that the schedule is not final, and may change.

Note that the schedule is not final, and may change.

Abstract 7: **Jesse Thomason in Visually Grounded Interaction and Language**, Thomason 11:30 AM
Note that the schedule is not final, and may change.

Abstract 9: **Lisa Anne Hendricks in Visually Grounded Interaction and Language**, Hendricks 01:50 PM
Note that the schedule is not final, and may change.

Abstract 10: **Linda Smith in Visually Grounded Interaction and Language**, Smith 02:30 PM
Note that the schedule is not final, and may change.

Abstract 13: **Josh Tenenbaum in Visually Grounded Interaction and Language**, Tenenbaum 04:40 PM
Note that the schedule is not final, and may change.

Robust AI in Financial Services: Data, Fairness, Explainability, Trustworthiness, and Privacy

Alina Oprea, Avigdor Gal, Isabelle Moulinier, Jiahao Chen, Manuela Veloso, Senthil Kumar, Tanveer Faruque

Fri Dec 13, 08:00 AM

The financial services industry has unique needs for robustness when adopting artificial intelligence and machine learning (AI/ML). Many challenges can be described as intricate relationships between algorithmic fairness, explainability, privacy, data management, and trustworthiness. For example, there are ethical and regulatory needs to prove that models used for activities such as credit decisioning and lending are fair and unbiased, or that machine reliance does not cause humans to miss critical pieces of data. The use and protection of customer data necessitates secure and privacy-aware computation, as well as explainability around the use of sensitive data. Some challenges like entity resolution are exacerbated because of scale, highly nuanced data points and missing information.

On top of these fundamental requirements, the financial industry is ripe with adversaries who purport fraud, resulting in large-scale data breaches and loss of confidential information in the financial industry. The need to counteract malicious actors therefore calls for robust methods that can tolerate noise and adversarial corruption of data. However, recent advances in adversarial attacks of AI/ML systems demonstrate how often generic solutions for robustness and security fail, thus highlighting the need for further advances. The challenge of robust AI/ML is further complicated by constraints on data privacy and fairness, as imposed by ethical and regulatory concerns like GDPR.

This workshop aims to bring together researchers and practitioners to discuss challenges for AI/ML in financial services, and the opportunities such challenges represent to research communities. The workshop will consist of invited talks, panel discussions and short paper presentations, which will showcase ongoing research and novel algorithms resulting from collaboration of AI/ML and cybersecurity communities, as well as the challenges that arise from applying these ideas in domain-specific contexts.
## Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 AM</td>
<td>Opening Remarks</td>
</tr>
<tr>
<td>09:15 AM</td>
<td>Invited talk #1</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Invited talk #2</td>
</tr>
<tr>
<td>09:55 AM</td>
<td>Invited talk #3</td>
</tr>
<tr>
<td>10:15 AM</td>
<td>Coffee break + poster set-up</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Invited talk #4</td>
</tr>
<tr>
<td>10:50 AM</td>
<td>Panel discussing how to increase transparency and dissemination of ‘soft knowledge’ in ML</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Lunch break</td>
</tr>
<tr>
<td>01:50 PM</td>
<td>Retrospective: An Intriguing Failing of Convolutional Neural Networks and the CoordConv Solution Liu</td>
</tr>
<tr>
<td>01:55 PM</td>
<td>Retrospective: Learning the structure of deep sparse graphical models Ghahramani</td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Invited talk #4</td>
</tr>
<tr>
<td>02:05 PM</td>
<td>Retrospective: FiLM: Visual Reasoning with a General Conditioning Layer</td>
</tr>
<tr>
<td>02:10 PM</td>
<td>Retrospective: Deep Ptych: Subsampled Fourier Ptychography via Generative Priors</td>
</tr>
<tr>
<td>02:15 PM</td>
<td>Retrospective: Markov games that people play Littman</td>
</tr>
<tr>
<td>02:25 PM</td>
<td>Retrospective: Deep Reinforcement Learning That Matters Islam</td>
</tr>
<tr>
<td>02:30 PM</td>
<td>Smarter prototyping for neural learning</td>
</tr>
<tr>
<td>02:35 PM</td>
<td>Advances in deep learning for skin cancer detection</td>
</tr>
<tr>
<td>02:40 PM</td>
<td>Unsupervised Minimax: Adversarial Curiosity, Generative Adversarial Networks, and Predictability Minimization Schmidhuber</td>
</tr>
<tr>
<td>02:45 PM</td>
<td>Posters + Coffee Break</td>
</tr>
<tr>
<td>04:10 PM</td>
<td>Invited talk #5</td>
</tr>
</tbody>
</table>

---

**Retrospectives: A Venue for Self-Reflection in ML Research**

Ryan Lowe, Yoshua Bengio, Joelle Pineau, Michela Paganini, Jessica Forde, Shagun Sodhani, Abhishek Gupta, Joel Lehman, Peter Henderson, Kanika Madan

Fri Dec 13, 08:00 AM

The NeurIPS Workshop on Retrospectives in Machine Learning will kick-start the exploration of a new kind of scientific publication, called retrospectives. The purpose of a retrospective is to answer the question:

“What should readers of this paper know now, that is not in the original publication?”

Retrospectives provide a venue for authors to reflect on their previous publications, to talk about how their intuitions have changed, to identify shortcomings in their analysis or results, and to discuss resulting extensions that may not be sufficient for a full follow-up paper. A retrospective is written about a single paper, by that paper’s author, and takes the form of an informal paper. The overarching goal of retrospectives is to improve the science, openness, and accessibility of the machine learning field, by widening what is publishable and helping to identifying opportunities for improvement. Retrospectives will also give researchers and practitioners who are unable to attend top conferences access to the author’s updated understanding of their work, which would otherwise only be accessible to their immediate circle.
Retrospectives

brainstorming session: how do we produce impact?

Optimal Transport for Machine Learning

Marco Cuturi, Gabriel Peyré, Rémi Flamary, Alexandra Suvorikova

Fri Dec 13, 08:00 AM

Optimal transport (OT) provides a powerful and flexible way to compare, interpolate and morph probability measures. Originally proposed in the eighteenth century, this theory later led to Nobel Prizes for Koopmans and Kantorovich as well as C. Villani and A. Figalli Fields’ Medals in 2010 and 2018. OT is now used in challenging learning problems that involve high-dimensional data such as the inference of individual trajectories by looking at population snapshots in biology, the estimation of generative models for images, or more generally transport maps to transform samples in one space into another as in domain adaptation. With more than a hundred papers mentioning Wasserstein or transport in their title submitted at NeurIPS this year, and several dozens appearing every month across ML/stats/imaging and data sciences, this workshop’s aim will be to federate and advance current knowledge in this rapidly growing field.

Schedule

08:00 AM Facundo Memoli  Mémoli
08:40 AM Karren Dai
09:00 AM Jon Weed  Niles-Weed
10:30 AM Stefanie Jegelka  Jegelka
11:10 AM SPOTLIGHTS 5 x 10
12:00 PM Poster Session
02:00 PM Geoffrey Schiebinger  Schiebinger
02:40 PM Charlie Frogner  Frogner
03:00 PM Aude Genevay  Genevay
04:20 PM Daniel Kuhn  Kuhn
05:00 PM Alexei Kroshnin  Kroshnin
05:20 PM Poster Session

MLSys: Workshop on Systems for ML

Aparna Lakshmiratan, Siddhartha Sen, Joseph Gonzalez, Dan Crankshaw, Sarah Bird

Fri Dec 13, 08:00 AM

A new area is emerging at the intersection of artificial intelligence, machine learning, and systems design. This has been accelerated by the explosive growth of diverse applications of ML in production, the continued growth in data volume, and the complexity of large-scale learning systems. The goal of this workshop is to bring together experts working at the crossroads of machine learning, system design and software engineering to explore the challenges faced when building large-scale ML systems. In particular, we aim to elicit new connections among these diverse fields, identifying theory, tools and design principles tailored to practical machine learning workflows. We also want to think about best practices for research in this area and how to evaluate it. The workshop will cover state of the art ML and AI platforms and algorithm toolkits (e.g. TensorFlow, PyTorch1.0, MXNet etc.), as well as dive into machine learning-focused developments in distributed learning platforms, programming languages, data structures, hardware accelerators, benchmarking systems and other topics.

This workshop will follow the successful model we have previously run at ICML, NeurIPS and SOSP.

Our plan is to run this workshop annually co-located with one ML venue and one Systems venue, to help build a strong community which we think will complement newer conferences like SysML targeting research at the intersection of systems and machine learning. We believe this dual approach will help to create a low barrier to participation for both communities.

This workshop is part two of a two-part series with one day focusing on ML for Systems and the other on Systems for ML. Although the two workshops are being led by different organizers, we are coordinating our call for papers to ensure that the workshops complement each other and that submitted papers are routed to the appropriate venue.

Schedule

09:00 AM Welcome
09:10 AM Keynote
09:30 AM Contributed Talk
10:00 AM Contributed Talk
10:30 AM Coffee Break
Minding the Gap: Between Fairness and Ethics

Igor Rubinov, Risi Kondor, Jack Poulson, Manfred K. Warmuth, Emanuel Moss, Alexa Hagerty

Fri Dec 13, 08:00 AM

When researchers and practitioners, as well as policy makers and the public, discuss the impacts of deep learning systems, they draw upon multiple conceptual frames that do not sit easily beside each other. Questions of algorithmic fairness arise from a set of concerns that are similar, but not identical, to those that circulate around AI safety, which in turn overlap with, but are distinct from, the questions that motivate work on AI ethics, and so on. Robust bodies of research on privacy, security, transparency, accountability, interpretability, explainability, and opacity are also incorporated into each of these frames and conversations in variable ways. These frames reveal gaps that persist across both highly technical and socially embedded approaches, and yet collaboration across these gaps has proven challenging.

Fairness, Ethics, and Safety in AI each draw upon different disciplinary prerogatives, variously centering applied mathematics, analytic philosophy, behavioral sciences, legal studies, and the social sciences in ways that make conversation between these frames fraught with misunderstandings. These misunderstandings arise from a high degree of linguistic slippage between different frames, and reveal the epistemic fractures that undermine valuable synergy and productive collaboration. This workshop focuses on ways to translate between these ongoing efforts and bring them into necessary conversation in order to understand the profound impacts of algorithmic systems in society.

Schedule

**10:30 AM** Posters and Coffee

**12:30 PM** Open source software showcase (TF, PyTorch, TVM...)

**02:00 PM** Coffee Break

**02:10 PM** Keynote

**02:30 PM** Contributed Talk

**02:50 PM** Contributed Talk

**03:10 PM** Contributed Talk

**03:30 PM** Posters and Coffee

**04:30 PM** Panel

---

 CiML 2019: Machine Learning Competitions for All

Adrienne Mendrik, Wei-Wei Tu, Isabelle Guyon, Evelyne Viegas, Ming LI

Fri Dec 13, 08:00 AM

Challenges in machine learning and data science are open online competitions that address problems by providing datasets or simulated environments. They measure the performance of machine learning algorithms with respect to a given problem. The playful nature of challenges naturally attracts students, making challenges a great teaching resource. However, in addition to the use of challenges as educational tools, challenges have a role to play towards a better democratization of AI and machine learning. They function as cost effective problem-solving tools and a means of encouraging the development of re-usable problem templates and open-sourced solutions. However, at present, the geographic, sociological repartition of challenge participants and organizers is very biased. While recent successes in machine learning have raised much hopes, there is a growing concern that the societal and economical benefits might increasingly be in the power and under control of a few.

CiML (Challenges in Machine Learning) is a forum that brings together workshop organizers, platform providers, and participants to discuss best practices in challenge organization and new methods and application opportunities to design high impact challenges. Following the success of previous years’ workshops, we will reconvene and discuss new opportunities for broadening our community.

For this sixth edition of the CiML workshop at NeurIPS our objective is twofold: (1) We aim to enlarge the community, fostering diversity in the community of participants and organizers; (2) We aim to promote the organization of challenges for the benefit of more diverse communities.
The workshop provides room for discussion on these topics and aims to bring together potential partners to organize such challenges and stimulate "machine learning for good", i.e. the organization of challenges for the benefit of society. We have invited prominent speakers that have experience in this domain.

### Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker/Sponsors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00 AM</td>
<td>Welcome and Opening Remarks</td>
<td>Mendrik, Guyon, Viegas, Tu, LI</td>
</tr>
<tr>
<td>08:15 AM</td>
<td>Amir Banifatemi (XPrize) &quot;AI for Good via Machine Learning Challenges&quot;</td>
<td>Banifatemi</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Emily Bender (University of Washington) &quot;Making Stakeholder Impacts Visible in the Evaluation Cycle: Towards Fairness-Integrated Shared Tasks and Evaluation Metrics&quot;</td>
<td>Bender</td>
</tr>
<tr>
<td>09:45 AM</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Dina Machuve (Nelson Mandela African Institution of Science and Technology) &quot;Machine Learning Competitions: The Outlook from Africa&quot;</td>
<td>Machuve</td>
</tr>
<tr>
<td>11:15 AM</td>
<td>Dog Image Generation Competition on Kaggle</td>
<td>Kan</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Learning To Run a Power Network Competition</td>
<td>Donnot</td>
</tr>
<tr>
<td>11:45 AM</td>
<td>The AI Driving Olympics: Accessible Robot Learning Benchmark</td>
<td>Walter</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Conclusion on TrackML, a Particle Physics Tracking</td>
<td>Rousseau</td>
</tr>
<tr>
<td>12:15 PM</td>
<td>Catered Lunch and Poster Viewing (in Workshop Room)</td>
<td>Stolovitzky, Pradhan, Dubouve, Tang, Natekin, Bondi, Bouthillier, Milani, Müller, Holzinger, Harrer, Day, Ustyuzhanin</td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Yang Yu (Nanjing University) on Machine Learning Challenges to Advance AI in China</td>
<td>Yu</td>
</tr>
<tr>
<td>02:45 PM</td>
<td>Design and Analysis of Experiments: A Challenge Approach in Teaching</td>
<td>Pavao</td>
</tr>
</tbody>
</table>

**Abstracts (11):**

**Abstract 2:** Amir Banifatemi (XPrize) "AI for Good via Machine Learning Challenges" in CiML 2019: Machine Learning Competitions for All, Banifatemi 08:15 AM

"AI for Good" efforts (e.g., applications work in sustainability, education, health, financial inclusion, etc.) have demonstrated the capacity to simultaneously advance intelligent system research and the greater good. Unfortunately, the majority of research that could find motivation in real-world "good" problems still center on problems with industrial or toy problem performance baselines.

Competitions can serve as an important shaping reward for steering academia towards research that is simultaneously impactful on our state of knowledge and the state of the world. This talk covers three aspects of AI for Good competitions. First, we survey current efforts within the AI for Good application space as a means of identifying current and future opportunities. Next we discuss how more qualitative notions of "Good" can be used as benchmarks in addition to more quantitative competition objective functions. Finally, we will provide notes on building coalitions of domain experts to develop and guide socially-impactful competitions in machine learning.

**Abstract 3:** Emily Bender (University of Washington) "Making Stakeholder Impacts Visible in the Evaluation Cycle: Towards Fairness-Integrated Shared Tasks and Evaluation Metrics" in CiML 2019: Machine Learning Competitions for All, Bender 09:00 AM

In a typical machine learning competition or shared task, success is measured in terms of systems' ability to reproduce gold-standard labels. The potential impact of the systems being developed on stakeholder populations, if considered at all, is studied separately from system 'performance'. Given the tight train-eval cycle of both shared tasks and system development in general, we argue that making disparate impact on vulnerable populations visible in dataset and metric design will be key to making the potential for such impact present and salient to
developers. We see this as an effective way to promote the development of machine learning technology that is helpful for people, especially those who have been subject to marginalization. This talk will explore how to develop such shared tasks, considering task choice, stakeholder community input, and annotation and metric design desiderata.

Joint work with Hal Daumé III, University of Maryland, Bernease Herman, University of Washington, and Brandeis Marshall, Spelman College.


The current AI landscape in Africa mainly focuses on capacity building. The ongoing efforts to strengthen the AI capacity in Africa are organized in summer schools, workshops, meetups, competitions and one long-term program at the Masters level. The main AI initiatives driving the AI capacity building agenda in Africa include a) Deep Learning Indaba, b) Data Science Africa, c) Data Science Nigeria, d) Nairobi Women in Machine Learning and Data Science, e) Zindi and f) The African Master's in Machine Intelligence (AMMI) at AIMS. The talk will summarize our experience on low participation of African AI developers at machine learning competitions and our recommendations to address the current challenges.

Abstract 6: Dog Image Generation Competition on Kaggle in CiML 2019: Machine Learning Competitions for All, Kan 11:15 AM

We present a novel format of machine learning competitions where a user submits code that generates images trained on training samples, the code then runs on Kaggle, produces dog images, and user receives scores for the performance of their generative content based on 1. quality of images, 2. diversity of images, and 3. memorization penalty. This style of competition targets the usage of Generative Adversarial Networks (GAN)[4], but is open for all generative models. Our implementation addresses overfitting by incorporating two different pre-trained neural networks, as well as two separate “ground truth” image datasets, for the public and private leaderboards. We also have an enclosed compute environment to prevent submissions of non-generated images. In this paper, we describe both the algorithmic and system design of our competition, as well as sharing our lessons learned from running this competition [6] in July 2019 with 900+ teams participating and over 37,000 submissions and their code received.

Abstract 7: Learning To Run a Power Network Competition in CiML 2019: Machine Learning Competitions for All, Donnot 11:30 AM

We present the results of the first edition as well as some perspective for a next potential edition of the “Learning To Run a Power Network” (L2RPN) competition to test the potential of Reinforcement Learning to solve a real-world problem of great practical importance: controlling power transportation in power grids while keeping people and equipment safe.


Despite recent breakthroughs, the ability of deep learning and reinforcement learning to outperform traditional approaches to control physically embodied robotic agents remains largely unproven. To help bridge this gap, we have developed the “AI Driving Olympics” (AI-DO), a competition with the objective of evaluating the state-of-the-art in machine learning and artificial intelligence for mobile robotics. Based on the simple and well specified autonomous driving and navigation environment called “Duckietown,” AI-DO includes a series of tasks of increasing complexity—from simple lane-following to fleet management. For each task, we provide tools for competitors to use in the form of simulators, data logs, code templates, baseline implementations, and low-cost access to robotic hardware. We evaluate submissions in simulation online, on standardized hardware environments, and finally at the competition events. We have held successful AI-DO competitions at NeurIPS 2018 and ICRA 2019, and will be holding AI-DO 3 at NeurIPS 2020. Together, these competitions highlight the need for better benchmarks, which are lacking in robotics, as well as improved mechanisms to bridge the gap between simulation and reality.

Abstract 10: Catered Lunch and Poster Viewing (in Workshop Room) in CiML 2019: Machine Learning Competitions for All, Stolovitzky, Pradhan, Duboue, Tang, Natekin, Bondi, Bouthillier, Milani, Müller, Holzinger, Harrer, Day, Ustyuzhanin 12:15 PM

Accepted Posters

"Kandinsky Patterns: An open toolbox for creating explainable machine learning challenges" Heimo Muller, Andreas Holzinger

"MOCA: An Unsupervised Algorithm for Optimal Aggregation of Challenge Submissions" Robert Vogel, Mehmet Eren Ahsen, Gustavo Stolovitzky

"Smart(er) Machine Learning for Practitioners" Prabhu Pradhan

"Improving Reproducibility of Benchmarks" Xavier Bouthillier

"Guaranteeing Reproducibility in Deep Learning Competitions" Brandon Houghton

"WikiCities: a Feature Engineering Educational Resource" Pablo Duboue

"Reinforcement Learning Meets Information Seeking: Dynamic Search Challenge" Zhiwen Tang, Grace Hui Yang

"AI Journey 2019: School Tests Solving Competition" Alexey Natekin, Peter Romov, Valentyn Malykh


"FDL: Mission Support Challenge" Luis F. Simões, Ben Day, Vinutha M. Sreenath, Callum Wilson

"From data challenges to collaborative gig science. Coopetitive research process and platform." Andrey Ustyuzhanin, Mikhail Belous, Leyla Khatbullina, Giles Strong

"Organizing crowd-sourced AI challenges in enterprise environments: opportunities and challenges" Mahtab Mirmomeni, Isabell Kiral, Subhrajit Roy, Todd Mummet, Alan Braz, Jason Tsay, Jianbin Tang, Umar Asif, Thomas Schafter, Eren Mehtem, Bruno De Assis Marques, Stefan Maetschke, Rania Khalaf, Michal Rosen-Zvi, John Cohn, Gustavo
Over the past few years, we have explored the benefits of involving students both in organizing and in participating in challenges as a pedagogical tool, as part of an international collaboration. Engaging in the design and resolution of a competition can be seen as a hands-on means of learning proper design and analysis of experiments and gaining a deeper understanding of other aspects of Machine Learning. Graduate students of University Paris-Sud (Paris, France) are involved in class projects in creating a challenge end-to-end, from defining the research problem, collecting or formatting data, creating a starting kit, to implementing and testing the website. The application domains and types of data are extremely diverse: medicine, ecology, marketing, computer vision, recommendation, text processing, etc. The challenges thus created are then used as class projects of undergraduate students who have to solve them, both at University Paris-Sud, and at Rensselaer Polytechnic Institute (RPI, New York, USA), to provide rich learning experiences at scale. New this year, students are involved in creating challenges motivated by “AI for good” and will create re-usable templates to inspire others to create challenges for the benefit of humanity.

Data competitions often rely on the physical distribution of data to challenge participants, a significant limitation given that much data is proprietary, sensitive, and often non-shareable. To address this, the DREAM Challenges has advanced a challenge framework called model-to-data (MTD), requiring participants to submit re-usable algorithms instead of model predictions. The DREAM organization has successfully completed multiple MTD-based challenges, and is expanding this approach to unlock highly sensitive and non-distributable human data for use in biomedical data challenges.

The typical setup in machine learning competitions is to provide one or more datasets and a performance metric, leaving it entirely up to participants which approach to use, how to engineer better features, whether and how to pretrain models on related data, how to tune hyperparameters, how to combine multiple models in an ensemble, etc. The fact that work on each of these components often leads to substantial improvements has several consequences: (1) amongst several skilled teams, the one with the most manpower and engineering drive often wins; (2) it is often unclear “why” one entry performs better than another one; and (3) scientific insights remain limited.

Based on my experience in both participating in several challenges and also organizing some, I will propose a new competition design that instead emphasizes scientific insight by dividing the various ways in which teams could improve performance into (largely orthogonal) modular components, each of which defines its own competition. E.g., one could run a competition focussing only on effective hyperparameter tuning of a given pipeline (across private datasets). With the same code base and datasets, one could likewise run a competition focussing only on finding better neural architectures, or only better preprocessing methods, or only a better training pipeline, or only better pre-training methods, etc. One could also run multiple of these competitions in parallel, hot-swapping better components found in one competition into the other competitions. I will argue that the result would likely be substantially more valuable in terms of scientific insights than traditional competitions and may even lead to better final performance.


*Raj Parihar, Michael Goldfarb, Satyam Srivastava, TAO SHENG*

**Fri Dec 13, 08:00 AM**

Sample Call for Papers:

A new wave of intelligent computing, driven by recent advances in machine learning and cognitive algorithms coupled with process technology and new design methodologies, has the potential to usher unprecedented disruption in the way modern computing systems are designed and deployed. These new and innovative approaches often provide an attractive and efficient alternative not only in terms of performance but also power, energy, and area. This disruption is easily visible across the whole spectrum of computing systems -- ranging from low end mobile devices to large scale data centers and servers including intelligent infrastructures.

A key class of these intelligent solutions is providing real-time, on-device cognition at the edge to enable many novel applications including computer vision and image processing, language understanding, speech and gesture recognition, malware detection and autonomous driving. Naturally, these applications have diverse requirements for performance, energy, reliability, accuracy, and security that demand a holistic approach to designing the hardware, software, and intelligence algorithms to achieve the best power, performance, and area (PPA).

**Topics:**

- Architectures for the edge: IoT, automotive, and mobile
- Approximation, quantization reduced precision computing
- Hardware/software techniques for sparsity
- Neural network architectures for resource constrained devices
- Neural network pruning, tuning and automatic architecture search
- Novel memory architectures for machine learning
- Communication/computation scheduling for better performance and energy
- Load balancing and efficient task distribution techniques
- Exploring the interplay between precision, performance, power and energy
- Exploration of new and efficient applications for machine learning
- Characterization of machine learning benchmarks and workloads
- Performance profiling and synthesis of workloads
- Simulation and emulation techniques, frameworks and platforms for machine learning
- Power, performance and area (PPA) based comparison of neural networks
- Verification, validation and determinism in neural networks
- Efficient on-device learning techniques
- Security, safety and privacy challenges and building secure AI systems

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Speaker(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00 AM</td>
<td>TBD</td>
<td>LeCun</td>
</tr>
<tr>
<td>08:45 AM</td>
<td>Efficient Computing for AI and Robotics</td>
<td>Sze</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Abandoning the Dark Arts: New Directions in Efficient DNN Design</td>
<td>Keutzer</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Poster Session 1</td>
<td>Spasov, Diego Andilla, Nayak, Zhang</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Putting the &quot;Machine&quot; Back in Machine Learning: The Case for Hardware-ML Model Co-design</td>
<td></td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Adaptive Multi-Task Neural Networks for Efficient Inference</td>
<td>Feris</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Oral Session 1</td>
<td>Yu, Wong, Hartmann, Li, Li</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Qualcomm Industry Talk</td>
<td></td>
</tr>
<tr>
<td>12:30 PM</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Cheap, Fast, and Low Power Deep Learning: I need it now!</td>
<td>Delp</td>
</tr>
<tr>
<td>02:45 PM</td>
<td>Advances and Prospects for In-memory Computing</td>
<td>Verma</td>
</tr>
<tr>
<td>03:15 PM</td>
<td>Algorithm-Accelerator Co-Design for Neural Network Specialization</td>
<td>Zhang</td>
</tr>
<tr>
<td>03:45 PM</td>
<td>Poster Session 2</td>
<td>Prato, Thakker, Galindez, Olascoga, Zhang, Partovi, Nia, Park</td>
</tr>
<tr>
<td>04:15 PM</td>
<td>Efficient Algorithms to Accelerate Deep Learning on Edge Devices</td>
<td>Han</td>
</tr>
<tr>
<td>04:45 PM</td>
<td>Oral Session 2</td>
<td>Liao, McKinstry, Izsak, Li</td>
</tr>
<tr>
<td>05:30 PM</td>
<td>Microsoft Industry Talk</td>
<td>Darvish Rouhani</td>
</tr>
<tr>
<td>06:00 PM</td>
<td>EMCC Results</td>
<td></td>
</tr>
</tbody>
</table>

Abstracts (9):

Abstract 1: TBD in EMC2: Energy Efficient Machine Learning and Cognitive Computing (5th edition), LeCun 08:00 AM

TBD


Computing near the sensor is preferred over the cloud due to privacy and/or latency concerns for a wide range of applications including robotics/drones, self-driving cars, smart Internet of Things, and portable/wearable electronics. However, at the sensor there are often stringent constraints on energy consumption and cost in addition to the throughput and accuracy requirements of the application. In this talk, we will describe how joint algorithm and hardware design can be used to reduce energy consumption while delivering real-time and robust performance for applications including deep learning, computer vision, autonomous navigation/exploration and video/image processing. We will show how energy-efficient techniques that exploit correlation and sparsity to reduce compute, data movement and storage costs can be applied to various tasks including image classification, depth estimation, super-resolution, localization and mapping.


Deep Neural Net models have provided the most accurate solutions to a very wide variety of problems in vision, language, and speech; however, the design, training, and optimization of efficient DNNs typically requires resorting to the “dark arts” of ad hoc methods and extensive hyperparameter tuning. In this talk we present our progress on abandoning these dark arts by using Differential Neural Architecture Search to guide the design of efficient DNNs and by using Hessian-based methods to guide the processes of training and quantizing those DNNs.


Machine learning (ML) applications have entered and impacted our lives unlike any other technology advance from the recent past. Indeed, almost every aspect of how we live or interact with others relies on or uses ML for applications ranging from image classification and object detection, to processing multimodal and heterogeneous datasets. While the holy grail for judging the quality of a ML model has largely been serving accuracy, and only recently its resource usage, neither of these metrics translate directly to energy efficiency, runtime, or mobile device battery lifetime. This talk will uncover the need for building accurate, platform-specific power and latency models for convolutional neural networks (CNNs) and efficient hardware-aware CNN design methodologies, thus allowing machine learners and hardware designers to identify not just the best accuracy NN configuration, but also those that satisfy given hardware constraints. Our proposed modeling framework is applicable to both high-end and mobile platforms and achieves 88.24% accuracy for latency, 88.34% for power, and 97.21% for energy prediction. Using similar predictive models, we demonstrate a novel differentiable neural architecture search (NAS) framework, dubbed...
Single-Path NAS, that uses one single-path over-parameterized CNN to encode all architectural decisions based on shared convolutional kernel parameters. Single-Path NAS achieves state-of-the-art top-1 ImageNet accuracy (75.62%), outperforming existing mobile NAS methods for similar latency constraints (~80ms) and finds the final configuration up to 5,000× faster compared to prior work. Combined with our quantized CNNs (Flexible Lightweight CNNs or FLightNNs) that customize precision level in a layer-wise fashion and achieve almost iso-accuracy at 5-10× energy reduction, such a modeling, analysis, and optimization framework is poised to lead to true co-design of hardware and ML model, orders of magnitude faster than state of the art, while satisfying both accuracy and latency or energy constraints.


Very deep convolutional neural networks have shown remarkable success in many computer vision tasks, yet their computational expense limits their impact in domains where fast inference is essential. While there has been significant progress on model compression and acceleration, most methods rely on a one-size-fits-all network, where the same set of features is extracted for all images or tasks, no matter their complexity. In this talk, I will first describe an approach called BlockDrop, which learns to dynamically choose which layers of a deep network to execute during inference, depending on the image complexity, so as to best reduce total computation without degrading prediction accuracy. Then, I will show how this approach can be extended to design compact multi-task networks, where a different set of layers is executed depending on the task complexity, and the level of feature sharing across tasks is automatically determined to maximize both the accuracy and efficiency of the model. Finally, I will conclude the talk presenting an efficient multi-scale neural network model, which achieves state-of-the-art results in terms of accuracy and FLOPS reduction on standard benchmarks such as the ImageNet dataset.


In this talk I will describe the need for low power machine learning systems. I will motivate this by describing several current projects at Purdue University that have a need for energy efficient deep learning and in some cases the real deployment of these methods will not be possible without lower power solutions. The applications include precision farming, health care monitoring, and edge-based surveillance.


Edge AI applications retain the need for high-performing inference models, while driving platforms beyond their limits of energy efficiency and throughput. Digital hardware acceleration, enabling 10-100x gains over general-purpose architectures, is already widely deployed, but is ultimately restricted by data-motion and memory accessing that dominates deep-learning computations. In-memory computing, based on both SRAM and emerging memory, offers fundamentally new tradeoffs for overcoming these barriers, with the potential for 10x higher energy efficiency and area-normalized throughput demonstrated in recent designs. But, those tradeoffs instate new challenges, especially affecting scaling to the level of computations required, integration in practical heterogeneous architectures, and mapping of diverse software. This talk examines those tradeoffs to characterize the challenges. It then explores recent research that provides promising paths forward, making in-memory computing more of a practical reality than ever before.


In recent years, machine learning (ML) with deep neural networks (DNNs) has been widely deployed in diverse application domains. However, the growing complexity of DNN models, the slowdown of technology scaling, and the proliferation of edge devices are driving a demand for higher DNN performance and energy efficiency. ML applications have shifted from general-purpose processors to dedicated hardware accelerators in both academic and commercial settings. In line with this trend, there has been an active body of research on both algorithms and hardware architectures for neural network specialization.

This talk presents our recent investigation into DNN optimization and low-precision quantization, using a co-design approach featuring contributions to both algorithms and hardware accelerators. First, we review static network pruning techniques and show a fundamental link between group convolutions and circulant matrices – two previously disparate lines of research in DNN compression. Then we discuss channel gating, a dynamic, fine-grained, and trainable technique for DNN acceleration. Unlike static approaches, channel gating exploits input-dependent dynamic sparsity at run time. This results in a significant reduction in compute cost with a minimal impact on accuracy. Finally, we present outlier channel splitting, a technique to improve DNN weight quantization by removing outliers from the weight distribution without retraining.


Efficient deep learning computing requires algorithm and hardware co-design to enable specialization. However, the extra degree of freedom creates a much larger design space. We propose AutoML techniques to architect efficient neural networks. We investigate automatically designing small and fast models (ProxylessNAS), auto channel pruning (AMC), and auto mixed-precision quantization (HAQ). We demonstrate such learning-based, automated design achieves superior performance and efficiency than rule-based human design. Moreover, we shorten the design cycle by 200× than previous work to efficiently search efficient models, so that we can afford to design specialized neural network models for different hardware platforms. We accelerate computation-intensive AI applications including (TSM) for efficient video recognition and PVCNN for efficient 3D recognition on point clouds. Finally, we'll describe scalable distributed training and the potential security issues of efficient deep learning.

Graph Representation Learning

Will Hamilton, Rianne van den Berg, Michael Bronstein, Stefanie Jegelka, Thomas Kipf, Jure Leskovec, Renjie Liao, Yizhou Sun, Petar Veličković

Fri Dec 13, 08:00 AM
Graph-structured data is ubiquitous throughout the natural and social sciences, from telecommunication networks to quantum chemistry. Building relational inductive biases into deep learning architectures is crucial if we want systems that can learn, reason, and generalize from this kind of data. Furthermore, graphs can be seen as a natural generalization of simpler kinds of structured data (such as images), and therefore, they represent a natural avenue for the next breakthroughs in machine learning.

Recent years have seen a surge in research on graph representation learning, including techniques for deep graph embeddings, generalizations of convolutional neural networks to graph-structured data, and neural message-passing approaches inspired by belief propagation. These advances in graph neural networks and related techniques have led to new state-of-the-art results in numerous domains, including chemical synthesis, 3D-vision, recommender systems, question answering, and social network analysis.

The workshop will consist of contributed talks, contributed posters, and invited talks on a wide variety of methods and problems related to graph representation learning. We will welcome 4-page original research papers on work that has not previously been published in a machine learning conference or workshop. In addition to traditional research paper submissions, we will also welcome 1-page submissions describing open problems and challenges in the domain of graph representation learning. These open problems will be presented as short talks (5-10 minutes) immediately preceding a coffee break to facilitate and spark discussions.

The primary goal for this workshop is to facilitate community building; with hundreds of new researchers beginning projects in this area, we hope to bring them together to consolidate this fast-growing area of graph representation learning into a healthy and vibrant subfield.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 AM</td>
<td>Invited talk: Marco Gori</td>
<td>Gori</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Invited talk: Marinka Zitnik</td>
<td>Zitnik</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Open Challenges - Spotlight</td>
<td>Sumba Toral, Maron, Kolbeinsson</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Invited talk: Andrew McCallum</td>
<td>McCallum</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Poster Session #1</td>
<td>Sanborn, Yao, Cai, Duan, Shao, Belli, Boyarski, Ye, Sarkar, KHADEMI, Wang, Wang, Wu, Joshi, Bose, Ma, Brockschmidt, Zambon, Graber, Glorot, Cameron, Huang, Li, Sun, SALHA, Vlastelica Pogar, Huang</td>
</tr>
<tr>
<td>12:30 PM</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>01:30 PM</td>
<td>Outstanding Contribution Talk #1</td>
<td></td>
</tr>
</tbody>
</table>

Information Theory and Machine Learning

Shengjia Zhao, Jiaming Song, Yanjun Han, Kristy Choi, Pratyusha Kalluri, Ben Poole, Alex Dimakis, Jiantao Jiao, Tsachy Weissman, Stefano Ermon

Fri Dec 13, 08:00 AM

Information theory is deeply connected to two key tasks in machine learning: prediction and representation learning. Because of these connections, information theory has found wide applications in machine learning tasks, such as proving generalization bounds, certifying fairness and privacy, optimizing information content of unsupervised/supervised representations, and proving limitations to prediction performance. Conversely, progress in machine learning have been successfully applied to classical information theory tasks such as compression and transmission.

These recent progress have lead to new open questions and opportunities: to marry the simplicity and elegance of information theoretic analysis with the complexity of modern high dimensional machine learning setups. However, because of the diversity of information theoretic research, different communities often progress independently despite shared questions and tools. For example, variational bounds to mutual information are concurrently developed in information theory, generative model, and learning theory communities.

This workshop hopes to bring together researchers from different disciplines, identify common grounds, and spur discussion on how information theory can apply to and benefit from modern machine learning setups.
Machine learning (ML) has seen a tremendous amount of recent success and has been applied in a variety of applications. However, it comes with several drawbacks, such as the need for large amounts of training data and the lack of explainability and verifiability of the results. In many domains, there is structured knowledge (e.g., from electronic health records, laws, clinical guidelines, or common sense knowledge) which can be leveraged for reasoning in an informed way (i.e., including the information encoded in the knowledge representation itself) in order to obtain high quality answers. Symbolic approaches for knowledge representation and reasoning (KRR) are less prominent today - mainly due to their lack of scalability - but their strength lies in the verifiable and interpretable reasoning that can be accomplished. The KR2ML workshop aims at the intersection of these two subfields of AI. It will shine a light on the synergies that (could/should) exist between KRR and ML, and will initiate a discussion about the key challenges in the field.

---

**KR2ML - Knowledge Representation and Reasoning Meets Machine Learning**

*Veronika Thost, Christian Muise, Kartik Talamadupula, Sameer Singh, Chris Ré*

**Fri Dec 13, 08:00 AM**

Machine learning (ML) has seen a tremendous amount of recent success and has been applied in a variety of applications. However, it comes with several drawbacks, such as the need for large amounts of training data and the lack of explainability and verifiability of the results. In many domains, there is structured knowledge (e.g., from electronic health records, laws, clinical guidelines, or common sense knowledge) which can be leveraged for reasoning in an informed way (i.e., including the information encoded in the knowledge representation itself) in order to obtain high quality answers. Symbolic approaches for knowledge representation and reasoning (KRR) are less prominent today - mainly due to their lack of scalability - but their strength lies in the verifiable and interpretable reasoning that can be accomplished. The KR2ML workshop aims at the intersection of these two subfields of AI. It will shine a light on the synergies that (could/should) exist between KRR and ML, and will initiate a discussion about the key challenges in the field.
Learning with Rich Experience: Integration of Learning Paradigms

Zhiting Hu, Andrew Wilson, Chelsea Finn, Lisa Lee, Taylor Berg-Kirkpatrick, Ruslan Salakhutdinov, Eric Xing

Fri Dec 13, 08:00 AM

Machine learning is about computational methods that enable machines to learn concepts and improve performance from experience. Here, experience can take diverse forms, including data examples, abstract knowledge, interactions and feedback from the environment, other models, and so forth. Depending on different assumptions on the types and amount of experience available there are different learning paradigms, such as supervised learning, active learning, reinforcement learning, knowledge distillation, adversarial learning, and combinations thereof. On the other hand, a hallmark of human intelligence is the ability to learn from all sources of information. In this workshop, we aim to explore various aspects of learning paradigms, particularly theoretical properties and formal connections between them, and new algorithms combining multiple modes of supervisions, etc.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
<th>Speaker(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:50 AM</td>
<td>Opening Remarks</td>
<td></td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Contributed Oral</td>
<td></td>
</tr>
<tr>
<td>09:10 AM</td>
<td>Invited Talk</td>
<td>Hadsell</td>
</tr>
<tr>
<td>09:45 AM</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Invited Talk</td>
<td>Mitchell</td>
</tr>
<tr>
<td>11:05 AM</td>
<td>Invited Talk</td>
<td>Blimes</td>
</tr>
<tr>
<td>11:40 AM</td>
<td>1min Lightning Talks - I</td>
<td>Chourasia, Xu, Cortes, Chang, Nagano, Min, Boecking, Tran, Seyed Ghasempour, Ding, Mani, Voleli, Fakoor, Xu, Marino, Lee, Tresp, Kayy, Zhang, Poczko, Khandelwal, Bardes, Shelhamer, Zhu, Li, Li, Krasheninnikov, Wang, Jaiswal, Barsoum, Sanjeev, Wattanavekin, Xie</td>
</tr>
<tr>
<td>11:55 AM</td>
<td>Poster Session</td>
<td></td>
</tr>
<tr>
<td>12:30 PM</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>02:00 PM</td>
<td>1min Lightning Talks - II</td>
<td></td>
</tr>
<tr>
<td>02:20 PM</td>
<td>Invited Talk</td>
<td>Abbeel</td>
</tr>
<tr>
<td>02:55 PM</td>
<td>Invited Talk</td>
<td>Choi</td>
</tr>
<tr>
<td>03:30 PM</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>04:20 PM</td>
<td>Invited Talk</td>
<td>Griffiths</td>
</tr>
<tr>
<td>04:55 PM</td>
<td>Contributed Oral</td>
<td></td>
</tr>
<tr>
<td>05:10 PM</td>
<td>Panel Discussion</td>
<td></td>
</tr>
<tr>
<td>05:50 PM</td>
<td>Closing Remarks</td>
<td></td>
</tr>
</tbody>
</table>

Machine Learning for the Developing World (ML4D): Challenges and Risks

Maria De-Arteaga, Amanda Coston, Tejumade Afonja

Fri Dec 13, 08:00 AM

As the use of machine learning becomes ubiquitous, there is growing interest in understanding how machine learning can be used to tackle global development challenges. The possibilities are vast, and it is important that we explore the potential benefits of such technologies, which has driven the agenda of the ML4D workshop in the past. However, there is a risk that technology optimism and a categorization of ML4D research as inherently “social good” may result in initiatives failing to account for unintended harms or deviating scarce funds towards initiatives that appear exciting but have no demonstrated effect. Machine learning technologies deployed in developing regions have often been created for different contexts and are trained with data that is not
representative of the new deployment setting. Most concerning of all, companies sometimes make the deliberate choice to deploy new technologies in countries with little regulation in order to experiment.

This year’s program will focus on the challenges and risks that arise when deploying machine learning in developing regions. This one-day workshop will bring together a diverse set of participants from across the globe to discuss essential elements for ensuring ML4D research moves forward in a responsible and ethical manner. Attendees will learn about potential unintended harms that may result from ML4D solutions, technical challenges that currently prevent the effective use of machine learning in vast regions of the world, and lessons that may be learned from other fields.

The workshop will include invited talks, a poster session of accepted papers and panel discussions. We welcome paper submissions featuring novel machine learning research that characterizes or tackle challenges of ML4D, empirical papers that reveal unintended harms of machine learning technology in developing regions, and discussion papers that examine the current state of the art of ML4D and propose paths forward.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:45 AM</td>
<td>Opening Remarks</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Deborah Raji</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Anubha Sinha</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Kentaro Toyama</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Coffee Break</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Breakout sessions</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Poster session</td>
</tr>
<tr>
<td>12:30 PM</td>
<td>Lunch</td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Contributed Talk</td>
</tr>
<tr>
<td>02:15 PM</td>
<td>Contributed Talk</td>
</tr>
<tr>
<td>02:30 PM</td>
<td>Contributed Talk</td>
</tr>
<tr>
<td>02:45 PM</td>
<td>Contributed Talk</td>
</tr>
<tr>
<td>03:00 PM</td>
<td>Coffee and Posters</td>
</tr>
<tr>
<td>03:30 PM</td>
<td>Grace Mutung’u</td>
</tr>
<tr>
<td>04:00 PM</td>
<td>Elisa Celis</td>
</tr>
<tr>
<td>04:30 PM</td>
<td>Panel Discussion: Risks and Challenges in ML4D</td>
</tr>
</tbody>
</table>

Biological and Artificial Reinforcement Learning

Raymond Chua, Sara Zannone, Feryal Behbahani, Rui Ponte Costa, Claudia Clopath, Blake Richards, Doina Precup

Fri Dec 13, 08:00 AM

Reinforcement learning (RL) algorithms learn through rewards and a process of trial-and-error. This approach was strongly inspired by the study of animal behaviour and has led to outstanding achievements in machine learning (e.g. in games, robotics, science). However, artificial agents still struggle with a number of difficulties, such as sample efficiency, learning in dynamic environments and over multiple timescales, generalizing and transferring knowledge. On the other end, biological agents excel at these tasks. The brain has evolved to adapt and learn in dynamic environments, while integrating information and learning on different timescales and for different duration. Animals and humans are able to extract information from the environment in efficient ways by directing their attention and actively choosing what to focus on. They can achieve complicated tasks by solving sub-problems and combining knowledge as well as representing the environment in efficient ways and plan their decisions off-line. Neuroscience and cognitive science research has largely focused on elucidating the workings of these mechanisms. Learning more about the neural and cognitive underpinnings of these functions could be key to developing more intelligent and autonomous agents. Similarly, having a computational and theoretical framework, together with a normative perspective to refer to, could and does contribute to elucidate the mechanisms used by animals and humans to perform these tasks. Building on the connection between biological and artificial reinforcement learning, our workshop will bring together leading and emergent researchers from Neuroscience, Psychology and Machine Learning to share: (i) how neural and cognitive mechanisms can provide insights to tackle challenges in RL research and (ii) how machine learning advances can help further our understanding of the brain and behaviour.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 AM</td>
<td>Opening Remarks</td>
</tr>
<tr>
<td>09:15 AM</td>
<td>Invited Talk #1: Jane Wang</td>
</tr>
<tr>
<td>09:45 AM</td>
<td>Contributed Talk #1: Humans flexibly transfer options at multiple levels of abstractions</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Contributed Talk #2: Slow processes of neurons enable a biologically plausible approximation to policy gradient</td>
</tr>
</tbody>
</table>

Page 47 of 55
10:15 AM  Coffee Break & Poster Session  
Mohinta, Agostinelli, Moringen, Lee, Lo, Maass, Sheffer, Bredenberg, Eysenbach, Xia, Markou, Lichtenberg, Richemond, Zhang, Lanier, Lin, Fedus, Berset, Sarrico, Crosby, McAleer, Ghiasian, Scherr, Bellec, Salaj, Kolbeinsson, Rosenberg, Shin, Lee, Cecchi, Rish, Hajek

10:45 AM  Invited Talk #2: Jacqueline Gottlieb  
Gottlieb

11:15 AM  Contributed Talk #3 MEMENTO: Further Progress Through Forgetting  
Brunskill

11:30 AM  Invited Talk #3 Emma Brunskill  
Brunskill

12:00 PM  Lunch Break & Poster Session  
Mordatch

Beyond first order methods in machine learning systems

Anastasios Kyrillidis, Albert Berahas, Fred Roosta, Michael W Mahoney

First-order methods, however, have significant limitations: (1) they require fine hyper-parameter tuning, (2) they do not incorporate curvature information, and thus are sensitive to ill-conditioning, and (3) they are often unable to fully exploit the power of distributed computing architectures.

Higher-order methods, such as Newton, quasi-Newton and adaptive gradient descent methods, are extensively used in many scientific and engineering domains. At least in theory, these methods possess several nice features: they exploit local curvature information to mitigate the effects of ill-conditioning, they avoid or diminish the need for hyper-parameter tuning, and they have enough concurrency to take advantage of distributed computing environments. Researchers have even developed stochastic versions of higher-order methods, that feature speed and scalability by incorporating curvature information in an economical and judicious manner. However, often higher-order methods are "undervalued."

This workshop will attempt to shed light on this statement. Topics of interest include --but are not limited to-- second-order methods, adaptive gradient descent methods, regularization techniques, as well as techniques based on higher-order derivatives.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker/Sponsors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00 AM</td>
<td>Opening Remarks</td>
<td>Kyrillidis, Berahas, Roosta, Mahoney</td>
</tr>
<tr>
<td>08:15 AM</td>
<td>Talk 1</td>
<td>Granziol, Pedregosa, Asi, Adolphs, Meng</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Spotlight talks</td>
<td>Gorbunov, d'Aspremont, Wang, Wang, Ginsburg, Quaglino, Castera, Adya, Granzoil, Das, Bollapragada, Pedregosa, Takas, Jahanli, Karimireddy, Asi, Daroczy, Adolphs, Rawal, Brandt, Li, Ughi, Romero, Skorokhodov, Scieur, Bae, Mishchenko, Anil, Sharan, Balu, Chen, Yao, Ergen, Grigas, Li, Ba, Roberts, Vaswani, Eftekhari, Sharma</td>
</tr>
<tr>
<td>09:45 AM</td>
<td>Poster Session</td>
<td></td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Economical use of second-order information in training machine learning models</td>
<td>Goldfarb</td>
</tr>
<tr>
<td>11:15 AM</td>
<td>Adaptive gradient methods: efficient implementation and generalization</td>
<td></td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Lunch break</td>
<td></td>
</tr>
<tr>
<td>02:00 PM</td>
<td>K-FAC: Extensions, improvements, and applications</td>
<td>Martens</td>
</tr>
<tr>
<td>02:45 PM</td>
<td>Spotlight talks</td>
<td>Grigas, Yao, Anil, Mishchenko, Scieur</td>
</tr>
</tbody>
</table>

Beyond first order methods in machine learning systems

Anastasios Kyrillidis, Albert Berahas, Fred Roosta, Michael W Mahoney

Optimization lies at the heart of many exciting developments in machine learning, statistics and signal processing. As models become more complex and datasets get larger, finding efficient, reliable and provable methods is one of the primary goals in these fields.

In the last few decades, much effort has been devoted to the development of first-order methods. These methods enjoy a low per-iteration cost and have optimal complexity, are easy to implement, and have proven to be effective for most machine learning applications.
Abstracts (12):

Abstract 1: Opening Remarks in Beyond first order methods in machine learning systems, Kyrillidis, Berahas, Roosta, Mahoney 08:00 AM

Opening remarks for the workshop

Abstract 2: Talk 1 in Beyond first order methods in machine learning systems, 08:15 AM

Plenary talk #1

Abstract 3: Spotlight talks in Beyond first order methods in machine learning systems, Granziol, Pedregosa, Asi, Adolphs, Meng 09:00 AM

How does mini-batching affect Curvature information for second order deep learning optimization? Diego Granziol (Oxford); Stephen Roberts (Oxford); Xingchen Wan (Oxford University); Stefan Zohren (University of Oxford); Binxin Ru (University of Oxford); Michael A. Osborne (University of Oxford); Andrew Wilson (NYU); sebastien ehhardt (Oxford); Dmitry P Vetrov (Higher School of Economics); Timur Garipov (Samsung AI Center in Moscow)

Acceleration through Spectral Modeling. Fabian Pedregosa (Google); Damien Scieur (Princeton University)

Using better models in stochastic optimization. Hilal Asi (Stanford University); John Duchi (Stanford University)

Ellipsoidal Trust Region Methods for Neural Nets. Leonard Adolphs (ETHZ); Jonas Kohler (ETHZ)

Sub-sampled Newton Methods Under Interpolation. Si Yi Meng (University of British Columbia); Sharan Vaswani (Mila, Université de Montréal); Issam Laradji (University of British Columbia); Mark Schmidt (University of British Columbia); Simon Lacoste-Julien (Mila, Université de Montréal)


An Accelerated Method for Derivative-Free Smooth Stochastic Convex Optimization. Eduard Gorbanov (Moscow Institute of Physics and Technology); Pavel Dvurechenskii (WIAS Germany); Alexander Gasnikov (Moscow Institute of Physics and Technology)

Fast Bregman Gradient Methods for Low-Rank Minimization Problems. Radu-Alexandru Dragomir (Université Toulouse 1); Jérôme Bolte (Université Toulouse 1); Alexandre d’Aspremont (Ecole Normale Superieure)

Gluster: Variance Reduced Mini-Batch SGD with Gradient Clustering. Fartash Faghri (University of Toronto); David Duvenaud (University of Toronto); David Fleet (University of Toronto); Jimmy Ba (University of Toronto)

Neural Policy Gradient Methods: Global Optimality and Rates of Convergence. Lingxiao Wang (Northwestern University); Qi Cai (Northwestern University); Zhaoran Yang (Princeton University); Zhaoran Wang (Northwestern University)

A Gram-Gauss-Newton Method Learning Overparameterized Deep Neural Networks for Regression Problems. Tianle Cai (Peking University); Ruqi Gao (Peking University); Jikai Hou (Peking University); Siyu Chen (Peking University); Dong Wang (Peking University); Di He (Peking University); Zhihua Zhang (Peking University); Liwei Wang (Peking University)

Stochastic Gradient Methods with Layerwise Adaptive Moments for Training of Deep Networks. Boris Ginsburg (NVIDIA); Oleksii Hrinchuk (NVIDIA); Jason Li (NVIDIA); Vitaly Lavrukhin (NVIDIA); Ryan Leary (NVIDIA); Oleksii Kuchaiev (NVIDIA); Jonathan Cohen (NVIDIA); Huyen Nguyen (NVIDIA); Yang Zhang (NVIDIA)

Accelerating Neural ODEs with Spectral Elements. Alessio Quaglino (NNAISENSE SA); Marco Gallieri (NNAISENSE); Jonathan Masci (NNAISENSE); Jan Koutnik (NNAISENSE)

An Inertial Newton Algorithm for Deep Learning. Camille Castera (CNRS, IRIT); Jérôme Bolte (Université Toulouse 1); Cédric Févotte (CNRS, IRIT); Edouard Pauwels (Toulouse 3 University)

Nonlinear Conjugate Gradients for Scaling Synchronous Distributed DNN Training. Saurabh Adya (Apple); Vinay Palakkode (Apple Inc.); Oncel Tuzel (Apple Inc.)

* How does mini-batching affect Curvature information for second order deep learning optimization? Diego Granziol (Oxford); Stephen Roberts (Oxford); Xingchen Wan (Oxford University); Stefan Zohren (University of Oxford); Michael A. Osborne (University of Oxford); Andrew Wilson (NYU); sebastien ehhardt (Oxford); Dmitry P Vetrov (Higher School of Economics); Timur Garipov (Samsung AI Center in Moscow)

On the Convergence of a Biased Version of Stochastic Gradient Descent. Rudrajit Das (University of Texas at Austin); Jiong Zhang (UT-Austin); Inderjit S. Dhillon (UT Austin & Amazon)

Adaptive Sampling Quasi-Newton Methods for Derivative-Free Stochastic Optimization. Raghu Bollapragada (Argonne National Laboratory); Stefan Wild (Argonne National Laboratory)
* Acceleration through Spectral Modeling. Fabian Pedregosa (Google); Damien Scieur (Princeton University)

Accelerating Distributed Stochastic L-BFGS by sampled 2nd-Order Information. Jie Liu (Lehigh University); Yu Rong (Tencent AI Lab); Martin Takac (Lehigh University); Junzhou Huang (Tencent AI Lab)

Grow Your Samples and Optimize Better via Distributed Newton CG and Accumulating Strategy. Majid Jahanian (Lehigh University); Xi He (Lehigh University); Chenzin Ma (Lehigh University); Aryan Mokhtari (UT Austin); Dheevatsa Mudigere (Intel Labs); Alejandro Ribeiro (University of Pennsylvania); Martin Takac (Lehigh University)

Global linear convergence of trust-region Newton's method without strong-convexity or smoothness. Sai Praneeth Karimireddy (EPFL); Sebastian Stich (EPFL); Martin Jaggi (EPFL)

FD-Net with Auxiliary Time Steps: Fast Prediction of PDEs using Hessian-Free Trust-Region Methods. Nur Sila Gulgec (Lehigh University); Zheng Shi (Lehigh University); Neil Deshmukh (MIT BeaverWorks - Medlytics); Shamim Pakzad (Lehigh University); Martin Takac (Lehigh University)

* Using better models in stochastic optimization. Hilal Asi (Stanford University); John Duchi (Stanford University)

Tangent space separability in feedforward neural networks. Bálint Daróczy (Institute for Computer Science and Control, Hungarian Academy of Sciences); Rita Aleksziev (Institute for Computer Science and Control, Hungarian Academy of Sciences); Andras Benczur (Hungarian Academy of Sciences)

* Ellipsoidal Trust Region Methods for Neural Nets. Leonard Adolphs (ETHZ); Jonas Kohler (ETHZ)

Closing the K-FAC Generalisation Gap Using Stochastic Weight Averaging. Xingchen Wan (University of Oxford); Diego Granziol (Oxford); Stefan Zohren (University of Oxford); Stephen Roberts (Oxford)

* Sub-sampled Newton Methods Under Interpolation. Si Yi Meng (University of British Columbia); Sharan Vaswani (Mila, Université de Montréal); Issam Laradji (University of British Columbia); Mark Schmidt (University of British Columbia); Simon Lacoste-Julien (Mila, Université de Montréal)

Learned First-Order Preconditioning. Aditya Rawal (Uber AI Labs); Rui Wang (Uber AI); Theodore Moskovitz (Gatsby Computational Neuroscience Unit); Sanyam Kapoor (Uber); Janice Lan (Uber AI); Jason Yosinski (Uber AI Labs); Thomas Miconi (Uber AI Labs)

Iterative Hessian Sketch in Input Sparsity Time. Charlie Dickens (University of Warwick); Graham Cormode (University of Warwick)

Nonlinear matrix recovery. Florentin Goyens (University of Oxford); Coralia Cartis (Oxford University); Armin Eftekhari (EPFL)

Making Variance Reduction more Effective for Deep Networks. Nicolas Brandt (EPFL); Farnood Salehi (EPFL); Patrick Thiran (EPFL)

Novel and Efficient Approximations for Zero-One Loss of Linear Classifiers. Hiva Ghanbari (Lehigh University); Minhan Li (Lehigh University); Katya Scheinberg (Lehigh)

A Model-Based Derivative-Free Approach to Black-Box Adversarial Examples: BOBYQA. Giuseppe Ughi (University of Oxford)

Distributed Accelerated Inexact Proximal Gradient Method via System of Coupled Ordinary Differential Equations. Chhavi Sharma (IIT Bombay); Vishnu Narayanan (IIT Bombay); Balamurugan Palaniappan (IIT Bombay)

Finite-Time Convergence of Continuous-Time Optimization Algorithms via Differential Inclusions. Orlando Romero (Rensselaer Polytechnic Institute); Mouhacine Benosman (MERL)

Loss Landscape Sightseeing by Multi-Point Optimization. Ivan Skorokhodov (MIPT); Mikhail Burtsev (NI)

* Symmetric Multisecant quasi-Newton methods. Damien Scieur (Samsung AI Research Montreal); Thomas Pumir (Princeton University); Nicolas Boumal (Princeton University)

Does Adam optimizer keep close to the optimal point? Kiwook Bae (KAIST); Heechang Ryu (KAIST); Hayong Shin (KAIST)

* Stochastic Newton Method and its Cubic Regularization via Majorization-Minimization. Konstantin Mishchenko (King Abdullah University of Science & Technology (KAUST)); Peter Richtarik (KAUST); Dmitry Kaledin (KAUST)

* Full Matrix Preconditioning Made Practical. Rohan Anil (Google); Vineet Gupta (Google); Tomer Koren (Google); Kevin Regan (Google); Yoram Singer (Princeton)

Memory-Sample Tradeoffs for Linear Regression with Small Error. Vatsal Sharan (Stanford University); Aaron Sidford (Stanford); Gregory Valiant (Stanford University)

On the Higher-order Moments in Adam. Zhanhong Jiang (Johnson Controls International); Aditya Balu (Iowa State University); Sin Yong Tan (Iowa State University); Young M Lee (Johnson Controls International); Chinnmay Hegde (Iowa State University); Soumik Sarkar (Iowa State University)

h-matrix approximation for Gauss-Newton Hessian. Chao Chen (UT Austin)

* Hessian-Aware trace-Weighted Quantization. Zhen Dong (UC Berkeley); Zhewei Yao (University of California, Berkeley); Amir Gholami (UC Berkeley); Yaohui Cai (Peking University); Daiyaan Arfeen (UC Berkeley); Michael Mahoney ("University of California, Berkeley"); Kurt Keutzer (UC Berkeley)

Random Projections for Learning Non-convex Models. Tolga Ergen (Stanford University); Emmanuel Candés (Stanford University); Mert Pilanci (Stanford)

* New Methods for Regularization Path Optimization via Differential Equations. Paul Grigas (UC Berkeley); Heyuan Liu (University of California, Berkeley)

Hessian-Aware Zeroth-Order Optimization. Haishan Ye (HKUST)
Stochastic gradient descent (SGD) and variants such as Adagrad and Adam, are extensively used today to train modern machine learning models. In this talk we will discuss ways to economically use second-order information to modify both the step size (learning rate) used in SGD and the direction taken by SGD. Our methods adaptively control the batch sizes used to compute gradient and Hessian approximations and ensure that the steps that are taken decrease the loss function with high probability assuming that the latter is self-concordant, as is true for many problems in empirical risk minimization. For such cases we prove that our basic algorithm is globally linearly convergent. A slightly modified version of our method is presented for training deep learning models. Numerical results will be presented that show that it exhibits excellent performance without the need for learning rate tuning. If there is time, additional ways to efficiently make use of second-order information will be presented.

Abstract 6: Adaptive gradient methods: efficient implementation and generalization in Beyond first order methods in machine learning systems. 11:15 AM

Adaptive gradient methods have had a transformative impact in deep learning. We will describe recent theoretical and experimental advances in their understanding, including low-memory adaptive preconditioning, and insights into their generalization ability.

Abstract 7: K-FAC: Extensions, improvements, and applications in Beyond first order methods in machine learning systems. Martens 02:00 PM

Second order optimization methods have the potential to be much faster than first order methods in the deterministic case, or pre-asymptotically in the stochastic case. However, traditional second order methods have proven ineffective or impractical for neural network training, due in part to the extremely high dimension of the parameter space. Kronecker-factored Approximate Curvature (K-FAC) is second-order optimization method based on a tractable approximation to the Gauss-Newton/Fisher matrix that exploits the special structure present in neural network training objectives. This approximation is neither low-rank nor diagonal, but instead involves Kronecker-products, which allows for efficient estimation, storage and inversion of the curvature matrix. In this talk I will introduce the basic K-FAC method for standard MLPs and then present some more recent work in this direction, including extensions to CNNs and RNNs, both of which requires new approximations to the Fisher. For these I will provide mathematical intuitions and empirical results which speak to their efficacy in neural network optimization. Time permitting, I will also discuss some recent results on large-batch optimization with K-FAC, and the use of adaptive adjustment methods that can eliminate the need for costly hyperparameter tuning.

Abstract 8: Spotlight talks in Beyond first order methods in machine learning systems. Grigas, Yao, Anil, Mishchenko, Scieur 02:45 PM

Symmetric Multisecant quasi-Newton methods. Damien Scieur (Samsung AI Research Montreal); Thomas Pumir (Princeton University); Nicolas Boumal (Princeton University)

Stochastic Newton Method and its Cubic Regularization via Majorization-Minimization. Konstantin Mishchenko (King Abdullah University of Science & Technology (KAUST)); Peter Richtarik (KAUST); Dmitry Koldaev (KAUST)

Full Matrix Preconditioning Made Practical. Rohan Anil (Google); Vineet Gupta (Google); Tomer Koren (Google); Kevin Regan (Google); Yoram Singer (Princeton)

Hessian-Aware trace-Weighted Quantization. Zhen Dong (UC Berkeley); Zhewei Yao (University of California, Berkeley); Amir Gholami (UC Berkeley); Yaohui Cai (Peking University); Daiyaan Arfeen (UC Berkeley); Michael Mahoney ("University of California, Berkeley"); Kurt Keutzer (UC Berkeley)

New Methods for Regularization Path Optimization via Differential Equations. Paul Grigas (UC Berkeley); Heyuan Liu (University of California, Berkeley)

Abstract 9: Poster Session (same as above) in Beyond first order methods in machine learning systems. 03:30 PM

An Accelerated Method for Derivative-Free Smooth Stochastic Convex Optimization. Eduard Gorbunov (Moscow Institute of Physics and Technology); Pavel Dvurechenskii (WIAS Germany); Alexander Gasnikov (Moscow Institute of Physics and Technology)

Fast Bregman Gradient Methods for Low-Rank Minimization Problems. Radu-Alexandru Dragomir (Université Toulouse 1); Jérôme Bolte (Université Toulouse 1); Alexandre d'Aspremont (Ecole Normale Superieure)

Gluster: Variance Reduced Mini-Batch SGD with Gradient Clustering. Fartash Faghi (University of Toronto); David Duvenaud (University of Toronto); David Fleet (University of Toronto); Jimmy Ba (University of Toronto)

Neural Policy Gradient Methods: Global Optimality and Rates of Convergence. Lingxiao Wang (Northwestern University); Qi Cai (Northwestern University); Zhaoran Wang (Northwestern University)

A Gram-Gauss-Newton Method Learning Overparameterized Deep Neural Networks for Regression Problems. Tianle Cai (Peking University); Ruqi Gao (Peking University); Jikai Hou (Peking University); Siyu Chen (Peking University); Dong Wang (Peking University); Di He (Peking University); Zhihua Zhang (Peking University); Zhaoran Wang (Northwestern University)

A Gram-Gauss-Newton Method Learning Overparameterized Deep Neural Networks for Regression Problems. Tianle Cai (Peking University); Ruqi Gao (Peking University); Jikai Hou (Peking University); Siyu Chen (Peking University); Dong Wang (Peking University); Di He (Peking University); Zhihua Zhang (Peking University); Zhaoran Wang (Northwestern University)

Stochastic Gradient Methods with Layerwise Adaptive Moments for Training of Deep Networks. Boris Ginsburg (NVIDIA); Oleksii Hinrichuk (NVIDIA); Jason Li (NVIDIA); Vitaly Lavrukhin (NVIDIA); Ryan Leary (NVIDIA); Oleksii Kuchaiev (NVIDIA); Jonathan Cohen (NVIDIA); Huyen Nguyen (NVIDIA); Yang Zhang (NVIDIA)

Accelerating Neural ODEs with Spectral Elements. Alessio Quaglini (NNAISENSE SA); Marco Gallieri (NNAISENSE); Jonathan Masci
(NNAISENSE); Jan Koutník (NNAISENSE)

An Inertial Newton Algorithm for Deep Learning. Camille Castera (CNRS, IRIT); Jérôme Bolte (Université Toulouse 1); Cédric Fèvotte (CNRS, IRIT); Edouard Pauwels (Toulouse 3 University)

Nonlinear Conjugate Gradients for Scaling Synchronous Distributed DNN Training. Saurabh Adya (Apple); Vinay Palakkode (Apple Inc.); Oncel Tuzel (Apple Inc.)

* How does mini-batching affect Curvature information for second order deep learning optimization? Diego Granziol (Oxford); Stephen Roberts (Oxford); Xingchen Wan (Oxford University); Stefan Zohren (University of Oxford); Bincin Ru (University of Oxford); Michael A. Osborne (University of Oxford); Andrew Wilson (NYU); sebastien ehhardt (Oxford); Dmitry P Vetrov (Higher School of Economics); Timur Garipov (Samsung AI Center in Moscow)

On the Convergence of a Biased Version of Stochastic Gradient Descent. Rudrajit Das (University of Texas at Austin); Jiong Zhang (UT-Austin); Inderjit S. Dhillon (UT Austin & Amazon)

Adaptive Sampling Quasi-Newton Methods for Derivative-Free Stochastic Optimization. Raghu Bollapragada (Argonne National Laboratory); Stefan Wild (Argonne National Laboratory)

* Acceleration through Spectral Modeling. Fabian Pedregosa (Google); Damien Scieur (Princeton University)

Accelerating Distributed Stochastic L-BFGS by sampled 2nd-Order Information. Jie Liu (Lehigh University); Yu Rong (Tencent AI Lab); Martin Takac (Lehigh University); Junzhou Huang (Tencent AI Lab)

Grow Your Samples and Optimize Better via Distributed Newton CG and Accumulating Strategy. Majid Jahani (Lehigh University); Xi He (Lehigh University); Chenxin Ma (Lehigh University); Aryan Mokhtari (UT Austin); Dheevatsa Mudigere (Intel Labs); Alejandro Ribeiro (University of Pennsylvania); Martin Takac (Lehigh University)

Global linear convergence of trust-region Newton’s method without strong-convexity or smoothness. Sai Praneeth Karimireddy (EPFL); Sebastian Stich (EPFL); Martin Jaggi (EPFL)

FD-Net with Auxiliary Time Steps: Fast Prediction of PDEs using Hessian-Free Trust-Region Methods. Nuri Sila Gulgec (Lehigh University); Zheng Shi (Lehigh University); Neil Deshmukh (MIT BeaverWorks - Medlytics); Shamim Pakzad (Lehigh University); Martin Takac (Lehigh University)

* Using better models in stochastic optimization. Hilal Asi (Stanford University); John Duchi (Stanford University)

Tangent space separability in feedforward neural networks. Bálint Daróczy (Institute for Computer Science and Control, Hungarian Academy of Sciences); Rita Alekseiev (Institute for Computer Science and Control, Hungarian Academy of Sciences); Andras Benczur (Hungarian Academy of Sciences)

* Ellipsoidal Trust Region Methods for Neural Nets. Leonard Adolphs (ETHZ); Jonas Kohler (ETHZ)

Closing the K-FAC Generalisation Gap Using Stochastic Weight Averaging. Xingchen Wan (University of Oxford); Diego Granziol (Oxford); Stefan Zohren (University of Oxford); Stephen Roberts (Oxford)

* Sub-sampled Newton Methods Under Interpolation. Si Yi Meng (University of British Columbia); Sharan Vaswani (Mila, Université de Montréal); Issam Laradji (University of British Columbia); Mark Schmidt (University of British Columbia); Simon Lacoste-Julien (Mila, Université de Montréal)

Learned First-Order Preconditioning. Aditya Rawal (Uber AI Labs); Rui Wang (Uber AI); Theodore Moskovitz (Gatsby Computational Neuroscience Unit); Sanyam Kapoor (Uber); Janice Lan (Uber AI); Jason Yosinski (Uber AI Labs); Thomas Miconi (Uber AI Labs)

Iterative Hessian Sketch in Input Sparsity Time. Charlie Dickens (University of Warwick); Graham Cormode (University of Warwick)

Nonlinear matrix recovery. Florentin Goyens (University of Oxford); Coralia Cartis (Oxford University); Armin Eftekhari (EPFL)

Making Variance Reduction more Effective for Deep Networks. Nicolas Brandt (EPFL); Farnood Salehi (EPFL); Patrick Thiran (EPFL)

Novel and Efficient Approximations for Zero-One Loss of Linear Classifiers. Hiva Ghanbari (Lehigh University); Minhan Li (Lehigh University); Katya Scheinberg (Lehigh)

A Model-Based Derivative-Free Approach to Black-Box Adversarial Examples: BOBYQA. Giuseppe Ughi (University of Oxford)

Distributed Accelerated Inexact Proximal Gradient Method via System of Coupled Ordinary Differential Equations. Chhavi Sharma (IIT Bombay); Vishnu Narayanan (IIT Bombay); Balamurugan Palaniappan (IIT Bombay)

Finite-Time Convergence of Continuous-Time Optimization Algorithms via Differential Inclusions. Orlando Romero (Rensselaer Polytechnic Institute); Mouhacine Benosman (MERL)

Loss Landscape Sightseeing by Multi-Point Optimization. Ivan Skorokhodov (MIPT); Mikhail Burtsev (NI)

* Symmetric Multisecant quasi-Newton methods. Damien Scieur (Samsung AI Research Montreal); Thomas Pumir (Princeton University); Nicolas Boumal (Princeton University)

Does Adam optimizer keep close to the optimal point? Kiwook Bae (KAIST)*; Heechang Ryu (KAIST); Hayoug Shin (KAIST)

* Stochastic Newton Method and its Cubic Regularization via Majorization-Minimization. Konstantin Mischenko (King Abdullah University of Science & Technology (KAUST)); Peter Richtarik (KAUST); Dmitry Koralev (KAUST)

* Full Matrix Preconditioning Made Practical. Rohan Anil (Google); Vineet Gupta (Google); Tomer Koren (Google); Kevin Regan (Google); Yoram Singer (Princeton)

Memory-Sample Tradeoffs for Linear Regression with Small Error. Vatsal Sharan (Stanford University); Aaron Sidford (Stanford); Gregory Valiant
Bayesian Deep Learning

Yarin Gal, Jose Miguel Hernández-Lobato, Christos Louizos, Eric Nalisnick, Zoubin Ghahramani, Kevin Murphy, Max Welling

Fri Dec 13, 08:00 AM

Extending on the workshop’s success from the past 3 years, this workshop will study the developments in the field of Bayesian deep learning (BDL) over the past year. The workshop will be a platform to host the recent flourish of ideas using Bayesian approaches in deep learning, and using deep learning tools in Bayesian modelling. The program includes a mix of invited talks, contributed talks, and contributed posters. Future directions for the field will be debated in a panel discussion.

Speakers:
* Andrew Wilson
* Deborah Marks
* Jasper Snoek
* Roger Grosse
* Chelsea Finn
* Yingzhen Li
* Alexander Matthews

Workshop summary:
While deep learning has been revolutionary for machine learning, most modern deep learning models cannot represent their uncertainty nor take advantage of the well studied tools of probability theory. This has started to change following recent developments of tools and techniques combining Bayesian approaches with deep learning. The intersection of the two fields has received great interest from the community, with the introduction of new deep learning models that take advantage of Bayesian techniques, and Bayesian models that incorporate deep learning elements. Many ideas from the 1990s are now being revisited in light of recent advances in the fields of approximate inference and deep learning, yielding many exciting new results.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00 AM</td>
<td>Opening remarks</td>
</tr>
<tr>
<td>08:05 AM</td>
<td>Invited talk</td>
</tr>
<tr>
<td>08:25 AM</td>
<td>Contributed talk</td>
</tr>
</tbody>
</table>
AI for Humanitarian Assistance and Disaster Response

Ritwik Gupta, Robin Murphy, Trevor Darrell, Eric Heim, Zhangyang Wang, Bryce Goodman, Piotr Biliński

Fri Dec 13, 08:00 AM

Natural disasters are one of the oldest threats to not just individuals but to the societies they co-exist in. As a result, humanity has ceaselessly sought way to provide assistance to people in need after disasters have struck. Further, natural disasters are but a single, extreme example of the many possible humanitarian crises. Disease outbreak, famine, and oppression against disadvantaged groups can pose even greater dangers to people that have less obvious solutions.

In this proposed workshop, we seek to bring together the Artificial Intelligence (AI) and Humanitarian Assistance and Disaster Response (HADR) communities in order to bring AI to bear on real-world humanitarian crises.

Through this workshop, we intend to establish meaningful dialogue between the communities.

By the end of the workshop, the NeurIPS research community can come to understand the practical challenges of in aiding those in crisis, while the HADR can understand the landscape that is the state of art and practice in AI.

Through this, we seek to begin establishing a pipeline of transitioning the research created by the NeurIPS community to real-world humanitarian issues.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Event Description</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00 AM</td>
<td>Introduction and Welcome</td>
<td>Gupta, Sajeev</td>
</tr>
<tr>
<td>08:15 AM</td>
<td>Invited Talks (x4)</td>
<td>Matias, Adole, Brown</td>
</tr>
<tr>
<td>10:15 AM</td>
<td>Spotlight Talks (x6)</td>
<td>Kruspe, Dalmasso, Schrempf, Oh, Doshi, Lu</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>01:00 PM</td>
<td>Invited Talks (x4)</td>
<td>Rasmussen, Stromberg, Darrell</td>
</tr>
<tr>
<td>03:00 PM</td>
<td>Spotlight Talks (x6)</td>
<td>Wang, Seo, Veitch-Michaels, Sidrane, Kapadia, Nevo, Dubey</td>
</tr>
<tr>
<td>04:30 PM</td>
<td>Convergence: Two-Way Limitations in Taking Theory to Applications</td>
<td>Dzombak</td>
</tr>
<tr>
<td>05:15 PM</td>
<td>Poster Session</td>
<td></td>
</tr>
</tbody>
</table>

Abstracts (4):

Abstract 2: Invited Talks (x4) in AI for Humanitarian Assistance and Disaster Response, Matias, Adole, Brown 08:15 AM

* Yossi Matias
* Tracy Adole
* Col Jason Brown
* Yang Cai
Safety and Robustness in Decision-making

Mohammad Ghavamzadeh, Shie Mannor, Yisong Yue, Marek Petrik, Yinlam Chow

Fri Dec 13, 08:00 AM

Interacting with increasingly sophisticated decision-making systems is becoming more and more a part of our daily life. This creates an immense responsibility for designers of these systems to build them in a way to guarantee safe interaction with their users and good performance, in the presence of noise and changes in the environment, and/or of model misspecification and uncertainty. Any progress in this area will be a huge step forward in using decision-making algorithms in emerging high stakes applications, such as autonomous driving, robotics, power systems, health care, recommendation systems, and finance.

This workshop aims to bring together researchers from academia and industry in order to discuss main challenges, describe recent advances, and highlight future research directions pertaining to develop safe and robust decision-making systems. We aim to highlight new and emerging theoretical and applied research opportunities for the community that arise from the evolving needs for decision-making systems and algorithms that guarantee safe interaction and good performance under a wide range of uncertainties in the environment.

Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:35 AM</td>
<td>Opening</td>
</tr>
<tr>
<td></td>
<td>Ghosh, Shafiee, Boopathy, Tamkin</td>
</tr>
<tr>
<td>09:00 AM</td>
<td>Aviv Tamar</td>
</tr>
<tr>
<td>09:30 AM</td>
<td>Scott Niekum</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Break and Posters</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Marco Pavone</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Dimitar Filev</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Lunch</td>
</tr>
<tr>
<td>02:00 PM</td>
<td>Nathan Kallus</td>
</tr>
</tbody>
</table>