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

Schedule Highlights

June 14, 2019

204, **Negative Dependence: Theory and Applications in Machine Learning** *Gartrell, Gillenwater, Kulesza, Mariet*

Grand Ballroom A, **1st Workshop on Understanding and Improving Generalization in Deep Learning** *Krishnan, Mobahi, Neyshabur, Bartlett, Song, Srebro*

Grand Ballroom B, **6th ICML Workshop on Automated Machine Learning (AutoML 2019)** *Hutter, Vanschoren, Eggensperger, Feurer*

Hall A, **Generative Modeling and Model-Based Reasoning for Robotics and AI** *Rajeswaran, Todorov, Mordatch, Agnew, Zhang, Pineau, Chang, Erhan, Levine, Stachenfeld, Zhang*

Hall B, **Uncertainty and Robustness in Deep Learning** *Li, Lakshminarayanan, Hendrycks, Dietterich, Gilmer*

Room 101, **ICML 2019 Workshop on Computational Biology** *Pe'er, Prabhakaran, Azizi, Diallo, Kundaje, Engelhardt, Dhifli, MEPHU NGUIFO, Tansey, Vogt, Listgarten, Burdziak, CompBio*

Room 102, **ICML 2019 Time Series Workshop** *Kuznetsov, Yang, Yu, Tang, Wang*

Room 103, **Human In the Loop Learning (HILL)** *Wang, Wang, Yu, Zhang, Gonzalez, Jia, Bird, Varshney, Kim, Weller*

Room 104 A, **Climate Change: How Can AI Help?** *Rolnick, Lacoste, Maharaj, Chayes, Bengio*

Room 104 B, **Workshop on the Security and Privacy of Machine Learning** *Papernot, Tramer, Li, Boneh, Evans, Jha, Liang, McDaniel, Steinhardt, Song*

Room 104 C, **Theoretical Physics for Deep Learning** *Lee, Pennington, Bahri, Welling, Ganguli, Bruna*

Room 201, **AI in Finance: Applications and Infrastructure for Multi-Agent Learning** *Reddy, Balch, Wellman, Kumar, Stoica, Elkind*

Room 202, **The Third Workshop On Tractable Probabilistic Modeling (TPM)** *Lowd, Vergari, Molina, Rahman, Domingos, Vergari*

Room 203, **Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)** *Ravi, Kozareva, Fan, Welling, Chen, Bailer, Kulis, Hu, Dekhtiar, Lin, Marculescu*

Seaside Ballroom, **Reinforcement Learning for Real Life** *Li, Geramifard, Li, Szepesvari, Wang*

Seaside Ballroom, **Real-world Sequential Decision Making: Reinforcement Learning and Beyond** *Le, Yue, Swaminathan, Boots, Cheng*

June 15, 2019

204, **Machine Learning for Music Discovery** *Schmidt, Nieto, Gouyon, Kinnaird, Lanckriet*

Grand Ballroom A, **Workshop on Self-Supervised Learning** *van den Oord, Aytar, Doersch, Vondrick, Radford, Sermanet, Zamir, Abbeel*

Grand Ballroom B, **Learning and Reasoning with Graph-Structured Representations** *Fetaya, Hu, Kipf, Li, Liang, Liao, Urtasun, Wang, Welling, Xing, Zemel*

Hall A, **Exploration in Reinforcement Learning Workshop** *Bhupatiraju, Eysenbach, Gu, Edwards, White, Oudeyer, Stanley, Brunskill*

Hall B, **Identifying and Understanding Deep Learning Phenomena** *Sedghi, Bengio, Hata, Madry, Morcos, Neyshabur, Raghu, Rahimi, Schmidt, Xiao*

Room 101, **Workshop on AI for autonomous driving** *Choromanska, Jackel, Li, Niebles, Gaidon, Posner, Chao*

Room 102, **Workshop on Multi-Task and Lifelong Reinforcement Learning** *Chandar, Sodhani, Khetarpal, Zahavy, Mankowitz, Mannor, Ravindran, Precup, Finn, Gupta, Zhang, Cho, Rusu, Rob Fergus*

Room 103, **Invertible Neural Networks and Normalizing Flows** *Huang, Krueger, Van den Berg, Papamakarios, Gomez, Cremer, Chen, Courville, J. Rezende*

Room 104 A, **Stein’s Method for Machine Learning and Statistics** *Briol, Mackey, Oates, Liu, Goldstein*

Room 104 B, **AI For Social Good (AISG)** *Luck, Sankaran, Sylvain, McGregor, Penn, Tadesse, Sylvain, Côté, Mackey, Ghani, Bengio*

Room 104 C, **Synthetic Realities: Deep Learning for Detecting AudioVisual Fakes** *Biggio, Korshunov, Mensink, Patrini, Sadhu, Rao*

Room 201, **ICML Workshop on Imitation, Intent, and Interaction (I3)** *Rhinehart, Levine, Finn, He, Kostrikov, Fu, Reddy*

Room 202, **Coding Theory For Large-scale Machine Learning** *Cadambe, Grover, Papailiopoulos, Joshi*

Room 203, **The How2 Challenge: New Tasks for Vision & Language** *Metze, Specia, Elliot, Barrault, Sanabria, Palaskar*

Seaside Ballroom, **Adaptive and Multitask Learning: Algorithms & Systems** *Al-Shedivat, Platanios, Stretcu, Andreas, Talwalkar, Caruana, Mitchell, Xing*

June 14, 2019

Negative Dependence: Theory and Applications in Machine Learning

Mike Gartrell, Jennifer Gillenwater, Alex Kulesza, Zelda Mariet

204, Fri Jun 14, 08:30 AM

Models of negative dependence are increasingly important in machine learning. Whether selecting training data, finding an optimal experimental design, exploring in reinforcement learning, or making suggestions with recommender systems, selecting high-quality but diverse items has become a core challenge. This workshop aims to bring together researchers who, using theoretical or applied techniques, leverage negative dependence in their work. We will delve into the rich underlying mathematical theory, understand key applications, and discuss the most promising directions for future research.

Schedule

08:45 AM	**Opening Remarks**	
08:50 AM	**Victor-Emmanuel Brunel: Negative Association and Discrete Determinantal Point Processes**	*Brunel*
09:30 AM	**Aarti Singh: Experimental Design**	*Singh*
10:10 AM	**On Two Ways to use Determinantal Point Processes for Monte Carlo Integration**	*Gautier*
10:30 AM	**[Coffee Break]**	
11:00 AM	**Jeff Bilmes: Deep Submodular Synergies**	*Bilmes*
11:40 AM	**Submodular Batch Selection for Training Deep Neural Networks**	*N Balasubramanian*
12:00 PM	**[Lunch Break]**	
02:00 PM	**Michal Valko: How Negative Dependence Broke the Quadratic Barrier for Learning with Graphs and Kernels**	*Valko*
02:40 PM	**Exact Sampling of Determinantal Point Processes with Sublinear Time Preprocessing**	*Derezinski*
03:00 PM	**[Coffee Break]**	
03:30 PM	**Sergei Levine: Distribution Matching and Mutual Information in Reinforcement Learning**	*Levine*
04:10 PM	**Seq2Slate: Re-ranking and Slate Optimization with RNNs**	*Meshi*
04:30 PM	**[Mini Break]**	
04:40 PM	**Cheng Zhang: Active Mini-Batch Sampling using Repulsive Point Processes**	*Zhang*
05:20 PM	**Gaussian Process Optimization with Adaptive Sketching: Scalable and No Regret**	*Valko*
05:40 PM	**Towards Efficient Evaluation of Risk via Herding**	*Xu*
06:00 PM	**Closing Remarks**	

Abstracts (12):

Abstract 2: **Victor-Emmanuel Brunel: Negative Association and Discrete Determinantal Point Processes in Negative Dependence: Theory and Applications in Machine Learning**, *Brunel* 08:50 AM

Discrete Determinantal Point Processes (DPPs) form a class of probability distributions that can describe the random selection of items from a finite or countable collection. They naturally arise in many problems in probability theory, and they have gained a lot of attention in machine learning, due to both their modeling flexibility and their tractability. In the finite case, a DPP is parametrized by a matrix, whose principal minors are the weights given by the DPP to each possible subset of items. When the matrix is symmetric, the DPP has a very special property, called Negative Association. Thanks to this property, symmetric DPPs enforce diversity within the randomly selected items, which is a feature that is sought for in many applications of Machine Learning, such as recommendation systems.

Abstract 3: **Aarti Singh: Experimental Design in Negative Dependence: Theory and Applications in Machine Learning**, *Singh* 09:30 AM

TBD

Abstract 4: **On Two Ways to use Determinantal Point Processes for Monte Carlo Integration in Negative Dependence: Theory and Applications in Machine Learning**, *Gautier* 10:10 AM

This paper focuses on Monte Carlo integration with determinantal point processes (DPPs) which enforce negative dependence between quadrature nodes. We survey the properties of two unbiased Monte Carlo estimators of the integral of inter- est: a direct one proposed by Bardenet & Hardy (2016) and a less obvious 60-year-old estimator by Ermakov & Zolotukhin (1960) that actually also relies on DPPs. We provide an efficient implementation to sample exactly a particular multidimensional DPP called multivariate Jacobi ensemble. This let us investigate the behavior of both estimators on toy problems in yet unexplored regimes.

Abstract 6: **Jeff Bilmes: Deep Submodular Synergies in Negative Dependence: Theory and Applications in Machine Learning**, *Bilmes* 11:00 AM

Submodularity is an attractive framework in machine learning to model concepts such as diversity, dispersion, and cooperative costs, and is having an ever increasing impact on the field of machine learning. Deep learning is having a bit of success as well. In this talk, we will discuss synergies, where submodular functions and deep neural networks can be used together to their mutual benefit. First, we'll discuss deep submodular functions (DSFs), an expressive class of functions that include many widely used submodular functions and that are defined analogously to deep neural networks (DNN). We'll show that the class of DSFs strictly increases with depth and discuss applications. Second, we'll see how a modification to DNN autoencoders can produce features that can be used in DSFs. These DSF/DNN hybrids address an open problem which is how best to produce a submodular function for your application. Third, we'll see how submodular functions can speed up the training of models. In one case, submodularity can be used to produce a sequence of mini-batches that speeds up training of DNN systems. In another case, submodularity can produce a training data subset for which we can show faster convergence to the optimal solution in the convex case. Empirically, this method speeds up gradient methods by up to 10x for convex and 3x for non-convex (i.e., deep) functions.

The above discusses various projects that were performed jointly with Wenruo Bai, Shengjie Wang, Chandrashekhar Lavania, Baharan Mirzasoleiman, and Jure Leskovec.

Abstract 7: **Submodular Batch Selection for Training Deep Neural Networks in Negative Dependence: Theory and Applications in Machine Learning**, *N Balasubramanian* 11:40 AM

Mini-batch gradient descent based methods are the de facto algorithms for training neural network architectures today. We introduce a mini-batch selection strategy based on submodular function maximization. Our novel submodular formulation captures the informativeness of each sample and diversity of the whole subset. We design an efficient, greedy algorithm which can give high-quality solutions to this NP-hard combinatorial optimization problem. Our extensive experiments on standard datasets show that the deep models trained using the proposed batch selection strategy provide better generalization than Stochastic Gradient Descent as well as a popular baseline sampling strategy across different learning rates, batch sizes, and distance metrics.

Abstract 9: **Michal Valko: How Negative Dependence Broke the Quadratic Barrier for Learning with Graphs and Kernels in Negative Dependence: Theory and Applications in Machine Learning**, *Valko* 02:00 PM

As we advance with resources, we move from reasoning on entities to reasoning on pairs and groups. We have beautiful frameworks: graphs, kernels, DPPs. However, the methods that work with pairs, relationships, and similarity are just slow. Kernel regression, or second-order gradient methods, or sampling from DPPs do not scale to large data, because of the costly construction and storing of matrix K_n. Prior work showed that sampling points according to their ridge leverage scores (RLS) generates small dictionaries with strong spectral approximation guarantees for K_n. However, computing exact RLS requires building and storing the whole kernel matrix. In this talk, we start with SQUEAK, a new online approach for kernel approximations based on RLS sampling that sequentially processes the data, storing a dictionary with a number of points that only depends on the effective dimension d_eff(gamma) of the dataset. The beauty of negative dependence, that we estimate on the fly, makes it possible to exclude huge portions of dictionary. With the small dictionary, SQUEAK never constructs the whole matrix K_n, runs in linear time O(n*d_eff(gamma)^3) w.r.t. n, and requires only a single pass over the dataset. A distributed version of SQUEAK runs in as little as O(log(n)*d_eff(gamma)^3) time. This online tool opens out a range of possibilities to finally have scalable, adaptive, and provably accurate kernel methods: semi-supervised learning or Laplacian smoothing on large graphs, scalable GP-UCB, efficient second-order kernel online learning, and even fast DPP sampling, some of these being featured in this workshop.

Abstract 10: **Exact Sampling of Determinantal Point Processes with Sublinear Time Preprocessing in Negative Dependence: Theory and Applications in Machine Learning**, *Derezinski* 02:40 PM

We study the complexity of sampling from a distribution over all index subsets of the set {1,...,n} with the probability of a subset S proportional to the determinant of the submatrix L_S of some n x n p.s.d. matrix L, where L_S corresponds to the entries of L indexed by S. Known as a determinantal point process, this distribution is widely used in machine learning to induce diversity in subset selection. In practice, we often wish to sample multiple subsets S with small expected size k = E[|S|] << n from a very large matrix L, so it is important to minimize the preprocessing cost of the procedure (performed once) as well as the sampling cost (performed repeatedly). To that end, we propose a new algorithm which, given access to L, samples exactly from a determinantal point process while satisfying the following two properties: (1) its preprocessing cost is n x poly(k) (sublinear in the size of L) and (2) its sampling cost is poly(k) (independent of the size of L). Prior to this work, state-of-the-art exact samplers required O(n^3) preprocessing time and sampling time linear in n or dependent on the spectral properties of L.

Abstract 12: **Sergei Levine: Distribution Matching and Mutual Information in Reinforcement Learning in Negative Dependence: Theory and Applications in Machine Learning**, *Levine* 03:30 PM

Conventionally, reinforcement learning is considered to be a framework for optimization: the aim for standard reinforcement learning algorithms is to recover an optimal or near-optimal policy that maximizes the reward over time. However, when considering more advanced reinforcement learning problems, from inverse reinforcement learning to unsupervised and hierarchical reinforcement learning, we often encounter settings where it is desirable to learn policies that match target distributions over trajectories or states, covering all modes, or else to simply learn collections of behaviors that are as broad and varied as possible. Information theory and probabilistic inference offer is a powerful set of tools for developing algorithms for these kinds of distribution matching problems. In this talk, I will outline methods that combine reinforcement learning, inference, and information theory to learn policies that match target distributions and acquire diverse behaviors, and discuss the applications of such methods for a variety of problems in artificial intelligence and robotics.

Abstract 13: **Seq2Slate: Re-ranking and Slate Optimization with RNNs in Negative Dependence: Theory and Applications in Machine**

Learning, *Meshi* 04:10 PM

Ranking is a central task in machine learning and information retrieval. In this task, it is especially important to present the user with a slate of items that is appealing as a whole. This in turn requires taking into account interactions between items, since intuitively, placing an item on the slate affects the decision of which other items should be placed alongside it. In this work, we propose a sequence-to-sequence model for ranking called seq2slate. At each step, the model predicts the next "best" item to place on the slate given the items already selected. The sequential nature of the model allows complex dependencies between the items to be captured directly in a flexible and scalable way. We show how to learn the model end-to-end from weak supervision in the form of easily obtained click-through data. We further demonstrate the usefulness of our approach in experiments on standard ranking benchmarks as well as in a real-world recommendation system.

Abstract 15: **Cheng Zhang: Active Mini-Batch Sampling using Repulsive Point Processes in Negative Dependence: Theory and Applications in Machine Learning**, *Zhang* 04:40 PM

We explore active mini-batch selection using repulsive point processes for stochastic gradient descent (SGD). Our approach simultaneously introduces active bias and leads to stochastic gradients with lower variance. We show theoretically and empirically that our approach improves over standard SGD both in terms of convergence speed as well as final model performance.

Abstract 16: **Gaussian Process Optimization with Adaptive Sketching: Scalable and No Regret in Negative Dependence: Theory and Applications in Machine Learning**, *Valko* 05:20 PM

Gaussian processes (GP) are a popular Bayesian approach for the optimization of black-box functions. Despite their effectiveness in simple problems, GP-based algorithms hardly scale to complex high-dimensional functions, as their per-iteration time and space cost is at least quadratic in the number of dimensions $d$ and iterations~$t$. Given a set of $A$ alternative to choose from, the overall runtime $O(t^3A)$ quickly becomes prohibitive. In this paper, we introduce BKB (budgeted kernelized bandit), an approximate GP algorithm for optimization under bandit feedback that achieves near-optimal regret (and hence near-optimal convergence rate) with near-constant per-iteration complexity and no assumption on the input space or the GP's covariance.

Combining a kernelized linear bandit algorithm (GP-UCB) with randomized matrix sketching technique (i.e., leverage score sampling), we prove that selecting inducing points based on their posterior variance gives an accurate low-rank approximation of the GP, preserving variance estimates and confidence intervals. As a consequence, BKB does not suffer from variance starvation, an important problem faced by many previous sparse GP approximations. Moreover, we show that our procedure selects at most $\widetilde{O}(d_{eff})$ points, where $d_{eff}$ is the \emph{effective} dimension of the explored space, which is typically much smaller than both $d$ and $t$. This greatly reduces the dimensionality of the
problem, thus leading to a $\widetilde{O}(TAd_{eff}^2)$ runtime and $\widetilde{O}(A d_{eff})$ space complexity.

Abstract 17: **Towards Efficient Evaluation of Risk via Herding in Negative Dependence: Theory and Applications in Machine Learning**, *Xu* 05:40 PM

We introduce a novel use of herding to address the problem of selecting samples from a large unlabeled dataset to efficiently evaluate the risk of a given model. Herding is an algorithm which elaborately draws samples to approximate the underlying distribution. We use herding to select the most informative samples and show that the loss evaluated on $k$ samples produced by herding converges to the expected loss at a rate $\mathcal{O}(1/k)$, which is much faster than $\mathcal{O}(1/\sqrt{k})$ for iid random sampling. We validate our analysis on both synthetic data and real data, and further explore the empirical performance of herding-based sampling in different cases of high-dimensional data.

1st Workshop on Understanding and Improving Generalization in Deep Learning

Dilip Krishnan, Hossein Mobahi, Behnam Neyshabur, Peter Bartlett, Dawn Song, Nati Srebro

Grand Ballroom A, Fri Jun 14, 08:30 AM

The 1st workshop on Generalization in Deep Networks: Theory and Practice will be held as part of ICML 2019. Generalization is one of the fundamental problems of machine learning, and increasingly important as deep networks make their presence felt in domains with big, small, noisy or skewed data. This workshop will consider generalization from both theoretical and practical perspectives. We welcome contributions from paradigms such as representation learning, transfer learning and reinforcement learning. The workshop invites researchers to submit working papers in the following research areas:

Implicit regularization: the role of optimization algorithms in generalization
Explicit regularization methods
Network architecture choices that improve generalization
Empirical approaches to understanding generalization
Generalization bounds; empirical evaluation criteria to evaluate bounds
Robustness: generalizing to distributional shift a.k.a dataset shift
Generalization in the context of representation learning, transfer learning and deep reinforcement learning: definitions and empirical approaches

Schedule

08:30 AM	**Opening Remarks**	
08:40 AM	**Progress on Nonvacuous Generalization Bounds**	*Roy*
09:10 AM	**Training for Generalization**	*Finn*
09:40 AM	**A Meta-Analysis of Overfitting in Machine Learning**	
09:55 AM	**Uniform convergence may be unable to explain generalization in deep learning**	
10:10 AM	**Break and Poster Session 1**	
10:40 AM	**Learning, Memory, and Entropy**	*Kakade*

11:10 AM	**Adversarially Robust Generalization**	*Madry*
11:40 AM	**Towards Task and Architecture-Independent Generalization Gap Predictors**	
11:55 AM	**Data-Dependent Sample Complexity of Deep Neural Networks via Lipschitz Augmentation**	
12:10 PM	**Lunch and Poster Session 1**	
01:30 PM	**A Hard Look at Generalization and its Theories**	*Belkin*
02:00 PM	**Overparameterization without Overfitting: Jacobian-based Generalization Guarantees for Neural Networks**	
02:15 PM	**How Learning Rate and Delay Affect Minima Selection in AsynchronousTraining of Neural Networks: Toward Closing the Generalization Gap**	
02:30 PM	**Towards Large Scale Structure of the Loss Landscape of Neural Networks**	
02:45 PM	**Zero-Shot Learning from scratch: leveraging local compositional representations**	
03:00 PM	**Break and Poster Session 2**	
03:30 PM	**On the Foundations of Deep Learning: SGD, Overparametrization, and Generalization**	*Lee*
04:00 PM	**Panel Discussion**	
05:00 PM	**Poster Session 2**	

6th ICML Workshop on Automated Machine Learning (AutoML 2019)

Frank Hutter, Joaquin Vanschoren, Katharina Eggensperger, Matthias Feurer

Grand Ballroom B, Fri Jun 14, 08:30 AM

Machine learning has achieved considerable successes in recent years, but this success often relies on human experts, who construct appropriate features, design learning architectures, set their hyperparameters, and develop new learning algorithms. Driven by the demand for off-the-shelf machine learning methods from an ever-growing community, the research area of AutoML targets the progressive automation of machine learning aiming to make effective methods available to everyone. The workshop targets a broad audience ranging from core machine learning researchers in different fields of ML connected to AutoML, such as neural architecture search, hyperparameter optimization, meta-learning, and learning to learn, to domain experts aiming to apply machine learning to new types of problems.

# Keynote Speakers

* Jeff Dean
* Rachel Thomas
* Raquel Urtasun
* Charles Sutton

Schedule

09:00 AM	**Welcome**	*Hutter*
09:05 AM	**Keynote 1**	
09:40 AM	**Poster Session 1**	*Gargiani, Zur, Baskin, Zheltonozhskii, Li, Talwalkar, Shang, Behl, Baydin, Couckuyt, Dhaene, Lin, Wei, Sun, Majumder, Donini, Ozaki, Adams, Geißler, Luo, peng, , Zhang, Langford, Caruana, Dey, Weill, Gonzalvo, Yang, Yak, Hotaj, Macko, Mohri, Cortes, Webb, Chen, Jankowiak, Goodman, Klein, Hutter, Javaheripi, Samragh, Lim, Kim, KIM, Volpp, Drori, Krishnamurthy, Cho, Jastrzebski, de Laroussilhe, Tan, Ma, Houlsby, Gesmundo, Borsos, Maziarz, Petroski Such, Lehman, Stanley, Clune, Gijsbers, Vanschoren, Mohr, Hüllermeier, Xiong, Zhang, zhu, Shao, Faust, Valko, Li, Escalante, Wever, Khorlin, Javidi, Francis, Mukherjee, Kim, McCourt, Kim, You, Choi, Knudde, Tornede*
11:00 AM	**Keynote 2**	
11:35 AM	**Contributed Talk 1: A Boosting Tree Based AutoML System for Lifelong Machine Learning**	*Xiong*
12:00 PM	**Poster Session 2**	
12:50 PM	**Lunch Break**	

02:00 PM	**Keynote 3: An Overview of Google's Work on AutoML and Future Directions**	*Dean*
02:35 PM	**Contributed Talk 2: Transfer NAS: Knowledge Transfer between Search Spaces with Transformer Agents**	*Borsos*
03:00 PM	**Poster Session 3**	
04:00 PM	**Contributed Talk 3: Random Search and Reproducibility for Neural Architecture Search**	*Li*
04:25 PM	**Keynote 4**	
05:00 PM	**Panel Discussion**	*Zhang, Sutton, Li, Thomas*
06:00 PM	**Closing Remarks**	*Hutter*

Abstracts (6):

Abstract 3: **Poster Session 1 in 6th ICML Workshop on Automated Machine Learning (AutoML 2019)**, *Gargiani, Zur, Baskin, Zheltonozhskii, Li, Talwalkar, Shang, Behl, Baydin, Couckuyt, Dhaene, Lin, Wei, Sun, Majumder, Donini, Ozaki, Adams, Geißler, Luo, peng, , Zhang, Langford, Caruana, Dey, Weill, Gonzalvo, Yang, Yak, Hotaj, Macko, Mohri, Cortes, Webb, Chen, Jankowiak, Goodman, Klein, Hutter, Javaheripi, Samragh, Lim, Kim, KIM, Volpp, Drori, Krishnamurthy, Cho, Jastrzebski, de Laroussilhe, Tan, Ma, Houlsby, Gesmundo, Borsos, Maziarz, Petroski Such, Lehman, Stanley, Clune, Gijsbers, Vanschoren, Mohr, Hüllermeier, Xiong, Zhang, zhu, Shao, Faust, Valko, Li, Escalante, Wever, Khorlin, Javidi, Francis, Mukherjee, Kim, McCourt, Kim, You, Choi, Knudde, Tornede* 09:40 AM

Accepted papers are available at https://sites.google.com/view/automl2019icml/accepted-papers and all posters will be up all time. Includes a coffee break from 10:30AM-11:00AM.

Abstract 5: **Contributed Talk 1: A Boosting Tree Based AutoML System for Lifelong Machine Learning in 6th ICML Workshop on Automated Machine Learning (AutoML 2019)**, *Xiong* 11:35 AM

AutoML aims at automating the process of designing good machine learning pipelines to solve different kinds of problems. However, existing AutoML systems are mainly designed for isolated learning by training a static model on a single batch of data; while in many real-world applications, data may arrive continuously in batches, possibly with concept drift. This raises a lifelong machine learning challenge for AutoML, as most existing AutoML systems can not evolve over time to learn from streaming data and adapt to concept drift. In this paper, we propose a novel AutoML system for this new scenario, i.e. a boosting tree based AutoML system for lifelong machine learning, which won the second place in the NeurIPS 2018 AutoML Challenge.

Abstract 6: **Poster Session 2 in 6th ICML Workshop on Automated Machine Learning (AutoML 2019)**, 12:00 PM

Accepted papers are available at https://sites.google.com/view/automl2019icml/accepted-papers and all posters will be up all time. Contributors are the same as for Poster Session 1.

Abstract 8: **Keynote 3: An Overview of Google's Work on AutoML and Future Directions in 6th ICML Workshop on Automated Machine Learning (AutoML 2019)**, *Dean* 02:00 PM

In this talk I'll survey work by Google researchers over the past several years on the topic of AutoML, or learning-to-learn. The talk will touch on basic approaches, some successful applications of AutoML to a variety of domains, and sketch out some directions for future AutoML systems that can leverage massively multi-task learning systems for automatically solving new problems.

Abstract 9: **Contributed Talk 2: Transfer NAS: Knowledge Transfer between Search Spaces with Transformer Agents in 6th ICML Workshop on Automated Machine Learning (AutoML 2019)**, *Borsos* 02:35 PM

Recent advances in Neural Architecture Search (NAS) have produced state-of-the-art architectures on several tasks. NAS shifts the efforts of human experts from developing novel architectures directly to designing architecture search spaces and methods to explore them efficiently. The search space definition captures prior knowledge about the properties of the architectures and it is crucial for the complexity and the performance of the search algorithm. However, different search space definitions require restarting the learning process from scratch. We propose a novel agent based on the Transformer that supports joint training and efficient transfer of prior knowledge between multiple search spaces and tasks.

Abstract 10: **Poster Session 3 in 6th ICML Workshop on Automated Machine Learning (AutoML 2019)**, 03:00 PM

Accepted papers are available at https://sites.google.com/view/automl2019icml/accepted-papers and all posters will be up all time. Contributors are the same as for Poster Session 1. Includes a coffee break from 3:00PM until 3:30PM.

Generative Modeling and Model-Based Reasoning for Robotics and AI

Aravind Rajeswaran, Emanuel Todorov, Igor Mordatch, William Agnew, Amy Zhang, Joelle Pineau, Michael Chang, Dumitru Erhan, Sergey Levine, Kimberly Stachenfeld, Marvin Zhang

Hall A, Fri Jun 14, 08:30 AM

Workshop website: https://sites.google.com/view/mbrl-icml2019

In the recent explosion of interest in deep RL, “model-free” approaches based on Q-learning and actor-critic architectures have received the most attention due to their flexibility and ease of use. However, this generality often comes at the expense of efficiency (statistical as well as computational) and robustness. The large number of required samples and safety concerns often limit direct use of model-free RL for real-world settings.

Model-based methods are expected to be more efficient. Given accurate models, trajectory optimization and Monte-Carlo planning methods can efficiently compute near-optimal actions in varied contexts. Advances in generative modeling, unsupervised, and self-supervised learning provide methods for learning models and representations that support subsequent planning and reasoning. Against this backdrop, our workshop aims to bring together researchers in generative modeling and model-based control to discuss research questions at their intersection, and to advance the state of the art in model-based RL for robotics and AI. In particular, this workshops aims to make progress on questions related to:

1. How can we learn generative models efficiently? Role of data, structures, priors, and uncertainty.
2. How to use generative models efficiently for planning and reasoning? Role of derivatives, sampling, hierarchies, uncertainty, counterfactual reasoning etc.
3. How to harmoniously integrate model-learning and model-based decision making?
4. How can we learn compositional structure and environmental constraints? Can this be leveraged for better generalization and reasoning?

Schedule

08:45 AM	**Welcome and Introduction**	*Rajeswaran*
09:00 AM	**Yann LeCun**	
09:30 AM	**Jessica Hamrick**	
10:00 AM	**Spotlight Session 1**	
11:00 AM	**Stefan Schaal**	
02:00 PM	**David Silver**	
03:30 PM	**Byron Boots**	
04:20 PM	**Chelsea Finn**	
04:50 PM	**Abhinav Gupta**	

Uncertainty and Robustness in Deep Learning

Yixuan (Sharon) Li, Balaji Lakshminarayanan, Dan Hendrycks, Tom Dietterich, Justin Gilmer

Hall B, Fri Jun 14, 08:30 AM

There has been growing interest in rectifying deep neural network vulnerabilities. Challenges arise when models receive samples drawn from outside the training distribution. For example, a neural network tasked with classifying handwritten digits may assign high confidence predictions to cat images. Anomalies are frequently encountered when deploying ML models in the real world. Well-calibrated predictive uncertainty estimates are indispensable for many machine learning applications, such as self-driving cars and medical diagnosis systems. Generalization to unseen and worst-case inputs is also essential for robustness to distributional shift. In order to have ML models reliably predict in open environment, we must deepen technical understanding in the following areas: (1) learning algorithms that are robust to changes in input data distribution (e.g., detect out-of-distribution examples); (2) mechanisms to estimate and calibrate confidence produced by neural networks and (3) methods to improve robustness to adversarial and common corruptions, and (4) key applications for uncertainty such as in artificial intelligence (e.g., computer vision, robotics, self-driving cars, medical imaging) as well as broader machine learning tasks.

This workshop will bring together researchers and practitioners from the machine learning communities, and highlight recent work that contribute to address these challenges. Our agenda will feature contributed papers with invited speakers. Through the workshop we hope to help identify fundamentally important directions on robust and reliable deep learning, and foster future collaborations.

Schedule

08:30 AM	**Welcome**	*Li*
08:40 AM	**Spotlight**	*Scott, Koshy, Aigrain, Bidart, Panda, Yap, Yacoby, Gontijo Lopes, Marchisio, Englesson, Yang, Graule, Sun, Kang, Dusenberry, Du, Maennel, Menda, Edupuganti, Metz, Stutz, Srinivasan, Sämann, N Balasubramanian, Mohseni, Cornish, Butepage, Wang, Li, Han, Li, Andriushchenko, Ruff, Vadera, Ovadia, Thulasidasan, Ji, Niu, Mahloujifar, Kumar, CHUN, Yin, Xu*
09:30 AM	**Keynote by Max Welling: A Nonparametric Bayesian Approach to Deep Learning (without GPs)**	*Welling*
10:00 AM	**Poster Session 1 (all papers)**	
11:00 AM	**Keynote by Kilian Weinberger: On Calibration and Fairness**	*Weinberger*
11:30 AM	**Why ReLU networks yield high-confidence predictions far away from the training data and how to mitigate the problem**	*Andriushchenko*
11:40 AM	**Detecting Extrapolation with Influence Functions**	*Madras*
11:50 AM	**How Can We Be So Dense? The Robustness of Highly Sparse Representations**	*Ahmad*
12:00 PM	**Keynote: Suchi Saria**	*Saria*
02:00 PM	**Subspace Inference for Bayesian Deep Learning**	*Kirichenko, Izmailov, Wilson*
02:10 PM	**Quality of Uncertainty Quantification for Bayesian Neural Network Inference**	*Yao*
02:20 PM	**'In-Between' Uncertainty in Bayesian Neural Networks**	*Foong*

02:30 PM	**Keynote by Dawn Song: Adversarial machine learning: Challenges, lessons, and future directions**	*Song*
03:30 PM	**Keynote by Terry Boult: The Deep Unknown: on Open-set and Adversarial Examples in Deep Learning**	*Boult*
04:00 PM	**Panel Discussion (moderator: Tom Dietterich)**	*Welling, Weinberger, Boult, Song, Dietterich*
05:00 PM	**Poster Session 2 (all papers)**	

Abstracts (9):

Abstract 3: **Keynote by Max Welling: A Nonparametric Bayesian Approach to Deep Learning (without GPs) in Uncertainty and Robustness in Deep Learning**, *Welling* 09:30 AM

We present a new family of exchangeable stochastic processes suitable for deep learning. Our nonparametric Bayesian method models distributions over functions by learning a graph of dependencies on top of latent representations of the points in the given dataset. In doing so, they define a Bayesian model without explicitly positing a prior distribution over latent global parameters; they instead adopt priors over the relational structure of the given dataset, a task that is much simpler. We show how we can learn such models from data, demonstrate that they are scalable to large datasets through mini-batch optimization and describe how we can make predictions for new points via their posterior predictive distribution. We experimentally evaluate FNPs on the tasks of toy regression and image classification and show that, when compared to baselines that employ global latent parameters, they offer both competitive predictions as well as more robust uncertainty estimates.

Abstract 5: **Keynote by Kilian Weinberger: On Calibration and Fairness in Uncertainty and Robustness in Deep Learning**, *Weinberger* 11:00 AM

We investigate calibration for deep learning algorithms in classification and regression settings. Although we show that typically deep networks tend to be highly mis-calibrated, we demonstrate that this is easy to fix - either to obtain more trustworthy confidence estimates or to detect outliers in the data. Finally, we relate calibration with the recently raised tension between minimizing error disparity across different population groups while maintaining calibrated probability estimates. We show that calibration is compatible only with a single error constraint (i.e. equal false-negatives rates across groups), and show that any algorithm that satisfies this relaxation is no better than randomizing a percentage of predictions for an existing classifier. These unsettling findings, which extend and generalize existing results, are empirically confirmed on several datasets.

Abstract 6: **Why ReLU networks yield high-confidence predictions far away from the training data and how to mitigate the problem in Uncertainty and Robustness in Deep Learning**, *Andriushchenko* 11:30 AM

Classifiers used in the wild, in particular for safety-critical systems, should know when they don't know, in particular make low confidence predictions far away from the train- ing data. We show that ReLU type neural networks fail in this regard as they produce almost always high confidence predictions far away from the training data. For bounded domains we propose a new robust optimization technique similar to adversarial training which enforces low confidence pre- dictions far away from the training data. We show that this technique is surprisingly effective in reducing the confidence of predictions far away from the training data while maintaining high confidence predictions and test error on the original classification task com- pared to standard training. This is a short version of the corresponding CVPR paper.

Abstract 7: **Detecting Extrapolation with Influence Functions in Uncertainty and Robustness in Deep Learning**, *Madras* 11:40 AM

In this work, we explore principled methods for extrapolation detection. We define extrapolation as occurring when a model's conclusion at a test point is underdetermined by the training data. Our metrics for detecting extrapolation are based on influence functions, inspired by the intuition that a point requires extrapolation if its inclusion in the training set would significantly change the model's learned parameters. We provide interpretations of our methods in terms of the eigendecomposition of the Hessian. We present experimental evidence that our method is capable of identifying extrapolation to out-of-distribution points.

Abstract 8: **How Can We Be So Dense? The Robustness of Highly Sparse Representations in Uncertainty and Robustness in Deep Learning**, *Ahmad* 11:50 AM

Neural networks can be highly sensitive to noise and perturbations. In this paper we suggest that high dimensional sparse representations can lead to increased robustness to noise and interference. A key intuition we develop is that the ratio of the match volume around a sparse vector divided by the total representational space decreases exponentially with dimensionality, leading to highly robust matching with low interference from other patterns. We analyze efficient sparse networks containing both sparse weights and sparse activations. Simulations on MNIST, the Google Speech Command Dataset, and CIFAR-10 show that such networks demonstrate improved robustness to random noise compared to dense networks, while maintaining competitive accuracy. We propose that sparsity should be a core design constraint for creating highly robust networks.

Abstract 10: **Subspace Inference for Bayesian Deep Learning in Uncertainty and Robustness in Deep Learning**, *Kirichenko, Izmailov, Wilson* 02:00 PM

Bayesian inference was once a gold standard for learning with neural networks, providing accurate full predictive distributions and well calibrated uncertainty. However, scaling Bayesian inference techniques to deep neural networks is challenging due to the high dimensionality of the parameter space. In this pa- per, we construct low-dimensional subspaces of parameter space that contain diverse sets of models, such as the first principal components of the stochastic gradient descent (SGD) trajectory. In these subspaces, we are able to apply elliptical slice sampling and variational inference, which struggle in the full parameter space. We show that Bayesian model averaging over the induced posterior in these subspaces produces high accurate predictions and well-calibrated predictive uncertainty for both regression and image classification.

Abstract 11: **Quality of Uncertainty Quantification for Bayesian Neural Network Inference in Uncertainty and Robustness in Deep Learning**, *Yao* 02:10 PM

Bayesian Neural Networks (BNNs) place priors over the parameters in a neural network. Inference in BNNs, however, is difficult; all inference methods for BNNs are approximate. In this work, we empirically compare the quality of predictive uncertainty estimates for 10 common inference methods on both regression and classification tasks. Our experiments demonstrate that commonly used metrics (e.g. test log-likelihood) can be misleading. Our experiments also indicate that inference innovations designed to capture structure in the posterior do not necessarily produce high quality posterior approximations.

Abstract 12: **'In-Between' Uncertainty in Bayesian Neural Networks in Uncertainty and Robustness in Deep Learning**, *Foong* 02:20 PM

We describe a limitation in the expressiveness of the predictive uncertainty estimate given by mean- field variational inference (MFVI), a popular approximate inference method for Bayesian neural networks. In particular, MFVI fails to give calibrated uncertainty estimates in between separated regions of observations. This can lead to catastrophically overconfident predictions when testing on out-of-distribution data. Avoiding such over-confidence is critical for active learning, Bayesian optimisation and out-of-distribution robustness. We instead find that a classical technique, the linearised Laplace approximation, can handle ‘in-between’ uncertainty much better for small network architectures.

Abstract 14: **Keynote by Terry Boult: The Deep Unknown: on Open-set and Adversarial Examples in Deep Learning in Uncertainty and Robustness in Deep Learning**, *Boult* 03:30 PM

The first part of the talk will explore issues with deep networks dealing with "unknowns" inputs, and the general problems of open-set recognition in deep networks. We review the core of open-set recognition theory and its application in our first attempt at open-set deep networks, "OpenMax" We discuss is successes and limitations and why classic "open-set" approaches don't really solve the problem of deep unknowns. We then present our recent work from NIPS2018, on a new model we call the ObjectoSphere. Using ObjectoSphere loss begins to address the learning of deep features that can handle unknown inputs. We present examples of its use first on simple datasets sets (MNIST/CFAR) and then onto unpublished work applying it to the real-world problem of open-set face recognition. We discuss of the relationship between open set recognition theory and adversarial image generation, showing how our deep-feature adversarial approach, called LOTS can attack the first OpenMax solution, as well as successfully attack even open-set face recognition systems. We end with a discussion of how open set theory can be applied to improve network robustness.

ICML 2019 Workshop on Computational Biology

Donna Pe'er, Sandhya Prabhakaran, Elham Azizi, Abdoulaye Baniré Diallo, Anshul Kundaje, Barbara Engelhardt, Wajdi Dhifli, Engelbert MEPHU NGUIFO, Wesley Tansey, Julia Vogt, Jennifer Listgarten, Cassandra Burdziak, Workshop CompBio

Room 101, Fri Jun 14, 08:30 AM

The workshop will showcase recent research in the field of Computational Biology. There has been significant development in genomic sequencing techniques and imaging technologies. These approaches not only generate huge amounts of data but provide unprecedented resolution of single cells and even subcellular structures. The availability of high dimensional data, at multiple spatial and temporal resolutions has made machine learning and deep learning methods increasingly critical for computational analysis and interpretation of the data. Conversely, biological data has also exposed unique challenges and problems that call for the development of new machine learning methods. This workshop aims to bring together researchers working at the intersection of Machine Learning and Biology to present recent advances and open questions in Computational Biology to the ML community.

The workshop is a sequel to the WCB workshops we organized in the last three years Joint ICML and IJCAI 2018, Stockholm, ICML 2017, Sydney and ICML 2016, New York as well as Workshop on Bioinformatics and AI at IJCAI 2015 Buenos Aires, IJCAI 2016 New York, IJCAI 2017 Melbourne which had excellent line-ups of talks and were well-received by the community. Every year, we received 60+ submissions. After multiple rounds of rigorous reviewing around 50 submissions were selected from which the best set of papers were chosen for Contributed talks and Spotlights and the rest were invited as Posters. We have a steadfast and growing base of reviewers making up the Program Committee. For the past edition, a special issue of Journal of Computational Biology will be released in the following weeks with extended versions of 14 accepted papers.

We have two confirmed invited speakers and we will invite at least one more leading researcher in the field. Similar to previous years, we plan to request partial funding from Microsoft Research, Google, Python, Swiss Institute of Bioinformatics that we intend to use for student travel awards. In past years, we have also been able to provide awards for best poster/paper and partially contribute to travel expenses for at least 8 students per year.

The Workshop proceedings will be available through CEUR proceedings. We would also have an extended version to be included in a special issue of the Journal of Computational Biology (JCB) for which we already have an agreement with JCB.

Schedule

08:30 AM	**WCB Organisers**	
08:40 AM	**Caroline Uhler**	
09:20 AM	**Jacopo Cirrone, Marttinen Pekka , Elior Rahmani, Elior Rahmani, Harri Lähdesmäki**	
09:50 AM	**Ali Oskooei**	
11:00 AM	**Francisco LePort**	
11:40 AM	**Ali Oskooei, Zhenqin Wu, Karren Dai Yang, Mijung Kim**	
12:00 PM	**Poster Session & Lunch break**	*Wiese, Carter, DeBlasio*
02:00 PM	**Ngo Trong Trung**	

02:20 PM	**Achille Nazaret, François Fages, Brandon Carter, Ruishan Liu, Nicasia Beebe-Wang**
02:50 PM	**Poster Session & Coffee Break**
04:00 PM	**Daphne Koller**
04:40 PM	**Bryan He**
05:00 PM	**Award Ceremony & Closing Remarks**

ICML 2019 Time Series Workshop

Vitaly Kuznetsov, Scott Yang, Rose Yu, Cheng Tang, Yuyang Wang

Room 102, Fri Jun 14, 08:30 AM

Time series data is both quickly growing and already ubiquitous. In domains spanning as broad a range as climate, robotics, entertainment, finance, healthcare, and transportation, there has been a significant shift away from parsimonious, infrequent measurements to nearly continuous monitoring and recording. Rapid advances in sensing technologies, ranging from remote sensors to wearables and social sensing, are generating rapid growth in the size and complexity of time series data streams. Thus, the importance and impact of time series analysis and modelling techniques only continues to grow.

At the same time, while time series analysis has been extensively studied by econometricians and statisticians, modern time series data often pose significant challenges for the existing techniques both in terms of their structure (e.g., irregular sampling in hospital records and spatiotemporal structure in climate data) and size. Moreover, the focus on time series in the machine learning community has been comparatively much smaller. In fact, the predominant methods in machine learning often assume i.i.d. data streams, which is generally not appropriate for time series data. Thus, there is both a great need and an exciting opportunity for the machine learning community to develop theory, models and algorithms specifically for the purpose of processing and analyzing time series data.

We see ICML as a great opportunity to bring together theoretical and applied researchers from around the world and with different backgrounds who are interested in the development and usage of time series analysis and algorithms. This includes methods for time series prediction, classification, clustering, anomaly and change point detection, causal discovery, and dimensionality reduction as well as general theory for learning and analyzing stochastic processes. Since time series have been studied in a variety of different fields and have many broad applications, we plan to host leading academic researchers and industry experts with a range of perspectives and interests as invited speakers. Moreover, we also invite researchers from the related areas of batch and online learning, deep learning, reinforcement learning, data analysis and statistics, and many others to both contribute and participate in this workshop.

Human In the Loop Learning (HILL)

Xin Wang, Xin Wang, Fisher Yu, Shanghang Zhang, Joseph Gonzalez, Yangqing Jia, Sarah Bird, Kush Varshney, Been Kim, Adrian Weller

Room 103, Fri Jun 14, 08:30 AM

https://sites.google.com/view/hill2019/home

This workshop is a joint effort between the 4th ICML Workshop on Human Interpretability in Machine Learning (WHI) and the ICML 2019 Workshop on Interactive Data Analysis System (IDAS). We have combined our forces this year to run Human in the Loop Learning (HILL) in conjunction with ICML 2019!

The workshop will bring together researchers and practitioners who study interpretable and interactive learning systems with applications in large scale data processing, data annotations, data visualization, human-assisted data integration, systems and tools to interpret machine learning models as well as algorithm designs for active learning, online learning, and interpretable machine learning algorithms. The target audience for the workshop includes people who are interested in using machines to solve problems by having a human be an integral part of the process. This workshop serves as a platform where researchers can discuss approaches that bridge the gap between humans and machines and get the best of both worlds.

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

Systems for online and interactive learning algorithms,
Active/Interactive machine learning algorithm design,
Systems for collecting, preparing, and managing machine learning data,
Model understanding tools (verifying, diagnosing, debugging, visualization, introspection, etc),
Design, testing and assessment of interactive systems for data analytics,
Psychology of human concept learning,
Generalized additive models, sparsity and rule learning,
Interpretable unsupervised models (clustering, topic models, etc.),
Interpretation of black-box models (including deep neural networks),
Interpretability in reinforcement learning.

Schedule

08:25 AM	**Opening Remarks**
08:30 AM	**Invited Talk: James Philbin**
09:00 AM	**Invited Talk: Sanja Fidler**
09:30 AM	**Invited Talk: Bryan Catanzaro**
10:00 AM	**IDAS Poster Session & Coffee break**
11:30 AM	**Invited Talk: Yisong Yue**
12:00 PM	**Invited Talk: Vittorio Ferrari**
12:30 PM	**Lunch Break**

02:00 PM	**Interpretability Contributed Talks**
03:00 PM	**Coffee Break**
03:30 PM	**Interpretability Invited Discussion: California's Senate Bill 10 (SB 10) on Pretrial Release and Detention with Solon Barocas and Peter Eckersley**
04:45 PM	**Human in the Loop Learning Panel Discussion**

## Climate Change: How Can AI Help?

David Rolnick, Alexandre Lacoste, Tegan Maharaj, Jennifer Chayes, Yoshua Bengio

Room 104 A, Fri Jun 14, 08:30 AM

# Climate Change: How Can AI Help?

We invite submission of extended abstracts applying machine learning to the problems of climate change. There will be three tracks (Deployed, Research, and Ideas).

Workshop website:
Submission website:
Submission deadline: **April 30, 11:59 PM** Pacific Time
Notification: May 15 (early notification possible upon request)
Contact:

## Summary
Climate change is widely agreed to be one of the greatest challenges facing humanity. We already observe increased incidence and severity of storms, droughts, fires, and flooding, as well as significant changes to global ecosystems, including the natural resources and agriculture on which humanity depends. The 2018 UN report on climate change estimates that the world has only thirty years to eliminate greenhouse emissions completely if we are to avoid catastrophic consequences.

Many in the machine learning community want to address climate change but feel their skills are inapplicable. This workshop will showcase the many settings in which machine learning can be applied to reducing greenhouse emissions and helping society adapt to the effects of climate change. Climate change is a complex problem requiring simultaneous action from many directions. While machine learning is not a silver bullet, there is significant potential impact for research and implementation.

## Call for submissions

We invite submission of extended abstracts on machine learning applied to problems in climate mitigation, adaptation, or modeling, including but not limited to the following topics:

- Power generation and grids
- Transportation
- Smart buildings and cities
- Industrial optimization
- Carbon capture and sequestration
- Agriculture, forestry and other land use
- Climate modeling
- Extreme weather events
- Disaster management and relief
- Societal adaptation
- Ecosystems and natural resources
- Data presentation and management
- Climate finance

Accepted submissions will be invited to give poster presentations at the workshop, of which some will be selected for spotlight talks. Please contact with questions, or if visa considerations make earlier notification important.

Dual-submissions are allowed, and the workshop does not record proceedings. Submissions will be reviewed double-blind; do your best to anonymize your submission, and do not include identifying information for authors in the PDF. We encourage, but do not require, use of the ICML style template (please do not use the “Accepted” format).

## Submission tracks
Extended abstracts are limited to 3 pages for the Deployed and Research tracks, and 2 pages for the Ideas track, in PDF format. An additional page may be used for references. All machine learning techniques are welcome, from kernel methods to deep learning. Each submission should make clear why the application has (or could have) positive impacts regarding climate change. There are three tracks for submissions:

### Deployed
Work that is already having an impact

Submissions for the Deployed track are intended for machine learning approaches which are impacting climate-relevant problems through consumers or partner institutions. This could include implementations of academic research that have moved beyond the testing phase, as well as results from startups/industry. Details of methodology need not be revealed if they are proprietary, though transparency is encouraged.

### Research
Work that will have an impact when deployed

Submissions for the Research track are intended for machine learning research applied to climate-relevant problems. Submissions should provide experimental or theoretical validation of the method proposed, as well as specifying what gap the method fills. Algorithms need not be novel from a machine learning perspective if they are applied in a novel setting.

Datasets may be submitted to this track that are designed to permit machine learning research (e.g. formatted with clear benchmarks for evaluation). In this case, baseline experimental results on the dataset are preferred but not required.

### Ideas
Future work that could have an impact

Submissions for the Ideas track are intended for proposed applications of machine learning to solve climate-relevant problems. While the least constrained, this track will be subject to a very high standard of review. No results need be demonstrated, but ideas should be justified as extensively as possible, including motivation for the problem being solved, an explanation of why current tools are inadequate, and details of how tools from machine learning are proposed to fill the gap.

Schedule

08:30 AM	**Opening Remarks**	
08:45 AM	**John Platt (Google AI)**	*Platt*
09:20 AM	**Jack Kelly (Open Climate Fix)**	*Kelly*
09:45 AM	**Andrew Ng (Stanford)**	
10:10 AM	**TBA**	
10:30 AM	**Morning Coffee Break + Poster Session**	
11:00 AM	**Chad Frischmann (Project Drawdown)**	
12:00 PM	**Networking Lunch (provided) + Poster Session**	
01:30 PM	**Yoshua Bengio (Mila)**	*Bengio*
01:55 PM	**Claire Monteleoni (CU Boulder)**	*Monteleoni*
02:30 PM	**TBA.**	
03:30 PM	**Karthik Mukkavilli (Mila)**	*Mukkavilli*
03:45 PM	**Sims Witherspoon (DeepMind)**	*Witherspoon*
04:20 PM	**TBA..**	
05:15 PM	**Panel Discussion**	

Abstracts (1):

Abstract 8: **Networking Lunch (provided) + Poster Session in Climate Change: How Can AI Help?**, 12:00 PM

Catered sandwiches and snacks will be provided (majority vegetarian/vegan, with gluten-free options).

Workshop on the Security and Privacy of Machine Learning

Nicolas Papernot, Florian Tramer, Bo Li, Dan Boneh, David Evans, Somesh Jha, Percy Liang, Patrick McDaniel, Jacob Steinhardt, Dawn Song

Room 104 B, Fri Jun 14, 08:30 AM

As machine learning has increasingly been deployed in critical real-world applications, the dangers of manipulation and misuse of these models has become of paramount importance to public safety and user privacy. In applications such as online content recognition to financial analytics to autonomous vehicles all have shown the be vulnerable to adversaries wishing to manipulate the models or mislead models to their malicious ends.

This workshop will focus on recent research and future directions about the security and privacy problems in real-world machine learning systems. We aim to bring together experts from machine learning, security, and privacy communities in an attempt to highlight recent work in these area as well as to clarify the foundations of secure and private machine learning strategies. We seek to come to a consensus on a rigorous framework to formulate adversarial attacks targeting machine learning models, and to characterize the properties that ensure the security and privacy of machine learning systems. Finally, we hope to chart out important directions for future work and cross-community collaborations.

Schedule

09:00 AM	**Patrick McDaniel**
09:30 AM	**Una-May O'Reilly**
10:00 AM	**Enhancing Gradient-based Attacks with Symbolic Intervals**
10:20 AM	**Adversarial Policies: Attacking Deep Reinforcement Learning**
10:45 AM	**Le Song**
11:15 AM	**Allen Qi**
11:45 AM	**Private vqSGD: Vector-Quantized Stochastic Gradient Descent**
01:15 PM	**Ziko Kolter**
01:45 PM	**Provable Certificates for Adversarial Examples:Fitting a Ball in the Union of Polytopes**
02:05 PM	**Poster Session #1**
02:45 PM	**Alexander Madry**
03:15 PM	**Been Kim**
03:45 PM	**Theoretically Principled Trade-off between Robustness and Accuracy**
04:05 PM	**Model weight theft with just noise inputs: The curious case of the petulant attacker**
04:15 PM	**Panel**
05:15 PM	**Poster Sesson #2**

Theoretical Physics for Deep Learning

Jaehoon Lee, Jeffrey Pennington, Yasaman Bahri, Max Welling, Surya Ganguli, Joan Bruna

Room 104 C, Fri Jun 14, 08:30 AM

Though the purview of physics is broad and includes many loosely connected subdisciplines, a unifying theme is the endeavor to provide concise, quantitative, and predictive descriptions of the often large and complex systems governing phenomena that occur in the natural world. While one could debate how closely deep learning is connected to the natural world, it is undeniably the case that deep learning systems are large and complex; as such, it is reasonable to consider whether the rich body of ideas and powerful tools from theoretical physicists could be harnessed to improve our understanding of deep learning. The goal of this workshop is to investigate this question by bringing together experts in theoretical physics and deep learning in order to stimulate interaction and to begin exploring how theoretical physics can shed light on the theory of deep learning.

We believe ICML is an appropriate venue for this gathering as members from both communities are frequently in attendance and because deep learning theory has emerged as a focus at the conference, both as an independent track in the main conference and in numerous workshops over the last few years. Moreover, the conference has enjoyed an increasing number of papers using physics tools and ideas to draw insights into deep learning.

Schedule

08:30 AM	**Opening Remarks**	*Lee, Pennington, Bahri, Welling, Ganguli, Bruna*
08:40 AM	**Linearized two-layers neural networks in high dimension**	*Montanari*
09:10 AM	**Loss landscape and behaviour of algorithms in the spiked matrix-tensor model**	*Zdeborova*
09:40 AM	**Poster spotlights**	*Novak, Dreyer, Golkar, Higgins, Antognini, Karakida, Ghosh*
10:20 AM	**Break and poster discussion**	
11:00 AM	**Kyle Cranmer (NYU)**	*Cranmer*
11:30 AM	**Why Deep Learning Works: Traditional and Heavy-Tailed Implicit Self-Regularization in Deep Neural Networks**	*Mahoney*
12:00 PM	**Analyzing the dynamics of online learning in over-parameterized two-layer neural networks**	*Goldt*
12:15 PM	**Convergence Properties of Neural Networks on Separable Data**	*Tachet des Combes*
12:30 PM	**Lunch**	
02:00 PM	**Is Optimization a sufficient language to understand Deep Learning?**	*Arora*
02:30 PM	**Towards Understanding Regularization in Batch Normalization**	
02:45 PM	**How Noise during Training Affects the Hessian Spectrum**	
03:00 PM	**Break and poster discussion**	
03:30 PM	**Understanding overparameterized neural networks**	*Sohl-Dickstein*
04:00 PM	**Feynman Diagrams for Large Width Networks**	
04:15 PM	**A Mean Field Theory of Quantized Deep Networks: The Quantization-Depth Trade-Off**	
04:30 PM	**Deep Learning on the 2-Dimensional Ising Model to Extract the Crossover Region**	
04:45 PM	**Learning the Arrow of Time**	
05:00 PM	**Poster discussion**	*Novak, Gabella, Dreyer, Golkar, Tong, Higgins, Milletari, Antognini, Goldt, Ramírez Rivera, Bondesan, Karakida, Tachet des Combes, Mahoney, Walker, Fort, Smith, Rahaman, Ghosh, Baratin, Granziol*

Abstracts (8):

Abstract 2: **Linearized two-layers neural networks in high dimension in Theoretical Physics for Deep Learning**, *Montanari* 08:40 AM

Speaker: Andrea Montanari (Stanford)

Abstract: Abstract: We consider the problem of learning an unknown function f on the d-dimensional sphere with respect to the square loss, given i.i.d. samples (y_i,x_i) where x_i is a feature vector uniformly distributed on the sphere and y_i = f(x_i). We study two popular classes of models that can be regarded as linearizations of two-layers neural networks around a random initialization: (RF) The random feature model of Rahimi-Recht; (NT) The neural tangent kernel model of Jacot-Gabriel-Hongler. Both these approaches can also be regarded as randomized approximations of kernel ridge regression (with respect to different kernels), and hence enjoy universal approximation properties when the number of neurons N diverges, for a fixed dimension d.

We prove that, if both d and N are large, the behavior of these models is instead remarkably simpler.

If N is of smaller order than d^2, then RF performs no better than linear regression with respect to the raw features x_i, and NT performs no better than linear regression with respect to degree-one and two monomials in the x_i's. More generally, if N is of smaller order than d^{k+1} then RF fits at most a degree-k polynomial in the raw features, and NT fits at most a degree-(k+ 1) polynomial.
We then focus on the case of quadratic functions, and N= O(d). We show that the gap in generalization error between fully trained neural networks and the linearized models is potentially unbounded.
[based on joint work with Behrooz Ghorbani, Song Mei, Theodor Misiakiewicz]

Abstract 3: **Loss landscape and behaviour of algorithms in the spiked matrix-tensor model in Theoretical Physics for Deep Learning**, *Zdeborova* 09:10 AM

Speaker: Lenka Zdeborova (CEA/SACLAY)

Abstract: A key question of current interest is: How are properties of optimization and sampling algorithms influenced by the properties of the loss function in noisy high-dimensional non-convex settings? Answering this question for deep neural networks is a landmark goal of many ongoing works. In this talk I will answer this question in unprecedented detail for the spiked matrix-tensor model. Information theoretic limits, and Kac-Rice analysis of the loss landscapes, will be compared to the analytically studied performance of message passing algorithms, of the Langevin dynamics and of the gradient flow. Several rather non-intuitive results will be unveiled and explained.

Abstract 4: **Poster spotlights in Theoretical Physics for Deep Learning**, *Novak, Dreyer, Golkar, Higgins, Antognini, Karakida, Ghosh* 09:40 AM

Bayesian Deep Convolutional Networks with Many Channels are Gaussian Processes Roman Novak (Google Brain)*; Lechao Xiao (Google Brain); Jaehoon Lee (Google Brain); Yasaman Bahri (Google Brain); Greg Yang (Microsoft Research AI); Jiri Hron (University of Cambridge); Daniel Abolafia (Google Brain); Jeffrey Pennington (Google Brain); Jascha Sohl-Dickstein (Google Brain)

Jet grooming through reinforcement learning Frederic A Dreyer (University of Oxford)*

Inferring the quantum density matrix with machine learning Kyle Cranmer (New York University); Siavash Golkar (NYU)*; Duccio Pappadopulo (Bloomberg)

Towards a Definition of Disentangled Representations Irina Higgins (DeepMind)*; David Amos (DeepMind); Sebastien Racaniere (DeepMind); David Pfau (); Loic Matthey (DeepMind); Danilo Jimenez Rezende (Google DeepMind)

Covariance in Physics and Convolutional Neural Networks Miranda Cheng (University of Amsterdam)*; Vassilis Anagiannis (University of Amsterdam); Maurice Weiler (University of Amsterdam); Pim de Haan (University of Amsterdam); Taco S. Cohen (Qualcomm AI Research); Max Welling (University of Amsterdam)

Finite size corrections for neural network Gaussian processes Joseph M Antognini (Whisper AI)*

Pathological Spectrum of the Fisher Information Matrix in Deep Neural Networks Ryo Karakida (National Institute of Advanced Industrial Science and Technology)*; Shotaro Akaho (AIST); Shun-ichi Amari (RIKEN)

A Quantum Field Theory of Representation Learning Robert Bamler (University of California at Irvine)*; Stephan Mandt (University of California, Irivine)

Scale Steerable Filters for Locally Scale-Invariant Convolutional Neural Networks Rohan Ghosh (National University of Singapore)*; Anupam Gupta (National University of Singapore)

Abstract 5: **Break and poster discussion in Theoretical Physics for Deep Learning**, 10:20 AM

Bayesian Deep Convolutional Networks with Many Channels are Gaussian Processes Roman Novak (Google Brain)*; Lechao Xiao (Google Brain); Jaehoon Lee (Google Brain); Yasaman Bahri (Google Brain); Greg Yang (Microsoft Research AI); Jiri Hron (University of Cambridge); Daniel Abolafia (Google Brain); Jeffrey Pennington (Google Brain); Jascha Sohl-Dickstein (Google Brain)
Topology of Learning in Artificial Neural Networks Maxime Gabella (Magma Learning)*
Jet grooming through reinforcement learning Frederic Dreyer (University of Oxford)*; Stefano Carrazza (University of Milan)
Inferring the quantum density matrix with machine learning Kyle Cranmer (New York University); Siavash Golkar (NYU)*; Duccio Pappadopulo (Bloomberg)
Backdrop: Stochastic Backpropagation Siavash Golkar (NYU)*; Kyle Cranmer (New York University)
Explain pathology in Deep Gaussian Process using Chaos Theory Anh Tong (UNIST)*; Jaesik Choi (Ulsan National Institute of Science and Technology)
Towards a Definition of Disentangled Representations Irina Higgins (DeepMind)*; David Amos (DeepMind); Sebastien Racaniere (DeepMind); David Pfau (DeepMind); Loic Matthey (DeepMind); Danilo Jimenez Rezende (DeepMind)
Towards Understanding Regularization in Batch Normalization Ping Luo (The Chinese University of Hong Kong); Xinjiang Wang (); Wenqi Shao (The Chinese University of HongKong)*; Zhanglin Peng (SenseTime)
Covariance in Physics and Convolutional Neural Networks Miranda Cheng (University of Amsterdam)*; Vassilis Anagiannis (University of Amsterdam); Maurice Weiler (University of Amsterdam); Pim de Haan (University of Amsterdam); Taco S. Cohen (Qualcomm AI Research); Max Welling (University of Amsterdam)
Meanfield theory of activation functions in Deep Neural Networks Mirco Milletari (Microsoft)*; Thiparat Chotibut (SUTD) ; Paolo E. Trevisanutto (National University of Singapore)
Finite size corrections for neural network Gaussian processes Joseph M Antognini (Whisper AI)*
SWANN: Small-World Neural Networks and Rapid Convergence Mojan Javaheripi (UC San Diego)*; Bita Darvish Rouhani (UC San Diego); Farinaz Koushanfar (UC San Diego)
Analysing the dynamics of online learning in over-parameterised two-layer neural networks Sebastian Goldt (Institut de Physique théorique, Paris)*; Madhu Advani (Harvard University); Andrew Saxe (University of Oxford); Florent Krzakala (École Normale Supérieure); Lenka Zdeborova (CEA Saclay)
A Halo Merger Tree Generation and Evaluation Framework Sandra Robles (Universidad Autónoma de Madrid); Jonathan Gómez (Pontificia

Universidad Católica de Chile); Adín Ramírez Rivera (University of Campinas)*; Jenny Gonzáles (Pontificia Universidad Católica de Chile); Nelson Padilla (Pontificia Universidad Católica de Chile); Diego Dujovne (Universidad Diego Portales)
Learning Symmetries of Classical Integrable Systems Roberto Bondesan (Qualcomm AI Research)*, Austen Lamacraft (Cavendish Laboratory, University of Cambridge, UK)
Cosmology inspired generative models Uros Seljak (UC Berkeley)*; Francois Lanusse (UC Berkeley)
Pathological Spectrum of the Fisher Information Matrix in Deep Neural Networks Ryo Karakida (National Institute of Advanced Industrial Science and Technology)*; Shotaro Akaho (AIST); Shun-ichi Amari (RIKEN)
How Noise during Training Affects the Hessian Spectrum Mingwei Wei (Northwestern University); David Schwab (Facebook AI Research)*
A Quantum Field Theory of Representation Learning Robert Bamler (University of California at Irvine)*; Stephan Mandt (University of California, Irivine)
Convergence Properties of Neural Networks on Separable Data Remi Tachet des Combes (Microsoft Research Montreal)*; Mohammad Pezeshki (Mila & University of Montreal); Samira Shabanian (Microsoft, Canada); Aaron Courville (MILA, Université de Montréal); Yoshua Bengio (Mila)
Universality and Capacity Metrics in Deep Neural Networks Michael Mahoney (University of California, Berkeley)*; Charles Martin (Calculation Consulting)
Feynman Diagrams for Large Width Networks Guy Gur-Ari (Google)*; Ethan Dyer (Google)
Deep Learning on the 2-Dimensional Ising Model to Extract the Crossover Region Nicholas Walker (Louisiana State Univ - Baton Rouge)*
Large Scale Structure of the Loss Landscape of Neural Networks Stanislav Fort (Stanford University)*; Stanislaw Jastrzebski (New York University)
Momentum Enables Large Batch Training Samuel L Smith (DeepMind)*; Erich Elsen (Google); Soham De (DeepMind)
Learning the Arrow of Time Nasim Rahaman (University of Heidelberg)*; Steffen Wolf (Heidelberg University); Anirudh Goyal (University of Montreal); Roman Remme (Heidelberg University); Yoshua Bengio (Mila)
Scale Steerable Filters for Locally Scale-Invariant Convolutional Neural Networks Rohan Ghosh (National University of Singapore)*; Anupam Gupta (National University of Singapore)
A Mean Field Theory of Quantized Deep Networks: The Quantization-Depth Trade-Off Yaniv Blumenfeld (Technion)*; Dar Gilboa (Columbia University); Daniel Soudry (Technion)
Rethinking Complexity in Deep Learning: A View from Function Space Aristide Baratin (Mila, Université de Montréal)*; Thomas George (MILA, Université de Montréal); César Laurent (Mila, Université de Montréal); Valentin Thomas (MILA); Guillaume Lajoie (Université de Montréal, Mila); Simon Lacoste-Julien (Mila, Université de Montréal)
The Deep Learning Limit: Negative Neural Network eigenvalues just noise? Diego Granziol (Oxford)*; Stefan Zohren (University of Oxford); Stephen Roberts (Oxford); Dmitry P Vetrov (Higher School of Economics); Andrew Gordon Wilson (Cornell University); Timur Garipov (Samsung AI Center in Moscow)
Gradient descent in Gaussian random fields as a toy model for high-dimensional optimisation Mariano Chouza (Tower Research Capital); Stephen Roberts (Oxford); Stefan Zohren (University of Oxford)*
Deep Learning for Inverse Problems Abhejit Rajagopal (University of California, Santa Barbara)*; Vincent R Radzicki (University of California, Santa Barbara)

Abstract 7: **Why Deep Learning Works: Traditional and Heavy-Tailed Implicit Self-Regularization in Deep Neural Networks in Theoretical Physics for Deep Learning**, *Mahoney* 11:30 AM

Speaker: Michael Mahoney (ICSI and Department of Statistics, University of California at Berkeley)

Abstract:

Random Matrix Theory (RMT) is applied to analyze the weight matrices of Deep Neural Networks (DNNs), including both production quality, pre-trained models and smaller models trained from scratch. Empirical and theoretical results clearly indicate that the DNN training process itself implicitly implements a form of self-regularization, implicitly sculpting a more regularized energy or penalty landscape. In particular, the empirical spectral density (ESD) of DNN layer matrices displays signatures of traditionally-regularized statistical models, even in the absence of exogenously specifying traditional forms of explicit regularization. Building on relatively recent results in RMT, most notably its extension to Universality classes of Heavy-Tailed matrices, and applying them to these empirical results, we develop a theory to identify 5+1 Phases of Training, corresponding to increasing amounts of implicit self-regularization. For smaller and/or older DNNs, this implicit self-regularization is like traditional Tikhonov regularization, in that there appears to be a ``size scale'' separating signal from noise. For state-of-the-art DNNs, however, we identify a novel form of heavy-tailed self-regularization, similar to the self-organization seen in the statistical physics of disordered systems. This implicit self-regularization can depend strongly on the many knobs of the training process. In particular, by exploiting the generalization gap phenomena, we demonstrate that we can cause a small model to exhibit all 5+1 phases of training simply by changing the batch size. This demonstrates that---all else being equal---DNN optimization with larger batch sizes leads to less-well implicitly-regularized models, and it provides an explanation for the generalization gap phenomena. Coupled with work on energy landscapes and heavy-tailed spin glasses, it also suggests an explanation of why deep learning works. Joint work with Charles Martin of Calculation Consulting, Inc.

Abstract 11: **Is Optimization a sufficient language to understand Deep Learning? in Theoretical Physics for Deep Learning**, *Arora* 02:00 PM

Speaker: Sanjeev Arora (Princeton/IAS)

Abstract: There is an old debate in neuroscience about whether or not learning has to boil down to optimizing a single cost function. This talk will suggest that even to understand mathematical properties of deep learning, we have to go beyond the conventional view of "optimizing a single cost function". The reason is that phenomena occur along the gradient descent trajectory that are not fully captured in the value of the cost function. I will illustrate briefly with three new results that involve such phenomena:

(i) (joint work with Cohen, Hu, and Luo) How deep matrix factorization solves matrix completion better than classical algorithms https://arxiv.org/abs/1905.13655

(ii) (joint with Du, Hu, Li, Salakhutdinov, and Wang) How to compute (exactly) with an infinitely wide net ("mean field limit", in physics terms) https://arxiv.org/abs/1904.11955

(iii) (joint with Kuditipudi, Wang, Hu, Lee, Zhang, Li, Ge) Explaining mode-connectivity for real-life deep nets (the phenomenon that low-cost solutions found by gradient descent are interconnected in the parameter space via low-cost paths; see Garipov et al'18 and Draxler et al'18)

Abstract 15: **Understanding overparameterized neural networks in Theoretical Physics for Deep Learning**, *Sohl-Dickstein* 03:30 PM

Speaker: Jascha Sohl-Dickstein (Google Brain)

Abstract: As neural networks become highly overparameterized, their accuracy improves, and their behavior becomes easier to analyze theoretically. I will give an introduction to a rapidly growing body of work which examines the learning dynamics and prior over functions induced by infinitely wide, randomly initialized, neural networks. Core results that I will discuss include: that the distribution over functions computed by a wide neural network often corresponds to a Gaussian process with a particular compositional kernel, both before and after training; that the predictions of wide neural networks are linear in their parameters throughout training; and that this perspective enables analytic predictions for how trainability depends on hyperparameters and architecture. These results provide for surprising capabilities -- for instance, the evaluation of test set predictions which would come from an infinitely wide trained neural network without ever instantiating a neural network, or the rapid training of 10,000+ layer convolutional networks. I will argue that this growing understanding of neural networks in the limit of infinite width is foundational for future theoretical and practical understanding of deep learning.

Abstract 20: **Poster discussion in Theoretical Physics for Deep Learning**, *Novak, Gabella, Dreyer, Golkar, Tong, Higgins, Milletari, Antognini, Goldt, Ramírez Rivera, Bondesan, Karakida, Tachet des Combes, Mahoney, Walker, Fort, Smith, Rahaman, Ghosh, Baratin, Granziol* 05:00 PM

Bayesian Deep Convolutional Networks with Many Channels are Gaussian Processes Roman Novak (Google Brain)*; Lechao Xiao (Google Brain); Jaehoon Lee (Google Brain); Yasaman Bahri (Google Brain); Greg Yang (Microsoft Research AI); Jiri Hron (University of Cambridge); Daniel Abolafia (Google Brain); Jeffrey Pennington (Google Brain); Jascha Sohl-Dickstein (Google Brain)
Topology of Learning in Artificial Neural Networks Maxime Gabella (Magma Learning)*
Jet grooming through reinforcement learning Frederic Dreyer (University of Oxford)*; Stefano Carrazza (University of Milan)
Inferring the quantum density matrix with machine learning Kyle Cranmer (New York University); Siavash Golkar (NYU)*; Duccio Pappadopulo (Bloomberg)
Backdrop: Stochastic Backpropagation Siavash Golkar (NYU)*; Kyle Cranmer (New York University)
Explain pathology in Deep Gaussian Process using Chaos Theory Anh Tong (UNIST)*; Jaesik Choi (Ulsan National Institute of Science and Technology)
Towards a Definition of Disentangled Representations Irina Higgins (DeepMind)*; David Amos (DeepMind); Sebastien Racaniere (DeepMind); David Pfau (DeepMind); Loic Matthey (DeepMind); Danilo Jimenez Rezende (DeepMind)
Towards Understanding Regularization in Batch Normalization Ping Luo (The Chinese University of Hong Kong); Xinjiang Wang (); Wenqi Shao (The Chinese University of HongKong)*; Zhanglin Peng (SenseTime)
Covariance in Physics and Convolutional Neural Networks Miranda Cheng (University of Amsterdam)*; Vassilis Anagiannis (University of Amsterdam); Maurice Weiler (University of Amsterdam); Pim de Haan (University of Amsterdam); Taco S. Cohen (Qualcomm AI Research); Max Welling (University of Amsterdam)
Meanfield theory of activation functions in Deep Neural Networks Mirco Milletari (Microsoft)*; Thiparat Chotibut (SUTD) ; Paolo E. Trevisanutto (National University of Singapore)
Finite size corrections for neural network Gaussian processes Joseph M Antognini (Whisper AI)*
SWANN: Small-World Neural Networks and Rapid Convergence Mojan Javaheripi (UC San Diego)*; Bita Darvish Rouhani (UC San Diego); Farinaz Koushanfar (UC San Diego)
Analysing the dynamics of online learning in over-parameterised two-layer neural networks Sebastian Goldt (Institut de Physique théorique, Paris)*; Madhu Advani (Harvard University); Andrew Saxe (University of Oxford); Florent Krzakala (École Normale Supérieure); Lenka Zdeborova (CEA Saclay)
A Halo Merger Tree Generation and Evaluation Framework Sandra Robles (Universidad Autónoma de Madrid); Jonathan Gómez (Pontificia Universidad Católica de Chile); Adín Ramírez Rivera (University of Campinas)*; Jenny Gonzáles (Pontificia Universidad Católica de Chile); Nelson Padilla (Pontificia Universidad Católica de Chile); Diego Dujovne (Universidad Diego Portales)
Learning Symmetries of Classical Integrable Systems Roberto Bondesan (Qualcomm AI Research)*, Austen Lamacraft (Cavendish Laboratory, University of Cambridge, UK)
Cosmology inspired generative models Uros Seljak (UC Berkeley)*; Francois Lanusse (UC Berkeley)
Pathological Spectrum of the Fisher Information Matrix in Deep Neural Networks Ryo Karakida (National Institute of Advanced Industrial Science and Technology)*; Shotaro Akaho (AIST); Shun-ichi Amari (RIKEN)
How Noise during Training Affects the Hessian Spectrum Mingwei Wei (Northwestern University); David Schwab (Facebook AI Research)*
A Quantum Field Theory of Representation Learning Robert Bamler (University of California at Irvine)*; Stephan Mandt (University of California, Irivine)
Convergence Properties of Neural Networks on Separable Data Remi Tachet des Combes (Microsoft Research Montreal)*; Mohammad Pezeshki (Mila & University of Montreal); Samira Shabanian (Microsoft, Canada); Aaron Courville (MILA, Université de Montréal); Yoshua Bengio (Mila)
Universality and Capacity Metrics in Deep Neural Networks Michael Mahoney (University of California, Berkeley)*; Charles Martin (Calculation Consulting)
Feynman Diagrams for Large Width Networks Guy Gur-Ari (Google)*; Ethan Dyer (Google)
Deep Learning on the 2-Dimensional Ising Model to Extract the Crossover Region Nicholas Walker (Louisiana State Univ - Baton Rouge)*
Large Scale Structure of the Loss Landscape of Neural Networks Stanislav Fort (Stanford University)*; Stanislaw Jastrzebski (New York University)
Momentum Enables Large Batch Training Samuel L Smith (DeepMind)*; Erich Elsen (Google); Soham De (DeepMind)
Learning the Arrow of Time Nasim Rahaman (University of Heidelberg)*; Steffen Wolf (Heidelberg University); Anirudh Goyal (University of Montreal); Roman Remme (Heidelberg University); Yoshua Bengio (Mila)

Scale Steerable Filters for Locally Scale-Invariant Convolutional Neural Networks Rohan Ghosh (National University of Singapore)*; Anupam Gupta (National University of Singapore)
A Mean Field Theory of Quantized Deep Networks: The Quantization-Depth Trade-Off Yaniv Blumenfeld (Technion)*; Dar Gilboa (Columbia University); Daniel Soudry (Technion)
Rethinking Complexity in Deep Learning: A View from Function Space Aristide Baratin (Mila, Université de Montréal)*; Thomas George (MILA, Université de Montréal); César Laurent (Mila, Université de Montréal); Valentin Thomas (MILA); Guillaume Lajoie (Université de Montréal, Mila); Simon Lacoste-Julien (Mila, Université de Montréal)
The Deep Learning Limit: Negative Neural Network eigenvalues just noise? Diego Granziol (Oxford)*; Stefan Zohren (University of Oxford); Stephen Roberts (Oxford); Dmitry P Vetrov (Higher School of Economics); Andrew Gordon Wilson (Cornell University); Timur Garipov (Samsung AI Center in Moscow)
Gradient descent in Gaussian random fields as a toy model for high-dimensional optimisation Mariano Chouza (Tower Research Capital); Stephen Roberts (Oxford); Stefan Zohren (University of Oxford)*
Deep Learning for Inverse Problems Abhejit Rajagopal (University of California, Santa Barbara)*; Vincent R Radzicki (University of California, Santa Barbara)

AI in Finance: Applications and Infrastructure for Multi-Agent Learning

Prashant Reddy, Tucker Balch, Michael Wellman, Senthil Kumar, Ion Stoica, Edith Elkind

Room 201, Fri Jun 14, 08:30 AM

Finance is a rich domain for AI and ML research. Model-driven strategies for stock trading and risk assessment models for loan approvals are quintessential financial applications that are reasonably well-understood. However, there are a number of other applications that call for attention as well.

In particular, many finance domains involve ecosystems of interacting and competing agents. Consider for instance the detection of financial fraud and money-laundering. This is a challenging multi-agent learning problem, especially because the real world agents involved evolve their strategies constantly. Similarly, in algorithmic trading of stocks, commodities, etc., the actions of any given trading agent affects, and is affected by, other trading agents -- many of these agents are constantly learning in order to adapt to evolving market scenarios. Further, such trading agents operate at such a speed and scale that they must be fully autonomous. They have grown in sophistication to employ advanced ML strategies including deep learning, reinforcement learning, and transfer learning.

Financial institutions have a long history of investing in technology as a differentiator and have been key drivers in advancing computing infrastructure (e.g., low-latency networking). As more financial applications employ deep learning and reinforcement learning, there is consensus now on the need for more advanced computing architectures--for training large machine learning models and simulating large multi-agent learning systems--that balance scale with the stringent privacy requirements of finance.

Historically, financial firms have been highly secretive about their proprietary technology developments. But now, there is also emerging consensus on the need for (1) deeper engagement with academia to advance a shared knowledge of the unique challenges faced in FinTech, and (2) more open collaboration with academic and technology partners through intellectually sophisticated fora such as this proposed workshop.

Schedule

09:00 AM	**Opening Remarks**	*Reddy, Kumar*
09:10 AM	**Invited Talk 1: Adaptive Tolling for Multiagent Traffic Optimization**	*Stone*
09:30 AM	**Invited Talk 2: The Strategic Perils of Learning from Historical Data**	*Morgenstern*
09:50 AM	**Oral Paper Presentations 1**	
10:30 AM	**Coffee Break and Socialization**	
11:00 AM	**Invited Talk 3: Trend-Following Trading Strategies and Financial Market Stability**	*Wellman*
11:20 AM	**Poster Highlights - Lightning Round**	
12:00 PM	**Lunch**	
02:00 PM	**Invited Talk 4: Towards AI Innovation in the Financial Domain**	*Veloso*
02:20 PM	**Oral Paper Presentations 2**	
03:00 PM	**Coffee Break and Socialization**	
03:30 PM	**Invited Talk 5**	*Stoica*
03:50 PM	**Invited Talk 6: Intra-day Stock Price Prediction as a Measure of Market Efficiency**	*Balch*
04:10 PM	**Poster Session - All Accepted Papers**	
06:00 PM	**ICML Reception**	

Abstracts (3):

Abstract 4: **Oral Paper Presentations 1 in AI in Finance: Applications and Infrastructure for Multi-Agent Learning**, 09:50 AM

09:50-10:03 Risk-Sensitive Compact Decision Trees for Autonomous Execution in presence of Simulated Market Response, Svitlana Vyetrenko (JP Morgan Chase); Kyle Xu (Georgia Institute of Technology)

10:04-10:16 Robust Trading via Adversarial Reinforcement Learning Thomas Spooner (University of Liverpool); Rahul Savani (Univ. of

Liverpool)

10:17-10:30 Generating Realistic Stock Market Order Streams, Junyi Li (University of Michigan); Xintong Wang (University of Michigan); Yaoyang Lin (University of Michigan); Arunesh Sinha (University of Michigan); Michael Wellman (University of Michigan)

Abstract 7: **Poster Highlights - Lightning Round in AI in Finance: Applications and Infrastructure for Multi-Agent Learning**, 11:20 AM

Self Organizing Supply Chains for Micro-Prediction; Present and Future Uses of the ROAR Protocol, Peter D Cotton (JP Morgan Chase)

Learning-Based Trading Strategies in the Face of Market Manipulation, Xintong Wang (University of Michigan); Chris Hoang (University of Michigan); Michael Wellman (University of Michigan)

Multi-Agent Simulation for Pricing and Hedging in a Dealer Market, Sumitra Ganesh (JPMorgan AI Research); Nelson Vadori (JPMorgan AI Research); Mengda Xu (JPMorgan AI Research); Hua Zheng (JPMorgan Chase); Prashant Reddy (JPMorgan AI Research); Manuela Veloso (JPMorgan AI Research)

Multi-Agent Reinforcement Learning for Liquidation Strategy Analysis, Wenhang Bao (Columbia University); Xiao-Yang Liu (Columbia University)

Some people aren't worth listening to: periodically retraining classifiers with feedback from a team of end users, Joshua Lockhart (JPMorgan AI Research); Mahmoud Mahfouz (JPMorgan AI Research); Tucker Balch (JPMorgan AI Research); Manuela Veloso (JPMorgan AI Research)

Optimistic Bull or Pessimistic Bear: Adaptive Deep Reinforcement Learning for Stock Portfolio Allocation, Xinyi Li (Columbia University); Yinchuan Li (Beijing Institute of Technology); Yuancheng Zhan (University of Science and Technology of China); Xiao-Yang Liu (Columbia University)

How to Evaluate Trading Strategies: Backtesting or Agent-based Simulation?, Tucker Balch (JPMorgan AI Research); David Byrd (Georgia Tech); Mahmoud Mahfouz (JPMorgan AI Research)

Deep Reinforcement Learning for Optimal Trade Execution, Siyu Lin (University of Virginia)

Abstract 10: **Oral Paper Presentations 2 in AI in Finance: Applications and Infrastructure for Multi-Agent Learning**, 02:20 PM

02:20-02:33 Towards Inverse Reinforcement Learning for Limit Order Book Dynamics, Jacobo Roa Vicens (University College London); Cyrine Chtourou (JPMorgan Chase); Angelos Filos (University of Oxford); Francisco Rullan (University College of London); Yarin Gal (University of Oxford);
Ricardo Silva (University College London)

02:34-02:47 An Agent-Based Model of Financial Benchmark Manipulation, Megan J Shearer (University of Michigan); Gabriel Rauterberg (University of Michigan); Michael Wellman (University of Michigan)

02:47-03:00 The sharp, the flat and the shallow: Can weakly interacting agents learn to escape bad minima?, Panos Parpas (Imperial College London); Nikolas Kantas (Imperial College London); Grigorios Pavliotis (Imperial College London)

The Third Workshop On Tractable Probabilistic Modeling (TPM)

Daniel Lowd, Antonio Vergari, Alejandro Molina, Tahrima Rahman, Pedro Domingos, Antonio Vergari

Room 202, Fri Jun 14, 08:30 AM

Probabilistic modeling has become the de facto framework to reason about uncertainty in Machine Learning and AI. One of the main challenges in probabilistic modeling is the trade-off between the expressivity of the models and the complexity of performing various types of inference, as well as learning them from data.

This inherent trade-off is clearly visible in powerful -- but intractable -- models like Markov random fields, (restricted) Boltzmann machines, (hierarchical) Dirichlet processes and Variational Autoencoders. Despite these models’ recent successes, performing inference on them resorts to approximate routines. Moreover, learning such models from data is generally harder as inference is a sub-routine of learning, requiring simplifying assumptions or further approximations. Having guarantees on tractability at inference and learning time is then a highly desired property in many real-world scenarios.

Tractable probabilistic modeling (TPM) concerns methods guaranteeing exactly this: performing exact (or tractably approximate) inference and/or learning. To achieve this, the following approaches have been proposed: i) low or bounded-treewidth probabilistic graphical models and determinantal point processes, that exchange expressiveness for efficiency; ii) graphical models with high girth or weak potentials, that provide bounds on the performance of approximate inference methods; and iii) exchangeable probabilistic models that exploit symmetries to reduce inference complexity. More recently, models compiling inference routines into efficient computational graphs such as arithmetic circuits, sum-product networks, cutset networks and probabilistic sentential decision diagrams have advanced the state-of-the-art inference performance by exploiting context-specific independence, determinism or by exploiting latent variables. TPMs have been successfully used in numerous real-world applications: image classification, completion and generation, scene understanding, activity recognition, language and speech modeling, bioinformatics, collaborative filtering, verification and diagnosis of physical systems.

The aim of this workshop is to bring together researchers working on the different fronts of tractable probabilistic modeling, highlighting recent trends and open challenges. At the same time, we want to foster the discussion across similar or complementary sub-fields in the broader probabilistic modeling community. In particular, the rising field of neural probabilistic models, such as normalizing flows and autoregressive models that achieve impressive results in generative modeling. It is an interesting open challenge for the TPM community to keep a broad range of inference routines tractable while leveraging these models’ expressiveness. Furthermore, the rising field of probabilistic programming promises to be the new lingua franca of model-based learning. This offers the TPM community opportunities to push the expressiveness of the models used for general-purpose universal probabilistic languages, such as Pyro, while maintaining efficiency.

We want to promote discussions and advance the field both by having high quality contributed works, as well as high level invited speakers coming from the aforementioned tangent sub-fields of probabilistic modeling.

Schedule

09:00 AM	**Welcome**	
09:10 AM	**Testing Arithmetic Circuits**	*Darwiche*
09:50 AM	**Poster spotlights**	
10:30 AM	**Coffee Break**	
11:00 AM	**Invited Talk**	
11:40 AM	**Poster spotlights**	
12:00 PM	**Invited Talk**	*Peharz*
12:40 PM	**Lunch**	
02:20 PM	**Tensor Variable Elimination in Pyro**	*Bingham*
03:00 PM	**Coffee Break**	
03:30 PM	**Invertible Residual Networks and a Novel Perspective on Adversarial Examples**	*Jacobsen*
04:10 PM	**Poster session**	

Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)

Sujith Ravi, Zornitsa Kozareva, Lixin Fan, Max Welling, Yurong Chen, Werner Bailer, Brian Kulis, Haoji Hu, Jonathan Dekhtiar, Yingyan Lin, Diana Marculescu

Room 203, Fri Jun 14, 08:30 AM

This joint workshop aims to bring together researchers, educators, practitioners who are interested in techniques as well as applications of on-device machine learning and compact, efficient neural network representations. One aim of the workshop discussion is to establish close connection between researchers in the machine learning community and engineers in industry, and to benefit both academic researchers as well as industrial practitioners. The other aim is the evaluation and comparability of resource-efficient machine learning methods and compact and efficient network representations, and their relation to particular target platforms (some of which may be highly optimized for neural network inference). The research community has still to develop established evaluation procedures and metrics.

The workshop also aims at reproducibility and comparability of methods for compact and efficient neural network representations, and on-device machine learning. Contributors are thus encouraged to make their code available. The workshop organizers plan to make some example tasks and datasets available, and invite contributors to use them for testing their work. In order to provide comparable performance evaluation conditions, the use of a common platform (such as Google Colab) is intended.

Schedule

08:30 AM	**Welcome and Introduction**	
08:40 AM	**Hardware Efficiency Aware Neural Architecture Search and Compression**	*Han*
09:10 AM	**Structured matrices for efficient deep learning**	*Kumar*
09:40 AM	**DeepCABAC: Context-adaptive binary arithmetic coding for deep neural network compression**	*Wiedemann*
10:00 AM	**Poster spotlight presentations**	
10:30 AM	**Coffee Break AM**	
11:00 AM	**Understanding the Challenges of Algorithm and Hardware Co-design for Deep Neural Networks**	*Sze*
11:30 AM	**Dream Distillation: A Data-Independent Model Compression Framework**	
11:50 AM	**The State of Sparsity in Deep Neural Networks**	
12:10 PM	**Lunch break**	
12:40 PM	**Poster session**	*Lin, Li*
02:00 PM	**DNN Training and Inference with Hyper-Scaled Precision**	
02:30 PM	**Mixed Precision Training & Inference**	*Dekhtiar*
03:00 PM	**Coffee Break PM**	
03:30 PM	**Learning Compact Neural Networks Using Ordinary Differential Equations as Activation Functions**	
03:50 PM	**Triplet Distillation for Deep Face Recognition**	
04:10 PM	**Single-Path NAS: Device-Aware Efficient ConvNet Design**	*Stamoulis*
04:30 PM	**Panel discussion**	
05:30 PM	**Wrap-up and Closing**	

Abstracts (10):

Abstract 4: **DeepCABAC: Context-adaptive binary arithmetic coding for deep neural network compression in Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)**, *Wiedemann* 09:40 AM

Simon Wiedemann, Heiner Kirchhoffer, Stefan Matlage, Paul Haase, Arturo Marban Gonzalez, Talmaj Marinc, Heiko Schwarz, Detlev Marpe, Thomas Wiegand, Ahmed Osman and Wojciech Samek

http://arxiv.org/abs/1905.08318

Abstract 5: **Poster spotlight presentations in Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)**, 10:00 AM

2min presentations of the posters that will be presented in the poster session during lunch break

Abstract 7: **Understanding the Challenges of Algorithm and Hardware Co-design for Deep Neural Networks in Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)**, *Sze* 11:00 AM

The co-design of algorithm and hardware has become an increasingly important approach for addressing the computational complexity of Deep Neural Networks (DNNs). There are several open problems and challenges in the co-design process and application; for instance, what metrics should be used to drive the algorithm design, how to automate the process in a simple way, how to extend these approaches to tasks beyond image classification, and how to design flexible hardware to support these different approaches. In this talk, we highlight recent and ongoing work that aim to address these challenges, namely energy-aware pruning and NetAdapt that automatically incorporate direct metrics such as latency and energy into the training and design of the DNN; FastDepth that extends the co-design approaches to a depth estimation task; and a flexible hardware accelerator called Eyeriss v2 that is computationally efficient across a wide range of diverse DNNs.
BIO: Vivienne Sze is an Associate Professor at MIT in the Electrical Engineering and Computer Science Department. Her research interests include energy-aware signal processing algorithms, and low-power circuit and system design for portable multimedia applications, including computer vision, deep learning, autonomous navigation, and video process/coding. Prior to joining MIT, she was a Member of Technical Staff in the R&D Center at TI, where she designed low-power algorithms and architectures for video coding. She also represented TI in the JCT-VC committee of ITU-T and ISO/IEC standards body during the development of High Efficiency Video Coding (HEVC), which received a Primetime Engineering Emmy Award. She is a co-editor of the book entitled “High Efficiency Video Coding (HEVC): Algorithms and Architectures” (Springer, 2014).

Abstract 8: **Dream Distillation: A Data-Independent Model Compression Framework in Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)**, 11:30 AM

Kartikeya Bhardwaj, Naveen Suda and Radu Marculescu

http://arxiv.org/abs/1905.07072

Abstract 9: **The State of Sparsity in Deep Neural Networks in Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)**, 11:50 AM

Trevor Gale, Erich Elsen and Sara Hooker

https://arxiv.org/abs/1902.09574 (to be updated)

Abstract 11: **Poster session in Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)**, *Lin, Li* 12:40 PM

Xiaofan Zhang, Hao Cong, Yuhong Li, Yao Chen, Jinjun Xiong, Wen-Mei Hwu and Deming Chen. A Bi-Directional Co-Design Approach to Enable Deep Learning on IoT Devices
https://arxiv.org/abs/1905.08369

Kushal Datta, Aishwarya Bhandare, Deepthi Karkada, Vamsi Sripathi, Sun Choi, Vikram Saletore and Vivek Menon. Efficient 8-Bit Quantization of Transformer Neural Machine Language Translation Model
http://arxiv.org/abs/1906.00532

Bin Yang, Lin Yang, Xiaochun Li, Wenhan Zhang, Hua Zhou, Yequn Zhang, Yongxiong Ren and Yinbo Shi. 2-bit Model Compression of Deep Convolutional Neural Network on ASIC Engine for image retrieval
https://arxiv.org/abs/1905.03362

Zhong Qiu Lin, Brendan Chwyl and Alexander Wong. EdgeSegNet: A Compact Network for Semantic Segmentation
https://arxiv.org/abs/1905.04222

Sheng Lin, Xiaolong Ma, Shaokai Ye, Geng Yuan, Kaisheng Ma and Yanzhi Wang. Toward Extremely Low Bit and Lossless Accuracy in DNNs with Progressive ADMM
https://arxiv.org/abs/1905.00789

Chengcheng Li, Zi Wang, Dali Wang, Xiangyang Wang and Hairong Qi. Investigating Channel Pruning through Structural Redundancy Reduction - A Statistical Study
https://arxiv.org/abs/1905.06498

Wei Niu, Yanzhi Wang and Bin Ren. CADNN: Ultra Fast Execution of DNNs on Mobile Devices with Advanced Model Compression and Architecture-Aware Optimization
https://arxiv.org/abs/1905.00571

Jonathan Ephrath, Lars Ruthotto, Eldad Haber and Eran Treister. LeanResNet: A Low-cost yet Effective Convolutional Residual Networks
https://arxiv.org/abs/1904.06952

Dushyant Mehta, Kwang In Kim and Christian Theobalt. Implicit Filter Sparsification In Convolutional Neural Networks
https://arxiv.org/abs/1905.04967

Abstract 12: **DNN Training and Inference with Hyper-Scaled Precision in Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)**, 02:00 PM

Kailash Gopalakrishnan

Abstract 15: **Learning Compact Neural Networks Using Ordinary Differential Equations as Activation Functions in Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network**

Representations (ODML-CDNNR), 03:30 PM

Mohamadali Torkamani, Phillip Wallis, Shiv Shankar and Amirmohammad Rooshenas

https://arxiv.org/abs/1905.07685

Abstract 16: **Triplet Distillation for Deep Face Recognition in Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)**, 03:50 PM

Yushu Feng, Huan Wang and Haoji Hu

https://arxiv.org/abs/1905.04457

Abstract 17: **Single-Path NAS: Device-Aware Efficient ConvNet Design in Joint Workshop on On-Device Machine Learning & Compact Deep Neural Network Representations (ODML-CDNNR)**, *Stamoulis* 04:10 PM

Dimitrios Stamoulis, Ruizhou Ding, Di Wang, Dimitrios Lymberopoulos, Bodhi Priyantha, Jie Liu and Diana Marculescu

https://arxiv.org/abs/1905.04159

Reinforcement Learning for Real Life

Yuxi Li, Alborz Geramifard, Lihong Li, Csaba Szepesvari, Tao Wang

Seaside Ballroom, Fri Jun 14, 08:30 AM

Reinforcement learning (RL) is a general learning, predicting, and decision making paradigm. RL provides solution methods for sequential decision making problems as well as those can be transformed into sequential ones. RL connects deeply with optimization, statistics, game theory, causal inference, sequential experimentation, etc., overlaps largely with approximate dynamic programming and optimal control, and applies broadly in science, engineering and arts.

RL has been making steady progress in academia recently, e.g., Atari games, AlphaGo, visuomotor policies for robots. RL has also been applied to real world scenarios like recommender systems and neural architecture search. See a recent collection about RL applications at https://medium.com/@yuxili/rl-applications-73ef685c07eb. It is desirable to have RL systems that work in the real world with real benefits. However, there are many issues for RL though, e.g. generalization, sample efficiency, and exploration vs. exploitation dilemma. Consequently, RL is far from being widely deployed. Common, critical and pressing questions for the RL community are then: Will RL have wide deployments? What are the issues? How to solve them?

The goal of this workshop is to bring together researchers and practitioners from industry and academia interested in addressing practical and/or theoretical issues in applying RL to real life scenarios, review state of the arts, clarify impactful research problems, brainstorm open challenges, share first-hand lessons and experiences from real life deployments, summarize what has worked and what has not, collect tips for people from industry looking to apply RL and RL experts interested in applying their methods to real domains, identify potential opportunities, generate new ideas for future lines of research and development, and promote awareness and collaboration. This is not "yet another RL workshop": it is about how to successfully apply RL to real life applications. This is a less addressed issue in the RL/ML/AI community, and calls for immediate attention for sustainable prosperity of RL research and development.

Schedule

08:30 AM	**optional early-bird posters**	
08:50 AM	**opening remarks by organizers**	
09:00 AM	**invited talk by David Silver (Deepmind)**	*Silver*
09:20 AM	**invited talk by John Langford (Microsoft Research): How do we make Real World Reinforcement Learning revolution?**	*Langford*
09:40 AM	**invited talk by Craig Boutilier (Google Research): Reinforcement Learning in Recommender Systems: Some Challenges**	*Boutilier*
10:00 AM	**posters**	*Chen, Garau Luis, Albert Smet, Modi, Tomkins, Simmons-Edler, Mao, Irpan, Lu, Wang, Mukherjee, Raghu, Shihab, Ahn, Fakoor, Chaudhari, Smirnova, Oh, Tang, Qin, Li*
10:30 AM	**coffee break**	
11:00 AM	**panel discussion with Craig Boutilier (Google Research), Emma Brunskill (Stanford), Chelsea Finn (Google Brain, Stanford, UC Berkeley), Mohammad Ghavamzadeh (Facebook AI), John Langford (Microsoft Research) and David Silver (Deepmind)**	*Stone, Boutilier, Brunskill, Finn, Langford, Silver, Ghavamzadeh*
12:00 PM	**optional posters**	

Abstracts (2):

Abstract 4: **invited talk by John Langford (Microsoft Research): How do we make Real World Reinforcement Learning revolution? in Reinforcement Learning for Real Life**, *Langford* 09:20 AM

Abstract: Doing Real World Reinforcement Learning implies living with steep constraints on the sample complexity of solutions. Where is this viable? Where might it be viable in the near future? In the far future? How can we design a research program around identifying and building such solutions? In short, what are the missing elements we need to really make reinforcement learning more mundane and commonly applied than Supervised Learning? The potential is certainly there given the naturalness of RL compared to supervised learning, but the present is manifestly different.

Abstract 5: **invited talk by Craig Boutilier (Google Research): Reinforcement Learning in Recommender Systems: Some Challenges in Reinforcement Learning for Real Life**, *Boutilier* 09:40 AM

Abstract: I'll present a brief overview of some recent work on reinforcement learning motivated by practical issues that arise in the application of RL to online, user-facing applications like recommender systems. These include stochastic action sets, long-term cumulative effects, and combinatorial action spaces. I'll provide some detail on the last of these, describing SlateQ, a novel decomposition technique that allows value-based RL (e.g., Q-learning) in slate-based recommender to scale to commercial production systems, and briefly describe both small-scale simulation and a large-scale experiment with YouTube.

Bio: Craig is Principal Scientist at Google, working on various aspects of decision making under uncertainty (e.g., reinforcement learning, Markov decision processes, user modeling, preference modeling and elicitation) and recommender systems. He received his Ph.D. from the University of Toronto in 1992, and has held positions at the University of British Columbia, University of Toronto, CombineNet, and co-founded Granata Decision Systems.

Craig was Editor-in-Chief of JAIR; Associate Editor with ACM TEAC, JAIR, JMLR, and JAAMAS; Program Chair for IJCAI-09 and UAI-2000. Boutilier is a Fellow of the Royal Society of Canada (RSC), the Association for Computing Machinery (ACM) and the Association for the Advancement of Artificial Intelligence (AAAI). He was recipient of the 2018 ACM/SIGAI Autonomous Agents Research Award and a Tier I Canada Research Chair; and has received (with great co-authors) a number of Best Paper awards including: the 2009 IJCAI-JAIR Best Paper Prize; the 2014 AIJ Prominent Paper Award; and the 2018 NeurIPS Best Paper Award.

Real-world Sequential Decision Making: Reinforcement Learning and Beyond

Hoang Le, Yisong Yue, Adith Swaminathan, Byron Boots, Ching-An Cheng

Seaside Ballroom, Fri Jun 14, 14:00 PM

This workshop aims to bring together researchers from industry and academia in order to describe recent advances and discuss future research directions pertaining to real-world sequential decision making, broadly construed. We aim to highlight new and emerging research opportunities for the machine learning community that arise from the evolving needs for making decision making theoretically and practically relevant for realistic applications. We are particularly interested in understanding the current challenges that prevent broader adoption of current policy learning and evaluation algorithms in high-impact applications, across a broad range of domains.

We welcome attendance and contribution from researchers, with either theoretical or applied interests, among a diverse range of focus:
- Application areas: applications of learning-based techniques to real-life sequential decision making (robotics, healthcare, education, transportation & energy, smart-grids, sustainability, NLP, social media & advertising, agriculture, manufacturing, economics and policy)
- Methods that address real-world desiderata and concerns: safety, reliable decision making, theoretical guarantees, verifiability & interpretability, data-efficiency, data-heterogeneity, efficient exploration, counterfactual reasoning and off-policy evaluation, cost function design, efficient implementations to large-scale systems
- Cross-boundary research along the theme of RL+X where X indicate areas not commonly viewed as RL in contemporary research. We would like to encourage researchers to explore the interface between traditional RL with: (i) Other related areas in machine learning including but not limited to: imitation learning, transfer learning, active learning, structured prediction, off-policy learning, fairness in ML, privacy (ii) Areas outside of machine learning including but not limited to: control theory & dynamical systems, formal methods, causal inference, game-theory, operations research, systems research, human-computer interactions, human behavior modeling

Schedule

02:00 PM	**Talk by Emma Brunskill (Stanford)**	*Brunskill*
02:30 PM	**Talk by Miroslav Dudík (Microsoft Research)**	*Dudik*
03:00 PM	**Poster Session Part 1 and Coffee Break**	
04:00 PM	**Talk by Suchi Saria (John Hopkins)**	*Saria*
04:30 PM	**Talk by Dawn Woodard (Uber)**	*Woodard*
05:00 PM	**Panel Discussion with Invited Speakers**	
05:30 PM	**Poster Session Part 2**	

June 15, 2019

Machine Learning for Music Discovery

Erik Schmidt, Oriol Nieto, Fabien Gouyon, Katherine Kinnaird, Gert Lanckriet

204, Sat Jun 15, 08:30 AM

The ever-increasing size and accessibility of vast music libraries has created a demand more than ever for artificial systems that are capable of understanding, organizing, or even generating such complex data. While this topic has received relatively marginal attention within the machine learning community, it has been an area of intense focus within the community of Music Information Retrieval (MIR). While significant progress has been made, these problems remain far from solved.

Furthermore, the recommender systems community has made great advances in terms of collaborative feedback recommenders, but these approaches suffer strongly from the cold-start problem. As such, recommendation techniques often fall back on content-based machine learning systems, but defining musical similarity is extremely challenging as myriad features all play some role (e.g., cultural, emotional, timbral, rhythmic). Thus, for machines must actually understand music to achieve an expert level of music recommendation.

On the other side of this problem sits the recent explosion of work in the area of machine creativity. Relevant examples are both Google Magenta and the startup Jukedeck, who seek to develop algorithms capable of composing and performing completely original (and compelling) works of music. These algorithms require a similar deep understanding of music and present challenging new problems for the machine learning and AI community at large.

This workshop proposal is timely in that it will bridge these separate pockets of otherwise very related research. And in addition to making progress on the challenges above, we hope to engage the wide AI and machine learning community with our nebulous problem space, and connect them with the many available datasets the MIR community has to offer (e.g., Audio Set, AcousticBrainz, Million Song Dataset), which offer near commercial scale to the academic research community.

Workshop on Self-Supervised Learning

Aaron van den Oord, Yusuf Aytar, Carl Doersch, Carl Vondrick, Alec Radford, Pierre Sermanet, Amir Zamir, Pieter Abbeel

Grand Ballroom A, Sat Jun 15, 08:30 AM

Self-supervised learning is a promising alternative where proxy tasks are developed that allow models and agents to learn without explicit supervision in a way that helps with downstream performance on tasks of interest. One of the major benefits of self-supervised learning is increasing data efficiency: achieving comparable or better performance with less labeled data or fewer environment steps (in Reinforcement learning / Robotics).
The field of self-supervised learning (SSL) is rapidly evolving, and the performance of these methods is creeping closer to the fully supervised approaches. However, many of these methods are still developed in domain-specific sub-communities, such as Vision, RL and NLP, even though many similarities exist between them. While SSL is an emerging topic and there is great interest in these techniques, there are currently few workshops, tutorials or other scientific events dedicated to this topic. This workshop aims to bring together experts with different backgrounds and applications areas to share inter-domain ideas and increase cross-pollination, tackle current shortcomings and explore new directions. The focus will be on the machine learning point of view rather than the domain side.

https://sites.google.com/corp/view/self-supervised-icml2019

Learning and Reasoning with Graph-Structured Representations

Ethan Fetaya, Zhiting Hu Hu, Thomas Kipf, Yujia Li, Xiaodan Liang, Renjie Liao, Raquel Urtasun, Hao Wang, Max Welling, Eric Xing, Richard Zemel

Grand Ballroom B, Sat Jun 15, 08:30 AM

Graph-structured representations are widely used as a natural and powerful way to encode information such as relations between objects or entities, interactions between online users (e.g., in social networks), 3D meshes in computer graphics, multi-agent environments, as well as molecular structures, to name a few. Learning and reasoning with graph-structured representations is gaining increasing interest in both academia and industry, due to its fundamental advantages over more traditional unstructured methods in supporting interpretability, causality, transferability, etc. Recently, there is a surge of new techniques in the context of deep learning, such as graph neural networks, for learning graph representations and performing reasoning and prediction, which have achieved impressive progress. However, it can still be a long way to go to obtain satisfactory results in long-range multi-step reasoning, scalable learning with very large graphs, flexible modeling of graphs in combination with other dimensions such as temporal variation and other modalities such as language and vision. New advances in theoretical foundations, models and algorithms, as well as empirical discoveries and applications are therefore all highly desirable.

The aims of this workshop are to bring together researchers to dive deeply into some of the most promising methods which are under active exploration today, discuss how we can design new and better benchmarks, identify impactful application domains, encourage discussion and foster collaboration. The workshop will feature speakers, panelists, and poster presenters from machine perception, natural language processing, multi-agent behavior and communication, meta-learning, planning, and reinforcement learning, covering approaches which include (but are not limited to):

-Deep learning methods on graphs/manifolds/relational data (e.g., graph neural networks)
-Deep generative models of graphs (e.g., for drug design)
-Unsupervised graph/manifold/relational embedding methods (e.g.,

hyperbolic embeddings)
-Optimization methods for graphs/manifolds/relational data
-Relational or object-level reasoning in machine perception
-Relational/structured inductive biases for reinforcement learning, modeling multi-agent behavior and communication
-Neural-symbolic integration
-Theoretical analysis of capacity/generalization of deep learning models for graphs/manifolds/ relational data
-Benchmark datasets and evaluation metrics

Schedule

08:45 AM	**Opening remarks**	
09:00 AM	**William L. Hamilton, McGill University**	*Hamilton*
09:30 AM	**Evolutionary Representation Learning for Dynamic Graphs; Aynaz Taheri and Tanya Berger-Wolf**	*taheri*
09:45 AM	**Poster spotlights #1**	
10:00 AM	**Morning poster session and coffee break**	
11:00 AM	**Pushmeet Kohli, DeepMind**	*Kohli*
11:30 AM	**Yaron Lipman, Weizmann Institute of Science**	*Lipman*
12:00 PM	**PAN: Path Integral Based Convolution for Deep Graph Neural Networks; Zheng Ma, Ming Li and Yu Guang Wang**	*Ma*
12:15 PM	**Poster spotlights #2**	
12:30 PM	**Lunch break**	
02:00 PM	**Alex Polozov, Microsoft Research**	*Polozov*
02:30 PM	**Sanja Fidler, University of Toronto**	*Fidler*
03:00 PM	**On Graph Classification Networks, Datasets and Baselines; Enxhell Luzhnica, Ben Day and Pietro Lió**	
03:15 PM	**Poster spotlights #3**	
03:30 PM	**Afternoon poster session and coffee break**	
04:30 PM	**Caroline Uhler, MIT**	*Uhler*
05:00 PM	**Alexander Schwing, University of Illinois at Urbana-Champaign**	*Schwing*

Exploration in Reinforcement Learning Workshop

Surya Bhupatiraju, Benjamin Eysenbach, Shixiang Gu, Harrison Edwards, Martha White, Pierre-Yves Oudeyer, Kenneth Stanley, Emma Brunskill

Hall A, Sat Jun 15, 08:30 AM

Exploration is a key component of reinforcement learning (RL). While RL has begun to solve relatively simple tasks, current algorithms cannot complete complex tasks. Our existing algorithms often endlessly dither, failing to meaningfully explore their environments in search of high-reward states. If we hope to have agents autonomously learn increasingly complex tasks, these machines must be equipped with machinery for efficient exploration.

The goal of this workshop is to present and discuss exploration in RL, including deep RL, evolutionary algorithms, real-world applications, and developmental robotics. Invited speakers will share their perspectives on efficient exploration, and researchers will share recent work in spotlight presentations and poster sessions.

Identifying and Understanding Deep Learning Phenomena

Hanie Sedghi, Samy Bengio, Kenji Hata, Aleksander Madry, Ari Morcos, Behnam Neyshabur, Maithra Raghu, Ali Rahimi, Ludwig Schmidt, Ying Xiao

Hall B, Sat Jun 15, 08:30 AM

Our understanding of modern neural networks lags behind their practical successes. As this understanding gap grows, it poses a serious challenge to the future pace of progress because fewer pillars of knowledge will be available to designers of models and algorithms. This workshop aims to close this understanding gap in deep learning. It solicits contributions that view the behavior of deep nets as a natural phenomenon to investigate with methods inspired from the natural sciences, like physics, astronomy, and biology. We solicit empirical work that isolates phenomena in deep nets, describes them quantitatively, and then replicates or falsifies them.

As a starting point for this effort, we focus on the interplay between data, network architecture, and training algorithms. We are looking for contributions that identify precise, reproducible phenomena, as well as systematic studies and evaluations of current beliefs such as “sharp local minima do not generalize well” or “SGD navigates out of local minima”. Through the workshop, we hope to catalogue quantifiable versions of such statements, as well as demonstrate whether or not they occur reproducibly.

Schedule

08:45 AM	**Opening Remarks**	
09:00 AM	**Nati Srebro: Optimization’s Untold Gift to Learning: Implicit Regularization**	*Srebro*
09:30 AM	**Bad Global Minima Exist and SGD Can Reach Them**	

09:45 AM	**Deconstructing Lottery Tickets: Zeros, Signs, and the Supermask**	
10:00 AM	**Chiyuan Zhang: Are all layers created equal? -- Studies on how neural networks represent functions**	
10:30 AM	**Break and Posters**	
11:00 AM	**Line attractor dynamics in recurrent networks for sentiment classification**	
11:15 AM	**Do deep neural networks learn shallow learnable examples first?**	
11:30 AM	**Crowdsourcing Deep Learning Phenomena**	
12:00 PM	**Lunch and Posters**	
01:30 PM	**Aude Oliva: Reverse engineering neuroscience and cognitive science principles**	
02:00 PM	**On Understanding the Hardness of Samples in Neural Networks**	
02:15 PM	**On the Convex Behavior of Deep Neural Networks in Relation to the Layers' Width**	
02:30 PM	**Andrew Saxe: Intriguing phenomena in training and generalization dynamics of deep networks**	*Saxe*
03:00 PM	**Break and Posters**	
04:00 PM	**Olga Russakovsky**	
04:30 PM	**Panel Discussion**	

Abstracts (4):

Abstract 1: **Opening Remarks in Identifying and Understanding Deep Learning Phenomena**, 08:45 AM

Hanie Sedghi

Abstract 14: **Andrew Saxe: Intriguing phenomena in training and generalization dynamics of deep networks in Identifying and Understanding Deep Learning Phenomena**, *Saxe* 02:30 PM

In this talk I will describe several phenomena related to learning dynamics in deep networks. Among these are (a) large transient training error spikes during full batch gradient descent, with implications for the training error surface; (b) surprisingly strong generalization performance of large networks with modest label noise even with infinite training time; (c) a training speed/test accuracy trade off in vanilla deep networks; (d) the inability of deep networks to learn known efficient representations of certain functions; and finally (e) a trade off between training speed and multitasking ability.

Abstract 16: **Olga Russakovsky in Identifying and Understanding Deep Learning Phenomena**, 04:00 PM

Olga Russakovsky

Abstract 17: **Panel Discussion in Identifying and Understanding Deep Learning Phenomena**, 04:30 PM

Panelists:
Kevin Murphy, Nati Srebro, Aude Oliva, Andrew Saxe, Olga Russakovsky
Moderator:
Ali Rahimi

Workshop on AI for autonomous driving

Anna Choromanska, Larry Jackel, Li Erran Li, Juan Carlos Niebles, Adrien Gaidon, Ingmar Posner, Wei-Lun (Harry) Chao

Room 101, Sat Jun 15, 08:30 AM

A diverse set of methods have been devised to develop autonomous driving platforms. They range from modular systems, systems that perform manual decomposition of the problem, systems where the components are optimized independently, and a large number of rules are programmed manually, to end-to-end deep-learning frameworks. Today’s systems rely on a subset of the following: camera images, HD maps, inertial measurement units, wheel encoders, and active 3D sensors (LIDAR, radar). There is a general agreement that much of the self-driving software stack will continue to incorporate some form of machine learning in any of the above mentioned systems in the future.

Self-driving cars present one of today’s greatest challenges and opportunities for Artificial Intelligence (AI). Despite substantial investments, existing methods for building autonomous vehicles have not yet succeeded, i.e., there are no driverless cars on public roads today without human safety drivers. Nevertheless, a few groups have started working on extending the idea of learned tasks to larger functions of autonomous driving. Initial results on learned road following are very promising.

The goal of this workshop is to explore ways to create a framework that is capable of learning autonomous driving capabilities beyond road following, towards fully driverless cars. The workshop will consider the current state of learning applied to autonomous vehicles and will explore how learning may be used in future systems. The workshop will span both theoretical frameworks and practical issues especially in the area of deep learning.

Schedule

09:00 AM	**Opening Remarks**	
09:15 AM	**Invited Talk 1**	*Kreiss, Alahi*
09:40 AM	**Invited Talk 2**	*Bansal*

10:05 AM	**Invited Talk 3**	*Finn*
10:50 AM	**Invited Talk 4**	*Levine*
11:15 AM	**Invited Talk 5**	*Burgard*
11:40 AM	**Invited Talk 6**	*Sadigh*
01:30 PM	**Poster Session**	*Jeong, Philion*
02:30 PM	**Invited Talk 7**	*Amini*
02:55 PM	**Invited Talk 8**	*Yu, Darrell*
03:20 PM	**Invited Talk 9**	*Canziani*
04:05 PM	**Invited Talk 10**	*Xiao*
04:30 PM	**Invited Talk 11**	*Ros*
04:55 PM	**Invited Talk 12**	*Narayanan, Bagnell*
05:20 PM	**Best Paper Award and Panel Discussion**	

Workshop on Multi-Task and Lifelong Reinforcement Learning

Sarath Chandar, Shagun Sodhani, Khimya Khetarpal, Tom Zahavy, Daniel J. Mankowitz, Shie Mannor, Balaraman Ravindran, Doina Precup, Chelsea Finn, Abhishek Gupta, Amy Zhang, Kyunghyun Cho, Andrei Rusu, Facebook Rob Fergus

Room 102, Sat Jun 15, 08:30 AM

Website
https://sites.google.com/view/mtlrl

Abstract
Significant progress has been made in reinforcement learning, enabling agents to accomplish complex tasks such as Atari games, robotic manipulation, simulated locomotion, and Go. These successes have stemmed from the core reinforcement learning formulation of learning a single policy or value function from scratch. However, reinforcement learning has proven challenging to scale to many practical real world problems due to problems in learning efficiency and objective specification, among many others. Recently, there has been emerging interest and research in leveraging structure and information across multiple reinforcement learning tasks to more efficiently and effectively learn complex behaviors. This includes:

* curriculum and lifelong learning, where the problem requires learning a sequence of tasks, leveraging their shared structure to enable knowledge transfer
* goal-conditioned reinforcement learning techniques that leverage the structure of the provided goal space to learn many tasks significantly faster
* meta-learning methods that aim to learn efficient learning algorithms that can learn new tasks quickly
* hierarchical reinforcement learning, where the reinforcement learning problem might entail a compositions of subgoals or subtasks with shared structure

Multi-task and lifelong reinforcement learning has the potential to alter the paradigm of traditional reinforcement learning, to provide more practical and diverse sources of supervision, while helping overcome many challenges associated with reinforcement learning, such as exploration, sample efficiency and credit assignment. However, the field of multi-task and lifelong reinforcement learning is still young, with many more developments needed in terms of problem formulation, algorithmic and theoretical advances as well as better benchmarking and evaluation.

The focus of this workshop will be on both the algorithmic and theoretical foundations of multi-task and lifelong reinforcement learning as well as the practical challenges associated with building multi-tasking agents and lifelong learning benchmarks. Our goal is to bring together researchers that study different problem domains (such as games, robotics, language, and so forth), different optimization approaches (deep learning, evolutionary algorithms, model-based control, etc.), and different formalisms (as mentioned above) to discuss the frontiers, open problems and meaningful next steps in multi-task and lifelong reinforcement learning.

Confirmed Speakers

* Jacob Andreas
* Jeff Clune
* Karol Hausman
* Nicolas Heess
* Sergey Levine
* Natalia Rodriguez
* Benjamin Rosman
* Peter Stone
* Martha White

Schedule

09:00 AM	**Sergey Levine: Unsupervised Reinforcement Learning and Meta-Learning**	*Levine*
09:25 AM	**Martha White: TBD**	*White*
11:00 AM	**Jacob Andreas: Linguistic Scaffolds for Policy Learning**	*Andreas*
11:25 AM	**Karol Hausman: TBD**	*Hausman*
02:00 PM	**Peter Stone: Learning Curricula for Transfer Learning in RL**	*Stone*
02:25 PM	**Natalia Diaz-Rodriguez: Continual Learning and Robotics: an overview**	*Diaz Rodriguez*
03:30 PM	**Jeff Clune: TBD**	*Clune*
04:15 PM	**Nicolas Heess: TBD**	*Heess*
04:40 PM	**Benjamin Rosman: Exploiting Structure For Accelerating Reinforcement Learning**	*Rosman*

Invertible Neural Networks and Normalizing Flows

Chin-Wei Huang, David Krueger, Rianne Van den Berg, George Papamakarios, Aidan Gomez, Chris Cremer, Ricky T. Q. Chen, Aaron Courville, Danilo J. Rezende

Room 103, Sat Jun 15, 08:30 AM

Invertible neural networks have been a significant thread of research in the ICML community for several years. Such transformations can offer a range of unique benefits:

(1) They preserve information, allowing perfect reconstruction (up to numerical limits) and obviating the need to store hidden activations in memory for backpropagation.
(2) They are often designed to track the changes in probability density that applying the transformation induces (as in normalizing flows).
(3) Like autoregressive models, normalizing flows can be powerful generative models which allow exact likelihood computations; with the right architecture, they can also allow for much cheaper sampling than autoregressive models.

While many researchers are aware of these topics and intrigued by several high-profile papers, few are familiar enough with the technical details to easily follow new developments and contribute. Many may also be unaware of the wide range of applications of invertible neural networks, beyond generative modelling and variational inference.

Schedule

09:30 AM	**Tutorial on normalizing flows**	*Jang*
10:30 AM	**Poster Spotlights**	
10:50 AM	**poster session I**	*Rhinehart, Tang, Prabhu, Yap, Wang, Finzi, Kumar, Lu, Kumar, Lei, Przystupa, De Cao, Kirichenko, Izmailov, Wilson, Kruse, Mesquita, Lezcano Casado, Müller, Simmons, Atanov*
11:30 AM	**Building a tractable generator network**	
11:50 AM	**Glow: Generative Flow with Invertible 1x1 Convolutions**	
12:10 PM	**Contributed talk**	
02:00 PM	**Householder meets Sylvester: Normalizing flows for variational inference**	
02:20 PM	**Neural Ordinary Differential Equations for Continuous Normalizing Flows**	
02:40 PM	**Contributed talk**	
03:00 PM	**poster session II**	
04:00 PM	**Invited talk**	
04:20 PM	**Invertible Neural Networks for Understanding and Controlling Learned Representations**	
04:40 PM	**Contributed talk**	
05:00 PM	**Panel Session**	

Stein’s Method for Machine Learning and Statistics

Francois-Xavier Briol, Lester Mackey, Chris Oates, Qiang Liu, Larry Goldstein

Room 104 A, Sat Jun 15, 08:30 AM

Stein's method is a technique from probability theory for bounding the distance between probability measures using differential and difference operators. Although the method was initially designed as a technique for proving central limit theorems, it has recently caught the attention of the machine learning (ML) community and has been used for a variety of practical tasks. Recent applications include generative modeling, global non-convex optimisation, variational inference, de novo sampling, constructing powerful control variates for Monte Carlo variance reduction, and measuring the quality of (approximate) Markov chain Monte Carlo algorithms. Stein's method has also been used to develop goodness-of-fit tests and was the foundational tool in one of the NeurIPS 2017 Best Paper awards.

Although Stein's method has already had significant impact in ML, most of the applications only scratch the surface of this rich area of research in probability theory. There would be significant gains to be made by encouraging both communities to interact directly, and this inaugural workshop would be an important step in this direction. More precisely, the aims are: (i) to introduce this emerging topic to the wider ML community, (ii) to highlight the wide range of existing applications in ML, and (iii) to bring together experts in Stein's method and ML researchers to discuss and explore potential further uses of Stein's method.

AI For Social Good (AISG)

Margaux Luck, Kris Sankaran, Tristan Sylvain, Sean McGregor, Jonnie Penn, Girmaw Abebe Tadesse, Virgile Sylvain, Myriam Côté, Lester Mackey, Rayid Ghani, Yoshua Bengio

Room 104 B, Sat Jun 15, 08:30 AM

AI for Social Good

Important information

Contact information: aisg2019.icml.contact@gmail.com

Submission deadline: **EXTENDED to April 26th 2019 11:59PM ET**

Workshop website

Submission website

Poster Information:

- Poster Size - **36W x 48H inches or 90 x 122 cm**
- Poster Paper - **lightweight paper - not laminated**

Abstract

This workshop builds on our AI for Social Good workshop at NeurIPS 2018 and ICLR 2019.

Introduction: The rapid expansion of AI research presents two clear conundrums:

- the comparative lack of incentives for researchers to address social impact issues and
- the dearth of conferences and journals centered around the topic. Researchers motivated to help often find themselves without a clear idea of which fields to delve into.

Goals: Our workshop address both these issues by bringing together machine learning researchers, social impact leaders, stakeholders, policy leaders, and philanthropists to discuss their ideas and applications for social good. To broaden the impact beyond the convening of our workshop, we are partnering with AI Commons to expose accepted projects and papers to the broader community of machine learning researchers and engineers. The projects/research may be at varying degrees of development, from formulation as a data problem to detailed requirements for effective deployment. We hope that this gathering of talent and information will inspire the creation of new approaches and tools by the community, help scientists access the data they need, involve social and policy stakeholders in the framing of machine learning applications, and attract interest from philanthropists invited to the event to make a dent in our shared goals.

Topics: The UN Sustainable Development Goals (SDGs), a set of seventeen objectives whose completion is set to lead to a more equitable, prosperous, and sustainable world. In this light, our main areas of focus are the following: health, education, the protection of democracy, urban planning, assistive technology, agriculture, environmental protection and sustainability, social welfare and justice, developing world. Each of these themes presents unique opportunities for AI to reduce human suffering and allow citizens and democratic institutions to thrive.

Across these topics, we have dual goals: recognizing high-quality work in machine learning motivated by or applied to social applications, and creating meaningful connections between communities dedicated to solving technical and social problems. To this extent, we propose two research tracks:

- **Short Papers Track (Up to four page papers + unlimited pages for citations)** for oral and/or poster presentation. The short papers should focus on past and current research work, showcasing actual results and demonstrating beneficial effects on society. We also accept short papers of recently published or submitted journal contributions to give authors the opportunity to present their work and obtain feedback from conference attendees.
- **Problem Introduction Track (Application form, up to five page responses + unlimited pages for citations)** which will present a specific solution that will be shared with stakeholders, scientists, and funders. The workshop will provide a suite of questions designed to: (1) estimate the feasibility and impact of the proposed solutions, and (2) estimate the importance of data in their implementation. The application responses should highlight ideas that have not yet been implemented in practice but can lead to real impact. The projects may be at varying degrees of development, from formulation as a data problem to structure for effective deployment. The workshop provides a supportive platform for developing these early-stage or hobby proposals into real projects. This process is designed to foster sharing different points of view ranging from the scientific assessment of feasibility, discussion of practical constraints that may be encountered, and attracting interest from philanthropists invited to the event. Accepted submissions may be promoted to the wider AI solutions community following the workshop via the AI Commons, with whom we are partnering to promote the longer-term development of projects.

Schedule

08:45 AM	**Welcoming and Poster set-up**	
09:00 AM	**Opening remarks**	*Bengio*
09:05 AM	**Solving societal challenges with AI through partnerships**	*Anandan*
09:45 AM	**AI Commons**	*Bengio*
09:50 AM	**Detecting Waterborne Debris with Sim2Real and Randomization**	*Sankaran*
10:00 AM	**Conversational agents to address abusive online behaviors**	*Beauxis-Aussalet*
10:10 AM	**Poster Session 1**	
10:30 AM	**Break / Poster Session 1**	
11:00 AM	**AI for Ecology and Conservation**	*Sheldon*
11:40 AM	**Using AI for Economic Upliftment of Handicraft Industry**	*Damani*
11:50 AM	**Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening**	*Geras, Wu*
12:00 PM	**Lunch - on your own**	
02:00 PM	**Keynote 3 - Coming Soon**	
02:40 PM	**Poster Session 2**	*Fang, Balashankar, Damani, Beauxis-Aussalet, Wu, Bondi, Rußwurm, Ruhe, Saxena, Spoon*
03:00 PM	**Break / Poster Session 2**	
03:30 PM	**The History of AI and Mental Health**	
04:10 PM	**Learning Global Variations in Outdoor PM_2.5 Concentrations with Satellite Images**	*Oliveira Pinheiro*
04:20 PM	**Pareto Efficient Fairness for Skewed Subgroup Data**	*Balashankar*
04:30 PM	**Crisis Sub-Events on Social Media: A Case Study of Wildfires**	
04:40 PM	**Towards Detecting Dyslexia in Children's Handwriting Using Neural Networks**	
04:50 PM	**Bridging Critics and Designers of Socially Impactful AI**	
05:50 PM	**Open announcement and Best Paper Award**	

Abstracts (15):

Abstract 2: **Opening remarks in AI For Social Good (AISG)**, *Bengio* 09:00 AM

Speaker bio:
Yoshua Bengio is Full Professor of the Department of Computer Science and Operations Research, scientific director of Mila, CIFAR Program co-director of the CIFAR Learning in Machines and Brains program (formerly Neural Computation and Adaptive Perception), scientific director of IVADO and Canada Research Chair in Statistical Learning Algorithms. His main research ambition is to understand principles of learning that yield intelligence. He supervises a large group of graduate students and post-docs. His research is widely cited (over 130000 citations found by Google Scholar in August 2018, with an H-index over 120, and rising fast).

Abstract 3: **Solving societal challenges with AI through partnerships in AI For Social Good (AISG)**, *Anandan* 09:05 AM

Wadhwani AI was inaugurated a little more than a year ago with the mission of bringing the power of AI to address societal challenges, especially among underserved communities throughout the world. We aim to address problems all major domains including health, agriculture, education, infrastructure, and financial inclusion. We are currently working on three solutions (two in health and one in agriculture) and are exploring more areas where we can apply AI for social good.The most important lesson that we have learned during our short stint is the importance of working in close partnership with other stakeholders and players in the social sectors, especially NGOs and Government organizations. In this talk, I will use one case, namely that of developing an AI based approach for Integrated Pest Management (IPM) in Cotton Farming, to describe how this partnership based approach has evolved and been critical to our solution development and implementation.

Speaker bio:
Dr. P. Anandan is the CEO of Wadhwani Institute of Artificial Intelligence. His prior experience includes - Adobe Research Lab India (2016-2017) as a VP for Research and a Distinguished Scientist and Managing Director at Microsoft Research (1997-2014). He was also the founding director of Microsoft Research India which he ran from 2005-2014. Earlier stint was at Sarnoff Corporation (1991-1997) as a researcher and an Assistant Professor of Computer Science at Yale University (1987-1991). His primary research area is Computer vision where he is well known for his fundamental and lasting contributions to the problem of visual motion analysis. He received his PhD in Computer Science from University of Massachusetts, Amherst in 1987, a Masters in Computer Science from University of Nebraska, Lincoln in 1979 and his B.Tech in Electrical Engineering from IIT Madras, India in 1977. He is a distinguished alumnus of IIT Madras, and UMass, Amherst and is on the Nebraska Hall of Computing. His hobbies include playing African drums, writing poems (in Tamil) and travel which makes his work related travel interesting.

Abstract 4: **AI Commons in AI For Social Good (AISG)**, *Bengio* 09:45 AM

AI Commons is a collective project whose goal is to make the benefits of AI available to all. Since AI research can benefit from the input of a large range of talents across the world, the project seeks to develop ways for developers and organizations to collaborate more easily and effectively. As a community operating in an environment of trust and problem-solving, AI Commons can empower researchers to tackle the world's important problems using all the possibilities of cutting-edge AI.

Speaker bio:
Yoshua Bengio is Full Professor of the Department of Computer Science and Operations Research, scientific director of Mila, CIFAR Program co-director of the CIFAR Learning in Machines and Brains program (formerly Neural Computation and Adaptive Perception), scientific director of IVADO and Canada Research Chair in Statistical Learning Algorithms. His main research ambition is to understand principles of learning that yield intelligence. He supervises a large group of graduate students and post-docs. His research is widely cited (over 130000 citations found by Google Scholar in August 2018, with an H-index over 120, and rising fast).

Abstract 5: **Detecting Waterborne Debris with Sim2Real and Randomization in AI For Social Good (AISG)**, *Sankaran* 09:50 AM

Marine debris pollution is one of the most ubiquitous and pressing environmental issues affecting our oceans today. Clean up efforts such as the Great Pacific Garbage Patch project have been implemented across the planet to combat this problem. However, resources to accomplish this goal are limited, and the afflicted area is vast. To this end, unmanned vehicles that are capable of automatically detecting and removing small-sized debris would be a great complementary approach to existing large-scale garbage collectors. Due to the complexity of fully functioning unmanned vehicles for both detecting and removing debris, in this project, we focus on the detection task as a first step. From the perspective of machine learning, there is an unfortunate lack of sufficient labeled data for training a specialized detector, e.g., a classifier that can distinguish debris from other objects like wild animals. Moreover, pre-trained detectors on other domains would be ineffective while creating such datasets manually would be very costly. Due to the recent progress of training deep models with synthetic data and domain randomization, we propose to train a debris detector based on a mixture of real and synthetic images.

Speaker bio:
Kris is a postdoc at Mila working with Yoshua Bengio on problems related to Humanitarian AI. He is generally interested in ways to broaden the scope of problems studied by the machine learning community and am curious about the ways to bridge statistical and computational thinking.

Abstract 6: **Conversational agents to address abusive online behaviors in AI For Social Good (AISG)**, *Beauxis-Aussalet* 10:00 AM

Technologies to address cyber bullying are limited to detecting and hiding abusive messages. We propose to investigate the potential of conversational technologies for addressing abusers. We will outline directions for studying the effectiveness dialog strategies (e.g., to educate or deter abusers, or keep them busy with chatbots rather than their victims) and for initiating new research on chatbot-mediated mitigation of online abuse.

Speaker bio:
Emma Beauxis-Aussalet is a Senior Track Associate at the Digital Society School of Amsterdam University of Applied Science, where she investigates how data-driven technologies can be applied for the best interests of society. She holds a PhD on classification errors and biases from Utrecht University. Her interests include ethical and explainable AI, data literacy in the general public, and the synergy between human & artificial intelligence to tackle job automation.

Abstract 9: **AI for Ecology and Conservation in AI For Social Good (AISG)**, *Sheldon* 11:00 AM

AI can help solve big data and decision-making problems to understand and protect the environment. I’ll survey several projects the area and discuss how to approach environmental problems using AI. The Dark Ecology project uses weather radar and machine learning to unravel mysteries of bird migration. A surprising probabilistic inference problem arises when analyzing animal survey data to monitor populations. Novel optimization algorithms can help reason about dams, hydropower, and the ecology of river networks.

Speaker bio:
Daniel Sheldon is an Assistant Professor of Computer Science at the University of Massachusetts Amherst and Mount Holyoke College. His research investigates fundamental problems in machine learning and AI motived by large-scale environmental data, dynamic ecological processes, and real-world network phenomena.

Abstract 10: **Using AI for Economic Upliftment of Handicraft Industry in AI For Social Good (AISG)**, *Damani* 11:40 AM

The handicraft industry is a strong pillar of Indian economy which provides large-scale employment opportunities to artisans in rural and underprivileged communities. However, in this era of globalization, diverse modern designs have rendered traditional designs old and monotonous, causing an alarming decline of handicraft sales. In this talk, we will discuss our approach leveraging techniques like GANs, Color Transfer, Pattern Generation etc. to generate contemporary designs for two popular Indian handicrafts - Ikat and Block Print. The resultant designs are evaluated to be significantly more likeable and marketable than the current designs used by artisans.

Speaker bio:
Sonam Damani is an Applied Scientist in Microsoft, India where she has worked on several projects in the field of AI and Deep Learning, including Microsoft's human-like-chatbot Ruuh, Cortana personality, novel art generation using AI, Bing search relevance, among others. In the past year, she has co-authored a bunch of publications in the field of conversational AI and AI creativity that were presented in NeurIPS, WWW and CODS-COMAD.

Abstract 11: **Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening in AI For Social Good (AISG)**, *Geras, Wu* 11:50 AM

We present a deep CNN for breast cancer screening exam classification, trained and evaluated on over 200,000 exams (over 1,000,000 images). Our network achieves an AUC of 0.895 in predicting whether there is a cancer in the breast, when tested on the screening population. We attribute the high accuracy of our model to a two-stage training procedure, which allows us to use a very high-capacity patch-level network to learn from pixel-level labels alongside a network learning from macroscopic breast-level labels. To validate our model, we conducted a reader study with 14 readers, each reading 720 screening mammogram exams, and find our model to be as accurate as experienced radiologists when presented with the same data. Finally, we show that a hybrid model, averaging probability of malignancy predicted by a radiologist with a prediction of our neural network, is more accurate than either of the two separately.

Speaker bio:
- Krzysztof Geras is an assistant professor at NYU School of Medicine and an affiliated faculty at NYU Center for Data Science. His main interests are in unsupervised learning with neural networks, model compression, transfer learning, evaluation of machine learning models and applications of these techniques to medical imaging. He previously did a postdoc at NYU with Kyunghyun Cho, a PhD at the University of Edinburgh with Charles Sutton and an MSc as a visiting student at the University of Edinburgh with Amos Storkey. His BSc is from the University of Warsaw.

- Nan Wu is a PhD student at NYU Center for Data Science. She is interested in data science with application to healthcare and currently working on medical image analysis. Before joining NYU, she graduated from School for Gifted Young, University of Science and Technology of China, receiving B.S in Statistics and B.A. in Business Administration.

Abstract 13: **Keynote 3 - Coming Soon in AI For Social Good (AISG)**, 02:00 PM

Coming Soon

Abstract 16: **The History of AI and Mental Health in AI For Social Good (AISG)**, 03:30 PM

Coming Soon.

Speaker bio:
Luke Stark is a Postdoctoral Researcher in the Fairness, Accountability, Transparency and Ethics (FATE) Group at Microsoft Research Montreal, and an Affiliate of the Berkman Klein Center for Internet & Society at Harvard University. Luke holds a PhD from the Department of Media, Culture, and Communication at New York University, and an Honours BA and MA in History from the University of Toronto. Trained as a media historian, his scholarship centers on the interconnected histories of artificial intelligence (AI) and behavioral science, and on the ways the social and ethical contexts of AI are changing how we work, communicate, and participate in civic life.

Abstract 17: **Learning Global Variations in Outdoor PM_2.5 Concentrations with Satellite Images in AI For Social Good (AISG)**, *Oliveira Pinheiro* 04:10 PM

The World Health Organization identifies outdoor fine particulate air pollution (PM2.5) as a leading risk factor for premature mortality globally. As such, understanding the global distribution of PM2.5 is an essential precursor towards implementing pollution mitigation strategies and modelling global public health. Here, we present a convolutional neural network based approach for estimating annual average outdoor PM2.5 concentrations using only satellite images. The resulting model achieves comparable performance to current state-of-the-art statistical models.

Speaker bio:
- Kris Y Hong is a research assistant and prospective PhD student in the Weichenthal Lab at McGill University, in Montreal, Canada. His interests lie in applying current statistical and machine learning techniques towards solving humanitarian and environmental challenges. Prior to joining McGill, he was a data analyst at the British Columbia Centre for Disease Control while receiving his B.Sc. in Statistics from the University of British Columbia.

- Dr. Pedro O Pinheiro is a research scientist at Element AI, in Montreal, Canada. His research focuses on computer vision, machine learning, and their intersection. He finished his PhD in 2017 at École Polytechnique Fédérale de Lausanne (EPFL), in Switzerland, under the supervision of Ronan Collobert (and working closely with Piotr Dollar from FAIR). During his PhD, his work focused mostly on large-scale image segmentation problems using deep learning techniques. Currently, he is interested in learning with less supervision and how to leverage knowledge from one task to another.

- Dr. Scott Weichenthal is an Assistant Professor in the Department of Epidemiology, Biostatistics, and Occupational Health at McGill University in Montreal, Canada. His research program is dedicated to identifying and evaluating environmental risk factors for chronic diseases such as cancer and cardiovascular disease. His current research is focused on the use of deep learning models in estimating environmental exposures on both a local and global scale.

Abstract 18: **Pareto Efficient Fairness for Skewed Subgroup Data in AI For Social Good (AISG)**, *Balashankar* 04:20 PM

As awareness of the potential for learned models to amplify existing societal biases increases, the field of ML fairness has developed mitigation techniques. A prevalent method applies constraints, including equality of performance, with respect to subgroups defined over the intersection of sensitive attributes such as race and gender. Enforcing such constraints when the subgroup populations are considerably skewed with respect to a target can lead to unintentional degradation in performance, without benefiting any individual subgroup, counter to the United Nations Sustainable Development goals of reducing inequalities and promoting growth. In order to avoid such performance degradation while ensuring equitable treatment to all groups, we propose Pareto-Efficient Fairness (PEF), which identifies the operating point on the Pareto curve of subgroup performances closest to the fairness hyperplane. Specifically, PEF finds a Pareto Optimal point which maximizes multiple subgroup accuracy measures. The algorithm scalarizes using the adaptive weighted metric norm by iteratively searching the Pareto region of all models enforcing the fairness constraint. PEF is backed by strong theoretical results on discoverability and provides domain practitioners finer control in navigating both convex and non-convex accuracy-fairness trade-offs. Empirically, we show that PEF increases performance of all subgroups in skewed synthetic data and UCI datasets.

Speaker bio:
Ananth Balashnkar is a 2nd year Ph.D student in Computer Science advised by Prof. Lakshminarayanan Subramanian at NYU's Courant Institute of Mathematical Sciences. He is currently interested in Interpretable Machine Learning and the challenges involved in applying machine perception for the domains of policy, privacy, economics and healthcare.

Abstract 19: **Crisis Sub-Events on Social Media: A Case Study of Wildfires in AI For Social Good (AISG)**, 04:30 PM

Social media has been extensively used for crisis management. Recent work examines possible sub-events as a major crisis unfolds. In this project, we first propose a framework to identify sub-events from tweets. Then, leveraging 4 California wildfires in 2018-2019 as a case study, we investigate how sub-events cascade based on existing hypotheses drawn from the disaster management literature, and find that most hypotheses are supported on social media, e.g., fire induces smoke, which causes air pollution, which later harms health and eventually affects the healthcare system. In addition, we discuss other unexpected sub-events that emerge from social media.

Speaker bio:
Alejandro (Alex) Jaimes is Chief Scientist and SVP of AI at Dataminr. Alex has 15+ years of intl. experience in research and product impact at scale. He has published 100+ technical papers in top-tier conferences and journals in diverse topics in AI and has been featured widely in the press (MIT Tech review, CNBC, Vice, TechCrunch, Yahoo! Finance, etc.). He has given 80+ invited talks (AI for Good Global Summit (UN, Geneva), the Future of Technology Summit, O'Reilly (AI, Strata, Velocity), Deep Learning Summit, etc.). Alex is also an Endeavor Network mentor (which leads the high-impact entrepreneurship movement around the world), and was an early voice in Human-Centered AI (Computing). He holds a Ph.D. from Columbia U.

Abstract 20: **Towards Detecting Dyslexia in Children's Handwriting Using Neural Networks in AI For Social Good (AISG)**, 04:40 PM

Dyslexia is a learning disability that hinders a person's ability to read. Dyslexia needs to be caught early, however, teachers are not trained to detect dyslexia and screening tests are used inconsistently. We propose (1) two new data sets of handwriting collected from children with and without dyslexia amounting to close to 500 handwriting samples, and (2) an automated early screening technique to be used in conjunction with current approaches, to accelerate the detection process. Preliminary results suggest our system out-performs teachers.

Speaker bio:
Katie Spoon recently completed her B.S./M.S. in computer science from Indiana University with minors in math and statistics, and with research interests in anomaly detection, computer vision, data visualization, and applications of computer vision to health and education, like her senior thesis detecting dyslexia with neural networks. She worked at IBM Research in the summer of 2018 on neuromorphic computing, and will be returning there full-time. She hopes to potentially get a PhD and become a corporate research scientist.

Abstract 21: **Bridging Critics and Designers of Socially Impactful AI in AI For Social Good (AISG)**, 04:50 PM

We will explore how AI systems can be designed and evaluated in a ways that include voices beyond tech. How can systems be designed in a participatory way? How can real problems be addressed with AI technology, without lapsing into solutionism? How do best practices in development and governance vary across contexts? We will consider these questions and others, in an effort to better characterize and control the social impacts of AI.

Speaker bio:

- Mwiza Simbeye is a tech enthusiast and aficionado who is driven by passion for change. He is a computer science student at the African Leadership University and is currently leading an AI Lab at the university; to foster research and development for AI at the university to solve common problems on the continent under agriculture, healthcare and transport. He is also co-founder of (http://www.agripredict.com) and a mentee under the Google AI Mentorship Program.

- Phebe Vayanos is Assistant Professor of Industrial & Systems Engineering and Computer Science at the University of Southern California, and Associate Director of the CAIS Center for Artificial Intelligence in Society. Her research aims to address fundamental questions in data-driven optimization (aka prescriptive analytics) with aim to tackle real-world decision- and policy-making problems in uncertain and adversarial environments.

Synthetic Realities: Deep Learning for Detecting AudioVisual Fakes

Battista Biggio, Pavel Korshunov, Thomas Mensink, Giorgio Patrini, Arka Sadhu, Delip Rao

Room 104 C, Sat Jun 15, 08:30 AM

With the latest advances of deep generative models, synthesis of images and videos as well as of human voices have achieved impressive realism. In many domains, synthetic media are already difficult to distinguish from real by the human eye and ear. The potential of misuses of these technologies is seldom discussed in academic papers; instead, vocal concerns are rising from media and security organizations as well as from governments. Researchers are starting to experiment on new ways to integrate deep learning with traditional media forensics and security techniques as part of a technological solution. This workshop will bring together experts from the communities of machine learning, computer security and forensics in an attempt to highlight recent work and discuss future effort to address these challenges. Our agenda will alternate contributed papers with invited speakers. The latter will emphasize connections among the interested scientific communities and the standpoint of institutions and media organizations.

Schedule

09:00 AM	**Welcome Remarks**	*Patrini*
09:10 AM	**Invited Talk by Professor Alexei Efros (UC Berkeley)**	*Efros*
09:40 AM	**Invited Talk by Dr. Matt Turek (DARPA)**	*Turek*
10:10 AM	**Contributed Talk: Limits of Deepfake Detection: A Robust Estimation Viewpoint**	
10:30 AM	**Poster session 1 and Coffee break**	

11:30 AM	**Invited Talk by Professor Pawel Korus (NYU) Neural Imaging Pipelines - the Scourge or Hope of Forensics?**	
12:00 PM	**Contributed Talk: We Need No Pixels: Video Manipulation Detection Using Stream Descriptors**	*Güera*
12:15 PM	**Contributed Talk: A Utility-Preserving GAN for Face Obscuration**	*Hao*
12:30 PM	**Lunch Break**	
02:00 PM	**Invited Talk by Professor Luisa Verdoliva (University Federico II Naples)**	*Verdoliva*
02:30 PM	**Contributed Talk: Tampered Speaker Inconsistency Detection with Phonetically Aware Audio-visual Features**	*Korshunov*
02:45 PM	**Contributed Talk: Measuring the Effectiveness of Voice Conversion on Speaker Identification and Automatic Speech Recognition Systems**	*Keskin*
03:00 PM	**CVPR19 Media Forensics workshop: a Preview**	
03:30 PM	**Poster session 2 and Coffee break**	
04:30 PM	**Invited Talk by Tom Van de Weghe (Stanford & VRT)**	
05:00 PM	**Panel Discussion moderated by Delip Rao**	

Abstracts (5):

Abstract 4: **Contributed Talk: Limits of Deepfake Detection: A Robust Estimation Viewpoint in Synthetic Realities: Deep Learning for Detecting AudioVisual Fakes**, 10:10 AM

Deepfake detection is formulated as a hypothesis testing problem to classify an image as genuine or GAN-generated. A robust statistics view of GANs is considered to bound the error probability for various GAN implementations in terms of their performance. The bounds are further simplified using a Euclidean approximation for the low error regime. Lastly, relationships between error probability and epidemic thresholds for spreading processes in networks are established.

Abstract 7: **Contributed Talk: We Need No Pixels: Video Manipulation Detection Using Stream Descriptors in Synthetic Realities: Deep Learning for Detecting AudioVisual Fakes**, *Güera* 12:00 PM

Manipulating video content is easier than ever. Due to the misuse potential of manipulated content, multiple detection techniques that analyze the pixel data from the videos have been proposed. However, clever manipulators should also carefully forge the metadata and auxiliary header information, which is harder to do for videos than images. In this paper, we propose to identify forged videos by analyzing their multimedia stream descriptors with simple binary classifiers, completely avoiding the pixel space. Using well-known datasets, our results show that this scalable approach can achieve a high manipulation detection score if the manipulators have not done a careful data sanitization of the multimedia stream descriptors.

Abstract 8: **Contributed Talk: A Utility-Preserving GAN for Face Obscuration in Synthetic Realities: Deep Learning for Detecting AudioVisual Fakes**, *Hao* 12:15 PM

From TV news to Google StreetView, face obscuration has been used for privacy protection. Due to recent advances in the field of deep learning, obscuration methods such as Gaussian blurring and pixelation are not guaranteed to conceal identity. In this paper, we propose a utility-preserving generative model, UP-GAN, that is able to provide an effective face obscuration, while preserving facial utility. By utility-preserving we mean preserving facial features that do not reveal identity, such as age, gender, skin tone, pose, and expression. We show that the proposed method achieves a better performance than the common obscuration methods in terms of obscuration and utility preservation.

Abstract 11: **Contributed Talk: Tampered Speaker Inconsistency Detection with Phonetically Aware Audio-visual Features in Synthetic Realities: Deep Learning for Detecting AudioVisual Fakes**, *Korshunov* 02:30 PM

The recent increase in social media based propaganda, i.e., ‘fake news’, calls for automated methods to detect tampered content. In this paper, we focus on detecting tampering in a video with a person speaking to a camera. This form of manipulation is easy to perform, since one can just replace a part of the audio, dramatically changing the meaning of the video. We consider several detection approaches based on phonetic features and recurrent networks. We demonstrate that by replacing standard MFCC features with embeddings from a DNN trained for automatic speech recognition, combined with mouth landmarks (visual features), we can achieve a significant performance improvement on several challenging publicly available databases of speakers (VidTIMIT, AMI, and GRID), for which we generated sets of tampered data. The evaluations demonstrate a relative equal error rate reduction of 55% (to 4.5% from 10.0%) on the large GRID corpus based dataset and a satisfying generalization of the model on other datasets.

Abstract 12: **Contributed Talk: Measuring the Effectiveness of Voice Conversion on Speaker Identification and Automatic Speech Recognition Systems in Synthetic Realities: Deep Learning for Detecting AudioVisual Fakes**, *Keskin* 02:45 PM

This paper evaluates the effectiveness of a CycleGAN based voice converter (VC) on four speaker identification (SID) systems and an automated speech recognition (ASR) system for various purposes. Audio samples converted by the VC model are classified by the SID systems as the intended target at up to 46% top-1 accuracy among more than 250 speakers. This encouraging result in imitating the target styles led us to investigate if converted (synthetic) samples can be used to improve ASR training. Unfortunately, adding synthetic data to the ASR training set only marginally improves word and character error rates. Our results indicate that even though VC models can successfully mimic the style of target speakers as measured by SID systems, improving ASR training with synthetic data from VC systems needs further research to establish its efficacy.

ICML Workshop on Imitation, Intent, and Interaction (I3)

Nicholas Rhinehart, Sergey Levine, Chelsea Finn, He He, Ilya Kostrikov, Justin Fu, Siddharth Reddy

Room 201, Sat Jun 15, 08:30 AM

Website:
https://sites.google.com/view/icml-i3

Abstract: A key challenge for deploying interactive machine learning systems in the real world is the ability for machines to understand human intent. Techniques such as imitation learning and inverse reinforcement learning are popular data-driven paradigms for modeling agent intentions and controlling agent behaviors, and have been applied to domains ranging from robotics and autonomous driving to dialogue systems. Such techniques provide a practical solution to specifying objectives to machine learning systems when they are difficult to program by hand.

While significant progress has been made in these areas, most research effort has concentrated on modeling and controlling single agents from dense demonstrations or feedback. However, the real world has multiple agents, and dense expert data collection can be prohibitively expensive. Surmounting these obstacles requires progress in frontiers such as:
1) the ability to infer intent from multiple modes of data, such as language or observation, in addition to traditional demonstrations.
2) the ability to model multiple agents and their intentions, both in cooperative and adversarial settings.
3) handling partial or incomplete information from the expert, such as demonstrations that lack dense action annotations, raw videos, etc..

The workshop on Imitation, Intention, and Interaction (I3) seeks contributions at the interface of these frontiers, and will bring together researchers from multiple disciplines such as robotics, imitation and reinforcement learning, cognitive science, AI safety, and natural language understanding. Our aim will be to reexamine the assumptions in standard imitation learning problem statements (e.g., inverse reinforcement learning) and connect distinct application disciplines, such as robotics and NLP, with researchers developing core imitation learning algorithms. In this way, we hope to arrive at new problem formulations, new research directions, and the development of new connections across distinct disciplines that interact with imitation learning methods.

Schedule

08:45 AM	**Welcoming Remarks**
09:00 AM	**Invited talks 1, 2, and 3**
10:00 AM	**Contributed talk 1**

## Coding Theory For Large-scale Machine Learning

Viveck Cadambe, Pulkit Grover, Dimitris Papailiopoulos, Gauri Joshi

Room 202, Sat Jun 15, 08:30 AM

# Coding Theory For Large-scale Machine Learning
Coding theory involves the art and science of how to add redundancy to data to ensure that a desirable output is obtained at despite deviations from ideal behavior from the system components that interact with the data. Through a rich, mathematically elegant set of techniques, coding theory has come to significantly influence the design of modern data communications, compression and storage systems. The last few years have seen a rapidly growing interest in coding theory based approaches for the development of efficient machine learning algorithms towards robust, large-scale, distributed computational pipelines.

The CodML workshop brings together researchers developing coding techniques for machine learning, as well as researchers working on systems implementations for computing, with cutting-edge presentations from both sides. The goal is to learn about non-idealities in system components as well as approaches to obtain reliable and robust learning despite these non-idealities, and identify problems of future interest.

The workshop is co-located with ICML 2019, and will be held in Long Beach, California, USA on June 14th or 15th, 2019.

Please see the [website](https://sites.google.com/view/codml2019) for more details:

## Call for Posters

### Scope of the Workshop

In this workshop we solicit research papers focused on the application of coding and information-theoretic techniques for distributed machine learning. More broadly, we seek papers that address the problem of making machine learning more scalable, efficient, and robust. Both theoretical as well as experimental contributions are welcome. We invite authors to submit papers on topics including but not limited to:

* Asynchronous Distributed Training Methods
* Communication-Efficient Training

* Model Compression and Quantization
* Gradient Coding, Compression and Quantization
* Erasure Coding Techniques for Straggler Mitigation
* Data Compression in Large-scale Machine Learning
* Erasure Coding Techniques for ML Hardware Acceleration
* Fast, Efficient and Scalable Inference
* Secure and Private Machine Learning
* Data Storage/Access for Machine Learning Jobs
* Performance evaluation of coding techniques

### Submission Format and Instructions

The authors should prepare extended abstracts in the ICML paper format and submit via [CMT](https://cmt3.research.microsoft.com/CODMLW2019/). Submitted papers may not exceed three (3) single-spaced double-column pages excluding references. All results, proofs, figures, tables must be included in the 3 pages. The submitted manuscripts should include author names and affiliations, and an abstract that does not exceed 250 words. The authors may include a link to an extended version of the paper that includes supplementary material (proofs, experimental details, etc.) but the reviewers are not required to read the extended version.

### Dual Submission Policy

Accepted submissions will be considered non-archival and can be submitted elsewhere without modification, as long as the other conference allows it. Moreover, submissions to CodML based on work recently accepted to other venues are also acceptable (though authors should explicitly make note of this in their submissions).

### Key Dates

Paper Submission: May 3rd, 2019, 11:59 PM anywhere on earth

Decision Notification: May 12th, 2019.

Workshop date: June 14 or 15, 2019

Schedule

09:00 AM	**Salman Avestimehr**	
09:30 AM	**Vivienne Sze: "Exploiting redundancy for efficient processing of DNNs and beyond"**	
10:00 AM	**Spotlight paper: "Locality Driven Coded Computation," Michael Rudow, Rashmi Vinayak and Venkat Guruswami**	
10:10 AM	**Spotlight paper: "CodeNet: Training Large-Scale Neural Networks in Presence of Soft-Errors," Sanghamitra Dutta, Ziqian Bai, Tze Meng Low and Pulkit Grover**	
10:20 AM	**Spotlight paper: "Reliable Clustering with Redundant Data Assignment," Venkat Gandikota, Arya Mazumdar and Ankit Singh Rawat**	
10:30 AM	**Poster Session I**	*Draper, Aktas, Guler, Wang, Gandikota, Park, So, Tauz, Narra, Lin, Maddahali, Yang, Dutta, Reisizadeh, Wang, Balevi, Jain, McVay, Rudow, Soto, Li, Subramaniam, Demirhan, Gupta, Oktay, Barnes, Ballé, Haddadpour*
11:30 AM	**Rashmi Vinayak**	
12:00 PM	**Lunch Break**	
01:30 PM	**Markus Weimer: "A case for coded computing on elastic compute"**	*Weimer*
02:00 PM	**Alex Dimakis**	
02:30 PM	**Wei Zhang**	
03:30 PM	**Spotlight paper: "OverSketched Newton: Fast Convex Optimization for Serverless Systems," Vipul Gupta, Swanand Kadhe, Thomas Courtade, Michael Mahoney and Kannan Ramchandran**	
03:40 PM	**Spotlight paper: "Cooperative SGD: A Unified Framework for the Design and Analysis of Communication-Efficient SGD Algorithms", Jianyu Wang and Gauri Joshi**	
03:50 PM	**Spotlight paper: "Secure Coded Multi-Party Computation for Massive Matrices with Adversarial Nodes," Seyed Reza, Mohammad Ali Maddah-Ali and Mohammad Reza Aref**	
04:00 PM	**Poster Session II**	

The How2 Challenge: New Tasks for Vision & Language

Florian Metze, Lucia Specia, Desmond Elliot, Loic Barrault, Ramon Sanabria, Shruti Palaskar

Room 203, Sat Jun 15, 08:30 AM

Research at the intersection of vision and language has been attracting a lot of attention in recent years. Topics include the study of multi-modal representations, translation between modalities, bootstrapping of labels from one modality into another, visually-grounded question answering, segmentation and storytelling, and grounding the meaning of language in visual data. An ever-increasing number of tasks and datasets are appearing around this recently-established field.

At NeurIPS 2018, we released the How2 data-set, containing more than 85,000 (2000h) videos, with audio, transcriptions, translations, and textual summaries. We believe it presents an ideal resource to bring together researchers working on the previously mentioned separate tasks around a single, large dataset. This rich dataset will facilitate the comparison of tools and algorithms, and hopefully foster the creation of additional annotations and tasks. We want to foster discussion about useful tasks, metrics, and labeling techniques, in order to develop a better understanding of the role and value of multi-modality in vision and language. We seek to create a venue to encourage collaboration between different sub-fields, and help establish new research directions and collaborations that we believe will sustain machine learning research for years to come.

[Workshop Homepage](https://srvk.github.io/how2-challenge/)

Schedule

09:00 AM	**The How2 Database and Challenge**	*Specia, Sanabria*
10:15 AM	**Coffee Break**	
10:30 AM	**Forcing Vision + Language Models To Actually See, Not Just Talk**	*Parikh*
11:00 AM	**TBA (Bernt Schiele)**	
11:30 AM	**TBA**	
12:00 PM	**Lunch (on your own)**	
01:30 PM	**Multi-agent communication from raw perceptual input: what works, what doesn't and what's next**	*Lazaridou*
02:00 PM	**Overcoming Bias in Captioning Models**	*Hendricks*
02:30 PM	**TBA**	
03:00 PM	**Poster Session and Coffee**	*Sanabria, Srinivasan, Raunak, Zhou, Kundu, Patel, Specia, Choe, Belova*
04:30 PM	**TBA**	*Jin*
05:00 PM	**New Directions for Vision & Language**	*Metze, Palaskar*

Abstracts (2):

Abstract 7: **Multi-agent communication from raw perceptual input: what works, what doesn't and what's next in The How2 Challenge: New Tasks for Vision & Language**, *Lazaridou* 01:30 PM

Multi-agent communication has been traditionally used as a computational tool to study language evolution. Recently, it has attracted attention also as a means to achieve better coordination among multiple interacting agents in complex environments. However, is it easy to scale previous research in the new deep learning era? In this talk, I will first give a brief overview of some of the previous approaches that study emergent communication in cases where agents are given as input symbolic data. I will then move on to presenting some of the challenges that agents face when are placed in grounded environments where they receive raw perceptual information and how environmental or pre-linguistic conditions affect the nature of the communication protocols that they learn. Finally, I will discuss some potential remedies that are inspired from human language and communication.

Abstract 8: **Overcoming Bias in Captioning Models in The How2 Challenge: New Tasks for Vision & Language**, *Hendricks* 02:00 PM

Most machine learning models are known to capture and exploit bias. While this can be beneficial for many classification tasks (e.g., it might be easier to recognize a computer mouse given the context of a computer and a desk), exploiting bias can also lead to incorrect predictions. In this talk, I will first consider how over-reliance on bias might lead to incorrect predictions in a scenario where is inappropriate to rely on bias: gender prediction in image captioning. I will present the Equalizer model which more accurately describes people and their gender by considering appropriate gender evidence. Next, I will consider how bias is related to hallucination, an interesting error mode in image captioning. I will present a metric designed to measure hallucination and consider questions like what causes hallucination, which models are prone to hallucination, and do current metrics accurately capture hallucination?

Adaptive and Multitask Learning: Algorithms & Systems

Maruan Al-Shedivat, Anthony Platanios, Otilia Stretcu, Jacob Andreas, Ameet Talwalkar, Rich Caruana, Tom Mitchell, Eric Xing

Seaside Ballroom, Sat Jun 15, 08:30 AM

Driven by progress in deep learning, the machine learning community is now able to tackle increasingly more complex problems—ranging from multi-modal reasoning to dexterous robotic manipulation—all of which typically involve solving nontrivial combinations of tasks. Thus, designing adaptive models and algorithms that can efficiently learn, master, and combine multiple tasks is the next frontier. AMTL workshop aims to bring together machine learning researchers from areas ranging from theory to applications and systems, to explore and discuss:

* advantages, disadvantages, and applicability of different approaches to learning in multitask settings,
* formal or intuitive connections between methods developed for different problems that help better understand the landscape of multitask learning techniques and inspire technique transfer between research lines,
* fundamental challenges and open questions that the community needs to tackle for the field to move forward.

Webpage: www.amtl-workshop.org

Schedule

08:30 AM	**Opening Remarks**

08:40 AM	**Building and Structuring Training Sets for Multi-Task Learning (Alex Ratner)**	*Ratner*
09:10 AM	**Multi-Task and Meta Reinforcement Learning (Chelsea Finn)**	*Finn*
09:40 AM	**Contributed Talk 1**	
09:55 AM	**Contributed Talk 2**	
10:10 AM	**Tricks of the Trade 1**	*Caruana*
10:25 AM	**Coffee Break**	
11:00 AM	**Poster Session**	
12:00 PM	**Lunch Break**	
01:45 PM	**Invited Talk (Massimiliano Pontil)**	
02:15 PM	**ARUBA: Efficient and Adaptive Meta-Learning with Provable Guarantees (Ameet Talwalkar)**	*Talwalkar*
02:45 PM	**Tricks of Trade 2**	*Caruana*
03:00 PM	**Coffee Break**	
03:30 PM	**Multi-Task Learning in the Wilderness (Andrej Karpathy)**	
04:00 PM	**Invited Talk (Justin Basilico)**	
04:30 PM	**Contributed Talk 3**	
04:45 PM	**Contributed Talk 4**	
05:00 PM	**Invited Talk (Hannaneh Hajishirzi)**	
05:30 PM	**Closing Remarks**	