

ICML 2017

THE 34TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING



WORKSHOP PROGRAM

Thu Aug 10th - Fri Aug 11th

WORKSHOP CHAIRS: ANIMA ANANDKUMAR
FEI SHA

ICML @ SYDNEY

INTERNATIONAL CONFERENCE ON MACHINE LEARNING

Workshop organizers make last-minute changes to their schedule. Download this document again to get the latest changes, or use the [ICML mobile application](#).

Schedule Highlights

Aug. 10, 2017

C4.1, **Lifelong Learning: A Reinforcement Learning Approach**

Chandar, Ravindran, Mankowitz, Mannor, Zahavy

C4.10, **ICML Workshop on Machine Learning for Autonomous Vehicles 2017**

Li, Urtasun, Gray, Savarese

C4.11, **Learning to Generate Natural Language**

Miao, Ling, Wen, Cao, Gerz, Blunsom, Dyer

C4.3, **Workshop on Visualization for Deep Learning**

Jiang, Canny, Chau, Fan, Zhu

C4.4, **Workshop on Computational Biology**

Pe'er, Leslie, Engelhardt, Azizi, Prabhakaran, Kshirsagar, Carr

C4.5, **Principled Approaches to Deep Learning**

Pronobis, Gens, Kakade, Domingos

C4.6, **Video Games and Machine Learning**

Synnaeve, Togelius, Schaul, Vinyals, Usunier

C4.7, **ML on a budget: IoT, Mobile and other tiny-ML applications**

Varma, Saligrama, Jain

C4.8, **Workshop on Human Interpretability in Machine Learning (WHI)**

Varshney, Weller, Kim, Malioutov

C4.9, **Automatic Machine Learning (AutoML 2017)**

Vanschoren, Garnett

Parkside 1, **Implicit Generative Models**

Ranganath, Goodfellow, Tran, Blei, Lakshminarayanan, Mohamed

Aug. 11, 2017

C4.1, **Time Series Workshop**

Kuznetsov, Liu, Yang, Yu

C4.10, **Reproducibility in Machine Learning Research**

Ke, Goyal, Lamb, Pineau, Bengio, Bengio

C4.11, **Interactive Machine Learning and Semantic Information Retrieval**

Glowacka, Buntine, Myllymaki

C4.3, **Machine Learning in Speech and Language Processing**

Livescu, Sainath, Lu, Ragni

C4.4, **Private and Secure Machine Learning**

Honkela, Shimizu, Kaski

C4.5, **Deep Structured Prediction**

Augenstein, Chang, Chechik, Huang, Torres Martins, Meshi, Schwing, Miao

C4.6, **Picky Learners: Choosing Alternative Ways to Process Data.**

Cortes, Chaudhuri, DeSalvo, Zhang, Zhang

C4.7, **Reliable Machine Learning in the Wild**

Hadfield-Menell, Steinhart, Weller, Milli

C4.8, **Human in the Loop Machine Learning**

Nock, Ong

C4.9, **Machine Learning for Music Discovery**

Schmidt, Nieto, Gouyon, Lanckriet

Parkside 1, **Reinforcement Learning Workshop**

Precup, Ravindran, Bacon

	Sergey Levine:	
03:50 PM	Self-supervision as a path to lifelong learning	<i>Levine</i>

04:30 PM	Lifelong Learning - Panel Discussion	<i>Levine, Pineau, Ravindran, Rusu</i>
----------	---	--

Lifelong Learning: A Reinforcement Learning Approach

Sarath Chandar, Balaraman Ravindran, Daniel J. Mankowitz, Shie Mannor, Tom Zahavy

C4.1, Thu Aug 10, 08:30 AM

One of the most challenging and open problems in Artificial Intelligence (AI) is that of Lifelong Learning:

“Lifelong Learning is the continued learning of tasks, from one or more domains, over the course of a lifetime, by a lifelong learning system. A lifelong learning system efficiently and effectively (1) retains the knowledge it has learned; (2) selectively transfers knowledge to learn new tasks; and (3) ensures the effective and efficient interaction between (1) and (2).”

Lifelong learning is still in its infancy. Many issues currently exist such as learning general representations, catastrophic forgetting, efficient knowledge retention mechanisms and hierarchical abstractions. Much work has been done in the Reinforcement Learning (RL) community to tackle different elements of lifelong learning. Active research topics include hierarchical abstractions, transfer learning, multi-task learning and curriculum learning. With the emergence of powerful function approximators such as in Deep Learning, we feel that now is a perfect time to provide a forum to discuss ways to move forward and provide a truly general lifelong learning framework, using RL-based algorithms, with more rigour than ever before. This workshop will endeavour to promote interaction between researchers working on the different elements of lifelong learning to try and find a synergy between the various techniques.

Schedule

08:30 AM	Introduction and Overview	
----------	----------------------------------	--

08:40 AM	Marc G. Bellemare: The role of density models in reinforcement learning	<i>Bellemare</i>
----------	--	------------------

09:20 AM	Poster Spotlights	
----------	--------------------------	--

10:00 AM	Poster Session + break - I	
----------	-----------------------------------	--

10:30 AM	Joelle Pineau: A few modest insights from my lifelong learning	<i>Pineau</i>
----------	---	---------------

11:10 AM	Andrei Rusu: Sequential Learning in Complex Environments	<i>Rusu</i>
----------	---	-------------

12:00 PM	Lunch	
----------	--------------	--

02:00 PM	Contributed Talks	
----------	--------------------------	--

02:45 PM	Poster Session + break - II	
----------	------------------------------------	--

ICML Workshop on Machine Learning for Autonomous Vehicles 2017

Li Erran Li, Raquel Urtasun, Andrew Gray, Silvio Savarese

C4.10, Thu Aug 10, 08:30 AM

Although dramatic progress has been made in the field of autonomous driving, there are many major challenges in achieving full-autonomy. For example, how to make perception accurate and robust to accomplish safe autonomous driving? How to reliably track cars, pedestrians, and cyclists? How to learn long term driving strategies (known as driving policies) so that autonomous vehicles can be equipped with adaptive human negotiation skills when merging, overtaking and giving way, etc? How to achieve near-zero fatality?

These complex challenges associated with autonomy in physical world naturally suggest that we take a machine learning approach. Deep learning and computer vision have found many real-world applications such as face tagging. However, perception for autonomous driving has a unique set of requirements such as safety and explainability. Autonomous vehicles need to choose actions, e.g. steering commands which will affect the subsequent inputs (driving scenes) encountered. This setting is well-suited to apply reinforcement learning to determine the best actions to take. Many autonomous driving tasks such as perception and tracking requires large data sets of labeled examples to learn rich and high-performance visual representation. However, the progress is hampered by the sheer expenses of human labelling needed. Naturally we would like to employ unsupervised learning, transfer learning leveraging simulators, and techniques can learn efficiently. The goal of this workshop is to bring together researchers and practitioners from in the field of autonomous driving to address core challenges with machine learning. These challenges include, but are not limited to

accurate and efficient pedestrian detection, pedestrian intent detection, machine learning for object tracking, unsupervised representation learning for autonomous driving, deep reinforcement learning for learning driving policies, cross-modal and simulator to real-world transfer learning, scene classification, real-time perception and prediction of traffic scenes, uncertainty propagation in deep neural networks, efficient inference with deep neural networks

The workshop will include invited speakers, panels, presentations of accepted papers and posters. We invite papers in the form of short, long and position papers to address the core challenges mentioned above. We encourage researchers and practitioners on self-driving cars, transportation systems and ride-sharing platforms to participate. Since this is a topic of broad and current interest, we expect at least 200 participants from leading university researchers, auto-companies and ride-sharing companies.

Schedule

08:20 AM	Opening Remarks: Drew Gray and Li Erran Li (Uber ATG)
08:30 AM	Carl Wellington, Uber ATG
09:00 AM	Efficient deep neural networks for perception in autonomous driving (Jose M. Alvarez, TRI)
09:30 AM	Visual 3D Scene Understanding and Prediction for ADAS (Manmohan Chandraker, NEC Labs)
10:00 AM	Coffee
10:30 AM	2 x 15 Contributed Talks on Datasets and Occupancy Maps
11:00 AM	Beyond Hand Labeling: Simulation and Self-Supervision for Self-Driving Cars (Matt Johnson, University of Michigan)
11:30 AM	Learning Affordance for Autonomous Driving (JianXiong Xiao, AutoX)
02:30 PM	2 x 15 Contributed Talks on Reinforcement Learning
03:00 PM	Coffee and Posters
03:30 PM	Min Sun, National Tsing Hua University: Assessing Risk and Adapting Changes on the Road
04:00 PM	Are we over-engineering autonomous vehicles? (Amar Shah, University of Cambridge)
04:30 PM	2 x 5 min Lightening Talks
04:40 PM	Panel Discussion (Jose M. Alvarez, Manmohan Chandraker, Matt Johnson, Min Sun, Carl Wellington)
05:25 PM	Closing Remarks: Li Erran Li and Drew Gray (Uber ATG)

Abstracts (9):

Abstract 3: **Efficient deep neural networks for perception in autonomous driving (Jose M. Alvarez, TRI) in ICML Workshop on**

Machine Learning for Autonomous Vehicles 2017, 09:00 AM

Abstract

Convolutional neural networks have achieved impressive success in many tasks in computer vision such as image classification, object detection / recognition or semantic segmentation. While these networks have proven effective in all these applications, they come at a high memory and computational cost, thus not feasible for applications where power and computational resources are limited. In addition, the process to train the network reduces productivity as it not only requires large computer servers but also takes a significant amount of time (several weeks) with the additional cost of engineering the architecture. In this talk, I first introduce our efficient architecture based on filter-compositions and then, a novel approach to jointly learn the architecture and explicitly account for compression during the training process. Our results show that we can learn much more compact models and significantly reduce training and inference time.

Bio:

Dr. Jose Alvarez is a senior research scientist at Toyota Research Institute. His main research interests are in developing robust and efficient deep learning algorithms for perception with focus on autonomous vehicles. Previously, he was a researcher at Data61 / CSIRO (formerly NICTA), a Postdoctoral researcher at the Courant Institute of Mathematical Science, New York University, and visiting scholar at University of Amsterdam and Group Research Electronics at Volkswagen. Dr. Alvarez graduated in 2012 and he was awarded the best Ph.D. Thesis award. Dr. Alvarez serves as associate editor for IEEE Trans. on Intelligent Transportation Systems.

Abstract 4: Visual 3D Scene Understanding and Prediction for ADAS (Manmohan Chandraker, NEC Labs) in ICML Workshop on Machine Learning for Autonomous Vehicles 2017, 09:30 AM

Abstract:

Modern advanced driver assistance systems (ADAS) rely on a range of sensors including radar, ultrasound, LIDAR and cameras. Active sensors have found applications in detecting traffic participants (TPs) such as cars or pedestrians and scene elements (SEs) such as roads. However, camera-based systems have the potential to achieve or augment these capabilities at a much lower cost, while allowing new ones such as determination of TP and SE semantics as well as their interactions in complex traffic scenes.

In this talk, we present several technical advances for vision-based ADAS. A common theme is to overcome the challenges posed by lack of large-scale annotations in deep learning frameworks. We introduce approaches to correspondence estimation that are trained on purely synthetic data but adapt well to real data at test-time. We introduce object detectors that are light enough for ADAS, trained with knowledge distillation to retain accuracies of deeper architectures. Our semantic segmentation methods are trained on weak supervision that requires only a tenth of conventional annotation time. We propose methods for 3D reconstruction that use deep supervision to recover fine TP part locations while relying on purely synthetic 3D CAD models. We develop deep learning frameworks for multi-target tracking, as well as occlusion-reasoning in TP localization and SE layout estimation. Finally, we present a framework for TP behavior prediction in complex traffic scenes that accounts for TP-TP and TP-SE interactions. Our approach allows prediction of diverse multimodal outcomes and aims to account for long-term strategic behaviors in complex scenes.

Bio:

Manmohan Chandraker is an assistant professor at the CSE department of the University of California, San Diego and leads the computer vision research effort at NEC Labs America in Cupertino. He received a B.Tech. in Electrical Engineering at the Indian Institute of Technology, Bombay and a PhD in Computer Science at the University of California, San Diego. His personal research interests are 3D scene understanding and reconstruction, with applications to autonomous driving and human-computer interfaces. His works have received the Marr Prize Honorable Mention for Best Paper at ICCV 2007, the 2009 CSE Dissertation Award for Best Thesis at UCSD, a PAMI special issue on best papers of CVPR 2011 and the Best Paper Award at CVPR 2014.

Abstract 6: 2 x 15 Contributed Talks on Datasets and Occupancy Maps in ICML Workshop on Machine Learning for Autonomous Vehicles 2017, 10:30 AM

Jonathan Binas, Daniel Neil, Shih-Chii Liu, Tobi Delbruck, DDD17: End-To-End DAVIS Driving Dataset

Ransalu Senanayake and Fabio Ramos, Bayesian Hilbert Maps for Continuous Occupancy Mapping in Dynamic Environments

Abstract 7: Beyond Hand Labeling: Simulation and Self-Supervision for Self-Driving Cars (Matt Johnson, University of Michigan) in ICML Workshop on Machine Learning for Autonomous Vehicles 2017, 11:00 AM

Self-driving cars now deliver vast amounts of sensor data from large unstructured environments. In attempting to process and interpret this data there are many unique challenges in bridging the gap between prerecorded data sets and the field. This talk will present recent work addressing the application of deep learning techniques to robotic perception. We focus on solutions to several pervasive problems when attempting to deploy such techniques on fielded robotic systems. The themes of the talk revolve around alternatives to gathering and curating data sets for training. Are there ways of avoiding the labor-intensive human labeling required for supervised learning? These questions give rise to several lines of research based around self-supervision, adversarial learning, and simulation. We will show how these approaches applied to self-driving car problems have great potential to change the way we train, test, and validate machine learning-based systems.

Bio:

Matthew Johnson-Roberson is Assistant Professor of Engineering in the Department of Naval Architecture & Marine Engineering and the Department of Electrical Engineering and Computer Science at the University of Michigan. He received a PhD from the University of Sydney in 2010. He has held prior postdoctoral appointments with the Centre for Autonomous Systems - CAS at KTH Royal Institute of Technology in Stockholm and the Australian Centre for Field Robotics at the University of Sydney. He is a recipient of the NSF CAREER award (2015). He has worked in robotic perception since the first DARPA grand challenge and his group focuses on enabling robots to better see and understand their environment.

Abstract 8: Learning Affordance for Autonomous Driving (JianXiong Xiao, AutoX) in ICML Workshop on Machine Learning for Autonomous Vehicles 2017, 11:30 AM

Today, there are two major paradigms for vision-based autonomous driving systems: mediated perception approaches that parse an entire scene to make a driving decision, and behavior reflex approaches that directly map an input image to a driving action by a regressor. In this paper, we propose a third paradigm: a direct perception based approach to estimate the affordance for driving. We propose to map an input image to a small number of key perception indicators that directly relate to the affordance of a road/traffic state for driving. Our representation provides a set of compact yet complete descriptions of the scene to enable a simple controller to drive autonomously. Falling in between the two extremes of mediated perception and behavior reflex, we argue that our direct perception representation provides the right level of abstraction. We evaluate our approach in a virtual racing game as well as real world driving and show that our model can work well to drive a car in a very diverse set of virtual and realistic environments.

Jianxiong Xiao (a.k.a., Professor X) is the Founder and CEO of AutoX Inc., a high-tech startup working on A.I. software solution for self-driving vehicles. AutoX's mission is to democratize autonomy and make autonomous driving universally accessible to everyone. Its innovative camera-first self-driving solution amounts to only a tiny fraction of the cost of traditional LiDar-based approaches. Dr. Xiao has over ten years of research and engineering experience in Computer Vision, Autonomous Driving, and Robotics. In particular, he is a pioneer in the fields of 3D Deep Learning, RGB-D Recognition and Mapping, Big Data, Large-scale Crowdsourcing, and Deep Learning for Robotics. Jianxiong received a BEng. and MPhil. in Computer Science from the Hong Kong University of Science and Technology in 2009. He received his Ph.D. from the Computer Science and Artificial Intelligence Laboratory (CSAIL) at the Massachusetts Institute of Technology (MIT) in 2013. And he was an Assistant Professor at Princeton University and the founding director of the Princeton Computer Vision and Robotics Labs from 2013 to 2016. His work has received the Best Student Paper Award at the European Conference on Computer Vision (ECCV) in 2012 and the Google Research Best Papers Award for 2012, and has appeared in the popular press. He was awarded the Google U.S./Canada Fellowship in Computer Vision in 2012, the MIT CSW Best Research Award in 2011, NSF/Intel VEC Research Award in 2016, and two Google Faculty Awards in 2014 and in 2015 respectively. He co-lead the MIT+Princeton joint team to participate in the Amazon Picking Challenge in 2016, and won the 3rd and 4th place worldwide. More information can be found at: <http://www.jianxiongxiao.com>.

Abstract 9: 2 x 15 Contributed Talks on Reinforcement Learning in ICML Workshop on Machine Learning for Autonomous Vehicles 2017, 02:30 PM

David Isele, Akansel Cosgun, To Go or Not to Go: A Case for Q-Learning at Unsignalized Intersections

Tomoki Nishi, Prashant Doshi, Danil Prokhorov, Freeway Merging in Congested Traffic based on Multipolicy Decision Making with Passive Actor Critic

Abstract 11: Min Sun, National Tsing Hua University: Assessing Risk and Adapting Changes on the Road in ICML Workshop on Machine Learning for Autonomous Vehicles 2017, 03:30 PM

It is critical for a self-driving car in the wild to assess risk and adapt to changes on the road. In this talk, we will first go over our proposed

accident anticipation method which is tested on a large dataset consisting of real-world accident videos. Then, we will present our latest ICCV paper about how to adapt a semantic segmentation model across 4 cities in three continents.

Bio: Min Sun is an assistant professor at National Tsing Hua University in Taiwan. Before that, he was a postdoctoral researcher at Washington University in Seattle and he graduated from the University of Michigan with a Ph.D. degree in EE: System. He also won the best paper award of 3dRR in 2007 and best paper award of CVGIP in 2015 and 2016.

Abstract 12: **Are we over-engineering autonomous vehicles? (Amar Shah, University of Cambridge) in ICML Workshop on Machine Learning for Autonomous Vehicles 2017**, 04:00 PM

Research, media and corporate interest in autonomous road vehicles has exploded. Large industrial labs are hiring large teams to engineer solutions to the problem. As machine learners, I conjecture that it will be big data and cutting edge end-to-end reinforcement learning techniques that will have the best chance to achieve autonomy over the next decade and not the vastly hand engineered, rule based approaches of current teams. A crucial component of self driving car systems which is overlooked is model uncertainty. I claim that model uncertainty is as important as model accuracy, and encourage the community to pursue research in this direction. Good decisions require good predictions and well calibrated uncertainty estimates around those predictions.

Abstract 13: **2 x 5 min Lightning Talks in ICML Workshop on Machine Learning for Autonomous Vehicles 2017**, 04:30 PM

1. Kangwook Lee, Hoon Kim, Changho Suh, Crash To Not Crash: Playing Video Games To Predict Vehicle Collisions
2. Edward Schwalb, Bernhard Bieder, Daniel Wiesenhütter, What Makes It Testable? Conceptual Model for Safety Quantification

Ahmad El Sallab, Mahmoud Saeed, Omar Abdel Tawab, Mohammed Abdou, Meta learning Framework for Automated Driving

Learning to Generate Natural Language

Yishu Miao, Wang Ling, Tsung-Hsien Wen, Kris Cao, Daniela Gerz, Phil Blunsom, Chris Dyer

C4.11, Thu Aug 10, 08:30 AM

Research on natural language generation is rapidly growing due to the increasing demand for human-machine communication in natural language. This workshop aims to promote the discussion, exchange, and dissemination of ideas on the topic of text generation, touching several important aspects in this modality: learning schemes and evaluation, model design and structures, advanced decoding strategies, and natural language generation applications. This workshop aims to be a venue for the exchange of ideas regarding data-driven machine learning approaches for text generation, including mainstream tasks such as dialogue generation, instruction generation, and summarization; and for establishing new directions and ideas with potential for impact in the fields of machine learning, deep learning, and NLP.

Schedule

08:30 AM	Tim Baldwin: Learning to Label Documents
<hr/>	
09:15 AM	Dani Yogatama
<hr/>	
10:00 AM	Coffee Break & Poster session 1
<hr/>	
10:30 AM	Andre Martins: Beyond Softmax: Sparsemax, Constrained Softmax, Differentiable Easy-First
<hr/>	
11:15 AM	Spotlight Paper Presentation
<hr/>	
12:00 PM	Lunch Break & Poster session 2

Abstracts (6):

Abstract 1: **Tim Baldwin: Learning to Label Documents in Learning to Generate Natural Language**, 08:30 AM

Document labelling (incl. multimodal objects) is widely used in NLP and ML, in forms include classic document categorisation, single-document summarisation and image captioning. In this talk, I consider the question of what, intrinsically, is a suitable "label" for a given document type, and then discuss some recent work on automatically generating multimodal labels for textual topics.

Abstract 2: **Dani Yogatama in Learning to Generate Natural Language**, 09:15 AM

Invited Talk 2

Abstract 3: **Coffee Break & Poster session 1 in Learning to Generate Natural Language**, 10:00 AM

Coffee Break & Poster session

Abstract 4: **Andre Martins: Beyond Softmax: Sparsemax, Constrained Softmax, Differentiable Easy-First in Learning to Generate Natural Language**, 10:30 AM

In the first part of the talk, I will propose sparsemax, a new activation function similar to the traditional softmax, but able to output sparse probabilities. After deriving its properties, I will show how its Jacobian can be efficiently computed, enabling its use in a network trained with backpropagation. Then, I will propose a new smooth and convex loss function which is the sparsemax analogue of the logistic loss, revealing an unexpected connection with the Huber classification loss. I will show promising empirical results in multi-label classification problems and in attention-based neural networks for natural language inference.

In the second part, I will introduce constrained softmax, another activation function that allows imposing upper bound constraints on attention probabilities. Based on this activation, I will introduce a novel neural end-to-end differentiable easy-first decoder that learns to solve sequence tagging tasks in a flexible order. The decoder iteratively updates a "sketch" of the predictions over the sequence. The proposed models compare favourably to BiLSTM taggers on three sequence tagging tasks.

This is joint work with Ramon Astudillo and Julia Kreutzer.

Abstract 5: **Spotlight Paper Presentation in Learning to Generate Natural Language**, 11:15 AM

Workshop Paper Presentation

Abstract 6: **Lunch Break & Poster session 2 in Learning to Generate Natural Language**, 12:00 PM

Lunch Break & Poster session

Workshop on Visualization for Deep Learning

Biye Jiang, John Canny, Polo Chau, Xiangmin Fan, Junyan Zhu

C4.3, Thu Aug 10, 08:30 AM

Deep networks have had profound impact across machine learning research and in many application areas. DNNs are complex to design and train. They are non-linear systems that almost always have many local optima and are often sensitive to training parameter settings and initial state. Systematic optimization of structure and hyperparameters is possible e.g. with Bayesian optimization, but hampered by the expense of training each design on realistic datasets. Exploration is still ongoing for best design principles. We argue that visualization can play an essential role in understanding DNNs and in developing new design principles. With rich tools for visual exploration of networks during training and inference, one should be able to form closer ties between theory and practice: validating expected behaviors, and exposing the unexpected which can lead to new insights. With the rise of generative modeling and reinforcement learning, more interesting directions like understanding and visualization of generative models, visual explanation for driving policy could be explored as well.

As the second edition of this workshop, we are proposing changes based on the lessons we learned last year. We would like to organize a few domain specific tutorials, and panel discussions. We do think machine learning researchers need a lot of tutorials and advice from the visualization/HCI community and vice versa. Many audience in our workshop last year also suggested that more discussion can greatly help us better define such interdisciplinary area.

Schedule

08:30 AM	Opening remark
08:40 AM	Becoming friends with every pixel, Phillip Isola (UC Berkeley)
09:20 AM	SmoothGrad: removing noise by adding noise Daniel Smilkov, Nikhil Thorat, Been Kim, Fernanda Viegas, Martin M Wattenberg

09:40 AM	Towards Visual Explanations for Convolutional Neural Networks via Input Resampling, Benjamin J Lengerich, Sandeep Konam, Eric Xing, Stephanie Rosenthal, Manuela Veloso
10:00 AM	Coffee Breaks and Poster session 1
10:30 AM	Understanding Generative Models in Google Brain Magenta, Cinjon Resnick (Google)
11:00 AM	Deep saliency: What is learnt by a deep network about saliency? Sen He, Nicolas Pugeault
11:15 AM	Self-supervised attention for Deep Learning explanations, Nathan Hodas, (PNL)
11:45 AM	Skip-Frame Embeddings for Feature Adaptation and Visualization, Zain Shah
12:00 PM	Lunch Breaks
02:00 PM	Quantifying the Interpretability of Deep Visual Representations, Bolei Zhou (MIT)
02:40 PM	Visualizing Feature Maps in Deep Neural Networks using DeepResolve - A Genomics Case Study, Ge Liu, David Gifford
03:30 PM	Visual Explanations from Deep Networks, Dhruv Batra (Georgia Tech and Facebook AI Research)
04:00 PM	Evolutionary Visual Analysis of Deep Neural Networks, Wen Zhong, Cong Xie, Yuan Zhong, Yang Wang, Wei Xu, Shenghui Cheng, Klaus Mueller
04:20 PM	ActiVis: Visual Exploration of Industry-Scale Deep Neural Network Models, Pierre Andrews (Facebook)
04:50 PM	Brainstorming on deep learning visualization techniques and tools

Abstracts (7):

Abstract 3: SmoothGrad: removing noise by adding noise Daniel Smilkov, Nikhil Thorat, Been Kim, Fernanda Viegas, Martin M Wattenberg in Workshop on Visualization for Deep Learning, 09:20 AM

Explaining the output of a deep network remains a challenge. In the case of an image classifier, one type of explanation is to identify pixels that strongly influence the final decision. A starting point for this strategy is the gradient of the class score function with respect to the input image. This gradient can be interpreted as a sensitivity map, and there are several techniques that elaborate on this basic idea. This paper makes two contributions: it introduces SMOOTHGRAD, a simple method that can help visually sharpen gradient-based sensitivity maps, and it discusses lessons in the visualization of these maps. We publish the code for our experiments and a website with our results.

Abstract 4: Towards Visual Explanations for Convolutional Neural Networks via Input Resampling, Benjamin J Lengerich, Sandeep Konam, Eric Xing, Stephanie Rosenthal, Manuela Veloso in Workshop on Visualization for Deep Learning, 09:40 AM

The predictive power of neural networks often costs model interpretability. Several techniques have been developed for explaining model outputs in terms of input features; however, it is difficult to translate such interpretations into actionable insight. Here, we propose a framework to analyze predictions in terms of the model's internal features by inspecting information flow through the network. Given a trained network and a test image, we select neurons by two metrics, both measured over a set of images created by perturbations to the input image: (1) magnitude of the correlation between the neuron activation and the network output and (2) precision of the neuron activation. We show that the former metric selects neurons that exert large influence over the network output while the latter metric selects neurons that activate on generalizable features. By comparing the sets of neurons selected by these two metrics, our framework offers a way to investigate the internal attention mechanisms of convolutional neural networks.

Abstract 7: Deep saliency: What is learnt by a deep network about saliency? Sen He, Nicolas Pugeault in Workshop on Visualization for Deep Learning, 11:00 AM

Deep convolutional neural networks have achieved impressive performance on a broad range of problems, beating prior art on established benchmarks, but it often remains unclear what are the representations learnt by those systems and how they achieve such performance.

This article examines the specific problem of saliency detection, where benchmarks are currently dominated by CNN-based approaches, and investigates the properties of the learnt representation by visualizing the artificial neurons' receptive fields. We demonstrate that fine tuning a pre-trained network on the saliency detection task lead to a profound transformation of the network's deeper layers. Moreover we argue that this transformation leads to the emergence of receptive fields conceptually similar to the centre-surround filters hypothesized by early research on visual saliency.

Abstract 9: Skip-Frame Embeddings for Feature Adaptation and Visualization, Zain Shah in Workshop on Visualization for Deep Learning, 11:45 AM

We present an unsupervised method for visualizing the generalization and adaptation capabilities of pre-trained features on video. Like the skip-grams method for unsupervised learning of word vector representations, we exploit temporal continuity in the target media, namely that neighboring frames are qualitatively similar. By enforcing this continuity in the adapted feature space we can adapt features to a new target task, like house price prediction, without supervision. The domain-specific embeddings can be easily visualized for qualitative introspection and evaluation.

Abstract 12: Visualizing Feature Maps in Deep Neural Networks using DeepResolve - A Genomics Case Study, Ge Liu, David Gifford in Workshop on Visualization for Deep Learning, 02:40 PM

Although many powerful visualization tools have been developed to interpret neural network decisions in input space, methods to interpret feature map space remain limited. Most existing tools examine a network's response to a specific input sample and thus are locally faithful to that sample. We introduce DeepResolve, a gradient ascent based method that visualizes intermediate layer feature maps in an input independent manner. We examine DeepResolve's capability to 1) discover network linear and non-linear combinatorial logic and summarize overall knowledge of a class, 2) reveal key features for a target class, 3) assess a network's activeness in pattern learning and network's vulnerability in feature space, and 4) analyze multi-task class similarity at high resolution. We demonstrate the value of DeepResolve on synthetic and experimental genomic datasets, and DeepResolve reveals biologically interesting observations from the experimental data.

Abstract 14: Evolutionary Visual Analysis of Deep Neural Networks, Wen Zhong, Cong Xie, Yuan Zhong, Yang Wang, Wei Xu, Shenghui Cheng, Klaus Mueller in Workshop on Visualization for Deep Learning, 04:00 PM

Recently, deep learning visualization gained a lot of attentions for understanding deep neural networks. However, there is a missing focus on the visualization of deep model training process. To bridge the gap, in this paper, we firstly define a discriminability metric to evaluate neuron evolution and a density metric to investigate output feature maps. Based on these metrics, a level-of-detail visual analytics framework is proposed to locally and globally inspect the evolution of deep neural networks. Finally, we demonstrate the effectiveness of our system with two real world case studies.

Abstract 15: ActiVis: Visual Exploration of Industry-Scale Deep Neural Network Models, Pierre Andrews (Facebook) in Workshop on Visualization for Deep Learning, 04:20 PM

While deep learning models have achieved state-of-the-art accuracies for many prediction tasks, understanding these models remains a challenge. Despite the recent interest in developing visual tools to help users interpret deep learning models, the complexity and wide variety of models deployed in industry, and the large-scale datasets that they used, pose unique design challenges that are inadequately addressed by existing work. Through participatory design sessions with over 15 researchers and engineers at Facebook, we have developed, deployed, and iteratively improved ActiVis, an interactive visualization system for interpreting large-scale deep learning models and results. By tightly integrating multiple coordinated views, such as a computation graph overview of the model architecture, and a neuron activation view for pattern discovery and comparison, users can explore complex deep neural network models at both the instance- and subset-level. ActiVis has been deployed on Facebook's machine learning platform. We present case studies with Facebook researchers and engineers, and usage scenarios of how ActiVis may work with different models.

Workshop on Computational Biology

Dana Pe'er, Christina Leslie, Barbara Engelhardt, Elham Azizi, Sandhya Prabhakaran, Meghana Kshirsagar, Ambrose Carr

C4.4, Thu Aug 10, 08:30 AM

The workshop will showcase recent research in the field of Computational Biology. There has been significant development in genomic sequencing techniques as well as imaging technologies that not only generate huge amounts of data but provide unprecedented levels of resolution, that of a single cell and even subcellular resolution. This availability of high dimensional data, at multiple spatial and temporal resolutions and capturing several perspectives of biological phenomena has made machine learning methods increasingly relevant for computational analysis 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 at bringing in researchers working at the intersection of Machine Learning and Biology to present recent advances and open questions in computational biology to the ICML community.

Schedule

08:45 AM	Opening Remarks
08:50 AM	Stability and Aggregation of Experimental Results
09:30 AM	Spotlight Presentations
10:00 AM	Coffee Break
10:30 AM	Deep learning approaches to impute, integrate and interpret regulatory genomic data
11:10 AM	Uncovering the gene usage of human tissue cells with joint factorized embeddings
11:30 AM	Dilated Convolutions for Modeling Long-Distance Genomic Dependencies
12:00 PM	Poster Session I
02:00 PM	Reasoning from "Messy" Clinical Time Series for Individualizing Care -- Suchi Saria
02:40 PM	Ask the doctor – Improving drug sensitivity predictions through active expert knowledge elicitation
03:00 PM	Poster Session II
04:15 PM	Contrastive Principal Component Analysis
04:35 PM	Closing Remarks and Awards

Abstracts (1):

Abstract 2: Stability and Aggregation of Experimental Results in Workshop on Computational Biology, 08:50 AM

Spearman's correlation measures the association between ranked lists. Given a set of ranked lists, we study two tasks: aggregating the set of ranks into one single ranked list, and computing the agreement of the lists as we traverse it. Applications include the analysis of the stability of feature selection and integration of various sources of information. This is illustrated with two examples respectively: We study the stability of identifying variations in GWAS by considering replication studies. In another study, we aggregate genomic distance, 3D associations, and literature information to find promising disease associated variations. It turns out that these problems can be tackled by considering a multivariate Spearman's correlation.

Principled Approaches to Deep Learning

Andrzej Pronobis, Robert Gens, Sham Kakade, Pedro Domingos

C4.5, Thu Aug 10, 08:30 AM

The recent advancements in deep learning have revolutionized the field of machine learning, enabling unparalleled performance and many new real-world applications. Yet, the developments that led to this success have often been driven by empirical studies, and little is known about the theory behind some of the most successful approaches. While theoretically well-founded deep learning architectures had been proposed in the past, they came at a price of increased complexity and reduced tractability. Recently, we have witnessed considerable interest in principled deep learning. This led to a better theoretical understanding of existing architectures as well as development of more mature deep models with solid theoretical foundations. In this workshop, we intend to review the state of those developments and provide a platform for the exchange of ideas between the theoreticians and the practitioners of the growing deep learning community. Through a series of invited talks by the experts in the field, contributed presentations, and an interactive panel discussion, the workshop will cover recent theoretical developments, provide an overview of promising and mature architectures, highlight their challenges and unique benefits, and present the most exciting recent results.

Schedule

08:30 AM	Welcome and Opening Remarks
08:45 AM	Invited Talk 1 - Sanjeev Arora
09:15 AM	Contributed Presentation 1 - Towards a Deeper Understanding of Training Quantized Neural Networks
09:30 AM	Invited Talk 2 - Surya Ganguli
10:00 AM	Coffee Break and Poster Session
10:45 AM	Invited Talk 3 - Ruslan Salakhutdinov
11:15 AM	Invited Talk 4 - Pedro Domingos
11:45 AM	Contributed Presentation 2 - LibSPN: A Library for Learning and Inference with Sum-Product Networks and TensorFlow
12:00 PM	Lunch
01:30 PM	Invited Talk 5 - Tomaso Poggio
02:00 PM	Contributed Presentation 3 - Emergence of invariance and disentangling in deep representations
02:15 PM	Invited Talk 6 - Nathan Srebro

02:45 PM	Contributed Presentation 4 - The Shattered Gradients Problem: If resnets are the answer, then what is the question?
03:00 PM	Coffee Break 2 and Poster Session
03:45 PM	Contributed Presentation 5 - Towards Deep Learning Models Resistant to Adversarial Attacks
04:00 PM	Panel Discussion
05:20 PM	Closing Remarks and Awards

Abstracts (10):

Abstract 2: **Invited Talk 1 - Sanjeev Arora in Principled Approaches to Deep Learning**, 08:45 AM

Do GANs Actually Learn the Distribution? Some Theory and Empirics

The Generative Adversarial Nets or GANs framework (Goodfellow et al'14) for learning distributions differs from older ideas such as autoencoders and deep Boltzmann machines in that it scores the generated distribution using a discriminator net, instead of a perplexity-like calculation. It appears to work well in practice, e.g., the generated images look better than older techniques. But how well do these nets learn the target distribution?

Our paper 1 (ICML'17) shows GAN training may not have good generalization properties; e.g., training may appear successful but the trained distribution may be far from target distribution in standard metrics. We show theoretically that this can happen even though the 2-person game between discriminator and generator is in near-equilibrium, where the generator appears to have "won" (with respect to natural training objectives).

Paper2 (arxiv June 26) empirically tests the whether this lack of generalization occurs in real-life training. The paper introduces a new quantitative test for diversity of a distribution based upon the famous birthday paradox. This test reveals that distributions learnt by some leading GANs techniques have fairly small support (i.e., suffer from mode collapse), which implies that they are far from the target distribution.

Paper 1: "Equilibrium and Generalization in GANs" by Arora, Ge, Liang, Ma, Zhang. (ICML 2017)

Paper 2: "Do GANs actually learn the distribution? An empirical study." by Arora and Zhang (<https://arxiv.org/abs/1706.08224>)

Abstract 3: **Contributed Presentation 1 - Towards a Deeper Understanding of Training Quantized Neural Networks in Principled Approaches to Deep Learning**, 09:15 AM

Towards a Deeper Understanding of Training Quantized Neural Networks

Hao Li, Soham De, Zheng Xu, Christoph Studer, Hanan Samet, Tom Goldstein

Training neural networks with coarsely quantized weights is a key step towards learning on embedded platforms that have limited computing resources, memory capacity, and power consumption. Numerous recent publications have studied methods for training quantized networks, but these studies have been purely experimental. In this work, we investigate the theory of training quantized neural networks. We analyze the convergence properties of commonly used quantized training methods. We also show that training algorithms that exploit high-precision representations have an important annealing property that purely quantized training methods lack, which explains many of the observed empirical differences between these types of algorithms.

Abstract 4: Invited Talk 2 - Surya Ganguli in Principled Approaches to Deep Learning, 09:30 AM

On the Beneficial Role of Dynamic Criticality and Chaos in Deep Learning

What does a generic deep function "look like" and how can we understand and exploit such knowledge to obtain practical benefits in deep learning? By combining Riemannian geometry with dynamic mean field theory, we show that generic nonlinear deep networks exhibit an order to chaos phase transition as synaptic weights vary from small to large. In the chaotic phase, deep networks acquire very high expressive power: measures of functional curvature and the ability to disentangle classification boundaries both grow exponentially with depth, but not with width. Moreover, we apply tools from free probability theory to study the propagation of error gradients through generic deep networks. We find, at the phase transition boundary between order and chaos, that not only the norms of gradients, but also angles between pairs of gradients are preserved even in infinitely deep sigmoidal networks with orthogonal weights. In contrast, ReLU networks do not enjoy such isometric propagation of gradients. In turn, this isometric propagation at the edge of chaos leads to training benefits, where very deep sigmoidal networks outperform ReLU networks, thereby pointing to a potential path to resurrecting saturating nonlinearities in deep learning.

Abstract 6: Invited Talk 3 - Ruslan Salakhutdinov in Principled Approaches to Deep Learning, 10:45 AM

Neural Map: Structured Memory for Deep Reinforcement Learning

A critical component to enabling intelligent reasoning in partially observable environments is memory. Despite this importance, Deep Reinforcement Learning (DRL) agents have so far used relatively simple memory architectures, with the main methods to overcome partial observability being either a temporal convolution over the past k frames or an LSTM layer. In this talk, we will introduce a memory system with an adaptable write operator that is customized to the sorts of 3D environments that DRL agents typically interact with. This architecture, called the Neural Map, uses a spatially structured 2D memory image to learn to store arbitrary information about the environment over long time lags. We demonstrate empirically that the Neural Map surpasses previous DRL memories on a set of challenging 2D and 3D maze environments and show that it is capable of generalizing to environments that were not seen during training.

Joint work with Emilio Parisotto

Abstract 7: Invited Talk 4 - Pedro Domingos in Principled Approaches to Deep Learning, 11:15 AM

The Sum-Product Theorem: A Foundation for Learning Tractable Deep Models

Inference in expressive probabilistic models is generally intractable, which makes them difficult to learn and limits their applicability. Sum-product networks are a class of deep models where, surprisingly, inference remains tractable even when an arbitrary number of hidden layers are present. In this talk, I generalize this result to a much broader set of learning problems: all those where inference consists of summing a function over a semiring. This includes satisfiability, constraint satisfaction, optimization, integration, and others. In any semiring, for summation to be tractable it suffices that the factors of every product have disjoint scopes. This unifies and extends many previous results in the literature. Enforcing this condition at learning time thus ensures that the learned models are tractable. I illustrate the power and generality of this approach by applying it to a new type of structured prediction problem: learning a nonconvex function that can be globally optimized in polynomial time. I show empirically that this greatly outperforms the standard approach of learning without regard to the cost of optimization. (Joint work with Abram Friesen)

Abstract 8: Contributed Presentation 2 - LibSPN: A Library for Learning and Inference with Sum-Product Networks and TensorFlow in Principled Approaches to Deep Learning, 11:45 AM

LibSPN: A Library for Learning and Inference with Sum-Product Networks and TensorFlow

Andrzej Pronobis, Avinash Ranganath, Rajesh Rao

Sum-Product Networks (SPNs) are a probabilistic deep architecture with solid theoretical foundations, which demonstrated state-of-the-art performance in several domains. Yet, surprisingly, there are no mature, general-purpose SPN implementations that would serve as a platform for the community of machine learning researchers centered around SPNs. Here, we present a new general-purpose Python library called LibSPN, which aims to become such a platform. The library is designed to make it straightforward and effortless to apply various SPN architectures to large-scale datasets and problems. The library achieves scalability and efficiency, thanks to a tight coupling with TensorFlow, a framework already used by a large community of researchers and developers in multiple domains. We describe the design and benefits of LibSPN, give several use-case examples, and demonstrate the applicability of the library to real-world problems on the example of spatial understanding in mobile robotics.

Abstract 11: Contributed Presentation 3 - Emergence of invariance and disentangling in deep representations in Principled Approaches to Deep Learning, 02:00 PM

Emergence of invariance and disentangling in deep representations

Alessandro Achille, Stefano Soatto

We show that invariance in a deep neural network is equivalent to the information minimality of the representation it computes, and that stacking layers and injecting noise during training naturally bias the

network towards learning invariant representations. Then, we show that overfitting is related to the quantity of information stored in the weights, and derive a sharp bound between this information and the minimality and Total Correlation of the layers. This allows us to conclude that implicit and explicit regularization of the loss function not only help limit overfitting, but also foster invariance and disentangling of the learned representation. We also shed light on the properties of deep networks in relation to the geometry of the loss function.

Abstract 12: **Invited Talk 6 - Nathan Srebro in Principled Approaches to Deep Learning**, 02:15 PM

Geometry, Optimization and Generalization in Multilayer Networks

What is it that enables learning with multi-layer networks? What causes the network to generalize well despite the model class having extremely high capacity? In this talk I will explore these questions through experimentation, analogy to matrix factorization (including some new results on the energy landscape and implicit regularization in matrix factorization), and study of alternate geometries and optimization approaches.

Abstract 13: **Contributed Presentation 4 - The Shattered Gradients Problem: If resnets are the answer, then what is the question? in Principled Approaches to Deep Learning**, 02:45 PM

The Shattered Gradients Problem: If resnets are the answer, then what is the question?

David Balduzzi, Brian McWilliams, Marcus Frean, John Lewis, Lennox Leary, Kurt Wan Duo Ma

A long-standing obstacle to progress in deep learning is the problem of vanishing and exploding gradients. Although, the problem has largely been overcome via carefully constructed initializations and batch normalization, architectures incorporating skip-connections such as highway and resnets perform much better than standard feedforward architectures despite well-chosen initialization and batch normalization. In this paper, we identify the shattered gradients problem. Specifically, we show that the correlation between gradients in standard feedforward networks decays exponentially with depth resulting in gradients that resemble white noise whereas, in contrast, the gradients in architectures with skip-connections are far more resistant to shattering, decaying sublinearly. Detailed empirical evidence is presented in support of the analysis, on both fully-connected networks and convnets. Finally, we present a new "looks linear" (LL) initialization that prevents shattering, with preliminary experiments showing the new initialization allows to train very deep networks without the addition of skip-connections.

Abstract 15: **Contributed Presentation 5 - Towards Deep Learning Models Resistant to Adversarial Attacks in Principled Approaches to Deep Learning**, 03:45 PM

Towards Deep Learning Models Resistant to Adversarial Attacks

Aleksander M┼dry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, Adrian Vladu

Recent work has demonstrated that neural networks are vulnerable to adversarial examples, i.e., inputs that are almost indistinguishable from natural data and yet classified incorrectly by the network. To address this problem, we study the adversarial robustness of neural networks through

the lens of robust optimization. This approach provides a broad and unifying view on much of the prior work on this topic. Its principled nature also enables us to identify general methods for both training and attacking neural networks that are reliable and, in a certain sense, universal. These methods let us train networks with significantly improved resistance to a wide range adversarial attacks. This suggests that adversarially resistant deep learning models might be within our reach after all.

Video Games and Machine Learning

Gabriel Synnaeve, Julian Togelius, Tom Schaul, Oriol Vinyals, Nicolas Usunier

C4.6, Thu Aug 10, 08:30 AM

Good benchmarks are necessary for developing artificial intelligence. Recently, there has been a growing movement for the use of video games as machine learning benchmarks [1,2,3], and also an interest in the applications of machine learning from the video games community. While games have been used for AI research for a long time, only recently have we seen modern machine learning methods applied to video games.

This workshop focuses on complex games which provide interesting and hard challenges for machine learning. Going beyond simple toy problems of the past, and games which can easily be solved with search, we focus on games where learning is likely to be necessary to play well. This includes strategy games such as StarCraft [4,5], open-world games such as MineCraft [6,7,8], first-person shooters such as Doom [9,10], as well as hard and unsolved 2D games such as Ms. Pac-Man and Montezuma's Revenge [11,12,13]. While we see most of the challenges in game-playing, there are also interesting machine learning challenges in modeling and content generation [14]. This workshop aims at bringing together all researchers from ICML who want to use video games as a benchmark. We will have talks by invited speakers from machine learning, from the game AI community, and from the video games industry.

[1] Greg Brockman, Catherine Olsson, Alex Ray, et al. "OpenAI Universe", <https://openai.com/blog/universe/> (2016).

[2] Charles Beattie, Joel Z. Leibo, Denis Teplyashin, Tom Ward, Marcus Wainwright, Heinrich K┼ttler, Andrew Lefrancq, Simon Green, V┴ctor Vald┴s, Amir Sadiq, Julian Schrittwieser, Keith Anderson, Sarah York, Max Cant, Adam Cain, Adrian Bolton, Stephen Gaffney, Helen King, Demis Hassabis, Shane Legg, Stig Petersen, "DeepMind Lab", arXiv:1612.03801 (2016).

[3] Gabriel Synnaeve, Nantas Nardelli, Alex Auvolat, Soumith Chintala, Timoth┴e Lacroix, Zeming Lin, Florian Richoux, Nicolas Usunier, "TorchCraft: a Library for Machine Learning Research on Real-Time Strategy Games", arXiv:1611.00625 (2016).

[4] Santiago Ontanon, Gabriel Synnaeve, Alberto Uriarte, Florian Richoux, David Churchill, Mike Preuss, "A Survey of Real-Time Strategy Game AI Research and Competition in StarCraft", IEEE Transactions on Computational Intelligence and AI in games 5.4 (2013): 293-311.

[5] StarCraft AI Competition @ AIIDE 2016

[6] Junhyuk Oh, Valliappa Chockalingam, Satinder Singh, and Honglak Lee, "Control of Memory, Active Perception, and Action in Minecraft", ICML (2016).

[7] Chen Tessler, Shahar Givony, Tom Zahavy, Daniel J. Mankowitz,

Shie Mannor, "A Deep Hierarchical Approach to Lifelong Learning in Minecraft", arXiv preprint arXiv:1604.07255 (2016).

[8] Matthew Johnson, Katja Hofmann, Tim Hutton, David Bignell, "The Malmo Platform for Artificial Intelligence Experimentation", IJCAI (2016).

[9] Visual Doom AI Competition @ CIG 2016

[10] Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Tim Harley, Timothy P. Lillicrap, David Silver, Koray Kavukcuoglu, "Asynchronous Methods for Deep Reinforcement Learning", arXiv preprint arXiv:1602.01783 (2016).

[11] Tejas D. Kulkarni, Karthik R. Narasimhan, Ardavan Saeedi, Joshua B. Tenenbaum, "Hierarchical Deep Reinforcement Learning: Integrating Temporal Abstraction and Intrinsic Motivation", arXiv preprint arXiv:1604.06057 (2016).

[12] Marc G. Bellemare, Sriram Srinivasan, Georg Ostrovski, Tom Schaul, David Saxton, Remi Munos, "Unifying Count-Based Exploration and Intrinsic Motivation", arXiv preprint arXiv:1606.01868 (2016).

[13] Diego Perez-Liebana, Spyridon Samothrakis, Julian Togelius, Tom Schaul, Simon Lucas, "General Video Game AI: Competition, Challenges and Opportunities", AAAI (2016).

[14] Julian Togelius, Georgios N. Yannakakis, Kenneth O. Stanley and Cameron Browne, "Search-based Procedural Content Generation: a Taxonomy and Survey". IEEE TCIAIG (2011).

ML on a budget: IoT, Mobile and other tiny-ML applications

Manik Varma, Venkatesh Saligrama, Prateek Jain

C4.7, Thu Aug 10, 08:30 AM

We routinely encounter scenarios where at test-time we must predict on a budget. Feature costs in Internet, Healthcare, and Surveillance applications arise due to feature extraction time and feature/sensor acquisition~\cite{trapeznikov:2013b} costs. Data analytics applications in mobile devices are often performed on remote cloud services due to the limited device capabilities, which imposes memory/prediction time costs. Naturally, in these settings, one needs to carefully understand the trade-off between accuracy and prediction cost. Uncertainty in the observations, which is typical in such scenarios, further adds to complexity of the task and requires a careful understanding of both the uncertainty as well as accuracy-cost tradeoffs.

In this workshop, we aim to bring together researchers from various domains to discuss the key aspects of the above mentioned emerging and critical topic. The goal is to provide a platform where ML/statistics/optimization researchers can interact closely with domain experts who need to deploy ML models in resource-constrained settings (like an IoT device maker), and chart out the foundational problems in the area and key tools that can be used to solve them.

Motivation

=====
 Prediction under budget constraints is a critical problem that arise in several settings like medical diagnosis, search engines and surveillance. In these applications, budget constraints arise as a result of limits on computational cost, time, network-throughput and power-consumption. For instance, in search engines CPU cost during prediction-time must be budgeted to enable business models such as online advertising. Additionally, search engines have time constraints at prediction-time as users are known to abandon the service is the response time of the

search engine is not within a few tens of milliseconds. In another example, modern passenger screening systems impose constraints on throughput.

An extreme version of these problems appear in the Internet of Things (IoT) setting where one requires prediction on tiny IoT devices which might have at most 2KB of RAM and no floating point computation unit. IoT is considered to be the next multi-billion industry with "smart" devices being designed for production-line, cars, retail stores, and even for toothbrush and spoons. Given that IoT based solutions seem destined to significantly permeate our day-to-day lives, ML based predictions on the device become critical due to several reasons like privacy, battery, latency etc.

Learning under resource constraints departs from the traditional machine learning setting and introduces new exciting challenges. For instance, features are accompanied by costs (e.g. extraction time in search engines or true monetary values in medical diagnosis) and their amortized sum is constrained at test-time. Also, different search strategies in prediction can have widely varying computational costs (e.g., binary search, linear search, dynamic programming). In other settings, a system must maintain a throughput constraint to keep pace with arriving traffic.

In IoT setting, the model itself has to be deployed on a 2-16KB RAM, posing an extremely challenging constraint on the algorithm.

The common aspect of all of these settings is that we must seeks trade-offs between prediction accuracy and prediction cost. Studying this tradeoff is an inherent challenge that needs to be investigated in a principled fashion in order to invent practically relevant machine learning algorithms. This problems lies at the intersection of ML, statistics, stochastic control and information theory. We aim to draw researchers working on foundational, algorithmic and application problems within these areas. We plan on organizing a demo session which would showcase ML algorithms running live on various resource-constrained device, demonstrating their effectiveness on challenging real-world tasks. In addition, we plan to invite Ofer Dekel from Microsoft Research to present a new platform for deploying ML on tiny devices which should provide a easy way to deploy and compare various ML techniques on realistic devices and further spur multiple research directions in this area.

Workshop on Human Interpretability in Machine Learning (WHI)

Kush Varshney, Adrian Weller, Been Kim, Dmitry Malioutov

C4.8, Thu Aug 10, 08:30 AM

This workshop will bring together researchers who study the interpretability of predictive models, develop interpretable machine learning algorithms, and develop methodology to interpret black-box machine learning models (e.g., post-hoc interpretations). This is a very exciting time to study interpretable machine learning, as the advances in large-scale optimization and Bayesian inference that have enabled the rise of black-box machine learning are now also starting to be exploited to develop principled approaches to large-scale interpretable machine learning. Participants in the workshop will exchange ideas on these and allied topics, including:

- Quantifying and axiomatizing interpretability
- Psychology of human concept learning
- Rule learning, Symbolic regression and case-based reasoning
- Generalized additive models, sparsity and interpretability
- Visual analytics
- Interpretable unsupervised models (clustering, topic models, e.t.c)
- Interpretation of black-box models (including deep neural networks)
- Causality of predictive models
- Verifying, diagnosing and debugging machine learning systems
- Interpretability in reinforcement learning.

Doctors, judges, business executives, and many other people are faced with making critical decisions that can have profound consequences. For example, doctors decide which treatment to administer to patients, judges decide on prison sentences for convicts, and business executives decide to enter new markets and acquire other companies. Such decisions are increasingly being supported by predictive models learned by algorithms from historical data.

The latest trend in machine learning is to use highly nonlinear complex systems such as deep neural networks, kernel methods, and large ensembles of diverse classifiers. While such approaches often produce impressive, state-of-the-art prediction accuracies, their black-box nature offers little comfort to decision makers. Therefore, in order for predictions to be adopted, trusted, and safely used by decision makers in mission-critical applications, it is imperative to develop machine learning methods that produce interpretable models with excellent predictive accuracy. It is in this way that machine learning methods can have impact on consequential real-world applications.

Schedule

08:30 AM	A. Dhurandhar, V. Iyengar, R. Luss, and K. Shanmugam, "A Formal Framework to Characterize Interpretability of Procedures"	<i>Shanmugam</i>
08:45 AM	A. Henelius, K. Puolamäki, and A. Ukkonen, "Interpreting Classifiers through Attribute Interactions in Datasets"	<i>Henelius</i>
09:00 AM	S. Lundberg and S.-I. Lee, "Consistent Feature Attribution for Tree Ensembles"	<i>Hiranuma</i>
09:15 AM	Invited Talk: D. Sontag	
10:30 AM	S. Penkov and S. Ramamoorthy, "Program Induction to Interpret Transition Systems"	<i>Penkov</i>
10:45 AM	R. L. Phillips, K. H. Chang, and S. Friedler, "Interpretable Active Learning"	<i>Phillips</i>

11:00 AM	C. Rosenbaum, T. Gao, and T. Klinger, "e-QRAQ: A Multi-turn Reasoning Dataset and Simulator with Explanations"	
11:15 AM	Invited Talk: T. Jebara, "Interpretability Through Causality"	
02:00 PM	W. Tansey, J. Thomason, and J. G. Scott, "Interpretable Low-Dimensional Regression via Data-Adaptive Smoothing"	<i>Tansey</i>
02:15 PM	Invited Talk: P. W. Koh	
03:30 PM	I. Valera, M. F. Pradier, and Z. Ghahramani, "General Latent Feature Modeling for Data Exploration Tasks"	
03:45 PM	A. Weller, "Challenges for Transparency"	<i>Weller</i>
04:00 PM	ICML WHI 2017 Awards Ceremony	
04:05 PM	Panel Discussion: Human Interpretability in Machine Learning	

Abstracts (12):

Abstract 1: **A. Dhurandhar, V. Iyengar, R. Luss, and K. Shanmugam, "A Formal Framework to Characterize Interpretability of Procedures" in Workshop on Human Interpretability in Machine Learning (WHI), Shanmugam 08:30 AM**

We provide a novel notion of what it means to be interpretable, looking past the usual association with human understanding. Our key insight is that interpretability is not an absolute concept and so we define it relative to a target model, which may or may not be a human. We define a framework that allows for comparing interpretable procedures by linking it to important practical aspects such as accuracy and robustness. We characterize many of the current state-of-the-art interpretable methods in our framework portraying its general applicability.

Abstract 2: **A. Henelius, K. Puolamäki, and A. Ukkonen, "Interpreting Classifiers through Attribute Interactions in Datasets" in Workshop on Human Interpretability in Machine Learning (WHI), Henelius 08:45 AM**

In this work we present the novel ASTRID method for investigating which attribute interactions classifiers exploit when making predictions. Attribute interactions in classification tasks mean that two or more attributes together provide stronger evidence for a particular class label. Knowledge of such interactions makes models more interpretable by revealing associations between attributes. This has applications, e.g., in pharmacovigilance to identify interactions between drugs or in bioinformatics to investigate associations between single nucleotide

polymorphisms. We also show how the found attribute partitioning is related to a factorisation of the data generating distribution and empirically demonstrate the utility of the proposed method.

Abstract 3: **S. Lundberg and S.-I. Lee, "Consistent Feature Attribution for Tree Ensembles" in Workshop on Human Interpretability in Machine Learning (WHI)**, *Hiranuma* 09:00 AM

It is critical in many applications to understand what features are important for a model, and why individual predictions were made. For tree ensemble methods these questions are usually answered by attributing importance values to input features, either globally or for a single prediction. Here we show that current feature attribution methods are inconsistent, which means changing the model to rely more on a given feature can actually decrease the importance assigned to that feature. To address this problem we develop fast exact solutions for SHAP (SHapley Additive exPlanation) values, which were recently shown to be the unique additive feature attribution method based on conditional expectations that is both consistent and locally accurate. We integrate these improvements into the latest version of XGBoost, demonstrate the inconsistencies of current methods, and show how using SHAP values results in significantly improved supervised clustering performance. Feature importance values are a key part of understanding widely used models such as gradient boosting trees and random forests. We believe our work improves on the state-of-the-art in important ways, and may impact any current user of tree ensemble methods.

Abstract 5: **S. Penkov and S. Ramamoorthy, "Program Induction to Interpret Transition Systems" in Workshop on Human Interpretability in Machine Learning (WHI)**, *Penkov* 10:30 AM

Explaining and reasoning about processes which underlie observed black-box phenomena enables the discovery of causal mechanisms, derivation of suitable abstract representations and the formulation of more robust predictions. We propose to learn high level functional programs in order to represent abstract models which capture the invariant structure in the observed data. We introduce the π -machine (program-induction machine) -- an architecture able to induce interpretable LISP-like programs from observed data traces. We propose an optimisation procedure for program learning based on backpropagation, gradient descent and A* search. We apply the proposed method to two problems: system identification of dynamical systems and explaining the behaviour of a DQN agent. Our results show that the π -machine can efficiently induce interpretable programs from individual data traces.

Abstract 6: **R. L. Phillips, K. H. Chang, and S. Friedler, "Interpretable Active Learning" in Workshop on Human Interpretability in Machine Learning (WHI)**, *Phillips* 10:45 AM

Active learning has long been a topic of study in machine learning. However, as increasingly complex and opaque models have become standard practice, the process of active learning, too, has become more opaque. There has been little investigation into interpreting what specific trends and patterns an active learning strategy may be exploring. This work expands on the Local Interpretable Model-agnostic Explanations framework (LIME) to provide explanations for active learning recommendations. We demonstrate how LIME can be used to generate locally faithful explanations for an active learning strategy, and how these explanations can be used to understand how different models and datasets explore a problem space over time. In order to quantify the per-subgroup differences in how an active learning strategy queries

spatial regions, we introduce a notion of uncertainty bias (based on disparate impact) to measure the discrepancy in the confidence for a model's predictions between one subgroup and another. Using the uncertainty bias measure, we show that our query explanations accurately reflect the subgroup focus of the active learning queries, allowing for an interpretable explanation of what is being learned as points with similar sources of uncertainty have their uncertainty bias resolved. We demonstrate that this technique can be applied to track uncertainty bias over user-defined clusters or automatically generated clusters based on the source of uncertainty.

Abstract 7: **C. Rosenbaum, T. Gao, and T. Klinger, "e-QRAQ: A Multi-turn Reasoning Dataset and Simulator with Explanations" in Workshop on Human Interpretability in Machine Learning (WHI)**, 11:00 AM

In this paper we present a new dataset and user simulator e-QRAQ (explainable Query, Reason, and Answer Question) which tests an Agent's ability to read an ambiguous text; ask questions until it can answer a challenge question; and explain the reasoning behind its questions and answer. The User simulator provides the Agent with a short, ambiguous story and a challenge question about the story. The story is ambiguous because some of the entities have been replaced by variables. At each turn the Agent may ask for the value of a variable or try to answer the challenge question. In response the User simulator provides a natural language explanation of why the Agent's query or answer was useful in narrowing down the set of possible answers, or not. To demonstrate one potential application of the e-QRAQ dataset, we train a new neural architecture based on End-to-End Memory Networks to successfully generate both predictions and partial explanations of its current understanding of the problem. We observe a strong correlation between the quality of the prediction and explanation.

Abstract 8: **Invited Talk: T. Jebara, "Interpretability Through Causality" in Workshop on Human Interpretability in Machine Learning (WHI)**, 11:15 AM

While interpretability often involves finding more parsimonious or sparser models to facilitate human understanding, Netflix also seeks to achieve human interpretability by pursuing causal learning. Predictive models can be impressively accurate in a passive setting but might disappoint a human user who expects the recovered relationships to be causal. More importantly, a predictive model's outcomes may no longer be accurate if the input variables are perturbed through an active intervention. I will briefly discuss applications at Netflix across messaging, marketing and originals promotion which leverage causal modeling in order to achieve models that can be actionable as well as interpretable. In particular, techniques such as two stage least squares (2SLS), instrumental variables (IV), extensions to generalized linear models (GLMs), and other causal methods will be summarized. These causal models can surprisingly recover more interpretable and simpler models than their purely predictive counterparts. Furthermore, sparsity can potentially emerge when causal models ignore spurious relationships that might otherwise be recovered in a purely predictive objective function. In general, causal models achieve better results algorithmically in active intervention settings and enjoy broader adoption from human stakeholders.

Abstract 9: **W. Tansey, J. Thomason, and J. G. Scott, "Interpretable Low-Dimensional Regression via Data-Adaptive Smoothing" in Workshop on Human Interpretability in Machine Learning (WHI)**,

Tansey 02:00 PM

We consider the problem of estimating a regression function in the common situation where the number of features is small, where interpretability of the model is a high priority, and where simple linear or additive models fail to provide adequate performance. To address this problem, we present Maximum Variance Total Variation denoising (MVTV), an approach that is conceptually related both to CART and to the more recent CRISP algorithm, a state-of-the-art alternative method for interpretable nonlinear regression. MVTV divides the feature space into blocks of constant value and fits the value of all blocks jointly via a convex optimization routine. Our method is fully data-adaptive, in that it incorporates highly robust routines for tuning all hyperparameters automatically. We compare our approach against CART and CRISP via both a complexity-accuracy tradeoff metric and a human study, demonstrating that that MVTV is a more powerful and interpretable method.

Abstract 11: **I. Valera, M. F. Pradier, and Z. Ghahramani, "General Latent Feature Modeling for Data Exploration Tasks" in Workshop on Human Interpretability in Machine Learning (WHI)**, 03:30 PM

This paper introduces a general Bayesian non-parametric latent feature model suitable to perform automatic exploratory analysis of heterogeneous datasets, where the attributes describing each object can be either discrete, continuous or mixed variables. The proposed model presents several important properties. First, it accounts for heterogeneous data while can be inferred in linear time with respect to the number of objects and attributes. Second, its Bayesian nonparametric nature allows us to automatically infer the model complexity from the data, i.e., the number of features necessary to capture the latent structure in the data. Third, the latent features in the model are binary-valued variables, easing the interpretability of the obtained latent features in data exploration tasks.

Abstract 12: **A. Weller, "Challenges for Transparency" in Workshop on Human Interpretability in Machine Learning (WHI)**, Weller 03:45 PM

Transparency is often deemed critical to enable effective real-world deployment of intelligent systems. Yet the motivations for and benefits of different types of transparency can vary significantly depending on context, and objective measurement criteria are difficult to identify. We provide a brief survey, suggesting challenges and related concerns. We highlight and review settings where transparency may cause harm, discussing connections across privacy, multi-agent game theory, economics, fairness and trust.

Abstract 13: **ICML WHI 2017 Awards Ceremony in Workshop on Human Interpretability in Machine Learning (WHI)**, 04:00 PM

Join us in recognizing the best papers of the workshop.

Abstract 14: **Panel Discussion: Human Interpretability in Machine Learning in Workshop on Human Interpretability in Machine Learning (WHI)**, 04:05 PM

panelists: Tony Jebara, Bernhard Schölkopf, Been Kim, Kush Varshney
moderator: Adrian Weller

Automatic Machine Learning (AutoML 2017)

Joaquin Vanschoren, Roman Garnett

C4.9, Thu Aug 10, 08:30 AM

Machine learning has achieved considerable successes in recent years and an ever-growing number of disciplines rely on it. However, this success crucially relies on human machine learning experts, who select appropriate features, workflows, machine learning paradigms, algorithms, and their hyperparameters. As the complexity of these tasks is often beyond non-experts, the rapid growth of machine learning applications has created a demand for off-the-shelf machine learning methods that can be used easily and without expert knowledge. We call the resulting research area that targets progressive automation of machine learning AutoML.

Implicit Generative Models

Rajesh Ranganath, Ian Goodfellow, Dustin Tran, David Blei, Balaji Lakshminarayanan, Shakir Mohamed

Parkside 1, Thu Aug 10, 08:30 AM

Probabilistic models are a central implement in machine learning practice. They form the basis for models that generate realistic data, uncover hidden structure, and make predictions. Traditionally, probabilistic models in machine learning have focused on prescribed models. Prescribed models specify a joint density over observed and hidden variables that can be easily evaluated. The requirement of a tractable density simplifies their learning but limits their flexibility --- several real world phenomena are better described by simulators that do not admit a tractable density. Probabilistic models defined only via the simulations they produce are called implicit models.

Arguably starting with generative adversarial networks, research on implicit models in machine learning has exploded in recent years. This workshop's aim is to foster a discussion around the recent developments and future directions of implicit models.

Implicit models have many applications. They are used in ecology where models simulate animal populations over time; they are used in phylogeny, where simulations produce hypothetical ancestry trees; they are used in physics to generate particle simulations for high energy processes. Recently, implicit models have been used to improve the state-of-the-art in image and content generation. Part of the workshop's focus is to discuss the commonalities among applications of implicit models.

Of particular interest at this workshop is to unite fields that work on implicit models. For example:

+ Generative adversarial networks (a NIPS 2016 workshop) are implicit models with an adversarial training scheme.

+ Recent advances in variational inference (a NIPS 2015 and 2016 workshop) have leveraged implicit models for more accurate approximations.

+ Approximate Bayesian computation (a NIPS 2015 workshop) focuses on posterior inference for models with implicit likelihoods.

+ Learning implicit models is deeply connected to two sample testing and density ratio estimation.

We hope to bring together these different views on implicit models, identifying their core challenges and combining their innovations.

We invite submission of 4 page papers for posters, contributed talks, and travel awards. Topics of interests are: implicit models, approximate Bayesian computation, generative adversarial networks, learning and inference for implicit models, implicit variational approximations, evaluation of implicit models and two sample testing. We encourage both theoretical and applied submissions.

Time Series Workshop

Vitaly Kuznetsov, Yan Liu, Scott Yang, Rose Yu

C4.1, Fri Aug 11, 08:30 AM

Time series data is ubiquitous. In domains as diverse as finance, entertainment, transportation and health-care, we observe a fundamental shift away from parsimonious, infrequent measurement to nearly continuous monitoring and recording. Rapid advances in diverse sensing technologies, ranging from remote sensors to wearables and social sensing, are generating a rapid growth in the size and complexity of time series archives. Thus, although time series analysis has been studied extensively, its importance only continues to grow. Furthermore, modern time series data pose significant challenges to existing techniques both in terms of the structure (e.g., irregular sampling in hospital records and spatiotemporal structure in climate data) and size. These challenges are compounded by the fact that standard i.i.d. assumptions used in other areas of machine learning are not appropriate for time series and new theory, models and algorithms are needed to process and analyse this data.

The goal of this workshop is to bring together theoretical and applied researchers interested in the analysis of time series and development of new algorithms to process sequential data. This includes algorithms for time series prediction, classification, clustering, anomaly and change point detection, correlation discovery, dimensionality reduction as well as a general theory for learning and comparing stochastic processes. We invite researchers from the related areas of batch and online learning, reinforcement learning, data analysis and statistics, econometrics, and many others to contribute to this workshop.

Our workshop will build on the success of past two time series workshops that were held at NIPS and KDD (also co-organized by the proposers). The workshop will attract a broader audience from ICML community. In particular, when we have the KDD workshop on time series in 2015 held in Sydney, it attracts many local researchers in Australia who work on time series research or related applications. We expect the proposed workshop will be a hit given its large interest in the ICML community as well as the local interest in Sydney.

Reproducibility in Machine Learning Research

Rosemary Nan Ke, Parth Goyal, Alex Lamb, Joelle Pineau, Samy Bengio, Yoshua Bengio

C4.10, Fri Aug 11, 08:30 AM

This workshop focuses on issues of reproducibility and replication of results in the Machine Learning community. Papers from the Machine Learning community are supposed to be a valuable asset. They can help to inform and inspire future research. They can be a useful educational

tool for students. They can give guidance to applied researchers in industry. Perhaps most importantly, they can help us to answer the most fundamental questions about our existence - what does it mean to learn and what does it mean to be human? Reproducibility, while not always possible in science (consider the study of a transient astrological phenomenon like a passing comet), is a powerful criteria for improving the quality of research. A result which is reproducible is more likely to be robust and meaningful and rules out many types of experimenter error (either fraud or accidental).

There are many interesting open questions about how reproducibility issues intersect with the Machine Learning community:

- * How can we tell if papers in the Machine Learning community are reproducible even in theory? If a paper is about recommending news sites before a particular election, and the results come from running the system online in production - it will be impossible to reproduce the published results because the state of the world is irreversibly changed from when the experiment was ran.
- * What does it mean for a paper to be reproducible in theory but not in practice? For example, if a paper requires tens of thousands of GPUs to reproduce or a large closed-off dataset, then it can only be reproduced in reality by a few large labs.
- * For papers which are reproducible both in theory and in practice - how can we ensure that papers published in ICML would actually be able to replicate if such an experiment were attempted?
- * What does it mean for a paper to have successful or unsuccessful replications?
- * Of the papers with attempted replications completed, how many have been published?
- * What can be done to ensure that as many papers which are reproducible in theory fall into the last category?
- * On the reproducibility issue, what can the Machine Learning community learn from other fields?

Our aim in the following workshop is to raise the profile of these questions in the community and to search for their answers. In doing so we aim for papers focusing on the following topics:

- * Analysis of the current state of reproducibility in machine learning venues
- * Tools to help increase reproducibility
- * Evidence that reproducibility is important for science
- * Connections between the reproducibility situation in Machine Learning and other fields
- * Replications, both failed and successful, of influential papers in the Machine Learning literature.

Interactive Machine Learning and Semantic Information Retrieval

Dorota Glowacka, Wray Buntine, Petri Myllymaki

C4.11, Fri Aug 11, 08:30 AM

Retrieval techniques operating on text or semantic annotations have become the industry standard for retrieval from large document collections. However, traditional information retrieval techniques operate on the assumption that the user issues a single query and the system responds with a ranked list of documents. In recent years we have

witnessed a substantial growth in text data coming from various online resources, such as online newspapers, blogs, specialised document collections (e.g. arXiv). Traditional information retrieval approaches often fail to provide users with adequate support when browsing such online resources, hence in recent years there has been a growing interest in developing new algorithms and design methods that can support interactive information retrieval. The aim of this workshop is to explore new methods and related system design for interactive data analytics and management in various domains, including specialised text collections (e.g. legal, medical, scientific) as well as for various tasks, such as semantic information retrieval, conceptual organization and clustering of data collections for sense making, semantic expert profiling, and document recommender systems.

Of interest, also, is probabilistic and machine learning formulations of the interactive information retrieval task above and beyond the simple "stochastic language models" framework developed in the information retrieval community.

The primary audience of the workshop are researchers and practitioners in the area of interactive and personalised system design as well as interactive machine learning both from academia and industry.

Schedule

09:00 AM	Welcome
	Using Web Text Based
09:15 AM	Analytics to Gather Customer Insights
	Intent Driven Dynamic
09:35 AM	Product Ranking System for Fashion E-commerce
10:00 AM	Coffee Break
	Towards End-to-End
10:30 AM	Reinforcement Learning of Dialogue Agents for Information Access
	Learning to Select Relevant and Diverse Subsets
11:30 AM	
12:00 PM	Lunch Break
	Focused Reading:
02:00 PM	Reinforcement Learning for What Documents to Read
	CCA: Attention Based Cross
02:20 PM	Modal Retrieval on Webpages
	User Study for Measuring Linguistic Complexity and its Reduction by
02:40 PM	Technology on a Patent Website
03:00 PM	Poster Session

Machine Learning in Speech and Language Processing

Karen Livescu, Tara Sainath, lianglu Lu, Anton Ragni

C4.3, Fri Aug 11, 08:30 AM

This workshop continues a tradition of MLSLP workshops held as satellites of ICML, ACL, and Interspeech conferences. While research in speech and language processing has always involved machine learning (ML), current research is benefiting from even closer interaction between these fields. Speech and language processing is continually mining new ideas from ML and ML, in turn, is devoting more interest to speech and language applications. This workshop is a venue for locating and incubating the next waves of research directions for interaction and collaboration. The workshop will (1) discuss emerging research ideas with potential for impact in speech/language and (2) bring together relevant researchers from ML and speech/language who may not regularly interact at conferences. Example topics include new directions for deep learning in speech/language, reinforcement learning, unsupervised/semi-supervised learning, domain adaptation/transfer learning, and topics at the boundary of speech, text, and other modalities.

Private and Secure Machine Learning

Antti Honkela, Kana Shimizu, Samuel Kaski

C4.4, Fri Aug 11, 08:30 AM

There are two complementary approaches to private and secure machine learning: differential privacy can guarantee privacy of the subjects of the training data with respect to the output of a differentially private learning algorithm, while cryptographic approaches can guarantee secure operation of the learning process in a potentially distributed environment. The aim of this workshop is to bring together researchers interested in private and secure machine learning, to stimulate interactions to advance either perspective or to combine them.

Deep Structured Prediction

Isabelle Augenstein, Kai-Wei Chang, Gal Chechik, Bert Huang, Andre Filipe Torres Martins, Ofer Meshi, Alex Schwing, Yishu Miao

C4.5, Fri Aug 11, 08:30 AM

In recent years, deep learning has revolutionized machine learning. Most successful applications of deep learning involve predicting single variables (e.g., univariate regression or multi-class classification). However, many real problems involve highly dependent, structured variables. In such scenarios, it is desired or even necessary to model correlations and dependencies between the multiple input and output variables. Such problems arise in a wide range of domains, from natural language processing, computer vision, computational biology and others.

Some approaches to these problems directly use deep learning concepts, such as those that generate sequences using recurrent neural networks or that output image segmentations through convolutions. Others adapt the concepts from structured output learning. These structured output prediction problems were traditionally handled using linear models and hand-crafted features, with a structured optimization

such as inference. It has recently been proposed to combine the representational power of deep neural networks with modeling variable dependence in a structured prediction framework. There are numerous interesting research questions related to modeling and optimization that arise in this problem space.

This workshop will bring together experts in machine learning and application domains whose research focuses on combining deep learning and structured models. Specifically, we aim to provide an overview of existing approaches from various domains to distill from their success principles that can be more generally applicable. We will also discuss the main challenges that arise in this setting and outline potential directions for future progress. The target audience consists of researchers and practitioners in machine learning and application areas.

Picky Learners: Choosing Alternative Ways to Process Data.

Corinna Cortes, Kamalika Chaudhuri, Giulia DeSalvo, Ningshan Zhang, Chicheng Zhang

C4.6, Fri Aug 11, 08:30 AM

Picky Learners consists of a broad range of learning scenarios where the learner does not simply process every data point blindly, but instead can choose to incorporate them in alternative ways. Despite the growing costs of processing and labelling vast amounts of data, only isolated efforts have tackled this problem primarily in the areas of active learning, learning with rejection and on-line learning with feedback graphs.

In active learning, the learner can choose whether or not to query for a label of each data point, thereby paying different costs for each data point. A key advantage in this setting is that the number of examples queried to learn a concept may be much smaller than the number of examples needed in standard supervised learning. More recently, some have used variations of confidence-based models to determine which labels to query. Confidence-based models lie under the more general framework of learning with rejection, which is a key learning scenario where the algorithm can abstain from making a prediction, at the price of incurring a fixed cost. In this scenario, our picky learners can thus choose to abstain from providing a label. In the on-line setting, one can cast learning with rejection under the more general topic of on-line learning with feedback graphs, a setting that interpolates between bandit and full expert scenario in that the player observes a variety of different expert losses after choosing an action. On-line learning with feedback graphs can then in turn be connected back to active learning where depending on the feedback graph only certain labels are requested.

In short, our picky learners can choose to query for the label (active learning), choose to abstain on the label (learning with rejection) or choose to receive different expert losses (on-line learning with feedback graphs). All of three of these fields attempt in different ways to reduce the cost of processing the data by allowing for picky learners, but the connections between these topics has not been fully explored in terms of both theory and practice. The goal of this workshop is then to bring together researchers and practitioners in these three areas in order to bridge the gap between active learning, learning with rejection, and on-line learning with feedback graphs. We expect that the fruitful collaborations started in this workshop will result in novel research that will help develop each field.

Reliable Machine Learning in the Wild

Dylan Hadfield-Menell, Jacob Steinhardt, Adrian Weller, Smitha Milli

C4.7, Fri Aug 11, 08:30 AM

When can we trust that a system that has performed well in the past will continue to do so in the future? Designing systems that are reliable in the wild is essential for high stakes applications such as self-driving cars and automated surgical assistants. This workshop aims to bring together researchers in diverse areas such as reinforcement learning, human-robot interaction, game theory, cognitive science, and security to further the field of reliability in machine learning. We will focus on three aspects — robustness (to adversaries, distributional shift, model misspecification, corrupted data); awareness (of when a change has occurred, when the model might be miscalibrated, etc.); and adaptation (to new situations or objectives). We aim to consider each of these in the context of the complex human factors that impact the successful application or meaningful monitoring of any artificial intelligence technology. Together, these will aid us in designing and deploying reliable machine learning systems.

Human in the Loop Machine Learning

Richard Nock, Cheng Soon Ong

C4.8, Fri Aug 11, 08:30 AM

For details see:

<http://machlearn.gitlab.io/hitl2017/>

As machine learning systems become more ubiquitous in everybody's day-to-day life or work, society and industry is in an intermediate state between fully manual and fully automatic systems. The gradient undoubtedly points towards full automation, but moving forward in this direction is going to face increasing challenges due to the fact that current machine learning research tends to focus on end to end systems, which puts aside the fact that for practical applications there are still gaps or caveats in the automation. Parts of these come from the presence of (or the necessity to have) the Human in the Loop.

There are two main locations for the Human in the automated system: (i) upstream, in which case the focus is mainly in the inputs of the algorithm. This can be essential for personalised assistants, that describe environments where the machine learning method is tightly embedded into the system. Such environments pose additional challenges related to privacy at large; (ii) downstream: other domains have machine learning approaches analyse parts of the data, and human experts use the results and intuition to make decisions.

The Human dependences between these two locations is also neither straightforward nor acyclic — some applications tend to have feedback effects on data as actions or interventions are undertaken based on machine learning predictions. Furthermore there are often very few rounds of decision making in practice, but each round may affect the statement of the problems related to the Human presence, as witnessed for example by eventual privacy leakages.

This workshop aims to bring together people who are working on systems where machine learning is only part of the solution. Participants will exchange ideas and experiences on human in the loop machine learning.

Topics of interest include:

- System architectures that allow for human decision making
- User interfaces for interacting with machine learning systems
- Validation of human in the loop software systems
- Viewpoints from traditional fields such as reinforcement learning and Bayesian optimisation
- Challenges related to the human presence in the loop (privacy, bias, fairness, etc.)
- Case studies of deployed machine learning

Machine Learning for Music Discovery

Erik Schmidt, Oriol Nieto, Fabien Gouyon, Gert Lanckriet

C4.9, Fri Aug 11, 08:30 AM

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

Furthermore, the recommender systems community has made great progress 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).

We seek to use this workshop to bring together a group of world-class experts to discuss these challenges and share them with the greater machine learning community. In addition to making progress on these challenges, we hope to engage the machine learning community with our nebulous problem space, and connect them with the many available datasets the MIR community has to offer (e.g., AcousticBrainz, Million Song Dataset), which offer near commercial scale to the academic research community.

Reinforcement Learning Workshop

Doina Precup, Balaraman Ravindran, Pierre-Luc Bacon

Parkside 1, Fri Aug 11, 08:30 AM

The workshop will contain presentations of late-breaking reinforcement learning results in all areas of the field, including deep reinforcement learning, exploration, transfer learning and using auxiliary tasks, theoretical result etc, as well as applications of reinforcement learning to various domains. A panel discussion on the most interesting and challenging current research directions will conclude the workshop.

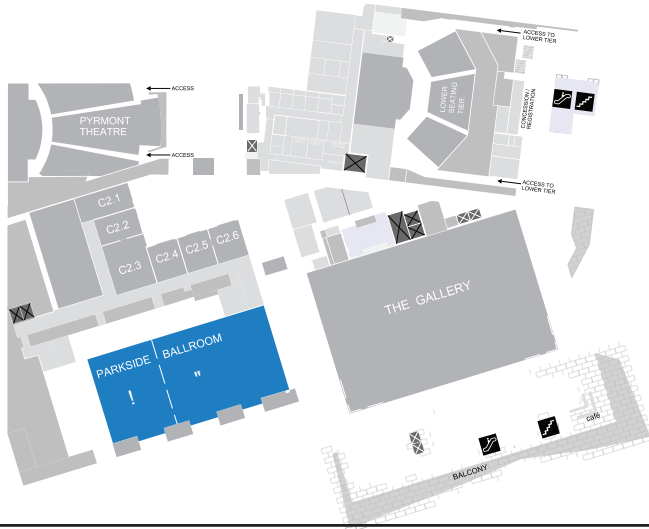
Schedule

08:45 AM	Introductory Remarks	
09:00 AM	Some experiments with learning hyperparameters, transfer, and multi-task learning	<i>Ravindran</i>
09:30 AM	Talk	<i>Finn</i>
10:00 AM	Break	
10:30 AM	Achieving Above-Human Performance on Ms. Pac-Man by Reward Decomposition	<i>van Seijen</i>
11:00 AM	Explorartion methods for options	<i>Lazaric</i>
11:30 AM	The effects of memory replay in reinforcement learning	<i>Liu</i>
12:00 PM	Lunch	
01:30 PM	Asynchronous data aggregation for training end-to-end visual control networks	<i>Monfort</i>
01:50 PM	On the reproducibility of policy gradient experiments	<i>Islam</i>
02:10 PM	Reinforcement learning for the MALMO Minecraft challenge	<i>Berariu</i>
02:30 PM	On the convergence of Tree backup and other reinforcement learning algorithms	<i>Touati</i>
02:50 PM	Transfer learning using successor state features	<i>Lehnert</i>
03:10 PM	Coffee Break	
03:30 PM	Panel Discussion	<i>Ravindran, Finn, Lazaric, Hofmann, Bellemare</i>

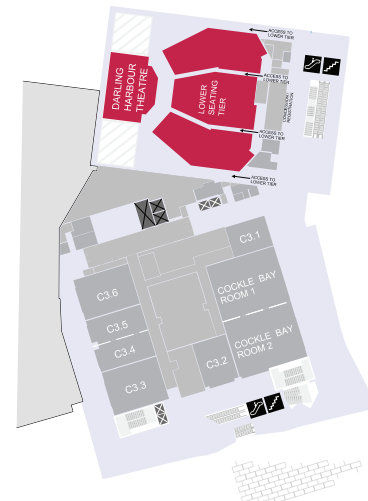


All Events Will
Take Place
In The Main
Building Here

Floor 2 - Parkside Ballrooms Only



Floor 3 - Darling Harbour Theater



Floor 4 - Most Events



Floor 4 - Workshops

