THE 33rd INTERNATIONAL CONFERENCE ON MACHINE LEARNING

NEW YORK CITY, NY, USA
JUNE 19 - JUNE 24, 2016

ICML 2016

ICML 2017

Sydney, Australia

CONFERENCE AT A GLANCE

SUNDAY JUNE 19th

Tutorial Sessions	8:30 am - 1 pm, Marriott & Crowne Plaza
	2:30 pm - 4:30 pm, Marriott & Crowne Plaza
Reception	6 pm - 8 pm, Broadway Ballroom, 8th Floor, Marriott

MONDAY JUNE 20th

Welcome	8:30 am, Marriott: Westside Ballroom 1-4
Invited Talk: Susan Athey	8:40 am, Marriott: Westside Ballroom 1-4
Conference Sessions	10:20 am - 12:20 pm, Marriott
	2 pm - 6:15 pm, Marriott
Poster Session	3 pm - 7 pm, Marriott

TUESDAY JUNE 21st

Invited Talk: Fei-Fei Li	8:30 am, Marriott: Westside Ballroom 1-4
Test Of Time Paper	9:30 am, Marriott: Westside Ballroom 1-4
Poster Session	10 am - 1 pm, Marriott
Conference Sessions	10:30 am - 12:30 pm, Marriott
	3:40 pm - 6:15 pm, Marriott
Invited Talk: Daniel Spielman	2 pm, Marriott: Westside Ballroom 1-4
Poster Session	3 pm - 7 pm, Marriott

WEDNESDAY JUNE 22nd

Conference Sessions	8:30 am - 12:20 pm, Marriott
	3:20 pm - 4:30 pm, Marriott
Poster Session	10 am - 1 pm, Marriott
Invited Talk: Christos Faloutsos	2 pm, Marriott: Westside Ballroom 1-4
Reception	U.S.S. Intrepid, 7 pm - 10 pm

THURSDAY JUNE 23rd

Workshop Sessions	8:30 am - 4:30 pm
	Marriott, Crowne Plaza, Microsoft, Westin

FRIDAY JUNE 24th

Workshop Sessions	8:30 am - 4:30 pm
	Marriott, Westin, Microsoft

For the Workshop Schedule & Maps See pages 55 - 56

Contents

WELCOME

A letter from the Program Chairs

Welcome to New York and the 33rd International Conference on Machine Learning (ICML 2016). The ICML committee has planned a wonderful conference.

Technical Program: We have 322 outstanding articles, selected from 1320 submissions. Each author will present their article to the community in a 15-minute talk, and present a poster at one of the poster sessions for discussion in smaller groups. All accepted articles are published in the Journal of Machine Learning Research (JMLR) as Volume 48 of their Workshop and Conference Proceedings series.

Keynote Speakers: We have four invited keynote speeches from some of the world's intellectual leaders: Susan Athey (Stanford University), Christos Faloutsos (Carnegie Mellon University), Fei-Fei Li (Stanford University), and Daniel Spielman (Yale University).

Tutorials: Nine tutorials spanning some of the most vital subjects in machine learning: deep learning, non-convex optimization, causal inference, stochastic gradient methods, convex optimization, adaptive data analysis, graph sketching, and reinforcement learning.

Workshops: 23 focused workshops for presenting late-breaking research and exploring new areas of machine learning.

Awards: We will present two best paper awards to honor some of the most promising research from the technical program. We will also present the ICML-2016 test of time award. This award is for the paper from the 2006 ICML (Pittsburgh, PA, USA) that has retrospectively had a significant impact on our field.

We would like to acknowledge all the people who made exceptional efforts and dedicated their time to bring this conference together; we were honored to work with them.

Reviewing and selecting papers for the technical program was a mammoth task. We worked with 97 wonderful area chairs and 909 dedicated reviewers to give each paper three high-quality reviews and make an informed (if sometimes difficult) decision. The entire program committee generously offered their time and expertise to the machine learning community, and we thank them. Some reviewers offered extra dedication; 31 are recognized with an ICML Outstanding Reviewer Award. The complete list of the program committee is available on the ICML web site.

In addition to the program committee, we would like to recognize and thank the entire organizing committee who put the conference together. Planning for the tutorials, workshops, volunteers, publications, and sponsorship was ably organized and executed by this team. Their efforts over the past year are the backbone of this fantastic event.

We would like to offer special recognition to several people. First, we thank John Langford, the General Chair, who provided leadership, direction, and advice throughout the planning process. Second, we thank Marek Petrik and Peder Olsen, the local organizers. Marek, Peder, and their team gave their time and energy to see to the many details around the day-to-day of this year's ICML. Last, we thank Jacob Gardner and Matthew Kusner, the workflow chairs. Their help was invaluable in nearly every aspect of our planning process; neither of us can imagine performing this task without them.

Finally, we want to acknowledge our sponsors (Inside Cover) and the IMLS board. ICML 2016 is not possible without their continued support.

On behalf of all of us at ICML 2016, enjoy the conference!

Nina Balcan and Kilian Weinberger, ICML 2016

A Warm Welcome From the Local Chairs

Thank you for attending the 33rd International Conference on Machine Learning organized by the International Machine Learning Society in New York City. On behalf of the entire conference organizing committee it is our honor and pleasure to be your hosts. It can be seen as symbolic that the conference this year takes place in Times Square in the city that never sleeps - arguably in the capital of the world. This year will be the best attended in the history of ICML and it will take place at a time when machine learning is undergoing tremendous growth and excitement.

We are confident that you will find the scientific program technically stimulating. With four exciting plenary speakers, 9 tutorials, 23 workshops and 322 papers, the attendees should be spoiled for choice. The city also has much to offer as does the highlight of our social program that takes place onboard a legendary aircraft carrier – the Intrepid Museum.

We trust that you will find ICML 2016 to be an enjoyable and memorable event.

With best wishes from the Local Chairs,

Peder Olsen and Marek Petrik

GENERAL INFORMATION

CONFERENCE VENUE
ICML will be held in the Marriott Marquis hotel located right in the middle of the iconic Times Square in New York City.

REGISTRATION HOURS
NY Marriott Marquis Hotel

Sunday, June 19:	7:30am – 3:00pm	7th floor
Monday, June 20:	7:30am – 6:00pm	5th floor
Tuesday, June 21:	8:00am – 6:00pm	5th floor
Wednesday, June 22:	8:00am – 4:30pm	5th floor
Thursday, June 23:	7:30am – 5:00pm	7th floor
Friday, June 24:	8:00am – noon	7th floor

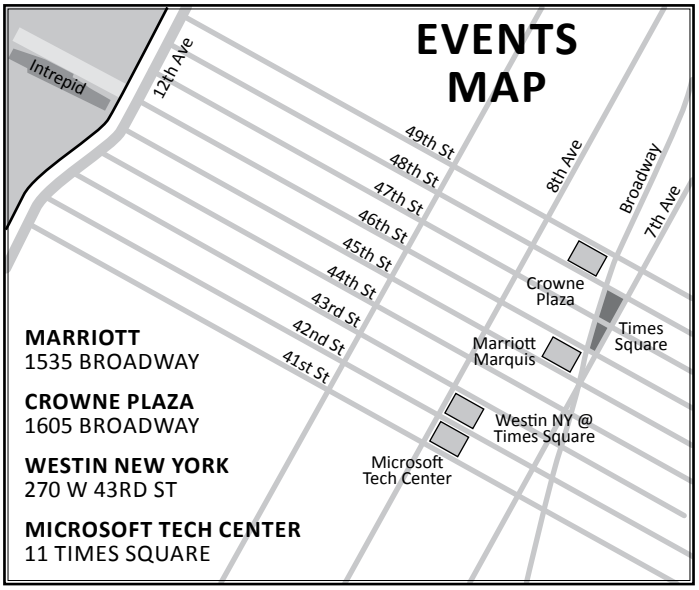
SUNDAY RECEPTION
Sunday in Broadway Lounge, 8th floor of Marriott, and takes place 6 pm - 8 pm. (ticket holders only)

WEDNESDAY RECEPTION
Join us at the Intrepid Sea, Air & Space Museum for ICML's Networking Reception on June 22nd from 7 pm - 10 pm on Hangar 2 & 3.
(For main conference registrants only)

The reception will be focused in Hangar 2 + Hangar 3 of the Hangar Deck

Guests are welcome to explore the Flight Deck throughout the evening (no food or drink).

Please see the map for walking directions from the NY Marriott Marquis on 7th Avenue to the Intrepid Museum on 12th Avenue



POSTER SESSIONS
The poster sessions will be in NY Marriott Marquis: Astor, Times Square, and Duffy.

Monday, June 20	3:00 pm – 7:00 pm
Tuesday, June 21	10:00 am – 1:00 pm
Tuesday, June 21	3:00 pm – 7:00 pm
Wednesday, June 22	10:00 am – 1:00 pm

MOBILE APP
Step 1: Download and install the Whova app from App Store (for iPhones) or Google Play (for Android phones).
Step 2: Sign up in the app using the email address you registered with.
Step 3: You're all set.

Now you will be able to:
- View the event agenda and plan your schedule.
- If you set up your own profile, you can send in-app messages and exchange contact information
- Receive update notifications from organizers.
- Access agenda, maps, and directions.

After downloading, sign up on Whova with the email address that you used to RSVP for our event, or sign up using your social media accounts. If you are asked to enter an invitation code to join the event, please use the following invitation code: "icml"

EXHIBITOR HOURS

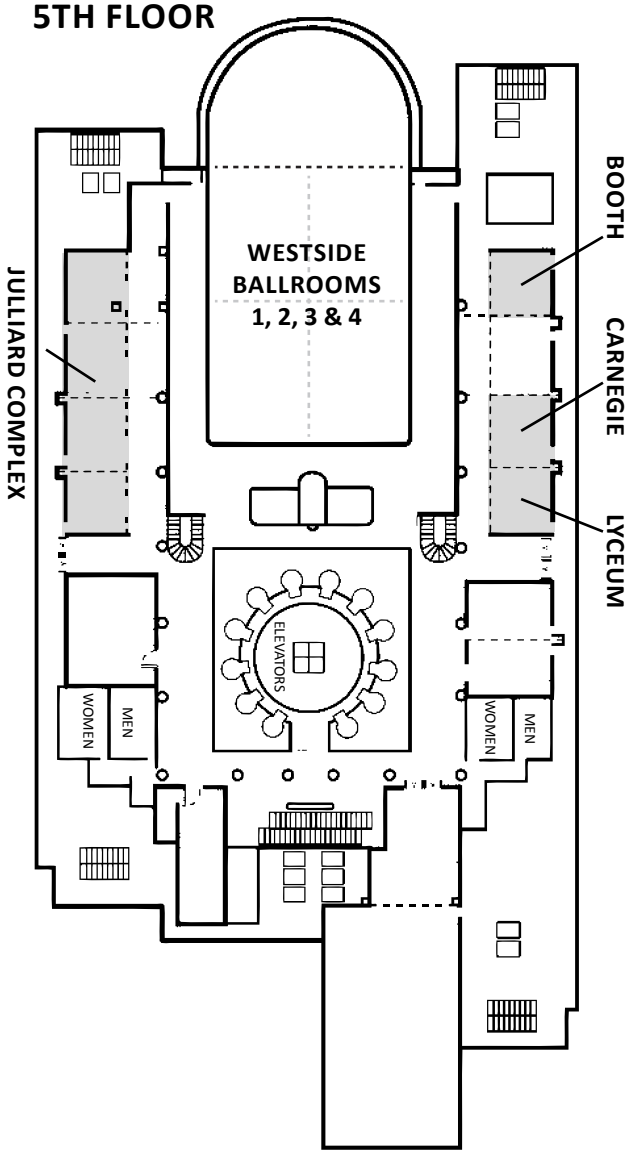
Monday	10:00 am - 7:00 pm
Tuesday	8:30 am - 7:00 pm
Wednesday	8:30 am - 6:00 pm
Thursday	8:30 am - 6:00 pm
Friday	8:30 am - 3:30 pm

LOCAL ATTRACTIONS
Please see http://www.nycgo.com/ for local NYC events and attractions.

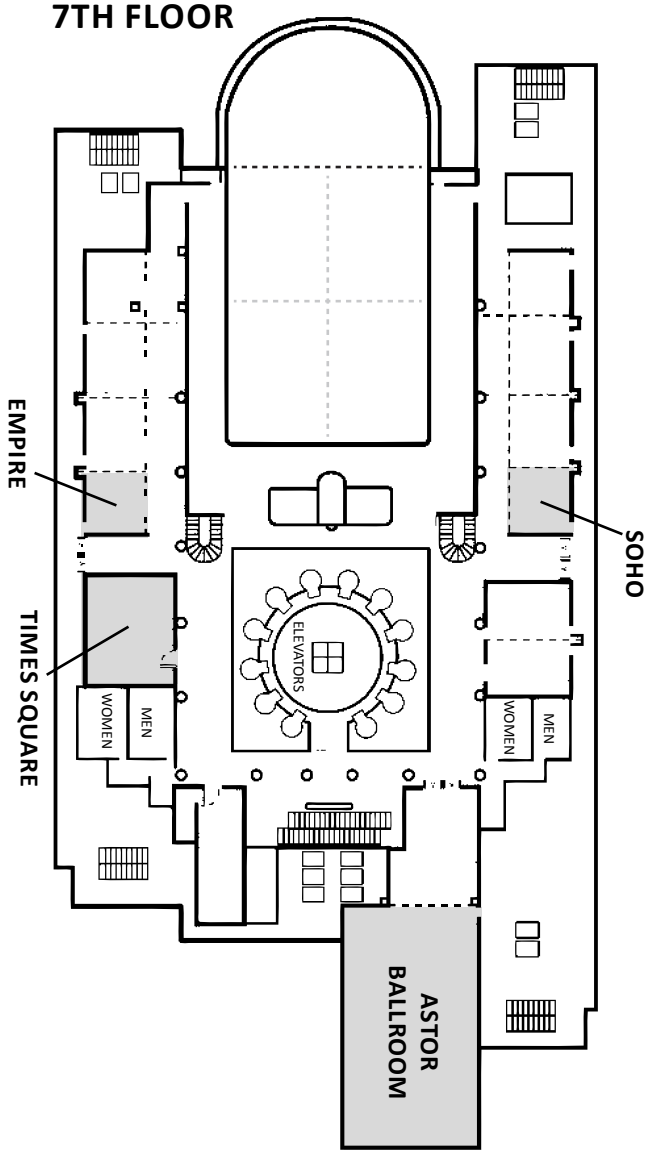
NEARBY RESTAURANTS
Manhattan has an unparalleled variety of restaurants of American and international cuisine. Hells Kitchen neighborhood near Times Square is very popular with locals. Some nearby restaurants include:

Toloache Mexican Grill	Trattoria Trecolori
Totto Ramen	Churrascaria Plataforma
Ippudo Westside	Uncle Vanya Cafe

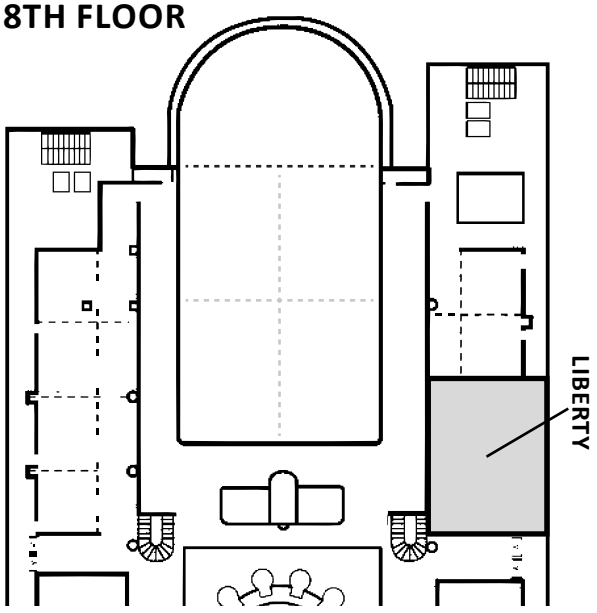
5TH FLOOR



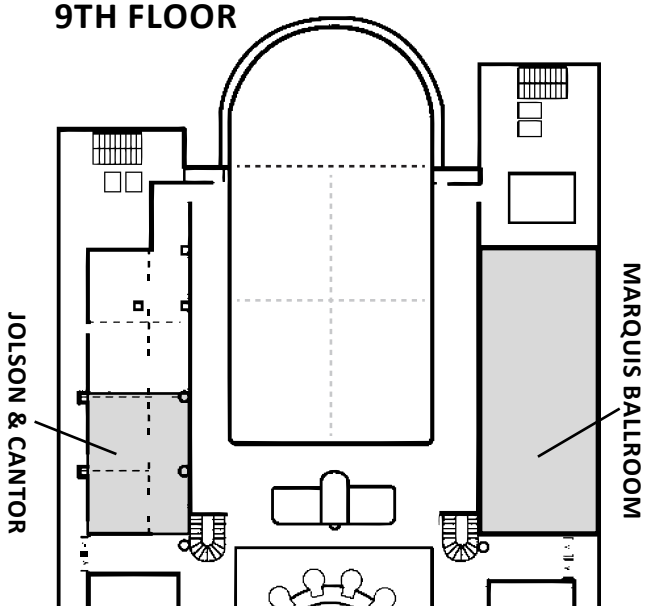
7TH FLOOR



8TH FLOOR



9TH FLOOR



ORGANIZING COMMITTEE

General chair: John Langford (Microsoft Research)

Program chairs: Nina Balcan (CMU) and Kilian Weinberger (Cornell University)

Local organization chairs: Peder Olsen (IBM Research) and Marek Petrik (IBM Research)

Tutorial chairs: Alina Beygelzimer (Yahoo! Labs) and Bernhard Schoelkopf (Max Planck Institute)

Workshop chair: Ruslan Salakhutdinov (University of Toronto) and Fei Sha (USC)

Financial chairs: John Cunningham (Columbia University), Gert Lanckriet (UCSD) and Robert Schapire (Microsoft Research)

Publication chairs: Dan Roy (University of Toronto) and David Sontag (NYU)

Workflow chairs: Jacob Gardner (Cornell) and Matthew Kusner (WUSTL)

Publicity chair: Jingrui He (Stevens Institute of Technology)

Webpage chair: Jérémie Mary (Univ. Lille / Inria)

AREA CHAIRS

Abernethy, Jacob — Univ. of Michigan
Adams, Ryan — Harvard University
Agarwal, Alekh — Microsoft Research
Anandkumar, Animashree — UC Irvine
Awasthi, Pranjal — Rutgers
Bach, Francis — INRIA-ENS
Bartlett, Peter — UC Berkeley
Bengio, Samy — Google
Bengio, Yoshua — University of Montreal
Beygelzimer, Alina — Yahoo! Labs
Bilmes, Jeff — Univ. of Washington
Blitzer, John — Google
Bubeck, Sebastien — Microsoft Rsrch.
Cesa-Bianchi, Nicolo — Univ. of Milano
Chudhuri, Kamalika — UC San Diego
Chechik, Gal — Bar Ilan University
Cho, Kyunghyun — New York University
Collobert, Ronan — Facebook
Cortes, Corinna — Google
Crammer, Koby — Technion
Cuturi, Marco — Kyoto University
Darrell, Trevor — UC Berkeley
Daume, Hal — University of Maryland
Blei, David — Columbia
Elkan, Charles — UC San Diego
Terzi, Evimaria — Boston University
Fergus, Rob — Facebook
Fuernkranz, Johannes — Technische Universität Darmstadt
Garnett, Roman — Washington University
Globerson, Amir — Hebrew University
Gordon, Geoff — Carnegie Mellon Univ.
Grangier, David — Facebook
Grauman, Kristen — UT Austin
Gupta, Maya — Google
Harchaoui, Zaid — INRIA-LEAR
Hardt, Moritz — Google
Heller, Katherine — Duke University
Hsu, Daniel — Columbia
Jegelka, Stefanie — MIT
Kalai, Adam — Microsoft
Kale, Satyen — Yahoo! Labs
Kegl, Balazs — Université Paris-Saclay
Kowk, James — Hong Kong University
Krause, Andreas — ETH Zurich
Kulesza, Alex — University of Michigan
Kulis, Brian — Boston University
Lacoste-Julien, Simon — INRIA
Larochelle, Hugo — Twitter
Le Roux, Nicolas — Criteo
Lee, Honglak — University of Michigan
Li, Lihong — Microsoft Research
Li, Ping — Rutgers
Liang, Yingyu — Princeton University
Lin, Chih-Jen — National Taiwan U
Long, Phil — Sentient Technologies
Mannor, Shie — Technion
McAuliffe, Jon — Berkeley
Meila, Marina — Univ. of Washington
Niculescu-Mizil, Alexandru — NEC Labs
Pineau, Joelle — McGill University
Pontil, Massimiliano — UC London
Precup, Doina — McGill University
Ranzato, Marc'Aurelio — Facebook
Ravikumar, Pradeep — UT Austin
Rish, Irina — IBM Research
Roth, Dan — University of Illinois U-C
Roth, Volker — University of Basel
Roy, Daniel — University of Toronto
Vishwanathan, S.V.N. — UC Santa Cruz
Salakhudinov, Ruslan — Univ. of Toronto
Saligrama, Venkatesh — Boston Univ
Saria, Suchi — Johns Hopkins Univ.
Seeger, Matthias — Amazon Research
Sha, Fei — USC
Shalev-Shwartz, Shai — Hebrew University
Shamir, Ohad — Weizman Institute
Song, Le — Georgia Tech
Sontag, David — New York University
Sra, Suvrit — MIT
Sridharan, Karthik — Cornell University
Srebro, Nati — Toyota Technological Institute / University of Chicago
Stone, Peter — UT Austin
Sutton, Rich — University of Alberta
Szsepesvari, Csaba — Univ. of Alberta
Urner, Ruth — Max Planck Institute
Urtasun, Raquel — University of Toronto
Van der Maaten, Laurens — Facebook
Varma, Manik — Microsoft Research
Wallach, Hanna — Microsoft Research
Williams, Chris — University of Edinburgh
Xing, Eric — Carnegie Mellon Univ.
Yue, Yisong — Caltech
Zhang, Tong — Baidu and Rutgers
Zheng, Alice — GraphLab
Zhu, Jun — Tsinghua University
Zhu, Xiaojin — University of Wisconsin
Yang, Yiming — Carnegie Mellon Univ.

LOCAL ORGANIZATION COMMITTEE

Naoki Abe
Aurelie Lozano
Dmitry Malioutov
Steven Rennie
Mary Ellen Perry
Priscila Rasmussen

EXECUTIVE EVENTS TEAM

Miki Hodge
Roxane Rose
Jody Anagnos
Shannon Cunningham

Sunday

JUNE 19TH | TUTORIALS

TUTORIAL SESSION 1 - 8:30 AM - 10:30 AM

Causal inference for Observational Studies

David Sontag - New York University
Uri Shalit - New York University
Location: Crowne Plaza - Broadway Ballroom

In many fields such as healthcare, education, and economics, policy makers have increasing amounts of data at their disposal. Making policy decisions based on this data often involves causal questions: Does medication X lead to lower blood sugar, compared with medication Y? Does longer maternity leave lead to better child social and cognitive skills? These questions have to be addressed in practice, every day, by scientists working across many different disciplines.

The goal of this tutorial is to bring machine learning practitioners closer to the vast field of causal inference as practiced by statisticians, epidemiologists and economists. We believe that machine learning has much to contribute in helping answer such questions, especially given the massive growth in the available data and its complexity. We also believe the machine learning community could and should be highly interested in engaging with such problems, considering the great impact they have on society in general.

We hope that participants in the tutorial will: a) learn the basic language of causal inference as exemplified by the two most dominant paradigms today: the potential outcomes framework, and causal graphs; b) understand the similarities and the differences between problems machine learning practitioners usually face and problems of causal inference; c) become familiar with the basic tools employed by practicing scientists performing causal inference, and d) be informed about the latest research efforts in bringing machine learning techniques to address problems of causal inference.

Deep Residual Networks: Deep Learning Gets Way Deeper

Kaiming He - Microsoft Research
Location: Marriott - Astor
Live Simulcast: Marriott (Empire & Cantor)

Deeper neural networks are more difficult to train. Beyond a certain depth, traditional deeper networks start to show severe underfitting caused by optimization difficulties. This tutorial will describe the recently developed residual learning framework, which eases the training of networks that are substantially deeper than those used previously. These residual networks are easier to converge, and can gain accuracy from considerably increased depth. On the ImageNet dataset we evaluate residual nets with depth of up to 152 layers—8x deeper than VGG nets but still having lower complexity. These deep residual networks are the foundations of our 1st-place winning entries in all five main tracks in ImageNet and COCO 2015 competitions, which cover image classification, object detection, and semantic segmentation.

In this tutorial we will further look into the propagation formulations of residual networks. Our latest work reveals that when the residual networks have identity mappings as skip connections and inter-block activations, the forward and backward signals can be directly propagated from one block to any other block. This leads us to promising results of 1001-layer residual networks. Our work suggests that there is much room to exploit the dimension of network depth, a key to the success of modern deep learning.

The Convex Optimization, Game-Theoretic Approach To Learning

Elad Hazan - Princeton University
Satyen Kale - Yahoo Research
Location: Marriott - Soho
Live Simulcast: Marriott (Duffy & Times Square)

In recent years convex optimization and the notion of regret minimization in games have been combined and applied to machine learning in a general framework called online convex optimization. We will survey the basics of this framework, its applications, main algorithmic techniques and future research directions.

TUTORIAL SESSION 2 - 11 AM - 1 PM

Memory Networks for Language Understanding

Jason Weston - Facebook
Location: Crowne Plaza - Broadway Ballroom

There has been a recent resurgence in interest in the use of the combination of reasoning, attention and memory for solving tasks, particularly in the field of language understanding. I will review some of these recent efforts, as well as focusing on one of my own group's contributions, memory networks, an architecture that we have applied to question answering, language modeling and general dialog. As we try to move towards the goal of true language understanding, I will also discuss recent datasets and tests that have been built to assess these models abilities to see how far we have come.

Stochastic Gradient Methods for Large-Scale Machine Learning

Leon Bottou - Facebook AI Research
Frank E. Curtis - Lehigh University
Jorge Nocedal - Northwestern University
Location: Marriott - Astor
Live Simulcast: Marriott (Empire & Cantor)

This tutorial provides an accessible introduction to the mathematical properties of stochastic gradient methods and their consequences for large scale machine learning. After reviewing the computational needs for solving optimization problems in two typical examples of large scale machine learning, namely, the training of sparse linear classifiers and deep neural networks, we present the theory of the simple, yet versatile stochastic gradient algorithm, explain its theoretical and practical behavior, and expose the opportunities available for designing improved algorithms. We then provide specific examples of advanced algorithms to illustrate the two essential directions for improving stochastic gradient methods, namely, managing the noise and making use of second order information.

Rigorous Data Dredging: Theory and Tools for Adaptive Data Analysis

Moritz Hardt - Google
Aaron Roth - University of Pennsylvania
Location: Marriott - Soho
Live Simulcast: Marriott (Duffy & Times Square)

Reliable tools for inference and model selection are necessary in all applications of machine learning and statistics. Much of the existing theory breaks down in the now common situation where the data analyst works interactively with the data, adaptively choosing which methods to use by probing the same data many times. We illustrate the problem through the lens of machine learning benchmarks, which currently all rely on the standard holdout method. After understanding why and when the standard holdout method fails, we will see practical alternatives to the holdout method that can be used many times without losing the guarantees of fresh data. We then transition into the emerging theory on this topic touching on deep connections to differential privacy, compression schemes, and hypothesis testing (although no prior knowledge will be assumed).

TUTORIAL SESSION 3 - 2:30 PM - 4:30 PM

Recent Advances in Non-Convex Optimization

Anima Anandkumar - University of California Irvine
Location: Crowne Plaza, Broadway Ballroom

Most machine learning tasks require solving non-convex optimization. The number of critical points in a non-convex problem grows exponentially with the data dimension. Local search methods such as gradient descent can get stuck in one of these critical points, and therefore, finding the globally optimal solution is computationally hard. Despite this hardness barrier, we have seen many advances in guaranteed non-convex optimization. The focus has shifted to characterizing transparent conditions under which the global solution can be found efficiently. In many instances, these conditions turn out to be mild and natural for machine learning applications. This tutorial will provide an overview of the recent theoretical success stories in non-convex optimization. This includes learning latent variable models, dictionary learning, robust principal component analysis, and so on. Simple iterative methods such as spectral methods, alternating projections, and so on, are proven to learn consistent models with polynomial sample and computational complexity. This tutorial will present main ingredients towards establishing these results. The tutorial with conclude with open challenges and possible path towards tackling them.

Deep Reinforcement Learning

David Silver - Google DeepMind
Location: Marriott - Astor
Live Simulcast: Marriott (Empire & Cantor)

A major goal of artificial intelligence is to create general-purpose agents that can perform effectively in a wide range of challenging tasks. To achieve this goal, it is necessary to combine reinforcement learning (RL) agents with powerful and flexible representations. The key idea of deep RL is to use neural networks to provide this representational power. In this tutorial we will present a family of algorithms in which deep neural networks are used for value functions, policies, or environment models. State-of-the-art results will be presented in a variety of domains, including Atari games, 3D navigation tasks, continuous control domains and the game of Go.

Graph Sketching, Streaming, and Space-Efficient Optimization

Sudipto Guha - University of Pennsylvania
Andrew McGregor - Univ. of Massachusetts Amherst
Location: Marriott - Cantor
Live Simulcast: Marriott (Times Square)

Graphs are one of the most commonly used data representation tools but existing algorithmic approaches are typically not appropriate when the graphs of interest are dynamic, stochastic, or do not fit into the memory of a single machine. Such graphs are often encountered as machine learning techniques are increasingly deployed to manage graph data and large-scale graph optimization problems. Graph sketching is a form of dimensionality reduction for graph data that is based on using random linear projections and exploiting connections between linear algebra and combinatorial structure. The technique has been studied extensively over the last five years and can be applied in many computational settings. It enables small-space online and data stream computation where we are permitted only a few passes (ideally only one) over an input sequence ofupdates to a large underlying graph. The technique parallelizes easily and can naturally be applied in various distributed settings. It can also be used in the context of convex programming to enable more efficient algorithms for combinatorial optimization problems such as correlation clustering.One of the main goals of the research on graph sketching is understanding and characterizing the types of graph structure and features that can be inferred from compressed representations of the relevant graphs.

INVITED SPEAKERS

Susan Athey – Stanford Graduate School of Business
Causal Inference for Policy Evaluation

Susan Athey is The Economics of Technology Professor at Stanford Graduate School of Business. She received her bachelor's degree from Duke University and her Ph.D. from Stanford, and she holds an honorary doctorate from Duke University. She previously taught at the economics departments at MIT, Stanford and Harvard. In 2007, Professor Athey received the John Bates Clark Medal, awarded by the American Economic Association to "that American economist under the age of forty who is adjudged to have made the most significant contribution to economic thought and knowledge." She was elected to the National Academy of Science in 2012 and to the American Academy of Arts and Sciences in 2008. Professor Athey's research focuses on the economics of the internet, online advertising, the news media, marketplace design, virtual currencies and the intersection of computer science, machine learning and economics. She advises governments and businesses on marketplace design and platform economics, notably serving since 2007 as a long-term consultant to Microsoft Corporation in a variety of roles, including consulting chief economist.

Abstract:
A variety of scientific problems require the researcher to evaluate the causal effect of a policy or intervention, such as giving a drug to a patient, changing a government policy such as the minimum wage, exposing a user to an advertisement, or releasing a new algorithm to users in an online service. This talk will review a series of recently developed statistical methods for causal inference in settings with many covariates. We consider approaches to estimating average effects of a policy in observational data as well as approaches for estimating heterogeneous treatment effects and personalized policies in randomized experiments. We show how popular methods such as regression trees and random forests can be adapted and optimized to produce estimates of treatment effects as well as confidence intervals.

Christos Faloutsos – Carnegie Mellon University
Mining Large Graphs: Patterns, Anomalies, and Fraud Detection

Christos Faloutsos is a Professor at Carnegie Mellon University. He has received the Presidential Young Investigator Award by the National Science Foundation (1989), the Research Contributions Award in ICDM 2006, the SIGKDD Innovations Award (2010), 22 "best paper" awards (including four "test of time" awards), and four teaching awards.

Six of his advisees have attracted KDD or SCS dissertation awards. He is an ACM Fellow, he has served as a member of the executive committee of SIGKDD; he has published over 300 refereed articles, 17 book chapters and two monographs. He holds nine patents and he has given over 40 tutorials and over 20 invited distinguished lectures. His research interests include large-scale data mining, for graphs and streams; networks, fractals, and multimedia databases.

Abstract:
Given a large graph, like who-calls-whom, or who-likes-whom, what behavior is normal and what should be surprising, possibly due to fraudulent activity? How do graphs evolve over time? We focus on these topics:
(a) anomaly detection in large static graphs and
(b) patterns and anomalies in large time-evolving graphs.

For the first, we present a list of static and temporal laws, we show how to use them to spot suspicious activities, in on-line buyer-and-seller settings, in FaceBook, in twitter-like networks. For the second, we show how to handle time-evolving graphs as tensors, as well as some discoveries such settings.

Fei-Fei Li – Stanford University
A Quest for Visual Intelligence in Computers

Dr. Fei-Fei Li is an Associate Professor in the Computer Science Department at Stanford, and the Director of the Stanford Artificial Intelligence Lab and the Stanford Vision Lab. She is also the Director of the recently established Stanford Toyota Center for Human-Centric AI Research. Dr. Fei-Fei Li's main research areas are in machine learning, computer vision and cognitive and computational neuroscience. She has published more than 100 scientific articles in top-tier journals and conferences, including Nature, PNAS, Journal of Neuroscience, CVPR, ICCV, NIPS, ECCV, IJCV, IEEE-PAMI, etc. Dr. Fei-Fei Li obtained her B.A. degree in physics from Princeton in 1999 with High Honors, and her PhD degree in electrical engineering from California Institute of Technology (Caltech) in 2005. She joined Stanford in 2009 as an assistant professor, and was promoted to associate professor with tenure in 2012. Prior to that, she was on faculty at Princeton University (2007-2009) and University of Illinois Urbana-Champaign (2005-2006). Dr. Fei-Fei Li is a speaker at the TED2015 main conference, a recipient of the 2014 IBM Faculty Fellow Award, 2011 Alfred Sloan Faculty Award, 2012 Yahoo Labs FREP award, 2009 NSF CAREER award, the 2006 Microsoft Research New Faculty Fellowship and a number of Google Research awards. Work from Dr. Li's lab have been featured in a variety of popular press magazines and newspapers including New York Times, Science, Wired Magazine, and New Scientists.

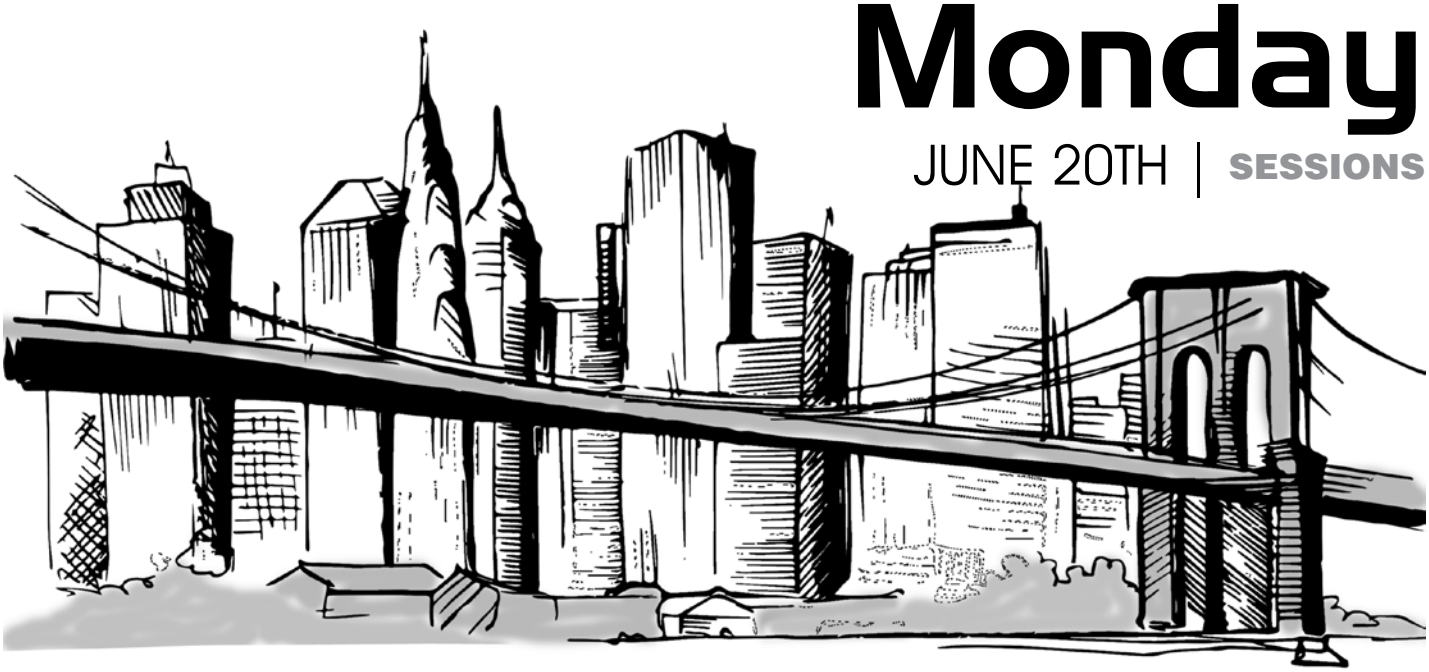
Abstract: It takes nature and evolution more than five hundred million years to develop a powerful visual system in humans. The journey for AI and computer vision is about fifty years. In this talk, I will briefly discuss the key ideas and the cutting edge advances in the quest for visual intelligences in computers. I will particularly focus on the latest work developed in my lab for both image and video understanding, powered by big data and the deep learning (a.k.a. neural network) architecture.

Daniel Spielman – Yale University
Laplacian Matrices of Graphs: Algorithms and Application

Daniel Alan Spielman received his B.A. in Mathematics and Computer Science from Yale in 1992, and his Ph.D in Applied Mathematics from M.I.T. in 1995. He spent a year as a NSF Mathematical Sciences Postdoc in the Computer Science Department at U.C. Berkeley, and then taught in the Applied Mathematics Department at M.I.T. until 2005. Since 2006, he has been a Professor at Yale University. He is presently the Henry Ford II Professor of Computer Science, Mathematics, and Applied Mathematics.

He has received many awards, including the 1995 ACM Doctoral Dissertation Award, the 2002 IEEE Information Theory Paper Award, the 2008 and 2015 Godel Prize, the 2009 Fulkerson Prize, the 2010 Nevanlinna Prize, the 2014 Polya Prize, an inaugural Simons Investigator Award, and a MacArthur Fellowship. He is a Fellow of the Association for Computing Machinery and a member of the Connecticut Academy of Science and Engineering. His main research interests include the design and analysis of algorithms, network science, machine learning, digital communications and scientific computing.

Abstract:
The Laplacian matrices of graphs arise in fields including Machine Learning, Computer Vision, Optimization, Computational Science, and of course Network Analysis. We will explain what these matrices are and why they arise in so many applications. In particular, we will show how Laplacian system solvers can be used to quickly solve linear programs arising from natural graph problems.



Monday

JUNE 20TH | SESSIONS

TIME	DESCRIPTION	MARRIOT HOTEL
8:30 am - 8:40 am	Welcome	
8:40 am - 9:40 am	**Invited Talk: Susan Athey: Causal Inference for Policy Evaluation**	
9:40 am - 10:20 am	Break	
10:20 am - 12:20 pm break from 11:10 - 11:30 am	Neural Networks & Deep Learning	Westside Ballroom 1&2
	Optimization (Continuous)	Westside Ballroom 3&4
	Reinforcement Learning	Marquis
	Online Learning	Lyceum
	Clustering	Empire
	Bayesian Nonparametric Methods	Soho
	Matrix Factorization / Neuroscience Applications	Liberty
12:20 pm - 2 pm	Lunch (On Your Own)	
2 pm - 4:00 pm break from 2:50 - 3:10 pm	Neural Networks & Deep Learning	Westside Ballroom 1&2
	Optimization / Online Learning	Westside Ballroom 3&4
	Machine Learning Applications	Marquis
	Matrix Factorization and Related Topics	Lyceum
	Bandit Problems	Empire
	Graphical Models	Soho
	Transfer Learning / Learning Theory	Liberty
3 pm - 7 pm	Poster Sessions	Astor, Duffy, & Times Square
4 pm - 4:15 pm	Break	
4:15 pm - 6:15 pm break from 5:05 - 5:25 pm	Neural Networks & Deep Learning I	Westside Ballroom 1&2
	Neural Networks & Deep Learning II	Westside Ballroom 3&4
	Approximate Inference	Marquis
	Metric and Manifold Learning / Kernel Methods	Lyceum
	Statistical Learning Theory	Empire
	Structured Prediction / Monte Carlo Methods	Soho
	Online Learning	Liberty

SESSIONS - 10:20 AM - 12:20 PM - MARRIOTT HOTEL

Neural Networks & Deep Learning
Location: Westside Ballroom 1 & 2 + Juliard

- **One-Shot Generalization in Deep Generative Models**
  Danilo Rezende, Shakir Mohamed, Ivo Danihelka, Karol Gregor, Daan Wierstra
- **Learning to Generate with Memory**
  Chongxuan Li, Jun Zhu, Bo Zhang
- **A Theory of Generative ConvNet**
  Jianwen Xie, Yang Lu, Song-Chun Zhu, Yingnian Wu
- **Deconstructing the Ladder Network Architecture**
  Mohammad Pezeshki, Linxi Fan, Philémon Brakel, Aaron Courville, Yoshua Bengio
- **Normalization Propagation: A Parametric Technique for Removing Internal Covariate Shift in Deep Networks**
  Devansh Arpit, Yingbo Zhou, Bhargava Kota, Venu Govindaraju
- **Unitary Evolution Recurrent Neural Networks**
  Martin Arjovsky, Amar Shah, Yoshua Bengio

Reinforcement Learning
Location: Westside Ballroom 3 & 4

- **Why Most Decisions Are Easy in Tetris—And Perhaps in Other Sequential Decision Problems, As Well**
  Özgür Şimşek, Simón Algorta, Amit Kothiyal
- **Opponent Modeling in Deep Reinforcement Learning**
  He He, Jordan Boyd-Graber, Kevin Kwok, Hal Daumé III
- **Memory-based Control of Active Perception and Action in Minecraft**
  Junhyuk Oh, Valliappa Chockalingam, Satinder Singh, Honglak Lee
- **Graying the black box: Understanding DQNs**
  Tom Zahavy, Nir Ben-Zrihem, Shie Mannor
- **Benchmarking Deep Reinforcement Learning for Continuous Control**
  Yan Duan, Xi Chen, Rein Houthooft, John Schulman, Pieter Abbeel
- **Dueling Network Architectures for Deep Reinforcement Learning**
  Ziyu Wang, Tom Schaul, Matteo Hessel, Hado van Hasselt, Marc Lanctot, Nando de Freitas

Optimization (Continuous)
Location: Marquis

- **SDCA without Duality, Regularization, and Individual Convexity**
  Shai Shalev-Shwartz
- **Stochastic Variance Reduction for Nonconvex Optimization**
  Sashank J. Reddi, Ahmed Hefny, Suvrit Sra, Barnabás Póczós, Alex Smola
- **A Fast Rate Analysis of Some Stochastic Optimization Algorithms**
  Chao Qu, Huan Xu, Chong jin Ong
- **Black-box optimization with a politician**
  Sebastien Bubeck, Yin Tat Lee
- **Starting Small - Learning with Adaptive Sample Sizes**
  Hadi Daneshmand, Aurelien Lucchi, Thomas Hofmann
- **Primal-Dual Rates and Certificates**
  Celestine Dünner, Simone Forte, Martin Takac, Martin Jaggi

Online Learning
Location: Lyceum

- **Online Learning with Feedback Graphs Without the Graphs**
  Alon Cohen, Tamir Hazan, Tomer Koren
- **Efficient Algorithms for Adversarial Contextual Learning**
  Vasilis Syrgkanis, Akshay Krishnamurthy, Robert Schapire
- **BISTRO: An Efficient Relaxation-Based Method for Contextual Bandits**
  Alexander Rakhlin, Karthik Sridharan
- **Online Stochastic Linear Optimization under One-bit Feedback**
  Lijun Zhang, Tianbao Yang, Rong Jin, Yichi Xiao, Zhi-hua Zhou
- **Tracking Slowly Moving Clairvoyant: Optimal Dynamic Regret of Online Learning with True and Noisy Gradient**
  Tianbao Yang, Lijun Zhang, Rong Jin, Jinfeng Yi
- **Adaptive Algorithms for Online Convex Optimization with Long-term Constraints**
  Rodolphe Jenatton, Jim Huang, Cedric Archambeau

SESSIONS - 10:20 AM - 12:20 PM - MARRIOTT HOTEL

Clustering
Location: Empire

- **Correlation Clustering and Biclustering with Locally Bounded Errors**
  Gregory Puleo, Olgica Milenkovic
- **K-Means Clustering with Distributed Dimensions**
  Hu Ding, Lingxiao Huang, Jian Li, Yu Liu
- **Speeding up k-means by approximating Euclidean distances via block vectors**
  Thomas Bottesch, Thomas Bühler, Markus Kächele
- **Fast k-means with accurate bounds**
  James Newling, François Fleuret
- **k-variates++: more pluses in the k-means++**
  Richard Nock, Raphaël Canyasse, Roksana Boreli, Frank Nielsen
- **Compressive Spectral Clustering**
  Nicolas Tremblay, Gilles Puy, Rémi Gribonval, Pierre Vandergheynst

Bayesian Nonparametric Methods
Location: Soho

- **Mixed membership modelling with hierarchical CRMs**
  Gaurav Pandey, Ambedkar Dukkipati
- **Hawkes Processes with Stochastic Excitations**
  Young Lee, Kar Wai Lim, Cheng Soon Ong
- **The Segmented iHMM: A Simple, Efficient Hierarchical Infinite HMM**
  Ardavan Saeedi, Matthew Hoffman, Matthew Johnson, Ryan Adams
- **Markov Latent Feature Models**
  Aonan Zhang, John Paisley
- **Diversity-Promoting Bayesian Learning of Latent Variable Models**
  Pengtao Xie, Jun Zhu, Eric Xing
- **Bayesian Poisson Tucker Decomposition for Learning the Structure of International Relations**
  Aaron Schein, Mingyuan Zhou, Blei David, Hanna Wallach

Matrix Factorization / Neuroscience Applications
Location: Liberty

- **The knockoff filter for FDR control in group-sparse and multitask regression**
  Ran Dai, Rina Barber
- **A Simple and Provable Algorithm for Sparse CCA**
  Megasthenis Asteris, Anastasios Kyrillidis, Oluwasanmi Koyejo, Russell Poldrack
- **Experimental Design on a Budget for Sparse Linear Models and Applications**
  Sathya Narayanan Ravi, Vamsi Ithapu, Sterling Johnson, Vikas Singh
- **Representational Similarity Learning with Application to Brain Networks**
  Urvashi Oswal, Christopher Cox, Matthew Lambon-Ralph, Timothy Rogers, Robert Nowak
- **Dictionary Learning for Massive Matrix Factorization**
  Arthur Mensch, Julien Mairal, Bertrand Thirion, Gaël Varoquaux
- **A Random Matrix Approach to Recurrent Neural Networks**
  Romain Couillet, Gilles Wainrib, Hafiz Tiomoko Ali, Harry Sevi

SESSIONS - 2:00 PM - 4:00 PM - MARRIOTT HOTEL

Neural Networks & Deep Learning
Location: Westside Ballroom 1 & 2 + Juliard

- **End-to-End Speech Recognition in English and Mandarin**
  Dario Amodei, Rishita Anubhai, Eric Battenberg, Carl Case, Jared Casper, Bryan Catanzaro, JingDong Chen, Mike Chrzanowski, Adam Coates, Greg Diamos, Erich Elsen, Jesse Engel, Linxi Fan, Christopher Fougner, Awni Hannun, Billy Jun, Tony Han, Patrick LeGresley, Xiangang Li, Libby Lin, Sharan Narang, Andrew Ng, Sherjil Ozair, Ryan Prenger, Sheng Qian, Jonathan Raiman, Sanjeev Satheesh, David Seetapun, Shubho Sengupta, Chong Wang, Yi Wang, Zhiqian Wang, Bo Xiao, Yan Xie, Dani Yogatama, Jun Zhan, zhenyao Zhu
- **Persistent RNNs: Stashing Recurrent Weights On-Chip**
  Greg Diamos, Shubho Sengupta, Bryan Catanzaro, Mike Chrzanowski, Adam Coates, Erich Elsen, Jesse Engel, Awni Hannun, Sanjeev Satheesh
- **Online Sequence Training of Recurrent Neural Networks with Connectionist Temporal Classification**
  Kyuyeon Hwang, Wonyong Sung
- **Analysis of Deep Neural Networks with Extended Data Jacobian Matrix**
  Shengjie Wang, Abdel-rahman Mohamed, Rich Caruana, Jeff Bilmes, Matthai Plilipose, Matthew Richardson, Krzysztof Geras, Gregor Urban, Ozlem Aslan
- **Understanding and Improving Convolutional Neural Networks via Concatenated Rectified Linear Units**
  Wenling Shang, Kihyuk Sohn, Diogo Almeida, Honglak Lee
- **Pixel Recurrent Neural Networks**
  Aäron van den Oord, Nal Kalchbrenner, Koray Kavukcuoglu

Optimization / Online Learning
Location: Westside Ballroom 3 & 4

- **Shifting Regret, Mirror Descent, and Matrices**
  András György, Csaba Szepesvari
- **Heteroscedastic Sequences: Beyond Gaussianity**
  Oren Anava, Shie Mannor
- **Convergence of Stochastic Gradient Descent for PCA**
  Ohad Shamir
- **Fast Stochastic Algorithms for SVD and PCA: Convergence Properties and Convexity**
  Ohad Shamir
- **Faster Eigenvector Computation via Shift-and-Invert Preconditioning**
  Dan Garber, Elad Hazan, Chi Jin, Sham M. Kakade, Cameron Musco, Praneeth Netrapalli, Aaron Sidford
- **Solving Ridge Regression using Sketched Preconditioned SVRG**
  Alon Gonen, Francesco Orabona, Shai Shalev-Shwartz

Machine Learning Applications
Location: Marquis

- **Bounded Off-Policy Evaluation with Missing Data for Course Recommendation and Curriculum Design**
  William Hoiles, Mihaela van der Schaar
- **Dealbreaker: A Nonlinear Latent Variable Model for Educational Data**
  Andrew Lan, Tom Goldstein, Richard Baraniuk, Christoph Studer
- **Estimating Cosmological Parameters from the Dark-Matter Distribution**
  Siamak Ravanbakhsh, Junier Oliva, Sebastian Fromenteau, Layne Price, Shirley Ho, Jeff Schneider, Barnabás Póczos
- **BASC: Applying Bayesian Optimization to the Search for Global Minima on Potential Energy Surfaces**
  Shane Carr, Roman Garnett, Cynthia Lo
- **Predictive Entropy Search for Multi-objective Bayesian Optimization**
  Daniel Hernández-Lobato, José Miguel Hernández-Lobato, Amar Shah, Ryan Adams
- **Pareto Frontier Learning with Expensive Correlated Objectives**
  Amar Shah, Zoubin Ghahramani

Matrix Factorization and Related Topics
Location: Lyceum

- **Complex Embeddings for Simple Link Prediction**
  Théo Trouillon, Johannes Welbl, Sebastian Riedel, Eric Gaussier, Guillaume Bouchard
- **PAC learning of Probabilistic Automaton based on the Method of Moments**
  Hadrien Glaude, Olivier Pietquin

SESSIONS - 2:00 PM - 4:00 PM - MARRIOTT HOTEL

- **Rich Component Analysis**
  Rong Ge, James Zou
- **Beyond CCA: Moment Matching for Multi-View Models**
  Anastasia Podosinnikova, Francis Bach, Simon Lacoste-Julien
- **Isotonic Hawkes Processes**
  Yichen Wang, Bo Xie, Nan Du, Le Song
- **Non-negative Matrix Factorization under Heavy Noise**
  Chiranjib Bhattacharya, Navin Goyal, Ravindran Kannan, Jagdeep Pani

Bandit Problems
Location: Empire

- **An optimal algorithm for the Thresholding Bandit Problem**
  Andrea Locatelli, Maurilio Gutzeit, Alexandra Carpentier
- **Anytime Exploration for Multi-armed Bandits using Confidence Information**
  Kwang-Sung Jun, Robert Nowak
- **Anytime optimal algorithms in stochastic multi-armed bandits**
  Rémy Degenne, Vianney Perchet
- **PAC Lower Bounds and Efficient Algorithms for The Max K-Armed Bandit Problem**
  Yahel David, Nahum Shimkin
- **Conservative Bandits**
  Yifan Wu, Roshan Shariff, Tor Lattimore, Csaba Szepesvári
- **No-Regret Algorithms for Heavy-Tailed Linear Bandits**
  Andres Munoz Medina, Scott Yang

Graphical Models
Location: Soho

- **Hierarchical Span-Based Conditional Random Fields for Labeling and Segmenting Events in Wearable Sensor Data Streams**
  Roy Adams, Nazir Saleheen, Edison Thomaz, Abhinav Parate, Santosh Kumar, Benjamin Marlin
- **Efficient Multi-Instance Learning for Activity Recognition from Time Series Data Using an Auto-Regressive Hidden Markov Model**
  Xinze Guan, Raviv Raich, Weng-Keen Wong
- **Topographical Features of High-Dimensional Categorical Data and Their Applications to Clustering**
  Chao Chen, Novi Quadrianto
- **Nonlinear Statistical Learning with Truncated Gaussian Graphical Models**
  Qinliang Su, xuejun Liao, changyou Chen, Lawrence Carin
- **Collapsed Variational Inference for Sum-Product Networks**
  Han Zhao, Tameem Adel, Geoff Gordon, Brandon Amos
- **Square Root Graphical Models: Multivariate Generalizations of Univariate Exponential Families which Allow Positive Dependencies**
  David Inouye, Pradeep Ravikumar, Inderjit S. Dhillon

Transfer Learning / Learning Theory
Location: Liberty

- **A New PAC-Bayesian Perspective on Domain Adaptation**
  Pascal Germain, Amaury Habrard, François Laviolette, Emilie Morvant
- **Domain Adaptation with Conditional Transferable Components**
  Mingming Gong, Kun Zhang, Tongliang Liu, Dacheng Tao, Clark Glymour, Bernhard Schölkopf
- **Train faster, generalize better: Stability of stochastic gradient descent**
  Moritz Hardt, Ben Recht, Yoram Singer
- **Accurate Robust and Efficient Error Estimation for Decision Trees**
  Lixin Fan
- **The Teaching Dimension of Linear Learners**
  Ji Liu, Xiaojin Zhu, Hrag Ohannessian
- **Loss factorization, weakly supervised learning and label noise robustness**
  Giorgio Patrini, Frank Nielsen, Richard Nock, Marcello Carioni

SESSIONS - 4:15 PM - 6:15 PM - MARRIOTT HOTEL

Neural Networks & Deep Learning
Location: Westside Ballroom 1 & 2 + Juliard

- **Texture Networks: Feed-forward Synthesis of Textures and Stylized Images**
  Dmitry Ulyanov, Vadim Lebedev, Andrea Vedaldi, Victor Lempitsky
- **Discrete Deep Feature Extraction: A Theory and New Architectures**
  Thomas Wiatowski, Michael Tschannen, Aleksandar Stanic, Philipp Grohs, Helmut Bölcskei
- **Deep Structured Energy Based Models for Anomaly Detection**
  Shuangfei Zhai, Yu Cheng, Weining Lu, Zhongfei Zhang
- **Noisy Activation Functions**
  Caglar Gulcehre, Marcin Moczulski, Misha Denil, Yoshua Bengio
- **A Kronecker-factored approximate Fisher matrix for convolution layers**
  Roger Grosse, James Martens
- **Recurrent Orthogonal Networks and Long-Memory Tasks**
  Mikael Henaff, Arthur Szlam, Yann LeCun

Neural Networks and Deep Learning II (Computer Vision)
Location: Westside Ballroom 3 & 4

- **Group Equivariant Convolutional Networks**
  Taco Cohen, Max Welling
- **Learning End-to-end Video Classification with Rank-Pooling**
  Basura Fernando, Stephen Gould
- **Learning Physical Intuition of Block Towers by Example**
  Adam Lerer, Sam Gross, Rob Fergus
- **Large-Margin Softmax Loss for Convolutional Neural Networks**
  Weiyang Liu, Yandong Wen, Zhiding Yu, Meng Yang
- **Network Morphism**
  Tao Wei, Changhu Wang, Yong Rui, Chang Wen Chen
- **MBA: Multi-Bias Non-linear Activation in Deep Neural Networks**
  Hongyang Li, Wanli Ouyang, Xiaogang Wang

Approximate Inference
Location: Marquis

- **Boolean Matrix Factorization and Noisy Completion via Message Passing**
  Siamak Ravanbakhsh, Barnabás Póczos, Russell Greiner
- **Stochastic Discrete Clenshaw-Curtis Quadrature**
  Nico Piatkowski, Katharina Morik
- **Beyond Parity Constraints: Fourier Analysis of Hash Functions for Inference**
  Tudor Achim, Ashish Sabharwal, Stefano Ermon
- **Variable Elimination in the Fourier Domain**
  Yexiang Xue, Stefano Ermon, Ronan Le Bras, Carla Gomes, Bart Selman
- **Learning and Inference via Maximum Inner Product Search**
  Stephen Mussmann, Stefano Ermon
- **Analysis of Variational Bayesian Factorizations for Sparse and Low-Rank Estimation**
  David Wipf

Metric and Manifold Learning / Kernel Methods
Location: Lyceum

- **Fast k-Nearest Neighbour Search via Dynamic Continuous Indexing**
  Ke Li, Jitendra Malik
- **Geometric Mean Metric Learning**
  Pourya Zadeh, Reshad Hosseini, Suvrit Sra
- **Low-rank tensor completion: a Riemannian manifold preconditioning approach**
  Hiroyuki Kasai, Bamdev Mishra
- **The Variational Nystrom method for large-scale spectral problems**
  Max Vladymyrov, Miguel Carreira-Perpiñán
- **Fast DPP Sampling for Nystrom with Application to Kernel Methods**
  Chengtao Li, Stefanie Jegelka, Suvrit Sra
- **Computationally Efficient Nystr\{o}m Approximation using Fast Transforms**
  Si Si, Cho-Jui Hsieh, Inderjit S. Dhillon

SESSIONS - 4:15 PM - 6:15 PM - MARRIOTT HOTEL

Statistical Learning Theory
Location: Empire

- **Barron and Covers' Theory in Supervised Learning and Its Application to Lasso**
  Masanori Kawakita, Jun'ichi Takeuchi
- **Exact Exponent in Optimal Rates for Crowdsourcing**
  Chao Gao, Yu Lu, Dengyong Zhou
- **Generalization Properties and Implicit Regularization for Multiple Passes SGM**
  Junhong Lin, Raffaello Camoriano, Lorenzo Rosasco
- **Generalized Direct Change Estimation in Ising Model Structure**
  Farideh Fazayeli, Arindam Banerjee
- **Gaussian process nonparametric tensor estimator and its minimax optimality**
  Heishiro Kanagawa, Taiji Suzuki, Hayato Kobayashi, Nobuyuki Shimizu, Yukihiro Tagami
- **Minimum Regret Search for Single- and Multi-Task Optimization**
  Jan Hendrik Metzen

Structured Prediction / Monte Carlo Methods
Location: Soho

- **The Sum-Product Theorem: A Foundation for Learning Tractable Models**
  Abram Friesen, Pedro Domingos
- **Train and Test Tightness of LP Relaxations in Structured Prediction**
  Ofer Meshi, Mehrdad Mahdavi, Adrian Weller, David Sontag
- **Evasion and Hardening of Tree Ensemble Classifiers**
  Alex Kantchelian, J. D. Tygar, Anthony Joseph
- **Importance Sampling Tree for Large-scale Empirical Expectation**
  Olivier Canévet, Cijo Jose, François Fleuret
- **Stratified Sampling Meets Machine Learning**
  Edo Liberty, Kevin Lang, Konstantin Shmakov
- **Scalable Discrete Sampling as a Multi-Armed Bandit Problem**
  Yutian Chen, Zoubin Ghahramani

Online Learning
Location: Liberty

- **Pricing a low-regret seller**
  Hoda Heidari, Mohammad Mahdian, Umar Syed, Sergei Vassilvitskii, Sadra Yazdanbo
- **Multi-Player Bandits -- a Musical Chairs Approach**
  Jonathan Rosenski, Ohad Shamir, Liran Szlak
- **Contextual Combinatorial Cascading Bandits**
  Shuai Li, Baoxiang Wang, Shengyu Zhang, Wei Chen
- **Copeland Dueling Bandit Problem: Regret Lower Bound, Optimal Algorithm, and Computationally Efficient Algorithm**
  Junpei Komiyama, Junya Honda, Hiroshi Nakagawa
- **DCM Bandits: Learning to Rank with Multiple Clicks**
  Sumeet Katariya, Branislav Kveton, Csaba Szepesvári, Zheng Wen
- **Distributed Clustering of Linear Bandits in Peer to Peer Networks**
  Nathan Korda, Balázs Szörényi, Shuai Li

MONDAY JUNE 20TH | POSTER SESSIONS

LOCATION: MARRIOTT - ASTOR, DUFFY + TIMES SQUARE - 3PM - 7PM

#1 One-Shot Generalization in Deep Generative Models
Danilo Rezende, Shakir Mohamed, Ivo Danihelka, Karol Gregor, Daan Wierstra

#2 Learning to Generate with Memory
Chongxuan Li, Jun Zhu, Bo Zhang

#3 A Theory of Generative ConvNet
Jianwen Xie, Yang Lu, Song-Chun Zhu, Yingnian Wu

#4 Deconstructing the Ladder Network Architecture
Mohammad Pezeshki, Linxi Fan, Philémon Brakel, Aaron Courville, Yoshua Bengio

#5 Normalization Propagation: A Parametric Technique for Removing Internal Covariate Shift in Deep Networks
Devansh Arpit, Yingbo Zhou, Bhargava Kota, Venu Govindaraju

#6 Unitary Evolution Recurrent Neural Networks
Martin Arjovsky, Amar Shah, Yoshua Bengio

#7 Sequential decision making under uncertainty: Are most decisions easy?
Özgür Şimşek, Simón Algorta, Amit Kothiyal

#8 Opponent Modeling in Deep Reinforcement Learning
He He, Jordan Boyd-Graber, Kevin Kwok, Hal Daumé III

#9 Memory-based Control of Active Perception and Action in Minecraft
Junhyuk Oh, Valliappa Chockalingam, Satinder Singh, Honglak Lee

#10 Graying the black box: Understanding DQNs
Tom Zahavy, Nir Ben-Zrihem, Shie Mannor

#11 Benchmarking Deep Reinforcement Learning for Continuous Control
Yan Duan, Xi Chen, Rein Houthooft, John Schulman, Pieter Abbeel

#12 Dueling Network Architectures for Deep Reinforcement Learning
Ziyu Wang, Tom Schaul, Matteo Hessel, Hado van Hasselt, Marc Lanctot, Nando de Freitas

#13 SDCA without Duality, Regularization, and Individual Convexity
Shai Shalev-Shwartz

#14 Stochastic Variance Reduction for Nonconvex Optimization
Sashank J. Reddi, Ahmed Hefny, Suvrit Sra, Barnabás Póczós, Alex Smola

#15 Fast Rate Analysis of Some Stochastic Optimization Algorithms
Chao Qu, Huan Xu, Chong jin Ong

#16 Black-box optimization with a politician
Sébastien Bubeck, Yin Tat Lee

#17 Starting Small: Learning with Adaptive Sample Sizes
Hadi Daneshmand, Aurelien Lucchi, Thomas Hofmann

#18 Primal-Dual Rates and Certificates
Celestine Dünner, Simone Forte, Martin Takac, Martin Jaggi

#19 Online Learning with Feedback Graphs Without the Graphs
Alon Cohen, Tamir Hazan, Tomer Koren

#20 Efficient Algorithms for Adversarial Contextual Learning
Vasilis Syrgkanis, Akshay Krishnamurthy, Robert Schapire

#21 BISTRO: An Efficient Relaxation-Based Method for Contextual Bandits
Alexander Rakhlin, Karthik Sridharan

#22 Online Stochastic Linear Optimization under One-bit Feedback
Lijun Zhang, Tianbao Yang, Rong Jin, Yichi Xiao, Zhi-hua Zhou

#23 Tracking Slowly Moving Clairvoyant: Optimal Dynamic Regret of Online Learning with True and Noisy Gradient
Tianbao Yang, Lijun Zhang, Rong Jin, Jinfeng Yi

#24 Adaptive Algorithms for Online Convex Optimization with Long-term Constraints
Rodolphe Jenatton, Jim Huang, Cedric Archambeau

#25 Correlation Clustering and Biclustering with Locally Bounded Errors
Gregory Puleo, Olgica Milenkovic

#26 K-Means Clustering with Distributed Dimensions
Hu Ding, Yu Liu, Lingxiao Huang, Jian Li

#27 Speeding up k-means by approximating Euclidean distances via block vectors
Thomas Bottesch, Thomas Bühler, Markus Kächele

#28 Fast k-means with accurate bounds
James Newling, François Fleuret

#29 k-variates++: more pluses in the k-means++
Richard Nock, Raphaël Canyasse, Roksana Boreli, Frank Nielsen

#30 Compressive Spectral Clustering
Nicolas TREMBLAY, Gilles Puy, Rémi Gribonval, Pierre Vandergheynst

#31 Mixed membership modelling with hierarchical CRMs
Gaurav Pandey, Ambedkar Dukkipati

#32 Hawkes Processes with Stochastic Excitations
Young Lee, Kar Wai Lim, Cheng Soon Ong

#33 The Segmented iHMM: A Simple, Efficient Hierarchical Infinite HMM
Ardavan Saeedi, Matthew Hoffman, Matthew Johnson, Ryan Adams

#34 Markov Latent Feature Models
Aonan Zhang, John Paisley

#35 Diversity-Promoting Bayesian Learning of Latent Variable Models
Pengtao Xie, Jun Zhu, Eric Xing

#36 Bayesian Poisson Tucker Decomposition for Learning the Structure of International Relations
Aaron Schein, Mingyuan Zhou, Blei David, Hanna Wallach

#37 The knockoff filter for FDR control in group-sparse and multitask regression
Ran Dai, Rina Barber

#38 A Simple and Provable Algorithm for Sparse CCA
Megasthenis Asteris, Anastasios Kyrillidis, Oluwasanmi Koyejo, Russell Poldrack

#39 Experimental Design on a Budget for Sparse Linear Models and Applications
Sathya Narayanan Ravi, Vamsi Ithapu, Sterling Johnson, Vikas Singh

#40 Representational Similarity Learning with Application to Brain Networks
Urvashi Oswal, Christopher Cox, Matthew Lambon-Ralph, Timothy Rogers, Robert Nowak

#41 Dictionary Learning for Massive Matrix Factorization
Arthur Mensch, Julien Mairal, Bertrand Thirion, Gaël Varoquaux

#42 A Random Matrix Approach to Recurrent Neural Networks
Romain Couillet, Gilles Wainrib, Hafiz Tiomoko Ali, Harry Sevi

#43 Strongly-Typed Recurrent Neural Networks
David Balduzzi, Muhammad Ghifary

#44 A Convolutional Attention Network for Extreme Summarization of Source Code
Miltiadis Allamanis, Hao Peng, Charles Sutton

#45 Ask Me Anything: Dynamic Memory Networks for Natural Language Processing
Ankit Kumar, Ozan Irsoy, Peter Ondruska, Mohit Iyyer, James Bradbury, Ishaan Gulrajani, Victor Zhong, Romain Paulus, Richard Socher

#46 Dynamic Memory Networks for Visual and Textual Question Answering
Caiming Xiong, Stephen Merity, Richard Socher

#47 Supervised and Semi-Supervised Text Categorization using One-Hot LSTM for Region Embeddings
Rie Johnson, Tong Zhang

#48 PHOG: Probabilistic Model for Code
Pavol Bielik, Veselin Raychev, Martin Vechev

#49 On the Analysis of Complex Backup Strategies in Monte Carlo Tree Search
Piyush Khandelwal, Elad Liebman, Scott Niekum, Peter Stone

#50 Generalization and Exploration via Randomized Value Functions
Ian Osband, Benjamin Van Roy, Zheng Wen

#51 Doubly Robust Off-policy Value Evaluation for Reinforcement Learning
Nan Jiang, Lihong Li

#52 Near Optimal Behavior via Approximate State Abstraction
David Abel, David Hershkowitz, Michael Littman

#53 Model-Free Trajectory Optimization for Reinforcement Learning of Motor Skills
Riad Akrour, Gerhard Neumann, Hany Abdulsamad, Abbas Abdolmaleki

#54 Model-Free Imitation Learning with Policy Optimization
Jonathan Ho, Jayesh Gupta, Stefano Ermon

#55 Algorithms for Optimizing the Ratio of Submodular Functions
Wenruo Bai, Rishabh Iyer, Kai Wei, Jeff Bilmes

#56 Horizontally Scalable Submodular Maximization
Mario Lucic, Olivier Bachem, Morteza Zadimoghaddam, Andreas Krause

#57 Learning Sparse Combinatorial Representations via Two-stage Submodular Maximization
Eric Balkanski, Baharan Mirzasoleiman, Andreas Krause, Yaron Singer

#58 Fast Constrained Submodular Maximization: Personalized Data Summarization
Baharan Mirzasoleiman, Ashwinkumar Badanidiyuru, Amin Karbasi

#59 A Box-Constrained Approach for Hard Permutation Problems
Cong Han Lim, Steve Wright

#60 A Convex Atomic-Norm Approach to Multiple Sequence Alignment and Motif Discovery
Ian En-Hsu Yen, Xin Lin, Jiong Zhang, Pradeep Ravikumar, Inderjit S. Dhillon

#61 Nonparametric canonical correlation analysis
Tomer Michaeli, Weiran Wang, Karen Livescu

#62 The Information Sieve
Greg Ver Steeg, Aram Galstyan

#63 Gromov-Wasserstein Barycenters of Similarity Matrices
Gabriel Peyré, Marco Cuturi, Justin Solomon

#64 Learning Representations for Counterfactual Inference
Fredrik Johansson, Uri Shalit, David Sontag

#65 Why Regularized Auto-Encoders learn Sparse Representation?
Devansh Arpit, Yingbo Zhou, Hung Ngo, Venu Govindaraju

#66 Robust Random Cut Forest Based Anomaly Detection on Streams
Sudipto Guha, Nina Mishra, Gourav Roy, Okke Schrijvers

#67 Mixing Rates for the Alternating Gibbs Sampler over Restricted Boltzmann Machines and Friends
Christopher Tosh

#68 Pliable Rejection Sampling
Akram Erraqabi, Michal Valko, Alexandra Carpentier, Odalric Maillard

#69 A Kernel Test of Goodness of Fit
Kacper Chwialkowski, Heiko Strathmann, Arthur Gretton

#70 A Kernelized Stein Discrepancy for Goodness-of-fit Tests and Model Evaluation
Qiang liu, Jason Lee, Michael Jordan

#71 Additive Approximations in High Dimensional Regression via the SALSA
Kirthevasan Kandasamy, Yaoliang Yu

#72 Doubly Decomposing Nonparametric Tensor Regression
Masaaki Imaizumi, Kohei Hayashi

#73 The Sample Complexity of Subspace Clustering with Missing Data
Daniel Pimentel-Alarcón, Robert Nowak

#74 Robust Principal Component Analysis with Side Information
Kai-Yang Chiang, Cho-Jui Hsieh, Inderjit S. Dhillon

#75 Online Low-Rank Subspace Clustering by Explicit Basis Modeling
Jie Shen, Ping Li, Huan Xu

#76 Provable Non-convex Phase Retrieval with Outliers: Median Truncated Wirtinger Flow
Huishuai Zhang, Yuejie Chi, Yingbin Liang

#77 Estimating Structured Vector Autoregressive Models
Igor Melnyk, Arindam Banerjee

#78 Towards Faster Rates and Oracle Property for Low-Rank Matrix Estimation
Huan Gui, Jiawei Han, Quanquan Gu

#79 Hierarchical Variational Models
Rajesh Ranganath, Dustin Tran, Blei David

#80 A Variational Analysis of Stochastic Gradient Algorithms
Stephan Mandt, Matthew Hoffman, Blei David

#81 Black-Box Alpha Divergence Minimization
José Miguel Hernández-Lobato, Yingzhen Li, Mark Rowland, Thang Bui, Daniel Hernández-Lobato, Richard Turner

#82 Variational inference for Monte Carlo objectives
Andriy Mnih, Danilo Rezende

#83 Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning
Yarin Gal, Zoubin Ghahramani

#84 Auxiliary Deep Generative Models
Lars Maaløe, Casper Kaae Sønderby, Søren Kaae Sønderby, Ole Winther

Tuesday

JUNE 21ST | SESSIONS

TIME	DESCRIPTION	MARRIOT HOTEL
8:30 am - 9:30 am	Invited Talk: Fei-Fei Li	Westside Ballroom 1&2
9:30 am - 9:45 am	Test of Time Award	Westside Ballroom 1&2
9:45 am - 10:30 am	Break	
10:00 am - 1:00 pm	Poster Session	Astor, Duffy, & Times Square
10:30 pm - 12:30 pm	Neural Networks & Deep Learning	Westside Ballroom 1&2
	Reinforcement Learning	Westside Ballroom 3&4
break from 11:20 - 11:40 am	Optimization (Combinatorial)	Marquis
	Unsupervised Learning / Representation Learning	Lyceum
	Sampling / Kernel Methods	Empire
	Sparsity and Compressed Sensing	Soho
	Approximate Inference	Liberty
12:30 pm - 2 pm	Lunch (On Your Own)	
2 pm - 3 pm	Invited Talk: Daniel Spielman	Westside Ballroom 1&2
3 pm - 7 pm	Poster Session	Astor, Duffy, & Times Square
3 pm - 3:40 pm	Break	
3:40 pm - 4:45 pm	Neural Networks & Deep Learning	Westside Ballroom 1&2
	Neural Networks & Deep Learning II	Westside Ballroom 3&4
	Optimization (Continuous)	Marquis
	Reinforcement Learning	Lyceum
	Matrix Factorization and Related Topics	Empire
	Unsupervised Learning / Applications	Soho
	Learning Theory	Liberty
4:45 pm - 5:10 pm	Break	
5:10 pm - 6:15 pm	Neural Networks & Deep Learning 1	Westside Ballroom 1&2
	Neural Networks & Deep Learning II	Westside Ballroom 3&4
	Reinforcement Learning	Marquis
	Optimization (Continuous)	Lyceum
	Large Scale Learning & Big Data	Empire
	Graphical Models	Soho
	Supervised Learning	Liberty

SESSIONS - 10:30 AM - 12:30 PM - MARRIOTT HOTEL

Neural Networks and Deep Learning
Location: Westside Ballroom 1 & 2 + Juliard

- **Strongly-Typed Recurrent Neural Networks**
  David Balduzzi, Muhammad Ghifary
- **A Convolutional Attention Network for Extreme Summarization of Source Code**
  Miltiadis Allamanis, Hao Peng, Charles Sutton
- **Ask Me Anything: Dynamic Memory Networks for Natural Language Processing**
  Ankit Kumar, Ozan Irsoy, Peter Ondruska, Mohit Iyyer, James Bradbury, Ishaan Gulrajani, Victor Zhong, Romain Paulus, Richard Socher
- **Dynamic Memory Networks for Visual and Textual Question Answering**
  Caiming Xiong, Stephen Merity, Richard Socher
- **Supervised and semi-supervised text categorization using LSTM for region embeddings**
  Rie Johnson, Tong Zhang
- **PHOG: Probabilistic Model for Code**
  Pavol Bielik, Veselin Raychev, Martin Vechev

Reinforcement Learning
Location: Westside Ballroon 3 & 4

- **On the Analysis of Complex Backup Strategies in Monte Carlo Tree Search**
  Piyush Khandelwal, Elad Liebman, Scott Niekum, Peter Stone
- **Generalization and Exploration via Randomized Value Functions**
  Ian Osband, Benjamin Van Roy, Zheng Wen
- **Doubly Robust Off-policy Value Evaluation for Reinforcement Learning**
  Nan Jiang, Lihong Li
- **Near Optimal Behavior via Approximate State Abstraction**
  David Abel, David Hershkowitz, Michael Littman
- **Model-Free Trajectory Optimization for Reinforcement Learning of Motor Skills**
  Riad Akrour, Gerhard Neumann, Hany Abdulsamad, Abbas Abdolmaleki
- **Model-Free Imitation Learning with Policy Optimization**
  Jonathan Ho, Jayesh Gupta, Stefano Ermon

Optimization (Combinatorial)
Location: Marquis

- **Algorithms for Optimizing the Ratio of Submodular Functions**
  Wenruo Bai, Rishabh Iyer, Kai Wei, Jeff Bilmes
- **Horizontally Scalable Submodular Maximization**
  Mario Lucic, Olivier Bachem, Morteza Zadimoghaddam, Andreas Krause
- **Learning Sparse Combinatorial Representations via Two-stage Submodular Maximization**
  Eric Balkanski, Baharan Mirzasoleiman, Andreas Krause, Yaron Singer
- **Fast Constrained Submodular Maximization: Personalized Data Summarization**
  Baharan Mirzasoleiman, Ashwinkumar Badanidiyuru, Amin Karbasi
- **A Box-Constrained Approach for Hard Permutation Problems**
  Cong Han Lim, Steve Wright
- **A Convex Atomic-Norm Approach to Multiple Sequence Alignment and Motif Discovery**
  Ian En-Hsu Yen, Xin Lin, Jiong Zhang, Pradeep Ravikumar, Inderjit S. Dhillon

Unsupervised Learning / Representation Learning
Location: Lyceum

- **A Nonparametric canonical correlation analysis**
  Tomer Michaeli, Weiran Wang, Karen Livescu
- **The Information Sieve**
  Greg Ver Steeg, Aram Galstyan
- **Gromov-Wasserstein Barycenters of Similarity Matrices**
  Gabriel Peyré, Marco Cuturi, Justin Solomon
- **Learning Representations for Counterfactual Inference**
  Fredrik Johansson, Uri Shalit, David Sontag
- **Why Regularized Auto-Encoders learn Sparse Representation?**
  Devansh Arpit, Yingbo Zhou, Hung Ngo, Venu Govindaraju
- **Robust Random Cut Forest Based Anomaly Detection on Streams**
  Sudipto Guha, Nina Mishra, Gourav Roy, Okke Schrijvers

SESSIONS - 10:30 AM - 12:30 PM - MARRIOTT HOTEL

Sampling / Kernel Methods
Location: Empire

- **Mixing Rates for the Alternating Gibbs Sampler over Restricted Boltzmann Machines and Friends**
  Christopher Tosh
- **Pliable Rejection Sampling**
  Akram Erraqabi, Michal Valko, Alexandra Carpentier, Odalric Maillard
- **A Kernel Test of Goodness of Fit**
  Kacper Chwialkowski, Heiko Strathmann, Arthur Gretton
- **A Kernelized Stein Discrepancy for Goodness-of-fit Tests and Model Evaluation**
  Qiang liu, Jason Lee, Michael Jordan
- **Additive Approximations in High Dimensional Regression via the SALSA**
  Kirthevasan Kandasamy, Yaoliang Yu
- **Doubly Decomposing Nonparametric Tensor Regression**
  Masaaki Imaizumi, Kohei Hayashi

Approximate Inference
Location: Liberty

- **Hierarchical Variational Models**
  Rajesh Ranganath, Dustin Tran, Blei David
- **A Variational Analysis of Stochastic Gradient Algorithms**
  Stephan Mandt, Matthew Hoffman, Blei David
- **Black-Box Alpha Divergence Minimization**
  José Miguel Hernández-Lobato, Yingzhen Li, Mark Rowland, Thang Bui, Daniel Hernández-Lobato, Richard Turner
- **Variational inference for Monte Carlo objectives**
  Andriy Mnih, Danilo Rezende
- **Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning**
  Yarin Gal, Zoubin Ghahramani
- **Auxiliary Deep Generative Models**
  Lars Maaløe, Casper Kaae Sønderby, Søren Kaae Sønderby, Ole Winther

Sparsity and Compressed Sensing
Location: Soho

- **The Sample Complexity of Subspace Clustering with Missing Data**
  Daniel Pimentel-Alarcón, Robert Nowak
- **Robust Principal Component Analysis with Side Information**
  Kai-Yang Chiang, Cho-Jui Hsieh, Inderjit S. Dhillon
- **Online Low-Rank Subspace Clustering by Explicit Basis Modeling**
  Jie Shen, Ping Li, Huan Xu
- **Provable Non-convex Phase Retrieval with Outliers: Median Truncated Wirtinger Flow**
  Huishuai Zhang, Yuejie Chi, Yingbin Liang
- **Estimating Structured Vector Autoregressive Models**
  Igor Melnyk, Arindam Banerjee
- **Towards Faster Rates and Oracle Property for Low-Rank Matrix Estimation**
  Huan Gui, Jiawei Han, Quanquan Gu

SESSIONS - 3:40 PM - 4:45 PM

Neural Networks and Deep Learning I
Location: Westside Ballroom 1 & 2 + Juliard

- **Factored Temporal Sigmoid Belief Networks for Sequence Learning**
  Jiaming Song, Zhe Gan, Lawrence Carin
- **Bidirectional Helmholtz Machines**
  Jörg Bornschein, Samira Shabanian, Asja Fischer, Yoshua Bengio
- **The Deep Neural Matrix Gaussian Process**
  Christos Louizos, Max Welling
- **Dropout distillation**
  Samuel Rota Bulò, Lorenzo Porzi, Peter Kontschieder

SESSIONS - 3:40 PM - 4:45 PM - MARRIOTT HOTEL

Neural Networks and Deep Learning II
Location: Westside Ballroom 3 & 4

- **Revisiting Semi-Supervised Learning with Graph Embeddings**
  Zhilin Yang, William Cohen, Ruslan Salakhudinov
- **ADIOS: Architectures Deep In Output Space**
  Moustapha Cissé, Maruan Al-Shedivat, Samy Bengio
- **Unsupervised Deep Embedding for Clustering Analysis**
  Junyuan Xie, Ross Girshick, Ali Farhadi
- **Learning Convolutional Neural Networks for Graphs**
  Mathias Niepert, Mohamed Ahmed, Konstantin Kutzkov

Reinforcement Learning
Location: Marquis

- **Inverse Optimal Control with Deep Networks via Policy Optimization**
  Chelsea Finn, Sergey Levine, Pieter Abbeel
- **Smooth Imitation Learning**
  Hoang Le, Andrew Kang, Yisong Yue, Peter Carr
- **Improving the Efficiency of Deep Reinforcement Learning with Normalized Advantage Functions and Synthetic Experience**
  Shixiang Gu, Timothy Lillicrap, Ilya Sutskever, Sergey Levine
- **Asynchronous Methods for Deep Reinforcement Learning**
  Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, Koray Kavukcuoglu

Optimization (Continuous)
Location: Lyceum

- **On the Statistical Limits of Convex Relaxations**
  Zhaoran Wang, Quanquan Gu, Han Liu
- **Faster Convex Optimization: Simulated Annealing with an Efficient Universal Barrier**
  Jacob Abernethy, Elad Hazan
- **A ranking approach to global optimization**
  Cédric Malherbe, Emile Contal, Nicolas Vayatis
- **Epigraph projections for fast general convex programming**
  Po-Wei Wang, Matt Wytock, J. Zico Kolter

Matrix Factorization and Related Topics
Location: Empire

- **Principal Component Projection Without Principal Component Analysis**
  Roy Frostig, Cameron Musco, Christopher Musco, Aaron Sidford
- **Recovery guarantee of weighted low-rank approximation via alternating minimization**
  Yuanzhi Li, Yingyu Liang, Andrej Risteski
- **Tensor Decomposition via Joint Matrix Schur Decomposition**
  Nicolò Colombo, Nikos Vlassis
- **Fast Methods for Estimating the Numerical Rank of Large Matrices**
  Shashanka Ubaru, Yousef Saad

Unsupervised Learning / Applications
Location: Soho

- **Markov-modulated marked Poisson processes for check-in data**
  Jiangwei Pan, Vinayak Rao, Pankaj Agarwal, Alan Gelfand
- **Hierarchical Compound Poisson Factorization**
  Mehmet Basbug, Barbara Engelhardt
- **Dirichlet Process Mixture Model for Correcting Technical Variation in Single-Cell Gene Expression Data**
  Sandhya Prabhakaran, Elham Azizi, Ambrose Carr, Dana Pe'er
- **The Automatic Statistician: A Relational Perspective**
  Yunseong Hwang, Anh Tong, Jaesik Choi

Learning Theory
Location: Liberty

- **Truthful Univariate Estimators**
  Ioannis Caragiannis, Ariel Procaccia, Nisarg Shah
- **Fast Algorithms for Segmented Regression**
  Jayadev Acharya, Ilias Diakonikolas, Jerry Li, Ludwig Schmidt
- **Stochastically Transitive Models for Pairwise Comparisons: Statistical and Computational Issues**
  Nihar Shah, Sivaraman Balakrishnan, Aditya Guntuboyina, Martin Wainwright
- **Provable Algorithms for Inference in Topic Models**
  Sanjeev Arora, Rong Ge, Frederic Koehler, Tengyu Ma, Ankur Moitra

SESSIONS - 5:10 PM - 6:15 PM - MARRIOTT HOTEL

Neural Networks and Deep Learning I
Location: Westside Ballroom 1 & 2 + Juliard

- **Expressiveness of Rectifier Neural Network**
  Xingyuan Pan, Vivek Srikumar
- **Convolutional Rectifier Networks as Generalized Tensor Decompositions**
  Nadav Cohen, Amnon Shashua
- **Fixed Point Quantization of Deep Convolutional Networks**
  Darryl Lin, Sachin Talathi, Sreekanth Annapureddy
- **CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy**
  Ran Gilad-Bachrach, Nathan Dowlin, Kim Laine, Kristin Lauter, Michael Naehrig, John Wernsing

Neural Networks and Deep Learning II
Location: Westside Ballroom 3 & 4

- **Correcting Forecasts with Multi-force Neural Attention**
  Matthew Riemer, Aditya Vempaty, Flavio Calmon, Fenno Heath, Richard Hull, Elham Khabiri
- **Meta-Learning with Memory-Augmented Neural Networks**
  Adam Santoro, Sergey Bartunov, Matthew Botvinick, Daan Wierstra, Timothy Lillicrap
- **Learning Simple Algorithms from Examples**
  Wojciech Zaremba, Tomas Mikolov, Armand Joulin, Rob Fergus
- **Associative Long Short-Term Memory**
  Ivo Danihelka, Greg Wayne, Benigno Uria, Nal Kalchbrenner, Alex Graves

Reinforcement Learning
Location: Marquis

- **Estimating Maximum Expected Value through Gaussian Approximation**
  Carlo D'Eramo, Marcello Restelli, Alessandro Nuara
- **Data-Efficient Off-Policy Policy Evaluation for Reinforcement Learning**
  Philip Thomas, Emma Brunskill
- **Cumulative Prospect Theory Meets Reinforcement Learning: Prediction and Control**
  Prashanth L.A., Cheng Jie, Michael Fu, Steve Marcus, Csaba Szepesvári
- **Softened Approximate Policy Iteration for Markov Games**
  Julien Pérolat, Bilal Piot, Matthieu Geist, Bruno Scherrer, Olivier Pietquin

Optimization (Continuous)
Location: Lyceum

- **Low-rank Solutions of Linear Matrix Equations via Procrustes Flow**
  Stephen Tu, Ross Boczar, Max Simchowitz, mahdi Soltanolkotabi, Ben Recht
- **Quadratic Optimization with Orthogonality Constraints: Explicit Lojasiewicz Exponent and Linear Convergence of Line-Search Methods**
  Huikang Liu, Weijie Wu, Anthony Man-Cho So
- **Efficient Algorithms for Large-scale Generalized Eigenvector Computation and CCA**
  Rong Ge, Chi Jin, Sham M. Kakade, Praneeth Netrapalli, Aaron Sidford
- **Matrix Eigendecomposition via Doubly Stochastic Riemannian Optimization**
  Zhiqiang Xu, Peilin Zhao, Jianneng Cao, Xiaoli Li

SESSIONS - 5:10 PM - 6:15 PM - MARRIOTT HOTEL

Large Scale Learning and Big Data
Location: Empire

- **Extreme F-measure Maximization using Sparse Probability Estimates**
  Kalina Jasinska, Krzysztof Dembczynski, Robert Busa-Fekete, Karlson Pfannschmidt, Timo Klerx, Eyke Hullermeier
- **Stochastic Optimization for Multiview Learning using Partial Least Squares**
  Raman Arora, Poorya Mianjy, Teodor Marinov
- **Gaussian quadrature for matrix inverse forms with applications**
  Chengtao Li, Suvrit Sra, Stefanie Jegelka
- **A Subspace Learning Approach for High Dimensional Matrix Decomposition with Efficient Column/Row Sampling**
  Mostafa Rahmani, Geroge Atia

Graphical Models
Location: Soho

- **Uprooting and Rerooting Graphical Models**
  Adrian Weller
- **Structure Learning of Partitioned Markov Networks**
  Song Liu, Taiji Suzuki, Masashi Sugiyama, Kenji Fukumizu
- **Ensuring Rapid Mixing and Low Bias for Asynchronous Gibbs Sampling**
  Christopher De Sa, Chris Re, Kunle Olukotun
- **Estimation from Indirect Supervision with Linear Moments**
  Aditi Raghunathan, Roy Frostig, John Duchi, Percy Liang

Supervised Learning
Location: Liberty

- **Early and Reliable Event Detection Using Proximity Space Representation**
  Maxime Sangnier, Jérôme Gauthier, Alain Rakotomamonjy
- **Meta--Gradient Boosted Decision Tree Model for Weight and Target Learning**
  Yury Ustinovskiy, Valentina Fedorova, Gleb Gusev, Pavel Serdyukov
- **Class Probability Estimation via Differential Geometric Regularization**
  Qinxun Bai, Steven Rosenberg, Zheng Wu, Stan Sclaroff
- **Linking losses for density ratio and class-probability estimation**
  Aditya Menon, Cheng Soon Ong

#1 **End-to-End Speech Recognition in English and Mandarin**
Dario Amodei, Rishita Anubhai, Eric Battenberg, Carl Case, Jared Casper, Bryan Catanzaro, JingDong Chen, Mike Chrzanowski, Adam Coates, Greg Diamos, Erich Elsen, Jesse Engel, Linxi Fan, Christopher Fougner, Awni Hannun, Billy Jun, Tony Han, Patrick LeGresley, Xiangang Li, Libby Lin, Sharan Narang, Andrew Ng, Sherjil Ozair, Ryan Prenger, Sheng Qian, Jonathan Raiman, Sanjeev Satheesh, David Seetapun, Shubho Sengupta, Chong Wang, Yi Wang, Zhiqian Wang, Bo Xiao, Yan Xie, Dani Yogatama, Jun Zhan, zhenyao Zhu

#2 **Persistent RNNs: Stashing Recurrent Weights On-Chip**
Greg Diamos, Shubho Sengupta, Bryan Catanzaro, Mike Chrzanowski, Adam Coates, Erich Elsen, Jesse Engel, Awni Hannun, Sanjeev Satheesh

#3 **Online Sequence Training of Recurrent Neural Networks with Connectionist Temporal Classification**
Kyuyeon Hwang, Wonyong Sung

#4 **Analysis of Deep Neural Networks with Extended Data Jacobian Matrix**
Shengjie Wang, Abdel-rahman Mohamed, Rich Caruana, Jeff Bilmes, Matthai Plilipose, Matthew Richardson, Krzysztof Geras, Gregor Urban, Ozlem Aslan

#5 **Understanding and Improving Convolutional Neural Networks via Concatenated Rectified Linear Units**
Wenling Shang, Kihyuk Sohn, Diogo Almeida, Honglak Lee

#6 **Pixel Recurrent Neural Networks**
Aäron van den Oord, Nal Kalchbrenner, Koray Kavukcuoglu

#7 **Shifting Regret, Mirror Descent, and Matrices**
András György, Csaba Szepesvari

#8 **Heteroscedastic Sequences: Beyond Gaussianity**
Oren Anava, Shie Mannor

#9 **Convergence of Stochastic Gradient Descent for PCA**
Ohad Shamir

#10 **Fast Stochastic Algorithms for SVD and PCA: Convergence Properties and Convexity**
Ohad Shamir

#11 **Faster Eigenvector Computation via Shift-and-Invert Preconditioning**
Dan Garber, Elad Hazan, Chi Jin, Sham M. Kakade, Cameron Musco, Praneeth Netrapalli, Aaron Sidford

#12 **Solving Ridge Regression using Sketched Preconditioned SVRG**
Alon Gonen, Francesco Orabona, Shai Shalev-Shwartz

#13 **Bounded Off-Policy Evaluation with Missing Data for Course Recommendation and Curriculum Design**
William Hoiles, Mihaela van der Schaar

#14 **Dealbreaker: A Nonlinear Latent Variable Model for Educational Data**
Andrew Lan, Tom Goldstein, Richard Baraniuk, Christoph Studer

#15 **Estimating Cosmological Parameters from the Dark-Matter Distribution**
Siamak Ravanbakhsh, Junier Oliva, Sebastian Fromenteau, Layne Price, Shirley Ho, Jeff Schneider, Barnabás Póczos

#16 **BASC: Applying Bayesian Optimization to the Search for Global Minima on Potential Energy Surfaces**
Shane Carr, Roman Garnett, Cynthia Lo

#17 **Predictive Entropy Search for Multi-objective Bayesian Optimization**
Daniel Hernández-Lobato, José Miguel Hernández-Lobato, Amar Shah, Ryan Adams

#18 **Pareto Frontier Learning with Expensive Correlated Objectives**
Amar Shah, Zoubin Ghahramani

#19 **Complex Embeddings for Simple Link Prediction**
Théo Trouillon, Johannes Welbl, Sebastian Riedel, Eric Gaussier, Guillaume Bouchard

#20 **PAC learning of Probabilistic Automaton based on the Method of Moments**
Hadrien Glaude, Olivier Pietquin

#21 **Rich Component Analysis**
Rong Ge, James Zou

#22 **Beyond CCA: Moment Matching for Multi-View Models**
Anastasia Podosinnikova, Francis Bach, Simon Lacoste-Julien

#23 **Isotonic Hawkes Processes**
Yichen Wang, Bo Xie, Nan Du, Le Song

#24 **Non-negative Matrix Factorization under Heavy Noise**
Chiranjib Bhattacharya, Navin Goyal, Ravindran Kannan, Jagdeep Pani

#25 **An optimal algorithm for the Thresholding Bandit Problem**
Andrea Locatelli, Maurilio Gutzeit, Alexandra Carpentier

#26 **Anytime Exploration for Multi-armed Bandits using Confidence Information**
Kwang-Sung Jun, Robert Nowak

#27 **Anytime optimal algorithms in stochastic multi-armed bandits**
Rémy Degenne, Vianney Perchet

#28 **PAC Lower Bounds and Efficient Algorithms for The Max K-Armed Bandit Problem**
Yahel David, Nahum Shimkin

#29 **Conservative Bandits**
Yifan Wu, Roshan Shariff, Tor Lattimore, Csaba Szepesvári

#30 **No-Regret Algorithms for Heavy-Tailed Linear Bandits**
Andres Munoz Medina, Scott Yang

#31 **Hierarchical Span-Based Conditional Random Fields for Labeling and Segmenting Events in Wearable Sensor Data Streams**
Roy Adams, Nazir Saleheen, Edison Thomaz, Abhinav Parate, Santosh Kumar, Benjamin Marlin

#32 **Efficient Multi-Instance Learning for Activity Recognition from Time Series Data Using an Auto-Regressive Hidden Markov Model**
Xinze Guan, Raviv Raich, Weng-Keen Wong

#33 **Topographical Features of High-Dimensional Categorical Data and Their Applications to Clustering**
Chao Chen, Novi Quadrianto

#34 **Nonlinear Statistical Learning with Truncated Gaussian Graphical Models**
Qinliang Su, xuejun Liao, changyou Chen, Lawrence Carin

#35 **Collapsed Variational Inference for Sum-Product Networks**
Han Zhao, Tameem Adel, Geoff Gordon, Brandon Amos

#36 **Square Root Graphical Models: Multivariate Generalizations of Univariate Exponential Families which Allow Positive Dependencies**
David Inouye, Pradeep Ravikumar, Inderjit S. Dhillon

#37 **A New PAC-Bayesian Perspective on Domain Adaptation**
Pascal Germain, Amaury Habrard, François Laviolette, Emilie Morvant

#38 **Domain Adaptation with Conditional Transferable Components**
Mingming Gong, Kun Zhang, Tongliang Liu, Dacheng Tao, Clark Glymour, Bernhard Schölkopf

#39 Train faster, generalize better: Stability of stochastic gradient descent
Moritz Hardt, Ben Recht, Yoram Singer

#40 Accurate Robust and Efficient Error Estimation for Decision Trees
Lixin Fan

#41 The Teaching Dimension of Linear Learners
Ji Liu, Xiaojin Zhu, Hrag Ohannessian

#42 Loss factorization, weakly supervised learning and label noise robustness
Giorgio Patrini, Frank Nielsen, Richard Nock, Marcello Carioni

#43 Texture Networks: Feed-forward Synthesis of Textures and Stylized Images
Dmitry Ulyanov, Vadim Lebedev, Andrea Vedaldi, Victor Lempitsky

#44 Discrete Deep Feature Extraction: A Theory and New Architectures
Thomas Wiatowski, Michael Tschannen, Aleksandar Stanic, Philipp Grohs, Helmut Bölcskei

#45 Deep Structured Energy Based Models for Anomaly Detection
Shuangfei Zhai, Yu Cheng, Weining Lu, Zhongfei Zhang

#46 Noisy Activation Functions
Caglar Gulcehre, Marcin Moczulski, Misha Denil, Yoshua Bengio

#47 A Kronecker-factored approximate Fisher matrix for convolution layers
Roger Grosse, James Martens

#48 Recurrent Orthogonal Networks and Long-Memory Tasks
Mikael Henaff, Arthur Szlam, Yann LeCun

#49 Group Equivariant Convolutional Networks
Taco Cohen, Max Welling

#50 Learning End-to-end Video Classification with Rank-Pooling
Basura Fernando, Stephen Gould

#51 Learning Physical Intuition of Block Towers by Example
Adam Lerer, Sam Gross, Rob Fergus

#52 Large-Margin Softmax Loss for Convolutional Neural Networks
Weiyang Liu, Yandong Wen, Zhiding Yu, Meng Yang

#53 Network Morphism
Tao Wei, Changhu Wang, Yong Rui, Chang Wen Chen

#54 MBA: Multi-Bias Non-linear Activation in Deep Neural Networks
Hongyang Li, Wanli Ouyang, Xiaogang Wang

#55 Boolean Matrix Factorization and Noisy Completion via Message Passing
Siamak Ravanbakhsh, Barnabás Póczos, Russell Greiner

#56 Stochastic Discrete Clenshaw-Curtis Quadrature
Nico Piatkowski, Katharina Morik

#57 Beyond Parity Constraints: Fourier Analysis of Hash Functions for Inference
Tudor Achim, Ashish Sabharwal, Stefano Ermon

#58 Variable Elimination in the Fourier Domain
Yexiang Xue, Stefano Ermon, Ronan Le Bras, Carla Gomes, Bart Selman

#59 Learning and Inference via Maximum Inner Product Search
Stephen Mussmann, Stefano Ermon

#60 Analysis of Variational Bayesian Factorizations for Sparse and Low-Rank Estimation
David Wipf

#61 Fast k-Nearest Neighbour Search via Dynamic Continuous Indexing
Ke Li, Jitendra Malik

#62 Geometric Mean Metric Learning
Pourya Zadeh, Reshad Hosseini, Suvrit Sra

#63 Low-rank tensor completion: a Riemannian manifold preconditioning approach
Hiroyuki Kasai, Bamdev Mishra

#64 The Variational Nystrom method for large-scale spectral problems
Max Vladymyrov, Miguel Carreira-Perpiñán

#65 Fast DPP Sampling for Nystrom with Application to Kernel Methods
Chengtao Li, Stefanie Jegelka, Suvrit Sra

#66 Computationally Efficient Nystr\{o}m Approximation using Fast Transforms
Si Si, Cho-Jui Hsieh, Inderjit S. Dhillon

#67 Barron and Covers' Theory in Supervised Learning and Its Application to Lasso
Masanori Kawakita, Jun'ichi Takeuchi

#68 Exact Exponent in Optimal Rates for Crowdsourcing
Chao Gao, Yu Lu, Dengyong Zhou

#69 Generalization Properties and Implicit Regularization for Multiple Passes SGM
Junhong Lin, Raffaello Camoriano, Lorenzo Rosasco

#70 Generalized Direct Change Estimation in Ising Model Structure
Farideh Fazayeli, Arindam Banerjee

#71 Gaussian process nonparametric tensor estimator and its minimax optimality
Heishiro Kanagawa, Taiji Suzuki, Hayato Kobayashi, Nobuyuki Shimizu, Yukihiro Tagami

#72 Minimum Regret Search for Single- and Multi-Task Optimization
Jan Hendrik Metzen

#73 The Sum-Product Theorem: A Foundation for Learning Tractable Models
Abram Friesen, Pedro Domingos

#74 Train and Test Tightness of LP Relaxations in Structured Prediction
Ofer Meshi, Mehrdad Mahdavi, Adrian Weller, David Sontag

#75 Evasion and Hardening of Tree Ensemble Classifiers
Alex Kantchelian, J. D. Tygar, Anthony Joseph

#76 Importance Sampling Tree for Large-scale Empirical Expectation
Olivier Canévet, Cijo Jose, François Fleuret

#77 Stratified Sampling Meets Machine Learning
Edo Liberty, Kevin Lang, Konstantin Shmakov

#78 Scalable Discrete Sampling as a Multi-Armed Bandit Problem
Yutian Chen, Zoubin Ghahramani

#79 Pricing a low-regret seller
Hoda Heidari, Mohammad Mahdian, Umar Syed, Sergei Vassilvitskii, Sadra Yazdanbod

#80 Multi-Player Bandits -- a Musical Chairs Approach
Jonathan Rosenski, Ohad Shamir, Liran Szlak

#81 Contextual Combinatorial Cascading Bandits
Shuai Li, Baoxiang Wang, Shengyu Zhang, Wei Chen

#82 Copeland Dueling Bandit Problem: Regret Lower Bound, Optimal Algorithm, and Computationally Efficient Algorithm
Junpei Komiyama, Junya Honda, Hiroshi Nakagawa

#83 DCM Bandits: Learning to Rank with Multiple Clicks
Sumeet Katariya, Branislav Kveton, Csaba Szepesvári, Zheng Wen

#84 Distributed Clustering of Linear Bandits in Peer to Peer Networks
Nathan Korda, Balázs Szörényi, Shuai Li

TUESDAY JUNE 21ST | AFTERNOON POSTER SESSIONS

LOCATION: MARRIOTT - ASTOR, DUFFY + TIMES SQUARE - 3 PM - 7 PM

#1 Neural Variational Inference for Text Processing
Yishu Miao, Lei Yu, Phil Blunsom

#2 A Deep Learning Approach to Unsupervised Ensemble Learning
Uri Shaham, Xiuyuan Cheng, Omer Dror, Ariel Jaffe, Boaz Nadler, Joseph Chang, Yuval Kluger

#3 From Softmax to Sparsemax: A Sparse Model of Attention and Multi-Label Classification
André Martins, Ramon Astudillo

#4 A Neural Autoregressive Approach to Collaborative Filtering
Yin Zheng, Bangsheng Tang, Wenkui Ding, Hanning Zhou

#5 Scalable Gradient-Based Tuning of Continuous Regularization Hyperparameters
Jelena Luketina, Tapani Raiko, Mathias Berglund, Klaus Greff

#6 SDNA: Stochastic Dual Newton Ascent for Empirical Risk Minimization
Zheng Qu, Peter Richtárik, Martin Takac, Olivier Fercoq

#7 Stochastic Block BFGS: Squeezing More Curvature out of Data
Robert Gower, Donald Goldfarb, Peter Richtárik

#8 A Primal and Dual Sparse Approach to Extreme Classification
Ian En-Hsu Yen, Xiangru Huang, Pradeep Ravikumar, Kai Zhong, Inderjit S. Dhillon

#9 Parallel and Distributed Block-Coordinate Frank-Wolfe Algorithms
Yu-Xiang Wang, Veeranjaneyulu Sadhanala, Wei Dai, Willie Neiswanger, Suvrit Sra, Eric Xing

#10 Minding the Gaps for Block Frank-Wolfe Optimization of Structured SVM
Anton Osokin, Jean-Baptiste Alayrac, Isabella Lukasewitz, Puneet Dokania, Simon Lacoste-Julien

#11 Asymmetric Multi-task Learning based on Task Relatedness and Confidence
Giwoong Lee, Eunho Yang, Sung ju Hwang

#12 Training Deep Neural Networks via Direct Loss Minimization
Yang Song, Alexander Schwing, Richard S. Zemel, Raquel Urtasun

#13 Structured Prediction Energy Networks
David Belanger, Andrew McCallum

#14 Conditional Bernoulli Mixtures for Multi-label Classification
Cheng Li, Bingyu Wang, Virgil Pavlu, Javed Aslam

#15 Training Neural Networks Without Gradients: A Scalable ADMM Approach
Gavin Taylor, Ryan Burmeister, Zheng Xu, Bharat Singh, Ankit Patel, Tom Goldstein

#16 Stability of Controllers for Gaussian Process Forward Models
Julia Vinogradska, Bastian Bischoff, Duy Nguyen-Tuong, Anne Romer, Henner Schmidt, Jan Peters

#17 A Distributed Variational Inference Framework for Unifying Parallel Sparse Gaussian Process Regression Models
Trong Nghia Hoang, Quang Minh Hoang, Bryan Kian Hsiang Low

#18 Deep Gaussian Processes for Regression using Approximate Expectation Propagation
Thang Bui, José Miguel Hernández-Lobato, Daniel Hernández-Lobato, Yingzhen Li, Richard Turner

#19 Preconditioning Kernel Matrices
Kurt Cutajar, Michael Osborne, John Cunningham, Maurizio Filippone

#20 Extended and Unscented Kitchen Sinks
Edwin Bonilla, Daniel Steinberg, Alistair Reid

#21 On the Consistency of Feature Selection With Lasso for Non-linear Targets
Yue Zhang, Weihong Guo, Soumya Ray

#22 No penalty no tears: Least squares in high-dimensional linear models
Xiangyu Wang, David Dunson, Chenlei Leng

#23 Simultaneous Safe Screening of Features and Samples in Doubly Sparse Modeling
Atsushi Shibagaki, Masayuki Karasuyama, Kohei Hatano, Ichiro Takeuchi

#24 Efficient Learning with Nonconvex Regularizers by Nonconvexity Redistribution
Quanming Yao, James Kwok

#25 How to Fake Multiply by a Gaussian Matrix
Michael Kapralov, Vamsi Potluru, David Woodruff

#26 Metadata-conscious anonymous messaging
Giulia Fanti, Peter Kairouz, Sewoong Oh, Kannan Ramchandran, Pramod Viswanath

#27 A Simple and Strongly-Local Flow-Based Method for Cut Improvement
Nate Veldt, David Gleich, Michael Jordan

#28 Community Recovery in Graphs with Locality
Yuxin Chen, Govinda Kamath, Changho Suh, David Tse

#29 Interactive Bayesian Hierarchical Clustering
Sharad Vikram, Sanjoy Dasgupta

#30 Cross-graph Learning of Multi-relational Associations
Hanxiao Liu, Yiming Yang

#31 Controlling the distance to a Kemeny consensus without computing it
Anna Korba, Yunlong Jiao, Eric Sibony

#32 Data-driven Rank Breaking for Efficient Rank Aggregation
Ashish Khetan, Sewoong Oh

#33 Parameter Estimation for Generalized Thurstone Choice Models
Milan Vojnovic, Seyoung Yun

#34 Learning Mixtures of Plackett-Luce Models
Zhibing Zhao, Peter Piech, Lirong Xia

#35 Recommendations as Treatments: Debiasing Learning and Evaluation
Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, Thorsten Joachims

#36 Generative Adversarial Text to Image Synthesis
Scott Reed, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran, Bernt Schiele, Honglak Lee

#37 Autoencoding beyond pixels using a learned similarity metric
Anders Boesen Lindbo Larsen, Søren Kaae Sønderby, Hugo Larochelle, Ole Winther

#38 Exploiting Cyclic Symmetry in Convolutional Neural Networks
Sander Dieleman, Jeffrey De Fauw, Koray Kavukcuoglu

#39 A Comparative Analysis and Study of Multiview Convolutional Neural Network Models for Joint Object Categorization and Pose Estimation
Mohamed Elhoseiny, Tarek El-Gaaly, Amr Bakry, Ahmed Elgammal

#40 Dynamic Capacity Networks
Amjad Almahairi, Nicolas Ballas, Tim Cooijmans, Yin Zheng, Hugo Larochelle, Aaron Courville

#41 Augmenting Neural Networks with Reconstructive Decoding Pathways for Large-scale Image Classification
Yuting Zhang, Kibok Lee, Honglak Lee

#42 On the Iteration Complexity of Oblivious First-Order Optimization Algorithms
Yossi Arjevani, Ohad Shamir

#43 Variance-Reduced and Projection-Free Stochastic Optimization
Elad Hazan, Haipeng Luo

#44 On Graduated Optimization for Stochastic Non-Convex Problems
Elad Hazan, Kfir Yehuda Levy, Shai Shalev-Shwartz

#45 A Self-Correcting Variable-Metric Algorithm for Stochastic Optimization
Frank Curtis

#46 A Superlinearly-Convergent Proximal Newton-type Method for the Optimization of Finite Sums
Anton Rodomanov, Dmitry Kropotov

#47 Stochastic Variance Reduced Optimization for Nonconvex Sparse Learning
Xingguo Li, Tuo Zhao, Raman Arora, Han Liu, Jarvis Haupt

#48 Hierarchical Decision Making In Electricity Grid Management
Gal Dalal, Elad Gilboa, Shie Mannor

#49 ForecastICU: A Prognostic Decision Support System for Timely Prediction of Intensive Care Unit Admission
Jinsung Yoon, Ahmed Alaa, Scott Hu, Mihaela van der Schaar

#50 Power of Ordered Hypothesis Testing
Lihua Lei, William Fithian

#51 Learning to Filter with Predictive State Inference Machines
Wen Sun, Arun Venkatraman, Byron Boots, J.Andrew Bagnell

#52 Learning population-level diffusions with generative RNNs
Tatsunori Hashimoto, David Gifford, Tommi Jaakkola

#53 Fast Parameter Inference in Nonlinear Dynamical Systems using Iterative Gradient Matching
Mu Niu, Simon Rogers, Maurizio Filippone, Dirk Husmeier

#54 Greedy Column Subset Selection: New Bounds and Distributed Algorithms
Jason Altschuler, Aditya Bhaskara, Gang Fu, Vahab Mirrokni, Afshin Rostamizadeh, Morteza Zadimoghaddam

#55 Efficient Private Empirical Risk Minimization for High-dimensional Learning
Shiva Prasad Kasiviswanathan, Hongxia Jin

#56 Binary embeddings with structured hashed projections
Anna Choromanska, Krzysztof Choromanski, Mariusz Bojarski, Tony Jebara, Sanjiv Kumar, Yann LeCun

#57 Differentially Private Policy Evaluation
Borja Balle, Maziar Gomrokchi, Doina Precup

#58 Learning from Multiway Data: Simple and Efficient Tensor Regression
Rose Yu, Yan Liu

#59 Low-Rank Matrix Approximation with Stability
Dongsheng Li, Chao Chen, Qin Lv, Junchi Yan, Li Shang, Stephen Chu

#60 Interacting Particle Markov Chain Monte Carlo
Tom Rainforth, Christian Naesseth, Fredrik Lindsten, Brooks Paige, Jan-Willem Vandemeent, Arnaud Doucet, Frank Wood

#61 Slice Sampling on Hamiltonian Trajectories
Benjamin Bloem-Reddy, John Cunningham

#62 Robust Monte Carlo Sampling using Riemannian Nos\'{e}-Poincar\'{e} Hamiltonian Dynamics
Anirban Roychowdhury, Brian Kulis, Srinivasan Parthasarathy

#63 Inference Networks for Sequential Monte Carlo in Graphical Models
Brooks Paige, Frank Wood

#64 Partition Functions from Rao-Blackwellized Tempered Sampling
David Carlson, Patrick Stinson, Ari Pakman, Liam Paninski

#65 Stochastic Quasi-Newton Langevin Monte Carlo
Umut Simsekli, Roland Badeau, Taylan Cemgil, Gaël Richard

#66 No Oops, You Won't Do It Again: Mechanisms for Self-correction in Crowdsourcing
Nihar Shah, Dengyong Zhou

#67 The Label Complexity of Mixed-Initiative Classifier Training
Jina Suh, Xiaojin Zhu, Saleema Amershi

#68 The Knowledge Gradient for Sequential Decision Making with Stochastic Binary Feedbacks
Yingfei Wang, Chu Wang, Warren Powell

#69 Estimating Accuracy from Unlabeled Data: A Bayesian Approach
Emmanouil Antonios Platanios, Avinava Dubey, Tom Mitchell

#70 Actively Learning Hemimetrics with Applications to Eliciting User Preferences
Adish Singla, Sebastian Tschiatschek, Andreas Krause

#71 Optimality of Belief Propagation for Crowdsourced Classification
Jungseul Ok, Sewoong Oh, Jinwoo Shin, Yung Yi

#72 Improved SVRG for Non-Strongly-Convex or Sum-of-Non-Convex Objectives
Zeyuan Allen-Zhu, Yang Yuan

#73 Variance Reduction for Faster Non-Convex Optimization
Zeyuan Allen-Zhu, Elad Hazan

#74 Even Faster Accelerated Coordinate Descent Using Non-Uniform Sampling
Zeyuan Allen-Zhu, Zheng Qu, Peter Richtárik, Yang Yuan

#75 False Discovery Rate Control and Statistical Quality Assessment of Annotators in Crowdsourced Ranking
QianQian Xu, Jiechao Xiong, Xiaochun Cao, Yuan Yao

#76 On the Power of Distance-Based Learning
Periklis Papakonstantinou, Jia Xu, Guang Yang

#77 Minimizing the Maximal Loss: How and Why
Shai Shalev-Shwartz, Yonatan Wexler

Wednesday

JUNE 22ND | SESSIONS

TIME	DESCRIPTION	MARRIOT HOTEL
8:30 am - 9:55 am	Neural Networks & Deep Learning	Westside Ballroom 1&2
	Optimization (Continuous)	Westside Ballroom 3&4
	Multi-label, Multi-task, & Neural Networks	Marquis
	Gaussian Processes	Lyceum
	Feature Selection & Dimensionality Reduction	Empire
	Graph Analysis/ Spectral Methods	Soho
	Ranking and Preference Learning	Liberty
9:55 am - 10:20 am	Break	
10 am - 1 pm	Poster Sessions	Astor, Duffy, & Times Square
10:20 am - 12:20 am	Neural Networks & Deep Learning	Westside Ballroom 1&2
	Optimization (Continuous)	Westside Ballroom 3&4
break from 11:10 - 11:30 am	Applications and Time-Series Analysis	Marquis
	Dimensionality Reduction / Private Learning	Lyceum
	Monte Carlo Methods	Empire
	Crowdsourcing and Interactive Learning	Soho
	Learning Theory	Liberty
12:20 pm - 2 pm	Lunch (On Your Own)	
2 pm - 3 pm	Invited Talk: Christos Faloutsos	Ballroom 1&2
3 pm - 3:40 pm	Break	
3:40 pm - 5:20 pm	Optimization (Continuous)	Westside Ballroom 1&2
	Supervised Learning	Westside Ballroom 3&4
	Kernel Methods	Marquis
	Matrix Factorization and Related Topics	Lyceum
	Privacy, Anonymity, and Security	Empire
	Causal Inference	Soho
	Optimization	Liberty

SESSIONS - 8:30 AM - 9:55 PM - MARRIOTT

Neural Networks and Deep Learning
Location: Westside Ballroom 1 & 2 + Juliard

- **Neural Variational Inference for Text Processing**
  Yishu Miao, Lei Yu, Phil Blunsom
- **A Deep Learning Approach to Unsupervised Ensemble Learning**
  Uri Shaham, Xiuyuan Cheng, Omer Dror, Ariel Jaffe, Boaz Nadler, Joseph Chang, Yuval Kluger
- **From Softmax to Sparsemax: A Sparse Model of Attention and Multi-Label Classification**
  André Martins, Ramon Astudillo
- **A Neural Autoregressive Approach to Collaborative Filtering**
  Yin Zheng, Bangsheng Tang, Wenkui Ding, Hanning Zhou
- **Scalable Gradient-Based Tuning of Continuous Regularization Hyperparameters**
  Jelena Luketina, Tapani Raiko, Mathias Berglund, Klaus Greff

Optimization (Continuous)
Location: Westside Ballroon 3 & 4

- **SDNA: Stochastic Dual Newton Ascent for Empirical Risk Minimization**
  Zheng Qu, Peter Richtárik, Martin Takac, Olivier Fercoq
- **Stochastic Block BFGS: Squeezing More Curvature out of Data**
  Robert Gower, Donald Goldfarb, Peter Richtárik
- **A Primal and Dual Sparse Approach to Extreme Classification**
  Ian En-Hsu Yen, Xiangru Huang, Pradeep Ravikumar, Kai Zhong, Inderjit S. Dhillon
- **Parallel and Distributed Block-Coordinate Frank-Wolfe Algorithms**
  Yu-Xiang Wang, Veeranjaneyulu Sadhanala, Wei Dai, Willie Neiswanger, Suvrit Sra, Eric Xing
- **Minding the Gaps for Block Frank-Wolfe Optimization of Structured SVM**
  Anton Osokin, Jean-Baptiste Alayrac, Isabella Lukasewitz, Puneet Dokania, Simon Lacoste-Julien

Multi-label, multi-task, and neural networks
Location: Marquis

- **Asymmetric Multi-task Learning based on Task Relatedness and Confidence**
  Giwoong Lee, Eunho Yang, Sung ju Hwang
- **Training Deep Neural Networks via Direct Loss Minimization**
  Yang Song, Alexander Schwing, Richard S. Zemel, Raquel Urtasun
- **Structured Prediction Energy Networks**
  David Belanger, Andrew McCallum
- **Conditional Bernoulli Mixtures for Multi-label Classification**
  Cheng Li, Bingyu Wang, Virgil Pavlu, Javed Aslam
- **Training Neural Networks Without Gradients: A Scalable ADMM Approach**
  Gavin Taylor, Ryan Burmeister, Zheng Xu, Bharat Singh, Ankit Patel, Tom Goldstein

Gaussian Processes
Location: Lyceum

- **Stability of Controllers for Gaussian Process Forward Models**
  Julia Vinogradska, Bastian Bischoff, Duy Nguyen-Tuong, Anne Romer, Henner Schmidt, Jan Peters
- **A Distributed Variational Inference Framework for Unifying Parallel Sparse Gaussian Process Regression Models**
  Trong Nghia Hoang, Quang Minh Hoang, Bryan Kian Hsiang Low
- **Deep Gaussian Processes for Regression using Approximate Expectation Propagation**
  Thang Bui, José Miguel Hernández-Lobato, Daniel Hernández-Lobato, Yingzhen Li, Richard Turner
- **Preconditioning Kernel Matrices**
  Kurt Cutajar, Michael Osborne, John Cunningham, Maurizio Filippone
- **Extended and Unscented Kitchen Sinks**
  Edwin Bonilla, Daniel Steinberg, Alistair Reid

SESSIONS - 8:30 AM - 9:55 PM - MARRIOTT

Feature Selection and Dimensionality Reduction
Location: Empire

- **On the Consistency of Feature Selection With Lasso for Non-linear Targets**
  Yue Zhang, Weihong Guo, Soumya Ray
- **No penalty no tears: Least squares in high-dimensional linear models**
  Xiangyu Wang, David Dunson, Chenlei Leng
- **Simultaneous Safe Screening of Features and Samples in Doubly Sparse Modeling**
  Atsushi Shibagaki, Masayuki Karasuyama, Kohei Hatano, Ichiro Takeuchi
- **Efficient Learning with Nonconvex Regularizers by Nonconvexity Redistribution**
  Quanming Yao, James Kwok
- **How to Fake Multiply by a Gaussian Matrix**
  Michael Kapralov, Vamsi Potluru, David Woodruff

Graph Analysis/ Spectral Methods
Location: Soho

- **Metadata-conscious anonymous messaging**
  Giulia Fanti, Peter Kairouz, Sewoong Oh, Kannan Ramchandran, Pramod Viswanath
- **A Simple and Strongly-Local Flow-Based Method for Cut Improvement**
  Nate Veldt, David Gleich, Michael Jordan
- **Community Recovery in Graphs with Locality**
  Yuxin Chen, Govinda Kamath, Changho Suh, David Tse
- **Interactive Bayesian Hierarchical Clustering**
  Sharad Vikram, Sanjoy Dasgupta
- **Cross-graph Learning of Multi-relational Associations**
  Hanxiao Liu, Yiming Yang

Ranking and Preference Learning
Location: Liberty

- **Controlling the distance to a Kemeny consensus without computing it**
  Anna Korba, Yunlong Jiao, Eric Sibony
- **Data-driven Rank Breaking for Efficient Rank Aggregation**
  Ashish Khetan, Sewoong Oh
- **Parameter Estimation for Generalized Thurstone Choice Models**
  Milan Vojnovic, Seyoung Yun
- **Learning Mixtures of Plackett-Luce Models**
  Zhibing Zhao, Peter Piech, Lirong Xia
- **Recommendations as Treatments: Debiasing Learning and Evaluation**
  Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, Thorsten Joachims

SESSIONS - 10:20 AM - 12:20 PM - MARRIOTT

Neural Networks and Deep Learning
Location: Westside Ballroom 1 & 2 + Juliard

- **Generative Adversarial Text to Image Synthesis**
  Scott Reed, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran, Bernt Schiele, Honglak Lee
- **Autoencoding beyond pixels using a learned similarity metric**
  Anders Boesen Lindbo Larsen, Søren Kaae Sønderby, Hugo Larochelle, Ole Winther
- **Exploiting Cyclic Symmetry in Convolutional Neural Networks**
  Sander Dieleman, Jeffrey De Fauw, Koray Kavukcuoglu
- **A Comparative Analysis and Study of Multiview Convolutional Neural Network Models for Joint Object Categorization and Pose Estimation**
  Mohamed Elhoseiny, Tarek El-Gaaly, Amr Bakry, Ahmed Elgammal
- **Dynamic Capacity Networks**
  Amjad Almahairi, Nicolas Ballas, Tim Cooijmans, Yin Zheng, Hugo Larochelle, Aaron Courville
- **Augmenting Neural Networks with Reconstructive Decoding Pathways for Large-scale Image Classification**
  Yuting Zhang, Kibok Lee, Honglak Lee

Optimization (Continuous)
Location: Westside Ballroon 3 & 4

- **On the Iteration Complexity of Oblivious First-Order Optimization Algorithms**
  Yossi Arjevani, Ohad Shamir
- **Variance-Reduced and Projection-Free Stochastic Optimization**
  Elad Hazan, Haipeng Luo
- **On Graduated Optimization for Stochastic Non-Convex Problems**
  Elad Hazan, Kfir Yehuda Levy, Shai Shalev-Shwartz
- **A Self-Correcting Variable-Metric Algorithm for Stochastic Optimization**
  Frank Curtis
- **A Superlinearly-Convergent Proximal Newton-type Method for the Optimization of Finite Sums**
  Anton Rodomanov, Dmitry Kropotov
- **Stochastic Variance Reduced Optimization for Nonconvex Sparse Learning**
  Xingguo Li, Tuo Zhao, Raman Arora, Han Liu, Jarvis Haupt

Applications and Time-Series Analysis
Location: Marquis

- **Hierarchical Decision Making In Electricity Grid Management**
  Gal Dalal, Elad Gilboa, Shie Mannor
- **ForecastICU: A Prognostic Decision Support System for Timely Prediction of Intensive Care Unit Admission**
  Jinsung Yoon, Ahmed Alaa, Scott Hu, Mihaela van der Schaar
- **Power of Ordered Hypothesis Testing**
  Lihua Lei, William Fithian
- **Learning to Filter with Predictive State Inference Machines**
  Wen Sun, Arun Venkatraman, Byron Boots, J.Andrew Bagnell
- **Learning population-level diffusions with generative RNNs**
  Tatsunori Hashimoto, David Gifford, Tommi Jaakkola
- **Fast Parameter Inference in Nonlinear Dynamical Systems using Iterative Gradient Matching**
  Mu Niu, Simon Rogers, Maurizio Filippone, Dirk Husmeier

Dimensionality Reduction / Private Learning
Location: Lyceum

- **Greedy Column Subset Selection: New Bounds and Distributed Algorithms**
  Jason Altschuler, Aditya Bhaskara, Gang Fu, Vahab Mirrokni, Afshin Rostamizadeh, Morteza Zadimoghaddam
- **Efficient Private Empirical Risk Minimization for High-dimensional Learning**
  Shiva Prasad Kasiviswanathan, Hongxia Jin
- **Binary embeddings with structured hashed projections**
  Anna Choromanska, Krzysztof Choromanski, Mariusz Bojarski, Tony Jebara, Sanjiv Kumar, Yann LeCun
- **Differentially Private Policy Evaluation**
  Borja Balle, Maziar Gomrokchi, Doina Precup
- **Learning from Multiway Data: Simple and Efficient Tensor Regression**
  Rose Yu, Yan Liu
- **Low-Rank Matrix Approximation with Stability**
  Dongsheng Li, Chao Chen, Qin Lv, Junchi Yan, Li Shang, Stephen Chu

SESSIONS - 10:20 AM - 12:20 PM - MARRIOTT

Monte Carlo Methods
Location: Empire

- **Interacting Particle Markov Chain Monte Carlo**
  Tom Rainforth, Christian Naesseth, Fredrik Lindsten, Brooks Paige, Jan-Willem Vandemeent, Arnaud Doucet, Frank Wood
- **Slice Sampling on Hamiltonian Trajectories**
  Benjamin Bloem-Reddy, John Cunningham
- **Robust Monte Carlo Sampling using Riemannian Nos\'{e}-Poincar\'{e} Hamiltonian Dynamics**
  Anirban Roychowdhury, Brian Kulis, Srinivasan Parthasarathy
- **Inference Networks for Sequential Monte Carlo in Graphical Models**
  Brooks Paige, Frank Wood
- **Partition Functions from Rao-Blackwellized Tempered Sampling**
  David Carlson, Patrick Stinson, Ari Pakman, Liam Paninski
- **Stochastic Quasi-Newton Langevin Monte Carlo**
  Umut Simsekli, Roland Badeau, Taylan Cemgil, Gaël Richard

Crowdsourcing and Interactive Learning
Location: Soho

- **No Oops, You Won't Do It Again: Mechanisms for Self-correction in Crowdsourcing**
  JNihar Shah, Dengyong Zhou
- **The Label Complexity of Mixed-Initiative Classifier Training**
  Jina Suh, Xiaojin Zhu, Saleema Amershi
- **The Knowledge Gradient for Sequential Decision Making with Stochastic Binary Feedbacks**
  Yingfei Wang, Chu Wang, Warren Powell
- **Estimating Accuracy from Unlabeled Data: A Bayesian Approach**
  Emmanouil Antonios Platanios, Avinava Dubey, Tom Mitchell
- **Actively Learning Hemimetrics with Applications to Eliciting User Preferences**
  Adish Singla, Sebastian Tschiatschek, Andreas Krause
- **Optimality of Belief Propagation for Crowdsourced Classification**
  Jungseul Ok, Sewoong Oh, Jinwoo Shin, Yung Yi

Learning Theory
Location: Liberty

- **Improved SVRG for Non-Strongly-Convex or Sum-of-Non-Convex Objectives**
  Zeyuan Allen-Zhu, Yang Yuan
- **Variance Reduction for Faster Non-Convex Optimization**
  Zeyuan Allen-Zhu, Elad Hazan
- **Even Faster Accelerated Coordinate Descent Using Non-Uniform Sampling**
  Zeyuan Allen-Zhu, Zheng Qu, Peter Richtárik, Yang Yuan
- **False Discovery Rate Control and Statistical Quality Assessment of Annotators in Crowdsourced Ranking**
  QianQian Xu, Jiechao Xiong, Xiaochun Cao, Yuan Yao
- **On the Power of Distance-Based Learning**
  Periklis Papakonstantinou, Jia Xu, Guang Yang
- **Minimizing the Maximal Loss: How and Why**
  Shai Shalev-Shwartz, Yonatan Wexler

SESSIONS - 3:40 PM - 4:30 PM - MARRIOTT

Optimization (Continuous)
Location: Westside Ballroom 1 & 2 + Juliard

- **Energetic Natural Gradient Descent**
  Philip Thomas;Bruno Castro da Silva;Christoph Dann;Emma Brunskill
- **On the Quality of the Initial Basin in Overspecified Neural Networks**
  Itay Safran;Ohad Shamir
- **L1-regularized Neural Networks are Improperly Learnable in Polynomial Time**
  Yuchen Zhang;Jason D. Lee;Michael Jordan

Supervised Learning
Location: Westside Ballroom 3 & 4

- **Sparse Nonlinear Regression: Parameter Estimation and Asymptotic Inference**
  Zhuoran Yang;Zhaoran Wang;Han Liu;Yonina Eldar;Tong Zhang
- **Polynomial Networks and Factorization Machines: New Insights and Efficient Training Algorithms**
  Mathieu Blondel;Masakazu Ishihata;Akinori Fujino;Naonori Ueda
- **Hyperparameter optimization with approximate gradient**
  Fabian Pedregosa

Kernel Methods
Location: Marquis

- **DR-ABC: Approximate Bayesian Computation with Kernel-Based Distribution Regression**
  Jovana Mitrovic;Dino Sejdinovic;Yee-Whye Teh
- **Persistence weighted Gaussian kernel for topological data analysis**
  Genki Kusano;Yasuaki Hiraoka;Kenji Fukumizu
- **Discriminative Embeddings of Latent Variable Models for Structured Data**
  Hanjun Dai;Bo Dai;Le Song

Matrix Factorization and Related Topics
Location: Lyceum

- **Recycling Randomness with Structure for Sublinear time Kernel Expansions**
  Krzysztof Choromanski;Vikas Sindhwani
- **Optimal Classification with Multivariate Losses**
  Nagarajan Natarajan;Oluwasanmi Koyejo;Pradeep Ravikumar;Inderjit S. Dhillon
- **Sparse Parameter Recovery from Aggregated Data**
  Avradeep Bhowmik;Joydeep Ghosh;Oluwasanmi Koyejo

Privacy, Anonymity, and Security
Location: Empire

- **Learning privately from multiparty data**
  Jihun Hamm;Yingjun Cao;Mikhail Belkin
- **Differentially Private Chi-Squared Hypothesis Testing: Goodness of Fit and Independence Testing**
  Ryan Rogers;Salil Vadhan;Hyun Lim;Marco Gaboardi
- **Discrete Distribution Estimation under Local Privacy**
  Peter Kairouz;Keith Bonawitz;Daniel Ramage

Causal Inference
Location: Soho

- **The Arrow of Time in Multivariate Time Series**
  Stefan Bauer;Bernhard Schölkopf;Jonas Peters
- **Causal Strength via Shannon Capacity: Axioms, Estimators and Applications**
  Weihao Gao;Sreeram Kannan;Sewoong Oh;Pramod Viswanath
- **Learning Granger Causality for Hawkes Processes**
  Hongteng Xu;Mehrdad Farajtabar;Hongyuan Zha

Optimization
Location: Liberty

- **Gossip Dual Averaging for Decentralized Optimization of Pairwise Functions**
  Igor Colin;Aurélien Bellet;Joseph Salmon;Stéphan Clémençon
- **Adaptive Sampling for SGD by Exploiting Side Information**
  Siddharth Gopal
- **Mixture Proportion Estimation via Kernel Embeddings of Distributions**
  Harish Ramaswamy;Clayton Scott;Ambuj Tewari

WEDNESDAY JUNE 22ND | POSTER SESSIONS

LOCATION: MARRIOTT - ASTOR, DUFFY + TIMES SQUARE - 10 AM - 1 PM

#1 Factored Temporal Sigmoid Belief Networks for Sequence Learning
Jiaming Song, Zhe Gan, Lawrence Carin

#2 Bidirectional Helmholtz Machines
Jörg Bornschein, Samira Shabanian, Asja Fischer, Yoshua Bengio

#3 The Deep Neural Matrix Gaussian Process
Christos Louizos, Max Welling

#4 Dropout distillation
Samuel Rota Bulò, Lorenzo Porzi, Peter Kontschieder

#5 Revisiting Semi-Supervised Learning with Graph Embeddings
Zhilin Yang, William Cohen, Ruslan Salakhudinov

#6 ADIOS: Architectures Deep In Output Space
Moustapha Cissé, Maruan Al-Shedivat, Samy Bengio

#7 Unsupervised Deep Embedding for Clustering Analysis
Junyuan Xie, Ross Girshick, Ali Farhadi

#8 Learning Convolutional Neural Networks for Graphs
Mathias Niepert, Mohamed Ahmed, Konstantin Kutzkov

#9 Inverse Optimal Control with Deep Networks via Policy Optimization
Chelsea Finn, Sergey Levine, Pieter Abbeel

#10 Smooth Imitation Learning
Hoang Le, Andrew Kang, Yisong Yue, Peter Carr

#11 Improving the Efficiency of Deep Reinforcement Learning with Normalized Advantage Functions and Synthetic Experience
Shixiang Gu, Timothy Lillicrap, Ilya Sutskever, Sergey Levine

#12 Asynchronous Methods for Deep Reinforcement Learning
Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, Koray Kavukcuoglu

#13 On the Statistical Limits of Convex Relaxations
Zhaoran Wang, Quanquan Gu, Han Liu

#14 Faster Convex Optimization: Simulated Annealing with an Efficient Universal Barrier
Jacob Abernethy, Elad Hazan

#15 A ranking approach to global optimization
Cédric Malherbe, Emile Contal, Nicolas Vayatis

#16 Epigraph projections for fast general convex programming
Po-Wei Wang, Matt Wytock, J. Zico Kolter

#17 Principal Component Projection Without Principal Component Analysis
Roy Frostig, Cameron Musco, Christopher Musco, Aaron Sidford

#18 Recovery guarantee of weighted low-rank approximation via alternating minimization
Yuanzhi Li, Yingyu Liang, Andrej Risteski

#19 Tensor Decomposition via Joint Matrix Schur Decomposition
Nicolò Colombo, Nikos Vlassis

#20 Fast Methods for Estimating the Numerical Rank of Large Matrices
Shashanka Ubaru, Yousef Saad

#21 Markov-modulated marked Poisson processes for check-in data
Jiangwei Pan, Vinayak Rao, Pankaj Agarwal, Alan Gelfand

#22 Hierarchical Compound Poisson Factorization
Mehmet Basbug, Barbara Engelhardt

#23 Dirichlet Process Mixture Model for Correcting Technical Variation in Single-Cell Gene Expression Data
Sandhya Prabhakaran, Elham Azizi, Ambrose Carr, Dana Pe'er

#24 The Automatic Statistician: A Relational Perspective
Yunseong Hwang, Anh Tong, Jaesik Choi

#25 Truthful Univariate Estimators
Ioannis Caragiannis, Ariel Procaccia, Nisarg Shah

#26 Fast Algorithms for Segmented Regression
Jayadev Acharya, Ilias Diakonikolas, Jerry Li, Ludwig Schmidt

#27 Stochastically Transitive Models for Pairwise Comparisons: Statistical and Computational Issues
Nihar Shah, Sivaraman Balakrishnan, Aditya Guntuboyina, Martin Wainwright

#28 Provable Algorithms for Inference in Topic Models
Sanjeev Arora, Rong Ge, Frederic Koehler, Tengyu Ma, Ankur Moitra

#29 Expressiveness of Rectifier Neural Network
Xingyuan Pan, Vivek Srikumar

#30 Convolutional Rectifier Networks as Generalized Tensor Decompositions
Nadav Cohen, Amnon Shashua

#31 Fixed Point Quantization of Deep Convolutional Networks
Darryl Lin, Sachin Talathi, Sreekanth Annapureddy

#32 CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy
Ran Gilad-Bachrach, Nathan Dowlin, Kim Laine, Kristin Lauter, Michael Naehrig, John Wernsing

#33 Correcting Forecasts with Multi-force Neural Attention
Matthew Riemer, Aditya Vempaty, Flavio Calmon, Fenno Heath, Richard Hull, Elham Khabiri

#34 Meta-Learning with Memory-Augmented Neural Networks
Adam Santoro, Sergey Bartunov, Matthew Botvinick, Daan Wierstra, Timothy Lillicrap

#35 Learning Simple Algorithms from Examples
Wojciech Zaremba, Tomas Mikolov, Armand Joulin, Rob Fergus

#36 Associative Long Short-Term Memory
Ivo Danihelka, Greg Wayne, Benigno Uria, Nal Kalchbrenner, Alex Graves

#37 Estimating Maximum Expected Value through Gaussian Approximation
Carlo D'Eramo, Marcello Restelli, Alessandro Nuara

#38 Data-Efficient Off-Policy Policy Evaluation for Reinforcement Learning
Philip Thomas, Emma Brunskill

#39 Cumulative Prospect Theory Meets Reinforcement Learning: Prediction and Control
Prashanth L.A., Cheng Jie, Michael Fu, Steve Marcus, Csaba Szepesvári

#40 Softened Approximate Policy Iteration for Markov Games
Julien Pérolat, Bilal Piot, Matthieu Geist, Bruno Scherrer, Olivier Pietquin

#41 Low-rank Solutions of Linear Matrix Equations via Procrustes Flow
Stephen Tu, Ross Boczar, Max Simchowitz, mahdi Soltanolkotabi, Ben Recht

#42 Quadratic Optimization with Orthogonality Constraints: Explicit Lojasiewicz Exponent and Linear Convergence of Line-Search Methods
Huikang Liu, Weijie Wu, Anthony Man-Cho So

#43 Efficient Algorithms for Large-scale Generalized Eigenvector Computation and CCA
Rong Ge, Chi Jin, Sham M. Kakade, Praneeth Netrapalli, Aaron Sidford

#44 Matrix Eigendecomposition via Doubly Stochastic Riemannian Optimization
Zhiqiang Xu, Peilin Zhao, Jianneng Cao, Xiaoli Li

#45 Extreme F-measure Maximization using Sparse Probability Estimates
Kalina Jasinska, Krzysztof Dembczynski, Robert Busa-Fekete, Karlson Pfannschmidt, Timo Klerx, Eyke Hullermeier

#46 Stochastic Optimization for Multiview Learning using Partial Least Squares
Raman Arora, Poorya Mianjy, Teodor Marinov

#47 Gaussian quadrature for matrix inverse forms with applications
Chengtao Li, Suvrit Sra, Stefanie Jegelka

#48 A Subspace Learning Approach for High Dimensional Matrix Decomposition with Efficient Column/Row Sampling
Mostafa Rahmani, Geroge Atia

#49 Uprooting and Rerooting Graphical Models
Adrian Weller

#50 Structure Learning of Partitioned Markov Networks
Song Liu, Taiji Suzuki, Masashi Sugiyama, Kenji Fukumizu

#51 Ensuring Rapid Mixing and Low Bias for Asynchronous Gibbs Sampling
Christopher De Sa, Chris Re, Kunle Olukotun

#52 Estimation from Indirect Supervision with Linear Moments
Aditi Raghunathan, Roy Frostig, John Duchi, Percy Liang

#53 Early and Reliable Event Detection Using Proximity Space Representation
Maxime Sangnier, Jérôme Gauthier, Alain Rakotomamonjy

#54 Meta--Gradient Boosted Decision Tree Model for Weight and Target Learning
Yury Ustinovskiy, Valentina Fedorova, Gleb Gusev, Pavel Serdyukov

#55 Class Probability Estimation via Differential Geometric Regularization
Qinxun Bai, Steven Rosenberg, Zheng Wu, Stan Sclaroff

#56 Linking losses for density ratio and class-probability estimation
Aditya Menon, Cheng Soon Ong

#57 Energetic Natural Gradient Descent
Philip Thomas, Bruno Castro da Silva, Christoph Dann, Emma Brunskill

#58 On the Quality of the Initial Basin in Overspecified Neural Networks
Itay Safran, Ohad Shamir

#59 L1-regularized Neural Networks are Improperly Learnable in Polynomial Time
Yuchen Zhang, Jason D. Lee, Michael Jordan

#60 Sparse Nonlinear Regression: Parameter Estimation and Asymptotic Inference
Zhuoran Yang, Zhaoran Wang, Han Liu, Yonina Eldar, Tong Zhang

#61 Polynomial Networks and Factorization Machines: New Insights and Efficient Training Algorithms
Mathieu Blondel, Masakazu Ishihata, Akinori Fujino, Naonori Ueda

#62 Hyperparameter optimization with approximate gradient
Fabian Pedregosa

#63 DR-ABC: Approximate Bayesian Computation with Kernel-Based Distribution Regression
Jovana Mitrovic, Dino Sejdinovic, Yee-Whye Teh

#64 Persistence weighted Gaussian kernel for topological data analysis
Genki Kusano, Yasuaki Hiraoka, Kenji Fukumizu

#65 Discriminative Embeddings of Latent Variable Models for Structured Data
Hanjun Dai, Bo Dai, Le Song

#66 Recycling Randomness with Structure for Sublinear time Kernel Expansions
Krzysztof Choromanski, Vikas Sindhwani

#67 Optimal Classification with Multivariate Losses
Nagarajan Natarajan, Oluwasanmi Koyejo, Pradeep Ravikumar, Inderjit S. Dhillon

68 Sparse Parameter Recovery from Aggregated Data
Avradeep Bhowmik, Joydeep Ghosh, Oluwasanmi Koyejo

#69 Learning privately from multiparty data
Jihun Hamm, Yingjun Cao, Mikhail Belkin

#70 Differentially Private Chi-Squared Hypothesis Testing: Goodness of Fit and Independence Testing
Ryan Rogers, Salil Vadhan, Hyun Lim, Marco Gaboardi

#71 Discrete Distribution Estimation under Local Privacy
Peter Kairouz, Keith Bonawitz, Daniel Ramage

#72 The Arrow of Time in Multivariate Time Series
Stefan Bauer, Bernhard Schölkopf, Jonas Peters

#73 Causal Strength via Shannon Capacity: Axioms, Estimators and Applications
Weihao Gao, Sreeram Kannan, Sewoong Oh, Pramod Viswanath

#74 Learning Granger Causality for Hawkes Processes
Hongteng Xu, Mehrdad Farajtabar, Hongyuan Zha

#75 Gossip Dual Averaging for Decentralized Optimization of Pairwise Functions
Igor Colin, Aurélien Bellet, Joseph Salmon, Stéphan Clémençon

#76 Adaptive Sampling for SGD by Exploiting Side Information
Siddharth Gopal

#77 Mixture Proportion Estimation via Kernel Embeddings of Distributions
Harish Ramaswamy, Clayton Scott, Ambuj Tewari renner, Alex Graves

Thursday

JUNE 23RD | WORKSHOPS

THURSDAY WORKSHOPS 8:00 AM - 4:30 PM

Gimli: Geometry in Machine Learning

Location: Crowne Plaza - Times Square

Søren Hauberg *(Technical University of Denmark)*
Oren Freifeld *(MIT)*
Michael Schober *(Max Plack Institute)*

Many machine learning (ML) problems are fundamentally geometric in nature, e.g. finding optimal subspaces can be recast as finding point estimates on the Grassmannian; multi-metric learning can be recast as the learning of a Riemannian tensor; and covariance estimation entails optimization over a nonlinear cone. In spite of this, most practitioners neglect the geometry, only to find suboptimal models. Furthermore, many difficult problems that involve both geometry and statistical learning are usually ignored by the ML community. This workshop will raise these discussion points through a series of invited talks from experts on both geometry and machine learning.

https://sites.google.com/site/gimliworkshop

Machine Learning for Digital Education and Assessment Systems

Location: Marriott - Times Square

Alina A. von Davier *(Educational Testing Service)*
Mihaela van der Schaar *(UCLA)*
Richard Baraniuk *(Rice University)*

The focus of this workshop is on multidisciplinary research in the area of machine learning to enable new forms of digital education and assessment tools.

Recent developments indicate that the society is interested in redesigning learning and assessment systems (LAS) and not merely improving the systems we have. There is a renewed interest in performance assessments that are individualized and adaptive, which are developed in virtual settings. However, virtual LASs come with a number of psychometric and operational challenges. Advances in ML provide opportunities to address these challenges.

This workshop provides a platform for the sharing of knowledge and ideas across disciplines including ML, computational psychometrics, adaptive learning and testing, and natural language processing.

http://medianetlab.ee.ucla.edu/ICML-Education2016.html

Human Interpretability in Machine Learning

Location: Microsoft, Central Park (6th floor)
Entrance is between 42nd and 41st on 8th Avenue

Been Kim *(Allen Institute for Artificial Intelligence)*
Dmitry Malioutov *(IBM T. J. Watson Research Center)*
Kush Varshney *(IBM T. J. Watson Research Center)*

The goal of this workshop is to bring together researchers who study interpretable machine learning. 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 blackbox machine learning (e.g., deep 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, but not limited to, developing interpretability of predictive models, interpretable machine learning algorithms, methodology to interpret blackbox machine learning models (e.g., post hoc interpretations), and visual analytics.

https://sites.google.com/site/2016whi

Multi-View Representation Learning

Location: Marriott: Carnegie-Booth

Xiaodong He *(Microsoft Research)*
Karen Livescu *(TTI-Chicago)*
Weiran Wang *(TTI-Chicago)*
Scott Wen-tau Yih *(Microsoft Research)*

The workshop will bring together researchers and practitioners in this area, and discuss both theoretical and practical aspects of representation/ feature learning in the presence of multi-view data.

http://ttic.uchicago.edu/~wwang5/ICML2016_MVRL/

Theory and Practice of Differential Privacy (TPDP 2016)

Location: Marriott: O'Neil

Gilles Barthe *(IMDEA Software)*
Christos Dimitra kakis *(Chalmers University)*
Marco Gaboardi *(University at Buffalo, SUNY)*
Andreas Haeberlen *(University of Pennsylvania)*
Aaron Roth *(University of Pennsylvania)*
Aleksandra Slavkovic *(Penn State University)*

Differential privacy is a promising approach to the privacy-preserving release of data: it offers a strong guaranteed bound on the increase in harm that a user incurs as a result of participating in a differentially private data analysis. Several mechanisms and software tools have been developed to ensure differential privacy for a wide range of data analysis task.

Researchers in differential privacy come from several disciplines such as computer science, data analysis, statistics, security, law and privacy making, social science. The workshop is an occasion for researchers to discuss the recent developments in the theory and practice of differential privacy and applications.

http://tpdp16.cse.buffalo.edu/

Visualization for Deep Learning

Location: Marriott: Astor

Biye Jiang *(UC Berkeley)*
John Canny *(UC Berkeley)*
Polo Chau *(Georgia Tech)*
Aditya Khosla *(MIT)*

Deep neural networks are complex to design and train. They are non-linear systems that have many local optima and are sensitive to hyper-parameters. Systematic optimization of structure and hyper-parameters is possible, but hampered by the expense of training each design on realistic datasets. 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.

http://icmlviz.github.io/

Reliable Machine Learning in the Wild

Location: Marriott: Empire

Jacob Steinhardt (Stanford)
Tom Dietterich (OSU)
Percy Liang (Stanford)
Andrew Critch (MIRI)
Jessica Taylor (MIRI)
Adrian Weller (Cambridge)

How can we be confident that a system that performed well in the past will do so in the future, in the presence of novel and potentially adversarial input distributions? Answering these questions is critical for high stakes applications such as autonomous driving, as well as for building reliable large-scale machine learning systems. This workshop explores approaches that are principled or can provide performance guarantees, ensuring AI systems are robust and beneficial in the long run. We will focus on three aspects — robustness, adaptation, and monitoring — that can aid us in designing and deploying reliable machine learning systems.

https://sites.google.com/site/wildml2016/

Neural Networks Back To The Future

Location: Crowne Plaza - Broadway

Léon Bottou (Facebook)
David Grangier (Facebook)
Tomas Mikolov (Facebook)
John Platt (Google)

As research in deep learning is extremely active today, we could take a step back and examine its foundations. We propose to have a critical look at previous work on neural networks, and try to have a better understanding of the differences with today's work. Previous work can point at promising directions to follow, pitfalls to avoid, ideas and assumptions to revisit. Similarly, today's progress can allow a critical examination of what should still be investigated, what has been answered…

https://sites.google.com/site/nnb2tf

Deep Learning Workshop

Location: Marriott: Westside Ballroom 3 & 4

Antoine Bordes (Facebook AI Research),
Kyunghyun Cho (New York University),
Emily Denton (New York University),
Nando de Freitas (Google DeepMind, University of Oxford),
Rob Fergus (Facebook AI Research, New York University)

Deep learning is a fast-growing field of machine learning concerned with the study and design of computer algorithms for learning good representations of data, at multiple levels of abstraction. There has been rapid progress in this area in recent years, both in terms of methods and in terms of applications, which are attracting the major IT companies as well as major research labs. Many challenges remain, however, in aspects like large sample complexity of deep learning approaches, generative modeling, learning representations for reinforcement learning and symbolic reasoning, modeling of temporal data with long-term dependencies, efficient Bayesian inference for deep learning and multi-modal data and models. This workshop aims at tackling two major challenges in deep learning, which are unsupervised learning in the regime of small data, and simulation-based learning and its transferability to the real world, by bringing together researchers in the field of deep learning.

https://sites.google.com/site/dlworkshop16/

Abstraction in Reinforcement Learning

Location: Marriott: Marquis

Daniel Mankowitz,
Shie Mannor (Technion Israel Institute of Technology),
Timothy Mann (Google Deepmind)

Many real-world domains can be modeled using some form of abstraction. An abstraction is an important tool that enables an agent to focus less on the lower level details of a task and more on solving the task at hand. Temporal abstraction (i.e., options or skills) as well as spatial abstraction (i.e., state space representation) are two important examples. The goal of this workshop is to provide a forum to discuss the current challenges in designing as well as learning abstractions in real-world Reinforcement Learning (RL).

http://rlabstraction2016.wix.com/icml

Advances in non-convex analysis and optimization

Location: Westin - Majestic

Animashree Anandkumar	*(UCI)*
Sivaraman Balakrishnan	*(CMU)*
Srinadh Bhojanapalli	*(TTI)*
Kamalika Chaudhuri	*(UCSD)*
Yudong Chen	*(Cornell)*
Anastasios Kyrillidis	*(UT Austin)*
Percy Liang	*(Stanford)*
Praneeth Netrapalli	*(Microsoft)*
Sewoong Oh	*(UIUC)*
Zhaoran Wang	*(Princeton)*

This workshop will attempt to present some of the very recent developments on non-convex analysis and optimization, as reported in diverse research fields: from machine learning and mathematical programming to statistics and theoretical computer science. We believe that this workshop can bring researchers closer, in order to facilitate a discussion regarding why tackling non-convexity is important, where it is found, why non-convex schemes work well in practice and, how we can progress further with interesting research directions and open problems.

https://sites.google.com/site/noncvxicml16/

Machine Learning for Music Discovery

Location: Marriott: Wilder

Erik Schmidt	*(Pandora)*
Fabien Gouyon	*(Pandora)*
Oriol Nieto	*(Pandora)*
Gert Lanckriet	*(Amazon/UC San Diego)*

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. Collaborative filtering provides excellent music recommendations when the necessary user data is available, but these approaches also suffer heavily from the cold-start problem. Furthermore, defining musical similarity directly is extremely challenging as myriad features play some role (e.g., cultural, emotional, timbral, rhythmic). The topics discussed will span a variety of music recommender systems challenges including cross-cultural recommendation, content-based audio processing and representation learning, automatic music tagging, and evaluation.

https://sites.google.com/site/ml4md2016/

Friday

JUNE 24TH | WORKSHOPS

FRIDAY WORKSHOPS 8:00 AM - 4:30 PM

Data-Efficient Machine Learning

Location: Marriott: Astor

Marc Deisenroth *(Imperial College London)*
Shakir Mohamed *(Google Deepmind)*
Finale Doshi-Velez *(Harvard University)*
Andreas Krause *(ETH Zürich)*
Max Welling *(University of Amsterdam)*

Recent efforts in machine learning have addressed the problem of learning from massive amounts data. We now have highly scalable solutions for problems in object detection and recognition, machine translation, text-to-speech, recommender systems, and information retrieval, all of which attain state-of-the-art performance when trained with large amounts of data. In these domains, the challenge we now face is how to learn efficiently with the same performance in less time and with less data. Other problem domains, such as personalized healthcare, robot reinforcement learning, sentiment analysis, and community detection, are characterized as either small-data problems, or big-data problems that are a collection of small-data problems. The ability to learn in a sample-efficient manner is a necessity in these data-limited domains. Collectively, these problems highlight the increasing need for data-efficient machine learning: the ability to learn in complex domains without requiring large quantities of data.

This workshop will discuss the diversity of approaches that exist for data-efficient machine learning, and the practical challenges that we face. There are many approaches that demonstrate that data-efficient machine learning is possible, including methods that consider trade-offs between incorporating explicit domain knowledge and more general-purpose approaches, exploit structural knowledge of our data, such as symmetry and other invariance properties, apply bootstrapping and data augmentation techniques that make statistically efficient reuse of available data, use semi-supervised learning techniques, e.g., where we can use generative models to better guide the training of discriminative models, generalize knowledge across domains (transfer learning), use active learning and Bayesian optimization for experimental design and data-efficient black-box optimization, apply non-parametric methods, one-shot learning and Bayesian deep learning.

The objective of this interdisciplinary workshop is to provide a platform for researchers from a variety of areas, spanning transfer learning, Bayesian optimization, bandits, deep learning, approximate inference, robot learning, healthcare, computational neuroscience, active learning, reinforcement learning, and social network analysis, to share insights and perspectives on the problem of data-efficient machine learning, discuss challenges and to debate the roadmap towards more data-efficient machine learning.

https://sites.google.com/site/dataefficientml/

Computational Biology

Location: Marriott: Cantor/Jolson

Dana Pe'er *(Columbia University)*
Elham Azizi *(Columbia University)*
Sandhya Prabhakaran *(Columbia University)*
Olga Troyanskaya *(Princeton University)*
Edoardo Airoldi *(Harvard University)*
Volker Roth *(University of Basel)*

The application of Machine Learning in Computational biology has advanced significantly in recent years. In computational biology, there has been credible developments in many high-throughput technologies like next-generation sequencing, CyToF and single-cell sequencing that enable data generation from many interesting biological systems. The gamut of novel algorithms in Machine Learning makes it very attractive to apply these methods to the challenging biological questions. It therefore only seems befitting to bring together researchers engaged in applying ML in Computational biology to discuss recent advances in this interdisciplinary field and ongoing developments.

https://sites.google.com/site/compbioworkshopicml2016

Anomaly Detection 2016

Location: Marriott - Soho

Nico Goernitz *(Berlin Institute of Technology)*
Marius Kloft *(Humboldt University of Berlin)*
Vitaly Kuznetsov *(Courant Institute)*

Anomaly, outlier and novelty detection methods are crucial tools in any data scientist's inventory and are critical components of many real-world applications. Abnormal user activities can be used to detect credit card fraud, network intrusions or other security breaches. In computational biology, characterization of systematic anomalies in gene expression can be translated into clinically relevant information. With the rise of Internet-of-Things, the task of monitoring and diagnostics of numerous autonomous systems becomes intractable for a human and needs to be outsources to a machine. Early detection of an upcoming earthquake or tsunami can potentially save human lives. These applications make anomaly detection methods increasingly relevant in the modern world.

However, with the advent of Big Data, new challenges and questions are introduced, which will need to be addressed by the next generation of the anomaly and outlier detection algorithms. The goal of our workshop is to survey the existing techniques and discuss new research directions in this area.

https://sites.google.com/site/icmlworkshoponanomalydetection

Automatic Machine Learning (AutoML)

Location: Marriott: Empire

Frank Hutter *(University of Freiburg)*
Lars Kotthoff *(University of British Columbia)*
Joaquin Vanschoren *(Eindhoven University)*

Machine learning has been very successful, but its successes rely on human machine learning experts to define the learning problem, select, collect and preprocess the training data, choose appropriate ML architectures (deep learning, random forests, SVMs, …) and their hyperparameters, and finally evaluate the suitability of the learned models for deployment. 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 are more bullet-proof and can be used easily without expert knowledge. We call the resulting research area that targets progressive automation of machine learning AutoML.

See also ChaLearn's AutoML challenge:
http://automl.chalearn.org/
http://icml2016.automl.org/

Machine Learning Systems

Location: Microsoft, Central Park (6th floor)
Entrance is between 42nd and 41st on 8th Avenue

Aparna Lakshmi Ratan (Facebook)
Joaquin Quiñonero Candela (Facebook)
Hussein Mehanna (Facebook)
Joseph Gonzalez (UC Berkeley)

The diverse use of machine learning, the explosive growth in data, and the complexity of large-scale learning systems have fueled an interesting area at intersection of Machine Learning and large scale System Design. The goal of this workshop is to bring together experts working in the intersection of machine learning, system design, software engineering to explore the challenges needed to address real world, large scale machine learning problems. In particular, we aim to elicit new connections among these diverse fields, identify tools, best practices and design principles. The workshop will cover ML and AI platforms and algorithm toolkits (Caffe, Torch, MXNet and parameter server, Theano etc), as well as dive into Machine learning focused developments in distributed learning platforms, programming languages, data structures and general purpose GPU programming.

The workshop will have a mix of invited speakers and reviewed papers to facilitate the flow of new ideas as well as best practices which can benefit those looking to implement large ML systems in academia or industry.

https://sites.google.com/site/mlsys2016/

#data4good: Machine Learning in Social Good Applications

Location: Marriott: Wilder

James Faghmous (Mount Sinai)
Matt Gee (University of Chicago)
Rayid Ghani (University of Chicago)
Gideon Mann (Bloomberg)
Aleksandra Mojsilović (IBM Research)
Kush Varshney (IBM Research)

This workshop will bring together experts from different fields to explore the opportunities for machine learning in applications with social impact. Our goal is to raise awareness among ML practitioners about the opportunities in Data-for-Good movement and push the boundaries on addressing tough humanitarian challenges. The workshop will consist of: 1) invited presentations from the leading practitioners in the field and 2) a series of presentations on research that fits the theme of machine learning for social good; broadly construed, this could be machine learning related social good applications, or machine learning methods/theory of particular interest for social good applications.

https://sites.google.com/site/icml2016data4goodworkshop

Theory of Deep Learning

Location: Marriott: Westside Ballroom 3 & 4

Rene Vidal (the John Hopkins University)
Alex M. Bronstein (Technion – IIT)
Raja Giryes (Tel Aviv University)

Deep learning led to a significant breakthrough in many applications in computer vision and machine learning. However, only little is known about the theory behind this successful paradigm. This workshop will discuss the recent achievements with respect to the theoretical understanding of deep networks.

https://sites.google.com/site/deeplearningtheory

On-Device Intelligence

Location: Marriott: Odets

Vikas Sindhwani
Daniel Ramage
Keith Bonawitz (Google)
Suyog Gupta (IBM)
Sachin Talathi (Qualcomm)

Consumer adoption of mobile devices has created a new normal in computing: there are now more mobile devices on the planet than people, and exabytes of mobile data per month now dominates global internet traffic. As computing systems, these pocket-sized devices are more powerful in many ways than

vintage supercomputers. They come packed with an ever growing array of sensors. They are "always-on", and becoming increasingly capable of rich contextual understanding and natural interaction with their users.

This workshop will focus on research themes emerging at the intersection of machine learning and mobile systems. The topics of interest range from the design of new machine learning algorithms under storage and power constraints, new on-device learning mechanisms, the interaction between devices and cloud resources for privacy-aware distributed training, and opportunities for machine learning in the nascent area of "Internet of Things." The scope of the workshop also extends to real-time learning and optimization in the context of novel form-factors: wearable computers, home intelligence devices, and consumer robotics systems. We are also interested in hardware-software co-design for mobile machine learning applications.

https://sites.google.com/site/ondeviceintelligence/icml2016

Online advertising systems

Location: Marriott: Carnegie/Booth

Sharat Chikkerur *(Nanigans Inc)*
Hossein Azari *(Google Research)*
Edoardo Airoldi *(Harvard)*

Online advertising is a multi-billion dollar industry driven by the confluence of machine learning, optimization, control systems, auction algorithms, econometrics and software engineering. The goal of this workshop is to discuss how machine learning systems operate within the context of an advertising system.

https://sites.google.com/site/admlsystemsworkshop

Optimization Methods for the Next Generation of Machine Learning

Location: Westin - Majestic

Katya Scheinberg (Lehigh University)
Frank E. Curtis (Lehigh University)
Jorge Nocedal (Northwestern University)
Yoshua Bengio (University of Montreal)

The future of optimization for machine learning, lies in the design of methods for nonconvex optimization problems, such as those arising through the use of deep neural networks. Nonconvex formulations lead to more powerful predictive models, but are much more complex in the sense that they result in much more challenging optimization problems. This workshop will bring together experts from the machine learning and optimization communities whose research focuses on the design of optimization methodologies that combine recent trends of optimization in machine learning—stochasticity, parallel and distributed computing, and second order information—but do so in nonconvex settings.

http://optml.lehigh.edu/ICML2016

Computational Frameworks for Personalization

Location: Marriott: O'Neil

Suchi Saria *(Johns Hopkins University)*
Yisong Yue *(Caltech)*
Khalid El-Arini *(Facebook)*
Ambuj Tewari *(University of Michigan)*

This workshop aims to bring together researchers from industry and academia in order to describe recent advances and discuss future research directions pertaining to computational frameworks for personalization, broadly construed. Personalization has already made a huge impact in online recommender systems. Furthermore, there are many emerging applications where personalization has begun to show great promise, such as education and medicine. We are particularly interested in understanding what are the common computational challenges that underlie all these applications, with the goal of accelerating the development of personalization frameworks across a broad range of domains.

https://sites.google.com/site/icml2016ersonalization/

SPONSOR SCHOLARS

Funding for our Sponsor Scholars program was generously provided by our platinum sponsors, American Express and Baidu, as well as the National Science Foundation and the Artificial Intelligence Journal. Their exemplary support provided scholarships to help 129 student researchers travel to ICML to present their research. We gratefully acknowledge these sponsors, and congratulate all of our Sponsor Scholars, listed below.

Miltiadis Allamanis, University of Edinburgh, UK
Ehsaneddin Asgari, University of California, Berkeley
Megasthenis Asteris, The University of Texas at Austin
Saleh Babak, Rutgers University
Björn Barz, Friedrich Schiller University Jena
Thang Bui, University of Cambridge
Joan Capdevila Pujol, UPC BarcelonaTech
Kai-Yang Chiang, University of Texas at Austin
Taco Cohen, University of Amsterdam
Nadav Cohen, The Hebrew University of Jerusalem
Kurt Cutajar, EURECOM
Carlo D'Eramo, Politecnico di Milano
Ran Dai, University of Chigago
Gal Dalal, Technion
Yahel David, Technion
Mohamed Elhoseiny, Rutgers University
Akram Erraqabi, MILA
Farideh Fazayeli, University of Minnesota
Yarin Gal, University of Cambridge
Weihao Gao, University of Illinois at Urbana-Champaign
Hadrien Glaude, University of Lille / INRIA
Maziar Gomrokchi, McGill
Robert Gower, University of Edinburgh
Felix Grün, TU München
Huan Gui, University of Illinois at Urbana-Champaign
He He, University of Maryland, College Park
Masaaki Imaizumi, University of Tokyo
Song Jiaming, Tsinghua University
Fredrik Johansson, Chalmers University of Technology
Deepak Kadetotad, Arizona State University
Kirthevasan Kandasamy, Carnegie Mellon University
Alex Kantchelian, UC Berkeley
Sumeet Katariya, University of Wisconsin Madison
Piyush Khandelwal, University of Texas at Austin
Ashish Kumar Khetan, University of Illinois Urbana-Champaign
Andrew Lan, Rice University
Lihua Lei, University of California, Berkeley
Yehuda Kfir Levy, Technion
Shuai Li, University of Insubria
Xingguo Li, University of Minnesota
Chengtao Li, MIT
Hongyang Li, The Chinese University of Hong Kong
Cheng Li, Northeastern University
Chongxuan Li, Tsinghua University
Shuai Li, The Chinese University of Hong Kong
Cong Han Lim, University of Wisconsin-Madison
Xin Lin, The University of Texas at Austin
Huikang Liu, The Chinese University of Hong Kong
Andrea Locatelli, Potsdam University
Christos Louizos, University of Amsterdam
Ines Marusic, University of Oxford
Arthur Mensch, Inria - Université Paris Saclay
Poorya Mianjy, Johns Hopkins University
Baharan Mirzasoleimanbarzi, ETH Zurich
Jovana Mitrovic, University of Oxford
James Newling, EPFL
Junhyuk Oh, University of Michigan
Jungseul Ok, KAIST (Korea Advanced Institute of Science and Technology)
Jiangwei Pan, Duke University
Xingyuan Pan, University of Utah
Gaurav Pandey, Indian Institute of Science
Giorgio Patrini, Australian National University / NICTA
Julien Perolat, Univ. Lille
Kairouz Peter, University of Illinois at Urbana Champaign
Daniel Pimentel-Alarcon, University of Wisconsin-Madison
Anastasia Podosinnikova, Ecole Normale Superieure Paris
Lorenzo Porzi, University of Perugia, Fundazione Bruno Kessler
Bhanu Pratap Singh, Malaviya National Institute of Technology, Jaipur, India. (Alma-mater)
Chao Qu, NATIONAL UNIVERSITY OF SINGAPORE
Aditi Raghunathan, Indian Institute of Technology Madras
Mostafa Rahmani, University of Central Florida
Thomas Rainforth, University of Oxford
Sathya Narayanan Ravi, University of Wisconsin, Madison
Ramin Raziperchikolaei, University of California, Merced
Anton Rodomanov, Higher School of Economics
Anirban Roychowdhury, Ohio State University
Okke Schrijvers, Stanford University
Nisarg Shah, Carnegie Mellon University
Amar Shah, University of Cambridge
Nihar Shah, UC Berkeley
Uri Shaham, Yale
Jie Shen, Rutgers University
Si Si, University of Texas at Austin
Maximilian Soelch, Technical University of Munich
Yang Song, Tsinghua University
Wen Sun, Carnegie Mellon University
Schnabel Tobias, Cornell University
Aristide Charles Yedia Tossou, Chalmers University of Technology
Shashanka Ubaru, University of Minnesota, Twin Cities
Yichen Wang, Georgia Institute of Technology
Xiangyu Wang, Duke University
Yu-Xiang Wang, Carnegie Mellon University
Zhaoran Wang, Princeton University
Yifan Wu, University of Alberta
Hongteng Xu, Georgia Institute of Technology
Yingzhen Yang, University of Illinois at Urbana-Champaign
Quanming Yao, Hong Kong University of Science and Technology
Ian En-Hsu Yen, University of Texas at Austin
Shihui Yin, Arizona State University
Qi (Rose) Yu, University of Southern California
Zhiding Yu, Carnegie Mellon University
Tom Zahavy, Technion
Shuangfei Zhai, Binghamton University
Huishuai Zhang, Syracuse University
Yue Zhang, Case Western Reserve University
Hao Zhang, Carnegie Mellon University
Zhibing Zhao, Rensselaer Polytechnic Institute
Han Zhao, Carnegie Mellon University
Tuo Zhao, Johns Hopkins University
Konrad Żołna, Jagiellonian University

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

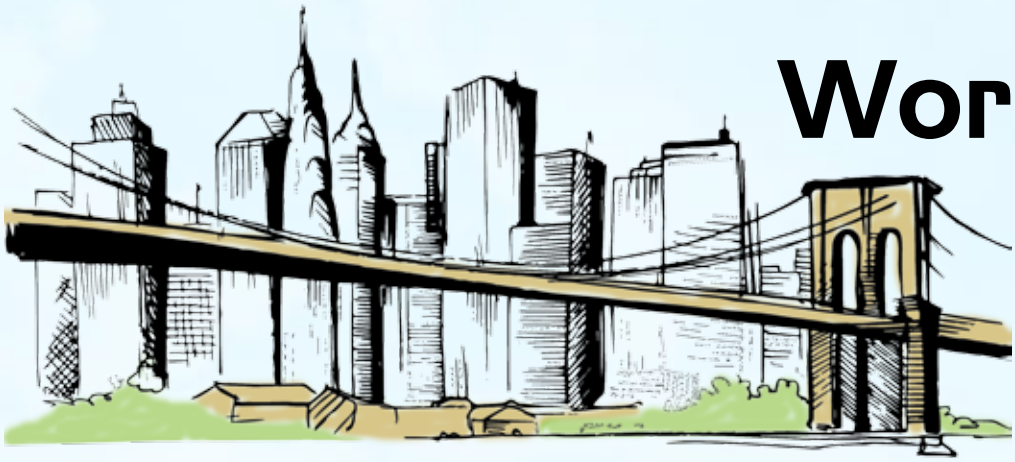
AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

outstanding reviewer awards.

Akshay Balsubramani
Gilles Blanchard
Josip Djolonga
Miroslav Dudik
Søren Hauberg
Qirong Ho
Nikos Karampatziakis
Vladimir Kolmogorov
Edo Liberty
Daniel Lowd
Tengyu Ma
Lester Mackey
Julien Mairal
Gergely Neu
Jennifer Neville
Pascal Poupart
Nikhil Rao
Afshin Rostamizadeh
Sivan Sabato
Peter Sadowski
Lawrence Saul
Mark Schmidt
Yefgeny Seldin
Uri Shalit
Ryan Tibshirani
Robert Tillman
Guy Van Den Broeck
Theophane Weber
David Weiss
Minjie Xu
Yuchen Zhang



Workshops

JUNE 23RD & 24TH

Maps On Next Page

Thursday

- **Gimli: Geometry in Machine Learning**
  Crowne Plaza: Times Square
- **Machine Learning for Digital Education and Assessment Systems**
  Marriott: Times Square
- **Human Interpretability in Machine Learning**
  Microsoft; Central Park
- **Multi-View Representation Learning**
  Marriott: Carnegie-Booth
- **Visualization for Deep Learning**
  Marriott: Astor
- **Reliable Machine Learning in the Wild**
  Marriott: Empire
- **Neural Networks Back To The Future**
  Crowne Plaza: Broadway
- **Deep Learning Workshop**
  Marriott: Westside Ballroom 3,4
- **Abstraction in Reinforcement Learning**
  Marriott: Marquis
- **Advances in non-convex analysis and optimization**
  Westin: Majestic
- **Machine Learning for Music Discovery**
  Marriott: Wilder
- **Theory and Practice of Differential Privacy (TPDP 2016)**
  Marriott: O'Neil

Friday

- **Data-Efficient Machine Learning**
  Marriott: Astor
- **Computational Biology**
  Marriott: Cantor/Jolson
- **Anomaly Detection 2016**
  Microsoft: Central Park
- **Automatic Machine Learning (AutoML)**
  Marriott: Empire
- **Machine Learning Systems**
  Marriott: Soho
- **#data4good: Machine Learning in Social Good Applications**
  Marriott: Wilder
- **Theory of Deep Learning**
  Marriott: Westside Ballroom 3,4
- **On-Device Intelligence**
  Marriott: Odets
- **Online advertising systems**
  Marriott: Carnegie-Booth
- **Optimization Methods for the Next Generation of Machine Learning**
  Westin: Majestic
- **Computational Frameworks for Personalization**
  Marriott: O'Neil

Workshop Maps

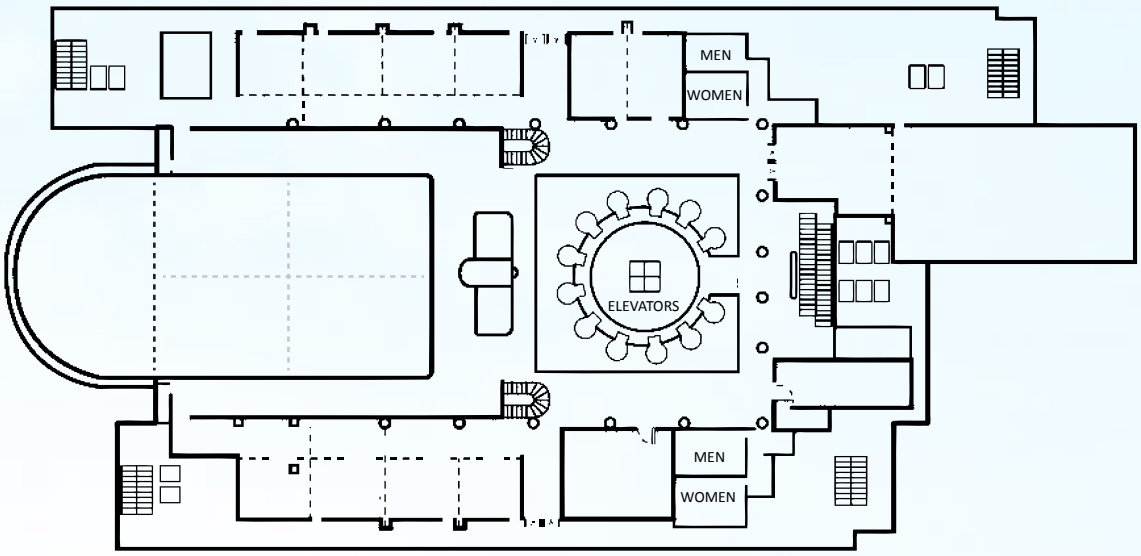
MARRIOTT MARQUIS

4TH FLOOR
- ODETS
- O'NEIL
- WILDER

5TH FLOOR
- CARNEGIE-BOOTH
- WESTSIDE BALLROOMS

7TH FLOOR
- ASTOR BALLROOM
- EMPIRE
- SOHO
- TIMES SQUARE

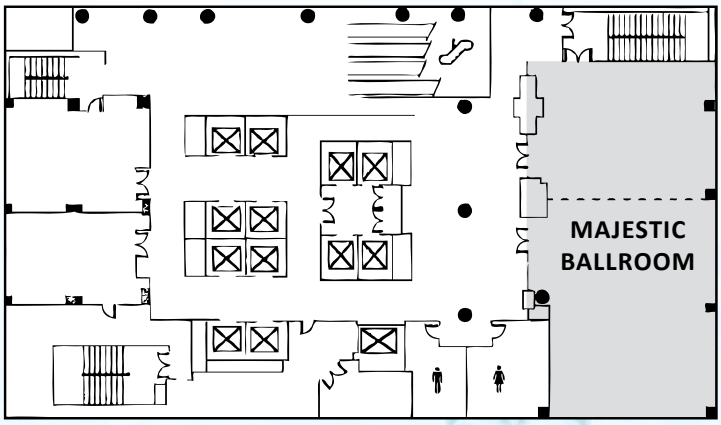
9TH FLOOR
- CANTOR/JOLSON
- MARQUIS



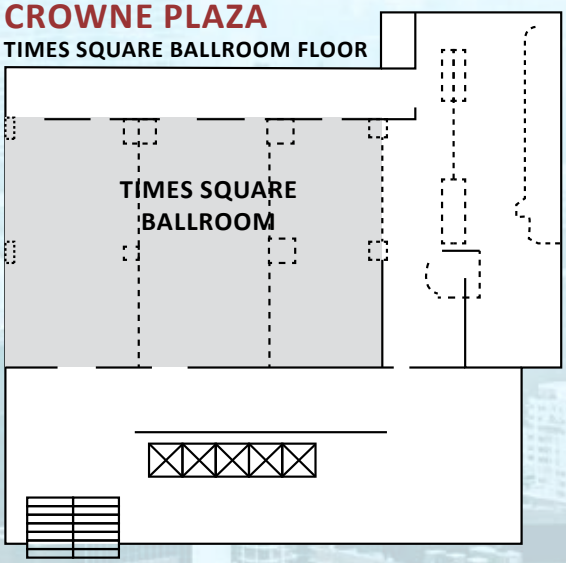
Microsoft Office
Entrance is between 42nd and 41st on 8th Ave.

Central Park Conference Room is Located on the 6th floor

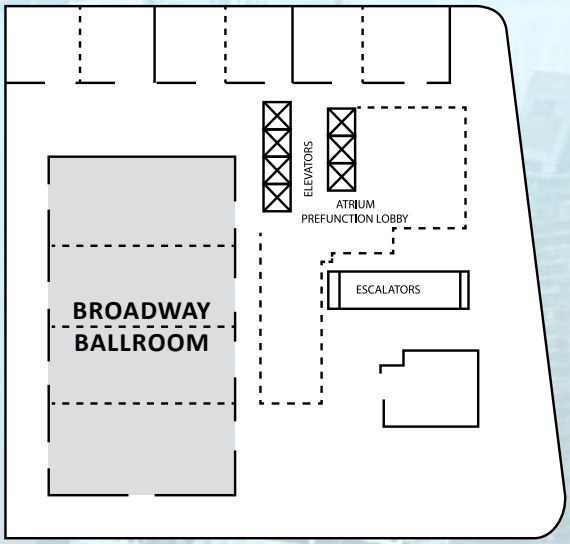
WESTIN NEW YORK - MAJESTIC BALLROOM



CROWNE PLAZA
TIMES SQUARE BALLROOM FLOOR



CROWNE PLAZA - 4TH FLOOR



ICML would like to thank the generous sponsors who help make this conference happen each year.

PLATINUM SPONSORS

Baidu Research

AMERICAN EXPRESS

GOLD SPONSORS

Adobe

Alibaba Group 阿里巴巴集团

amazon.com

Bloomberg

BOSCH

Cubist Systematic Strategies

D-Wave The Quantum Computing Company™

facebook

Google

IBM

intel

Jump Trading

Microsoft

NETFLIX

NVIDIA

Panasonic

QUALCOMM

Shift Technology

TOYOTA RESEARCH INSTITUTE

UBER ADVANCED TECHNOLOGIES CENTER

THE VOLEON GROUP

XILINX

SILVER SPONSORS

accenture High performance. Delivered.

ansatz

ARM

Artificial Intelligence www.elsevier.com/locate/artint

CISCO

sas

TERALYTICS Signal to Information to Value

TWO SIGMA

WORLDQUANT

Yandex

ICML would like to thank the generous sponsors who help make this conference happen each year.

PLATINUM SPONSORS

Baidu Research

American Express

GOLD SPONSORS

Adobe

Alibaba Group 阿里巴巴集团

amazon.com

Bloomberg

BOSCH

Cubist Systematic Strategies

D-Wave The Quantum Computing Company

facebook

Google

IBM

intel

Jump Trading

Microsoft

NETFLIX

NVIDIA

Panasonic

QUALCOMM

Shift Technology

TOYOTA RESEARCH INSTITUTE

UBER ADVANCED TECHNOLOGIES CENTER

THE VOLEON GROUP

XILINX

SILVER SPONSORS

accenture High performance. Delivered.

ansatz

ARM

Artificial Intelligence www.elsevier.com/locate/artint

CISCO

sas

TERALYTICS Signal to Information to Value

TWO SIGMA

WORLDQUANT

Yandex