NIPS 2017

LONG BEACH CA | DEC 4 - 9 | NIPS.CC

Neural Information
Processing Systems

CONFERENCE AT A GLANCE

MONDAY DECEMBER 4TH

7:00 - 8:00 AM	Coffee
8:00 - 10:15 AM	Tutorial 1
10:15 - 10:45 AM	Coffee break
10:45 - 1:00 PM	Tutorial 2
1:00 - 2:30 PM	Lunch on your own
2:30 - 4:45 PM	Tutorial 3
4:45 - 5:00 PM	Quick break
5:00 - 5:30 PM	Opening Remarks
5:30 - 6:20 PM	**Invited talk, John Platt:** *Energy Strategies to Decrease CO2 Emissions*
6:30 - 10:30 PM	Opening Reception and Posters

TUESDAY DECEMBER 5TH

7:30 - 9:00 AM	Coffee
9:00 - 9:50 AM	**Invited talk, Brendan Frey:** *Why AI Will Make it Possible to Reprogram the Human Genome*
9:50 - 10:10 AM	Test Of Time Award
10:10 - 10:40 AM	Coffee break
10:40 - 12:00 PM	**Parallel Tracks on Algorithms and Optimization**
12:00 - 1:50 PM	Lunch on your own
1:50 - 2:40 PM	**Invited talk, Kate Crawford:** *The Trouble with Bias*
2:40 - 2:50 PM	Quick break
2:50 - 3:50 PM	**Parallel Tracks on Algorithms, Optimization & Theory**
3:50 - 4:20 PM	Coffee break
4:20 - 6:00 PM	**Parallel Tracks on Deep Learning, Applications and Algorithms**
6:00 - 7:00 PM	Light snack
7:00 - 10:30 PM	Poster session and Demos

WEDNESDAY DECEMBER 6TH

7:30 - 9:00 AM	Coffee
9:00 - 9:50 AM	**Invited talk, Lise Getoor:** *The Unreasonable Effectiveness of Structure*
9:50 - 10:20 AM	Coffee break
10:20 - 12:00 PM	**Parallel Tracks on Theory, Probabilistic Methods and Deep Learning**
12:00 - 1:50 PM	Lunch on your own
1:50 - 2:40 PM	**Invited talk, Pieter Abbeel:** *Deep Learning for Robotics*
2:40 - 2:50 PM	Quick break
2:50 - 3:50 PM	**Parallel Tracks on Reinforcement Learning, Deep Learning and Optimization**
3:50 - 4:20 PM	Coffee break
4:20 - 6:00 PM	**Parallel Tracks on Reinforcement Learning, Algorithms, Applications and Probabilistic Methods, Applications**
6:00 - 7:00 PM	Light snack
7:00 - 10:30 PM	Poster session and Demos

Table Of Contents

THURSDAY DECEMBER 7TH

7:30 - 9:00 AM	Coffee
9:00 - 9:50 AM	**Invited talk, Yael Niv:** *Learning State Representations*
9:50 - 10:40 AM	**Invited talk, Yee Whye Teh:** *On Bayesian Deep Learning and Deep Bayesian Learning*
10:40 - 11:10 AM	Coffee break
11:10 - 12:30 PM	**Parallel Tracks on Neuroscience and Deep Learning, Algorithms**
12:30 - 2:00 PM	Lunch on your own
2:00 - 4:00 PM	**SYMPOSIA**
4:00 - 4:30 PM	Coffee break
4:30 - 6:30 PM	Symposia
6:30 - 7:30 PM	Light Dinner
7:30 - 9:30 PM	Symposia

FRIDAY & SATURDAY DECEMBER 8TH & 9TH

Each workshop has its own schedule, check the website

7:00 - 8:30 AM	Coffee
10:30 - 11:00 AM	Coffee break
12:00 - 2:00 PM	Lunch on your own
3:00 - 3:30 PM	Coffee Break
6:30 - 10:30 PM	Saturday Closing Reception

2017 SPONSORS

NIPS gratefully acknowledges the generosity of those individuals and organizations who have provided financial support for the NIPS 2017 conference. Their financial support enables us to sponsor student travel and participation, general support to host the conference, and the volunteers who assist during NIPS.

DIAMOND SPONSORS

Intel Nervana™ - Leveraging Intel's world leading position in silicon innovation and proven history in creating the compute standards that power our world, is transforming Artificial Intelligence (AI) with the Intel Nervana platform. Harnessing silicon designed specifically for AI, end-to-end solutions that broadly span from the data center to the edge, and tools that enable customers to quickly deploy and scale up, Intel Nervana is inside AI and leading the next evolution of compute. To learn more, visit: www.intelnervana.com

NVIDIA - We are the pioneer of GPU computing — a supercharged form of computing loved by the most demanding computer users in the world. Fueled by the insatiable demand for better 3D graphics and the massive scale of the gaming market, NVIDIA evolved the GPU into a computer brain at the exciting intersection of virtual reality, high performance computing, and artificial intelligence. In 2006, the creation of the CUDA programming model opened up the parallel-processing capabilities of the GPU to general-purpose computing. And when GPU computing was applied to deep learning in 2012, it ignited the era of modern AI. Today, more than 600,000 developers have embraced the NVIDIA platform to solve some of the most complex problems in computer science. NVIDIA GPU computing has become the essential tool of the da Vincis and Einsteins of our time. For them, we've built the equivalent of a time machine.

AUDI - Audi continues to shape the national conversation about the future of mobility and advancing the state of the art in connected vehicle technology. This year, Audi introduced the first level three automated vehicle with the next A8. By 2020, accelerated by artificial intelligence and deep learning, Audi will deliver highly and fully automated vehicles.

IBM RESEARCH - At IBM Research, we invent things that matter to the world. Today, we are pioneering promising and disruptive technologies that will transform industries and society, including the future of AI, blockchain and quantum computing. We are driven to discover. We are home to 3,000+ researchers including 5 Nobel Laureates, 9 US National Medals of Technology, 5 US National Medals of Science, 6 Turing Awards and 13 Inductees in the National Inventors Hall of Fame.

MICROSOFT - At Microsoft, we aim to empower every person and every organization on the planet to achieve more. We care deeply about having a global perspective and making a difference in lives and organizations in all corners of the planet. This involves playing a small part in the most fundamental of human activities: Creating tools that enable each of us along our journey to become something more. Our mission is grounded in both the world in which we live and the future we strive to create. Today, we live in a mobile-first, cloud-first world, and we aim to enable our customers to thrive in this world.

PLATINUM SPONSORS
















FACEBOOK - Giving people the power to share and connect requires constant innovation. At Facebook, research permeates everything we do. We believe the most interesting research questions are derived from real world problems. Working on cutting edge research with a practical focus, we push product boundaries every day. At the same time, we publish papers, give talks, and collaborate broadly with the academic community.

BAIDU - Baidu was founded in 2000 by Internet pioneer Robin Li, creator of visionary search technology Hyperlink Analysis, with the mission of providing people with the best way to find information and connect users with services.

CITADEL - Analyzing global markets with discipline and focus, we employ a diverse range of strategies in an effort to bring capital to its fullest potential and deliver consistent investment returns to our clients. We do this by working harder, seeing farther, and by empowering the world's most talented minds with the tools and culture they need to achieve peak performance.

NAVER LINE - NAVER, Korea's No. 1 internet company, and LINE, the Japan-based global messenger platform, have teamed up to use their extensive databases of contents and user information to develop an AI platform called Clova (Cloud-based Virtual Assistant). A Clova-powered mobile app and AI speakers have already launched in Korea and Japan, and the companies continue to explore machine learning tech and Applications.

GOOGLE - Research at Google tackles the most challenging problems in Computer Science and related fields. Being bold and taking risks is essential to what we do, and research teams are embedded throughout Google allowing our discoveries to affect billions of users each day.

ALIBABA - Alibaba is the largest retail commerce company in the world. We are building the future infrastructure of commerce to help merchants, brands, and small businesses all over the world reach and engage consumers and operating a range of online marketplaces that connect buyers and sellers, as well as various synergistic services such as cloud computing, and digital media and entertainment.

DEEPMIND - At DeepMind, our mission is to solve intelligence and then use that to make the world a better place. Our motivation in all we do is to maximize the positive and transformative impact of AI. We believe that AI should ultimately belong to the world, in order to benefit the many and not the few, and we steadfastly research, publish and implement our work to that end.

AMAZON - Amazon strives to be Earth's most customer-centric company where people can find and discover virtually anything they want to buy online. The world's brightest technology minds come to Amazon.com to research and develop technology that improves the lives of shoppers, sellers and developers. For more information about Machine Learning at Amazon, visit: amazon.jobs/NIPS

GRAPHCORE - Graphcore has created a completely new processor that's the first to be specifically designed for machine intelligence workloads – an Intelligence Processing Unit (IPU) that will set a new pace of innovation.

CAPITAL ONE - At Capital One, we dare to dream, disrupt and deliver a better way. Our goal is simple – bring ingenuity, simplicity, and humanity to an industry ripe for change. Founder-led, Capital One is on a mission to help people live their best lives and build one of America's leading information-based technology companies.

DIDI - Didi Chuxing is the world's leading mobile transportation platform. We are committed to working with communities and partners to solve the world's transportation, environmental and employment challenges by using big data-driven deep- learning algorithms that optimize resource allocation. By continuously improving user experience and creating social value, we strive to build an open, efficient, collaborative, and sustainable transportation ecosystem.

KLA-TENCOR - At KLA-Tencor, we research, develop, and manufacture the world's most advanced inspection and measurement equipment for the semiconductor industry. We use machine learning and computational physics to drive data analytics. By enabling the manufacture of ICs at next generation technology nodes and collaborating with institutes all over the world, we're leading the semiconductor industry and we've enabled Moore's Law for more than 40 years.

UBER - AI that moves the world. Uber's mission is to make transportation as reliable as running water—everywhere, for everyone. At AI Labs, we drive this mission by developing cutting-edge machine learning algorithms that extend the state of the art. By blending a wide variety of approaches across the field, we deliver innovation to application.

APPLE INC - Apple revolutionized personal technology with the introduction of the Macintosh in 1984. Today, Apple leads the world in innovation with iPhone, iPad, Mac, Apple Watch and Apple TV. Apple's four software platforms — iOS, macOS, watchOS and tvOS — provide seamless experiences across all Apple devices and empower people with breakthrough services including the App Store, Apple Music, Apple Pay and iCloud. Apple's more than 100,000 employees are dedicated to making the best products on earth, and to leaving the world better than we found it.

DATA COLLECTIVE - Data Collective (DCVC) is a venture capital fund that backs entrepreneurs applying deep tech, including cutting edge ML and AI, to transform giant industries. DCVC and its principals have supported brilliant people changing global-scale businesses for over twenty years, helping create tens of billions of dollars of wealth while also making the world a markedly better place.

TENCENT - Tencent AI Lab is a leading AI research and application lab of Tencent, China's largest internet company. It was founded in 2016 and backed by 70 world-class research scientists and 300 experienced engineers in China and US. Its research focuses on: machine learning, computer vision, speech recognition and natural language processing. To serve the needs of Tencent's core business, its application focuses on: content, game, social and platform AI.

GOLD SPONSORS

BOREALIS AI - Borealis AI, a RBC Institute for Research, is a curiosity-driven research centre dedicated to achieving state-of-the-art in machine learning. Established in 2016, and with labs in Toronto and Edmonton, we support open academic collaborations and partner with world-class research centres in artificial intelligence. With a focus on ethical AI that will help communities thrive, our machine learning scientists perform fundamental and applied research in areas such as Reinforcement learning, natural language processing, deep learning, and unsupervised learning to solve ground-breaking problems in diverse fields.

JD.COM - JD.com is both the largest ecommerce company in China and the country's largest retailer by revenue. The company strives to offer consumers the best online shopping experience. The company has the largest fulfillment infrastructure of any e-commerce company in China. JD.com is a member of the NASDAQ100 and a Fortune Global 500 company.

OPTIVER - Optiver's story began over 30 years ago, when we started business as a single trader on the floor of Amsterdam's options exchange. Today, we are at the forefront of trading and technology as a leading global electronic market maker, focused on pricing, execution and risk management. With over 1,000 employees globally, our mission to improve the market unites us.

PREFERRED NETWORKS - Preferred Networks Inc. (PFN) is a Tokyo-based startup that applies the latest artificial intelligence technologies to emerging industrial problems. PFN develops Chainer, a Python deep learning framework. PFN works with many world-leading companies, such as FANUC for manufacturing robots, Toyota Motors for autonomous driving, and National Cancer Research Center in Japan for healthcare. Our team consists of experts in various fields not only in machine learning but also computer vision, robotics, natural language processing, bioinformatics and many others. The company is expanding globally, and its subsidiary is located in California.

SALESFORCE - Salesforce, the global CRM leader, empowers companies to connect with their customers in a whole new way. The company has been democratizing technology since 1999, making the cloud, mobile, social, IoT, and now AI available to all companies, regardless of size and scale. Salesforce's Customer Success Platform includes cloud-based Applications for sales, service, marketing, commerce and more.

VOLEON - Founded in 2007 by leading machine learning scientists, The Voleon Group designs, develops, and implements advanced technology for investment management. We are committed to solving large-scale financial prediction problems with statistical machine learning.

DE SHAW - The D. E. Shaw group is a global investment and technology development firm with more than $43 billion in investment capital as of July 1, 2017, and offices in North America, Europe, and Asia. Since our founding in 1988, our firm has earned an international reputation for successful investing based on innovation, careful risk management, and the quality and depth of our staff. We have a significant presence in the world's capital markets, investing in a wide range of companies and financial instruments in both developed and developing economies.

AIRBNB - Founded in 2008, Airbnb's mission is to create a world where people can belong when they travel by being connected to local cultures and having unique travel experiences. Its community marketplace provides access to millions of unique accommodations and experiences in more than 65,000 cities and 191 countries

MERCEDES-BENZ - Mercedes-Benz Research and Development North America, Inc. is a place for exceptional people with outstanding ideas and the absolute willingness to bring them to life. With R&D facilities across the U.S., engineers, software developers, and designers push technology boundaries to shape the future – it's not just about cars at MBRDNA, it's also about the latest and greatest software, cutting-edge technology, and groundbreaking innovation.

WINTON - Winton is an investment management and data science company. The firm was founded in 1997 on the belief that the application of the scientific method offers the best approach to investing. Winton's investment decisions are driven by statistical inference based on the empirical analysis of data, rather than instinct or intuition.

CRITEO RESEARCH - Criteo Research is pioneering innovations in computational advertising. As the center of scientific excellence in the company, Criteo Research delivers both fundamental and applied scientific leadership through published research, product innovations and new technologies powering the company's products. We are looking for outstanding machine learning research scientists whose skills span the entire spectrum of scientific research and are interested in revolutionizing the world of online and computational advertising.

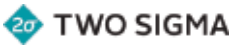
RENAISSANCE - Renaissance Technologies is a quantitative hedge fund management company founded by James Simons in 1982 and located in East Setauket, NY. Renaissance has 300 employees, 90 of whom have PhDs in mathematics, physics, statistics, or computer science. The firm's trading is based on models developed through the application of machine learning to massive quantities of financial data.

NETFLIX - Netflix is the world's leading internet entertainment service with 104 million members in over 190 countries enjoying more than 125 million hours of TV shows and movies per day, including original series, documentaries and feature films. Members can watch as much as they want, anytime, anywhere, on nearly any internet-connected screen.

MALUUBA - Maluuba is a Microsoft company teaching machines to think, reason and communicate. We demonstrate progress by focusing on challenging problems, driving new techniques in Reinforcement learning with a focus on machine comprehension and conversational interfaces. Our vision is to solve 'Artificial General Intelligence' by creating literate machines that can think, reason and communicate like humans.

ELEMENT AI - Element AI is the platform that enables organizations to embrace an AI-First world for today and tomorrow. Composed of a research lab and consultancy uniquely connected to the world's best academic ecosystems, Element AI creates novel AI-First solutions for large corporations.

SILVER SPONSORS




HUAWEI - Huawei is a leading global information and communications technology solutions provider whose business spans telecommunication infrastructure, consumer electronics and business IT solutions. Huawei's Noah's Ark Labs, located in Hong Kong, Shenzhen, Paris, London and other locations, focus on machine learning, data mining, speech and language processing, and computer vision.

TWO SIGMA - At Two Sigma, we imagine breakthroughs in investment management, insurance and related fields by pushing the boundaries of what open source and proprietary technology can do. Our engineers and modelers harness data at tremendous scale, using machine learning, distributed computing and other technologies to build powerful predictive models.

SIGOPT - SigOpt is an optimization platform that seamlessly tunes AI and ML model parameters through a state-of-the-art ensemble of Bayesian and global optimization algorithms behind a simple API. SigOpt can tune any predictive or machine learning model right in place, and the federated API design ensures no proprietary data leaves your premises.

EBAY - eBay is the world's most vibrant marketplace! Our vision for commerce is enabled by people, powered by technology and open to everyone. If you want write a chapter in the next part of our story please check out www.ebaycareers.com for all of our current opportunities.

ORACLE - Tackling Today's Biggest Challenges. The Mission of Oracle Labs is straightforward: Identify, explore, and transfer new technologies that have the potential to substantially improve Oracle's business. Oracle's commitment to R&D is a driving factor in the development of technologies that have kept Oracle at the forefront of the computer industry.

YANDEX - Yandex is one of the largest internet companies in Europe, operating Russia's most popular search engine. We provide user-centric products and services based on the latest innovations in information retrieval, machine learning and machine intelligence to a worldwide customer audience on all digital platforms and devices.

INVENIA LABS - Invenia Labs develops machine learning techniques to solve some of the world's most complex forecasting and decision problems. Located in Cambridge (UK), Invenia's current mission is to improve the efficiency of electrical grids, helping to reduce pollution and fight the global climate crisis.

SAP - As market leader in corporate application software, SAP now integrates machine learning technologies to empower business growth.SAP Leonardo,our digital innovation system offers an entire portfolio of automated Applications&services that allows companies to run better,rapidly innovate & scale.By bringing ML, IoT, blockchain, Big Data together on SAP cloud platform we enable the intelligent enterprise.

RECURSION - Recursion is an emerging biotechnology company that combines experimental biology and bioinformatics with artificial intelligence in a massively parallel system to efficiently identify treatments for genetic diseases utilizing massive, data-rich image sets. Recursion is becoming the first truly AI-enabled biotech company with a vision to ultimately create a map of human cellular biology.

JUMP TRADING - Jump Trading is a leading quantitative trading firm built upon cutting-edge research and machine learning, high-performance technology, and an entrepreneurial culture. We build predictive models from big data and develop algorithms to automatically execute trades in dozens of financial exchanges around the world.

DISNEY RESEARCH - Disney Research's objective is to drive value across The Walt Disney Company by injecting scientific & technological innovation. Our world-class research seeks to invent and transfer the most compelling technologies enabling the company to differentiate its content, services, and products. Disney Research combines the best of academia and industry, by doing basic and application-driven research.

QUORA - We want to connect the people who have knowledge to the people who need it, to bring together people with different perspectives so they can understand each other better, and to empower everyone to share their knowledge for the benefit of the rest of the world.

CIRRASCALE - Cirrascale Cloud Services is a premier provider of public and private dedicated, multi-GPU cloud solutions enabling deep learning. The company offers cloud-based solutions for large-scale deep learning operators, service providers, as well as HPC users. To learn more about Cirrascale Cloud Services and its unique dedicated, multi-GPU cloud solutions, please visit http://www.cirrascale.cloud

TOUTIAO - Founded in March 2012, Beijing Bytedance Technology Co., Ltd. runs Toutiao, a mobile application that makes personalized content recommendation based on data mining. Toutiao's mission is to help its users consume the most valuable information in the most convenient fashion.

BENEVOLENT AI - BenevolentAI is the global leader in the development and application of AI for scientific innovation and one of the world's top five private AI companies. It is transforming the process of scientific discovery by enabling previously unimaginable scientific advances, initially in the accelerated development of new medicines and more recently in energy storage.

SWISS DATA SCIENCE CENTER - The Swiss Data Science Center, a joint venture between EPFL and ETH Zurich, aims to accelerate the adoption of data science and machine learning techniques within academic disciplines, the academic community at large, and the industrial sector. It addresses the gap between those who create data, those who develop data analytics and systems, and those who could extract value from it.

SILVER SPONSORS

PDT PARTNERS - PDT Partners is a NYC-based quantitative hedge fund with a a stellar twenty-three-year track record and a reputation for excellence in the algorithmic trading space; we apply rigorous scientific investigation to model financial markets globally. We have a strong history of hiring, challenging and retaining engineers and scientists from a variety of industries and disciplines.

CROWDFLOWER - CrowdFlower is the essential human-in-the-loop AI platform for data science and machine learning teams. CrowdFlower's technology and expertise trains, tests, and tunes machine learning models for a wide range of use cases from autonomous vehicles, content moderation, and chatbots to medical image labeling, CRM management, and search relevance. For more information, visit www.crowdflower.com.

COGENT LABS - Cogent Labs is an AI company that aspires to shape the future of how people work and live using artificial intelligence. Cogent Labs bridges the gap between newly published academic research and real-world business solutions by leveraging expertise across many domains, including time series forecasting, information extraction, natural language and speech processing, and Reinforcement learning.

TARGET - Minneapolis-based Target Corporation (NYSE: TGT) serves guests at 1,816 stores and at Target.com. Since 1946, Target has given 5 percent of its profit to communities, which today equals millions of dollars a week. For more information, visit Target.com/Pressroom. For a behind-the-scenes look at Target, visit Target.com/abullseyeview or follow @TargetNews on Twitter.

ADOBE - Adobe is the global leader in digital marketing and digital media solutions. Our tools and services allow our customers to create groundbreaking digital content, deploy it across media and devices, measure and optimize it over time and achieve greater business success. We help our customers make, manage, measure and monetize their content across every channel and screen.

LAMBDA - Lambda DevBoxes and GPU servers are optimized for training Deep Neural Networks. Pre-installed with every Deep Learning library: TensorFlow, Torch/PyTorch, Caffe with NCCL, Keras, Theano, and MXNet. Buy it, plug it in, and start training.

XPRIZE - XPRIZE, a 501(c)(3) nonprofit organization, is the global leader in designing & implementing innovative competition models to solve the world's grandest challenges. The IBM Watson AI XPRIZE is a $5M global competition challenging teams to develop powerful Artificial Intelligence (AI) based Applications and demonstrate how humans can collaborate with AIs to tackle the world's grand challenges.

BLOOMBERG - Bloomberg's core product, the Terminal, is a must-have for the most influential people in finance. We are building machine learning models for predicting the impact of news stories on company prices, recommendation systems, semantic parsing for question answering, topic clustering and classification, sentiment analysis, anomaly detection in time series and a variety of other problems.

VATIC LABS - The word vatic means to describe or predict what will happen in the future, and that's exactly what we do. Vatic Labs is a quantitative trading firm where traders, AI researchers, and technologists collaborate to develop autonomous trading agents and cutting edge technology. We work together, building systems that boost market efficiency and transparency.

FEATUREX - FeatureX uses AI to discover meaningful patterns in the world's data. We use computer vision, deep learning and statistical machine learning on two projects. We process high-resolution satellite images to extract patterns in economic activity and we analyze financial datasets to build systematic trading models. We offer challenging problems, creative freedom, and a focus on research and learning.

AUTOX - AutoX is a startup creating full-stack A.I. software solution (perception, planning, and control) for level-5 full-autonomy self-driving vehicles. AutoX invented a camera-first self-driving solution that costs only a tiny fraction of traditional LiDar- based approaches. We believe that autonomous driving should not be a luxury, and we are making it universally available to every citizen.

ROSETTA ANALYTICS - Rosetta Analytics is an investment management startup committed to using artifi-cial intelligence and machine learning with traditional and nontraditional data sources to create scalable investment strategies for pension funds, endowments, foundations, and sovereign wealth funds.

OBSIDIAN SECURITY - Led by former founding team members of Cylance and Carbon Black, Obsidian Security is a data driven start-up living at the intersection of cybersecurity, artificial intelligence, and hybrid-cloud environments. Backed by Greylock Partners, Obsidian Security is focused on using applied mathematics and machine learning to bring insights to organizations using modern hybrid architectures.

MAN AHL - We are a quantitative investment manager. A pioneer of systematic trading since 1987, we mix machine learning, computer science and engineering with terabytes of data to invest billions of dollars every day. Our collaboration with academia – the Oxford-Man Institute of Quantitative Finance – celebrated its 10th anniversary in 2017. We are a flat-structured company that seeks the best.

HUDSON RIVER - Hudson River Trading brings a scientific approach to trading financial products. We have built one of the world's most sophisticated computing environments for research and development. Our researchers are at the forefront of innovation in the world of algorithmic trading.

SBERBANK - Sberbank is a powerful modern bank which is rapidly becoming one of the major global financial institutions. Accounting for one third of Russian banking system, Sberbank today is the only Russian bank featuring in the World's Top 50 Biggest Banks (total assets more than 400 bl US$). Sberbank has 145,6 mln retail clients worldwide and 1,8 mln corporate clients in Russia.

SILVER SPONSORS

JP MORGAN CHASE & CO - At JPMorgan Chase & Co, technology innovation is driven by a shared commitment to stay ahead of our customers' needs globally. In our worldwide tech centers, our team of 40,000 technology professionals collaborate to design, build, & deploy solutions that include strategic technology initiatives, big data, mobile solutions, electronic payments, machine learning, cyber security & cloud development.

SUSQUEHANNA INTERNATIONAL GROUP LLC - SIG is a global quantitative trading firm founded with an entrepreneurial mindset and a rigorous analytical approach to decision making. We commit our own capital to trade financial products around the world. Building virtually all of our own trading technology from scratch, we are leaders and innovators in high performance, low latency trading.

MONTREAL INTERNATIONAL - Montreal International is greater Montréal's economic development agency. Its mission is to contribute to the economic development of Greater Montréal and enhance its international status. Its mandates include attracting foreign direct investments, international organizations and international strategic workers as well as promoting the competitive and international environment of Greater Montréal.

CUBIST SYSTEMATIC STRATEGIES - Cubist Systematic Strategies, the quantitative investing business of Point72, deploys systematic, computer-driven trading strategies across multiple liquid asset classes, including equities, futures, and foreign exchange. The core of our effort is rigorous research into a wide range of market anomalies, fueled by our unparalleled access to a wide range of publicly available data sources.

TUDOR - Tudor is an investment firm with deep experience across market cycles gained over nearly four decades of trading worldwide financial markets. Driven by a commitment to innovation and excellence, we seek to generate consistent returns for our clients and a dynamic environment for our staff through persistent evolution of our technology, research methods and trading and investing techniques .

G-RESEARCH - G-Research is a leading quantitative research and technology company. We apply scientific techniques to find patterns in large, noisy and real-world data sets, using the latest statistical and "big data" analysis methodologies to predict global financial markets. Our technology, research and resources are combined to build a single, powerful platform for developing your ideas. We use rigorous...

WAVE COMPUTING - Wave Computing is the Silicon Valley start-up that is revolutionizing machine learning with its roadmap of WaveFlow compute appliances. The company's solutions leverage its native dataflow technology to outperform machine learning products available today. Wave Computing, one of CIO Application Magazine's Top 25 Artificial Intelligence Providers, is offering its solutions to customers globally.

AIG - American International Group, Inc. is a leading international insurance organization with the vision to become its clients' most valued insurer. AIG believes in harnessing the power of machine learning and deep learning techniques to generate new insights from data and to enhance human judgment in real business contexts. If you are passionate about evidence-based decision making, connect with AIG!

QUANTUMBLACK - QuantumBlack, a McKinsey company, is an advanced analytics firm operating at the intersection of strategy, technology & design to improve performance outcomes for organisations. With roots in Formula One, we now work across sector with some of the world's leading organisations in advanced industries, healthcare, and finance.

SNAP INC - Snap Inc. is a camera company. We believe that reinventing the camera represents our greatest opportunity to improve the way people live and communicate. Our products empower people to express themselves, live in the moment, learn about the world, and have fun together.

QUALCOMM - Qualcomm technologies powered the smartphone revolution, connecting billions of people. We pioneered 3G and 4G. Now we're leading the way to 5G - a new era of intelligent, connected devices. Our products are revolutionizing industries including automotive, computing, IoT, healthcare and data center.

BRONZE SPONSORS

PANASONIC - Panasonic is focusing on bringing new solutions to an ever-changing environment, full of cutting edge technologies. We apply Deep Learning as a tool to improve the real-life situations of today and the evolving situations of tomorrow. Deep Learning is just one of the key technologies we employ to understand more about each other and how to navigate through our lives: safely, honestly and happily.

QIHOO 360 - Qihoo 360 is a Chinese internet security company known for its antivirus software (360 Safeguard, 360 Mobile Safe), Web Browser (360 Browsers), and Mobile Application Store (360 Mobile Assistant).

CYLANCE - CylanceOPTICS™ is an artificial intelligence driven endpoint detection and response (EDR) solution designed to extend the prevention delivered by Cylance's award-winning product, CylancePROTECT®, through AI driven root cause analysis, scalable threat hunting, and immediate response with consistent visibility into threats against endpoints. Visit www.cylance.com.

PETUUM - Petuum, Inc. is a Machine Learning and Artificial Intelligence company built upon years of research in Machine Learning and Distributed Systems by leading Carnegie Mellon University faculty and students. Petuum aims to virtualize hardware and ubiquitize AI, by creating a powerful, programmable and distributed AI operating system with a familiar user experience - becoming an everyday tool for managers, professionals, programmers and data scientists in the enterprise. The Petuum operating system is omni-lingual (programmable in all languages), omni-mount (deployable on all hardware), and omni-source (compatible with all data formats). Built for the enterprise, Petuum is putting practical AI into the hands of everyone who needs it.

SENTIENT TECHNOLOGIES - Sentient Technologies builds disruptive products differentiated by scaled AI. Sentient's powerful distributed artificial intelligence (AI) platform combines multiple AI disciplines (deep learning, evolutionary AI, neuroevolution) with unmatched scale. Sentient uses its platform in finance (sentientim.com), retail (sentient.ai/aware), and digital media (sentient.ai/ascend).

NEXTAI - NextAI is a startup accelerator located in Toronto, Canada for entrepreneurs, researchers and scientists looking to launch AI-enabled ventures. We provide up to $200K seed funding, dedicated workspace and in-depth business and technical education taught by award-winning global faculty and entrepreneurs. Come build your company in one of the global hotspots for AI research and commercialization. Learn more at www.nextai.com

EXHIBITORS

Morgan & Claypool Publishers

MDPI

The MIT Press

Springer Machine Learning Journal

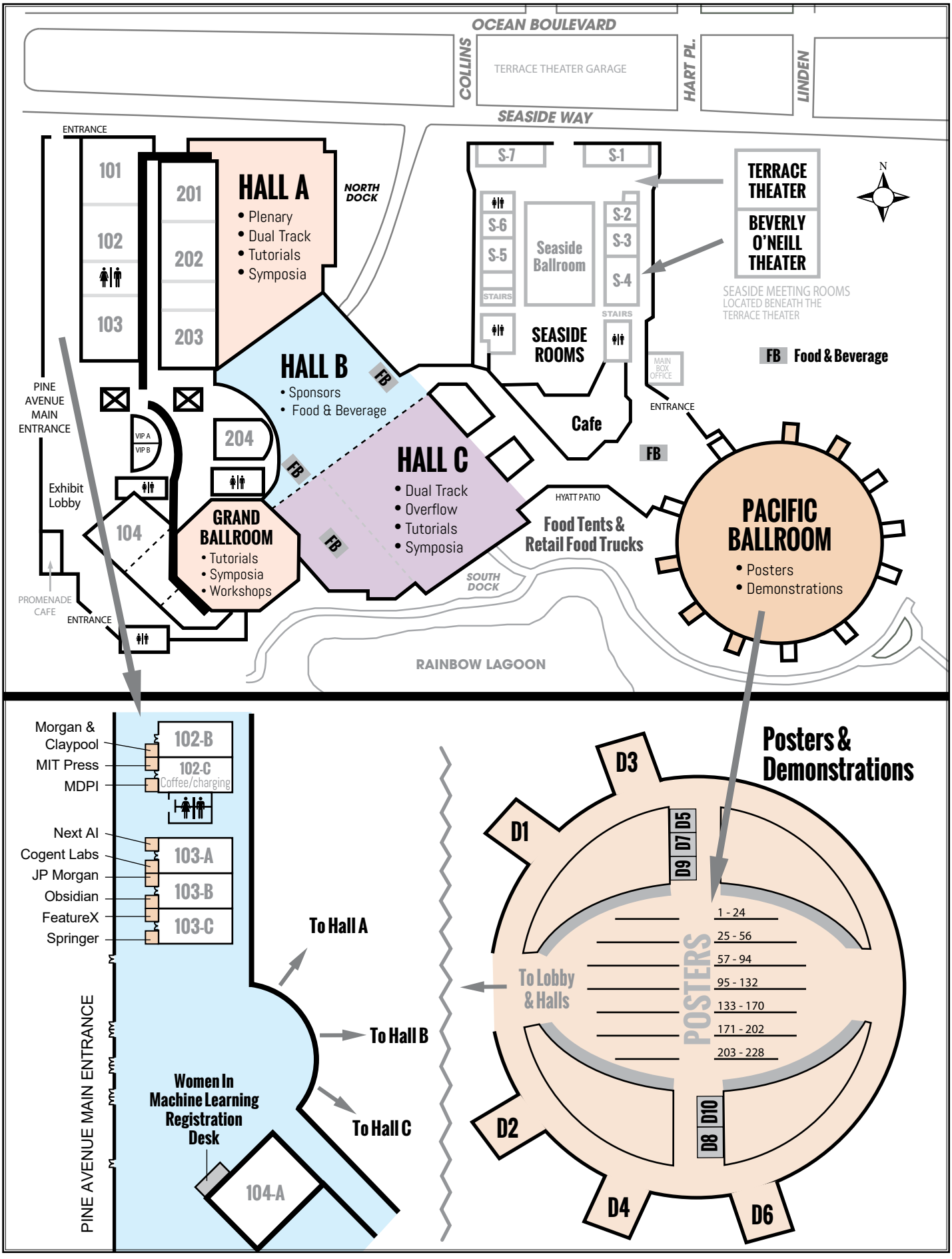
facebook

NIPS would like to especially thank Facebook for streaming services.

Microsoft

NIPS would like to especially thank Microsoft Research for their donation of Conference Management Toolkit (CMT) software and server space.

CONFERENCE MAP



SPONSOR MAP

HALL B - SPONSOR AREA

KLA-Tencor

AutoX

Borealis AI

Naver | Line

Preferred Networks

Apple Inc.

CrowdFlower

Disney Research

Didi Chuxing

Man AHL

PDT Partners

Google Inc

Montreal Intl.

Salesforce

DeepMind

Jump Trading

Facebook

Benevolent AI

Bloomberg

Citadel LLC

Cubist System

Food & Beverage Area

QuantumBlack	Vatic Labs
SigOpt	Toutiao
Rosetta Analytics	SAP SE
Recursion Pharmaceuticals	Yandex
Oracle	Invenia Labs
eBay Inc.	G Research
Sberbank	Two Sigma

SNAP Inc.	Target
Qualcomm	Adobe Research
Cirrascale Cloud Services	XPrize
Renaissance Technologies	Wave Computing
	Hudson River Trading
Hauwei	Swiss Data Science Center
Lambda Labs	Quora

Tencent

Amazon

Alibaba

Beijing Baidu

Graphcore

Uber

Capital One

Data Collective

JD.COM

Element AI

Optiver

AirBnB

Voleon Group

Criteo

Mercedes-Benz

D. E. Shaw

Maluuba Inc.

Netflix

Winton Group

DIAMOND SPONSOR AREA

IBM Research

Microsoft

NVIDIA

Audi Of America

intel Nervana™

FOOD & BEVERAGE AREA

GENERAL INFORMATION

REGISTRATION DESK

Sunday, December 3: 12:00 pm – 8:00 pm
Monday – Friday: 7:00 am – 6:00 pm
Saturday, December 9 : 7:00 am - 12:00 pm

OPENING RECEPTION AND POSTER SESSION

Monday, December 4 starting at 6:30 pm

Food service will be in many locations;
Please see the maps on the next page

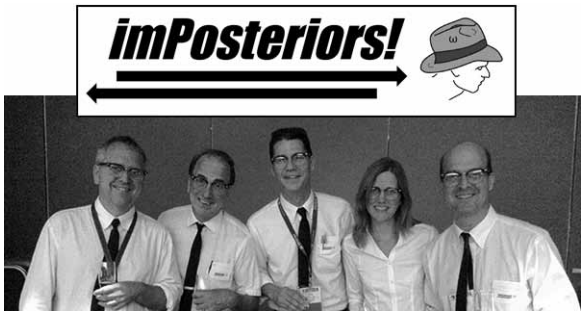
CLOSING RECEPTION PACIFIC BALLROOM

Saturday, December 9 at 6:30 pm
Performance By The Imposteriors (See Below)
Deep Learning Art Exhibit (See below)

POSTER SESSIONS PACIFIC BALLROOM

Monday, Dec. 4, 7:00 pm – 10:30 pm
Tuesday, Dec. 5, 7:00 pm – 10:30 pm
Wednesday, Dec. 6, 7:00 pm – 10:30 pm

Take down your poster at 10:30 pm or they will be discarded.



The Imposteriors (Brad Carlin, Don Hedeker, Mark Glickman, Jennifer Hill, and Michael Jordan) is a band made up of professors whose goal in life is to inspire even the most awkward music-lover to dance. We play a variety of musical genres from Motown to classic rock 'n roll to indie pop to 80's classics to punk/polka to current rock to "I'm-embarrassed-to-admit-I-like-that-song-from-the-radio" crowd pleasers. If it makes you want to dance, we want to play it.

We hail from five different parts of the country but get together several times a year at academic conferences and find a way to play together. In between gigs we live the lives of unassuming university professors whose students would never suspect we are really part-time rock legends. (In our show at NIPS, we will also also be joined by a handful of mystery guests, who are also unassuming university professors, and who also may be on their way to legendary status. Or not...)

WIFI

SSID: NIPS
Password: conference

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: **nips**

CHARGING TABLES

Located throughout the venue

SPONSORS & EXHIBITORS

Promenade lobby & Hall B.

DEEP LEARNING ART EXHIBIT

In the Pacific Ballroom, NIPS is organizing an art exhibit (sponsored by DeepArt, ChaLearn and Nvidia). It will display 50 posters that have been generated with a neural network program fusing the structure of a picture and the artistic style of another (an initial selection was made from submissions shown at the web address below according to popular votes). The posters most acclaimed by visitors of the NIPS exhibit will win a free dinner invitation. Please visit!

https://deepart.io/nips/submissions/votes

ORGANIZING COMMITTEE

General Chairs
Isabelle Guyon, U. Paris-Saclay & ChaLearn
Ulrike von Luxburg, U. of Tübingen

Program Chair
Samy Bengio, Google Brain

Program Co-Chairs
Hanna Wallach, MSR NYC
Rob Fergus, Facebook AI Research / NYU
S.V.N. Vishwanathan, UCSC and Amazon

Tutorials Chairs
Jenn Wortman Vaughan, Microsoft Research
Samuel Kaski, Aalto U. and U. of Helsinki

Workshop Chairs
Ralf Herbrich, Amazon
Suchi Saria, Johns Hopkins U.

Demo and Competition Chairs
Markus Weimer, Microsoft Research
Sergio Escalera, U. of Barcelona

Press Chairs
Neil Lawrence,U. of Sheffield / Amazon Cambridge
Katherine Gorman, Talking Machines / Collective Next

Publications Chair and Electronic Proceedings Chair
Roman Garnett,Washington U. in St. Louis

Program Managers
Choon Hui Teo, Amazon
Daniel Hill, Amazon

NIPS FOUNDATION OFFICERS & BOARD MEMBERS

PRESIDENT
Terrence Sejnowski, The Salk Institute

TREASURER
Marian Stewart Bartlett, Apple

SECRETARY
Michael Mozer, UC Boulder

EXECUTIVE DIRECTOR
Mary Ellen Perry, The Salk Institute

LEGAL ADVISOR
David Kirkpatrick

IT DIRECTOR
Lee Campbell, The Salk Institute

EXECUTIVE BOARD
Zoubin Ghahramani, Univ. of Cambridge
Corinna Cortes, Google Research
Léon Bottou, Microsoft Research
Chris J.C. Burges, Microsoft Research
Neil D. Lawrence, Univ. of Sheffield
Fernando Pereira, Google Research
Max Welling, Univ. of Amsterdam

ADVISORY BOARD
Peter Bartlett, Queensland Univ., UC Berkley
Sue Becker, McMaster Univ., Ontario, Canada
Yoshua Bengio, Univ. of Montreal, Canada
Jack Cowan, Univ. of Chicago
Thomas G. Dietterich, Oregon State Univ.
Stephen Hanson, Rutgers Univ.
Michael I. Jordan, Univ. of California, Berkeley
Michael Kearns, Univ. of Pennsylvania
Scott Kirkpatrick, Hebrew Univ., Jerusalem
Daphne Koller, Stanford Univ.
John Lafferty, Univ. of Chicago
Todd K. Leen, Oregon Health & Sciences Univ.
Richard Lippmann, MIT
Bartlett Mel, Univ. of Southern California
John Moody, UC Berkeley & Portland
John C. Platt, Google
Gerald Tesauro, IBM Watson Labs
Sebastian Thrun, Stanford Univ.
Dave Touretzky, Carnegie Mellon Univ.
Lawrence Saul, Univ. of California, San Diego
Bernhard Schölkopf, MP, Tübingen/Stuttgart
Dale Schuurmans, Univ. of Alberta, Canada
John Shawe-Taylor, Univ. College London
Sara A. Solla, Northwestern Univ. Med. School
Yair Weiss, Hebrew Univ. of Jerusalem
Chris Williams, Univ. of Edinburgh
Rich Zemel, Univ. of Toronto

SUPPORT TEAM

The organization and management of NIPS would not be possible without the help of many volunteers, students, researchers and administrators who donate their valuable time and energy to assist the conference in various ways. The support staffs' tireless efforts make the conference run smoothly and efficiently every year. NIPS would particularly like to acknowledge the exceptional work of:

Brad Brockmeyer
Terrance Gaines
Andrea Brown
Peggy Briggs
Jen Perry
Ramona Marchand

PROGRAM COMMITTEE

Senior Area Chairs
Aapo Hyvarinen, Aaron Courville, Alina Beygelzimer, Antoine Bordes, Claire Monteleoni, Claudio Gentile, Corinna Cortes, Dale Schuurmans, Eric Xing, Francis Bach, Gunnar Raetsch, Hugo Larochelle, Inderjit Dhillon, Irina Rish, Jeff Bilmes, Jennifer Neville, Jerry Zhu, John Langford, Katherine Heller, Kristen Grauman, Marc'Aurelio Ranzato, Maria-Florina Balcan, Marina Meila, Padhraic Smyth, Peter Stone, Raquel Urtasun

Area Chairs
Afshin Rostamizadeh, Akshay Krishnamurthy, Alain Rakotomamonjy, Alan Stocker, Alekh Agarwal, Aleksandrs Slivkins, Alexander Ihler, Alexandra Carpentier, Alexandros Dimakis, Ali Cemgil, Alp Kucukelbir, Ambuj Tewari, Amir Globerson, Angela Yu, Anima Anandkumar, Anitha Kaanan, Ariadna Quattoni, Armand Joulin, Arnak Dalalyan, Arthur Szlam, Barbara Engelhardt, Barnabas Poczos, Been Kim, Benjamin Marlin, Benjamin Van Roy, Bert Huang, Bharath Sriperumbudur, Bill Triggs, Brian Kulis, Cedric Archambeau, Cheng Soon Ong, Chris Pal, Cordelia Schmid, Daniel Hsu, Danilo Rezende, Dan Sheldon, David Duvenaud, David Mimno, David Sontag, David Wipf, Dhruv Batra, Dilan Gorur, D. Sculley, Dumitru Erhan, Elad Hazan, Elena Zheleva, Emma Brunskill, Emmanuel Abbe, Eunho Yang, Fei Sha, Fernando Perez-Cruz, Finale Doshi-Velez, Francesco Orabona, François Fleuret, Francois Laviolette, Frank Wood, Gal Chechik, Gilles Blanchard, Graham Taylor, Greg Mori, Guillaume Bouchard, Guillaume Obozinski, Hang Li, Hartmut Neven, Honglak Lee, Hugo Jair Escalante, Ian Goodfellow, Ichiro Takeuchi, Ivan Titov, Jacob Abernethy, Jakob Macke, Jasper Snoek, Jean-Philippe Vert, Jennifer Dy, Jennifer Listgarten, Jens Kober, Jianbo Shi, Jingyi Yu, Joan Bruna, John Duchi, John Wright, Jonathan Pillow, Jon Shlens, Julien Mairal, Jure Leskovec, Jyrki Kivinen, Karen Simonyan, Karsten Borgwardt, Karthik Sridharan, Kenji Fukumizu, Koby Crammer, Koji Tsuda, Koray Kavukcuoglu, Kun Zhang, Kyunghyun Cho, Laurens van der Maaten, Le Song, Lester Mackey, Li Deng, Lihong Li, Lorenzo Rosasco, Manuel Rodriguez, Marco Cuturi, Mario Marchand, Mark Schmidt, Matthew Blaschko, Matthias Hein, Matthias Seeger, Maxim Raginsky, Maya Gupta, Meg Mitchell, Mehreen Saeed, Michael Mahoney, Michele Sebag, Miguel Carreira-Perpinan, Mingyuan Zhou, Miro Dudik, Mohammad Emtiyaz Khan, Mohammad Ghavamzadeh, Moritz Hardt, Navdeep Jaitly, Nicolas LeRoux, Nicolas Usunier, Novi Quadrianto, Nuno Vasconcelos, Ofer Dekel, Oliver Stegle, Oriol Vinyals, Patrick Flaherty, Pavel Laskov, Peter Orbanz, Philipp Hennig, Prateek Jain, Purnamrita Sarkar, Raia Hadsell, Remi Munos, Ricardo Silva, Richard Wilkinson, Rita Cucchiara, Romer Rosales, Ronan Collobert, Ruslan Salakhutdinov, Ruth Urner, Ryan Adams, Sally Goldman, Sameer Singh, Samuel Kaski, Sanja Fidler, Sanjiv Kumar, Sanmi Koyejo, Sebastian Bubeck, Sebastian Nowozin, Seungjin Choi, Shakir Mohamed, Shie Mannor, Shinichi Nakajima, Simon Lacoste-Julien, Sinead Williamson, Stefanie Jegelka, Surya Ganguli, Suvrit Sra, Tamara Broderick, Tamir Hazan, Tara Sainath, Thomas Gaertner, Tieyan Liu, Tomoharu Iwata, Umar Syed, Venkatesh Saligrama, Viren Jain, Virginia de Sa, Vitaly Kuznetsov, Vlad Mnih, Wee Sun Lee, Xiaoli Fern, Xinhua Zhang, Yaron Singer, Y-Lan Boureau, Yuanqing Lin, Yves Grandvalet, Zaid Harchaoui, Zhi-Hua Zhou, Zico Kolter

MONDAY TUTORIALS

TIME & DESCRIPTION	LOCATION
8:00 am - 10:15 am - Tutorial Sessions	
A Primer on Optimal Transport Marco Cuturi, Justin M Solomon	Grand Ballroom
Deep Learning: Practice and Trends Nando de Freitas, Scott Reed, Oriol Vinyals	Hall A
Reinforcement Learning with People Emma Brunskill	Hall C
10:45 am -- 1:00 pm - Tutorial Sessions	
Fairness in Machine Learning Solon Barocas, Moritz Hardt	Grand Ballroom
Deep Probabilistic Modelling with Gaussian Processes Neil D Lawrence	Hall A
Statistical Relational Artificial Intelligence: Logic, Probability and Computation Luc De Raedt, David Poole, Kristian Kersting, Sriraam Natarajan	Hall C
1 pm - 2:30 pm - Lunch Break (On Your Own)	
2:30 pm - 4:45 pm - Tutorial Sessions	
Differentially Private Machine Learning: Theory, Algorithms and Applications Kamalika Chaudhuri, Anand D Sarwate	Grand Ballroom
Geometric Deep Learning on Graphs and Manifolds Michael Bronstein, Joan Bruna, Arthur Szlam, Xavier Bresson, Yann LeCun	Hall A
Engineering and Reverse-Engineering Intelligence Using Probabilistic Programs, Program Induction, and Deep Learning Josh Tenenbaum, Vikash K Mansinghka	Hall C
5:30 pm - 6:20 pm	
Invited Talk: Posner Lecture - Energy Strategies to Decrease CO2 Emissions John Platt	Hall A
6:30 pm - 10:30 pm Opening Reception & Posters	Pacific Ballroom

Coffee break - 10:15 am - 10:45 am

Tutorials: 8:00 - 10:15 AM

A Primer on Optimal Transport

Location: Grand Ballroom

Marco Cuturi (Univ. Paris-Saclay)
Justin M Solomon (MIT)

Optimal transport (OT) provides a powerful and flexible way to compare probability measures, discrete and continuous, which includes therefore point clouds, histograms, datasets, parametric and generative models. Originally proposed in the eighteenth century, this theory later led to Nobel Prizes for Koopmans and Kantorovich as well as Villani's Fields Medal in 2010. OT recently has reached the machine learning community, because it can tackle challenging learning scenarios including dimensionality reduction, structured prediction problems that involve histogram outputs, and estimation of generative models such as GANs in highly degenerate, high-dimensional problems. Despite very recent successes bringing OT from theory to practice, OT remains challenging for the machine learning community because of its mathematical formality. This tutorial will introduce in an approachable way crucial theoretical, computational, algorithmic and practical aspects of OT needed for machine learning Applications.

Deep Learning: Practice and Trends

Location: Hall A

Nando de Freitas (DeepMind)
Scott Reed (DeepMind)
Oriol Vinyals (DeepMind)

Deep Learning has become an essential toolbox which is used in a wide variety of Applications, research labs, industry, etc. In this tutorial, we will provide a set of guidelines which will help newcomers to the field understand the most recent and advanced models, their application to diverse data modalities (such as images, videos, waveforms, sequences, graphs,) and to complex tasks (such as learning to learn from a few examples, or generating molecules).

Reinforcement Learning with People

Location: Hall C

Emma Brunskill (Stanford)

There has been recent very exciting advances in (deep) Reinforcement learning, particularly in the areas of games and robotics. Yet perhaps the largest impact could come when Reinforcement learning systems interact with people. In this tutorial we will discuss work on Reinforcement learning for helping and assisting people, and frameworks and approaches for enabling people helping Reinforcement learning. We will cover Background on Reinforcement learning. Reinforcement learning for people-focused Applications Approaches for enabling people to assist Reinforcement learners. A number of the ideas presented here will also be relevant to many high stakes Reinforcement learning systems.

Tutorials: 10:45 - 1:00 PM

Fairness in Machine Learning

Location: Grand Ballroom

Solon Barocas (Cornell) Moritz Hardt (UC Berkeley)

Over the past few years, fairness has emerged as a matter of serious concern within machine learning. There is growing recognition that even models developed with the best of intentions may exhibit discriminatory biases, perpetuate inequality, or perform less well for historically disadvantaged groups. Considerable work is already underway within and outside machine learning to both characterize and address these problems. This tutorial will take a novel approach to parsing the topic, adopting three perspectives: statistics, causality, and measurement. Each viewpoint will shed light on different facets of the problem and help explain matters of continuing technical and normative debate. Rather than attempting to resolve questions of fairness within a single technical framework, the tutorial aims to equip the audience with a coherent toolkit to critically examine the many ways that machine learning implicates fairness.

Deep Probabilistic Modelling with Gaussian Processes

Location: Hall A

Neil D Lawrence (Amazon)

Neural network models are algorithmically simple, but mathematically complex. Gaussian process models are mathematically simple, but algorithmically complex. In this tutorial we will explore Deep Gaussian Process models. They bring advantages in their mathematical simplicity but are challenging in their algorithmic complexity. We will give an overview of Gaussian processes and highlight the algorithmic approximations that allow us to stack Gaussian process models: they are based on variational methods. In the last part of the tutorial will explore a use case exemplar: uncertainty quantification. We end with open questions.

Statistical Relational Artificial Intelligence: Logic, Probability and Computation

Location: Hall C

Luc De Raedt (KU Leuven) David Poole (U. of BC)
Kristian Kersting (Dortmund U.) Sriraam Natarajan (Indiana)

This tutorial will provide a gentle introduction into the foundations of statistical relational artificial intelligence, and will realize this by introducing the foundations of logic, of probability, of learning, and their respective combinations. Both predicate logic and probability theory extend propositional logic, one by adding relations, individuals and quantified variables, the other by allowing for measures over possible worlds and conditional queries. While logical and Probabilistic approaches have often been studied and used independently within artificial intelligence, they are not in conflict with each other but they are synergistic. This explains why there has been a considerable body of research in combining first-order logic and probability over the last 25 years, evolving into what has come to be called Statistical Relational Artificial Intelligence (StarAI). Relational Probabilistic models — we use this term in the broad sense, meaning any models that combine relations and probabilities — form the basis of StarAI, and can be seen as combinations of probability and predicate calculus that allow for individuals and relations as well as probabilities. In building on top of relational models, StarAI goes far beyond reasoning, optimization, learning and acting optimally in terms of a fixed number of features or variables, as it is typically studied in machine learning, constraint satisfaction, Probabilistic reasoning, and other areas of AI. Since StarAI draws upon ideas developed within many different fields, however, it can also be quite challenging for newcomers to get started and our tutorial precisely aims to provide this background.

Tutorials: 2:30 - 4:45 PM

Differentially Private Machine Learning: Theory, Algorithms and Applications

Location: Grand Ballroom

Kamalika Chaudhuri (UCSD)
Anand D Sarwate (Rutgers, U. of New Jersey)

Differential privacy has emerged as one of the de-facto standards for measuring privacy risk when performing computations on sensitive data and disseminating the results. Algorithms that guarantee differential privacy are randomized, which causes a loss in performance, or utility. Managing the privacy-utility tradeoff becomes easier with more data. Many machine learning algorithms can be made differentially private through the judicious introduction of randomization, usually through noise, within the computation. In this tutorial we will describe the basic framework of differential privacy, key mechanisms for guaranteeing privacy, and how to find differentially private approximations to several contemporary machine learning tools: convex optimization, Bayesian methods, and deep learning.

Geometric Deep Learning on Graphs and Manifolds

Location: Hall A

Michael Bronstein (USI Lugano, Tel Aviv U.& Intel)
Joan Bruna (NYU)
Arthur Szlam (Facebook)
Xavier Bresson (NTU)
Yann LeCun (Facebook AI Research, New York U.)

In the past years, deep learning methods have achieved unprecedented performance on a broad range of problems in various fields from computer vision to speech recognition. So far research has mainly focused on developing deep learning methods for Euclidean-structured data, while many important Applications have to deal with non-Euclidean structured data, such as graphs and manifolds. Such geometric data are becoming increasingly important in computer graphics and 3D vision, sensor networks, drug design, biomedicine, recommendation systems, and web Applications. The adoption of deep learning in these fields has been lagging behind until recently, primarily since the non-Euclidean nature of objects dealt with makes the very definition of basic operations used in deep networks rather elusive.

The purpose of the proposed tutorial is to introduce the emerging field of geometric deep learning on graphs and manifolds, overview existing solutions and Applications for this class of problems, as well as key difficulties and future research directions.

Engineering and Reverse-Engineering Intelligence Using Probabilistic Programs, Program Induction, and Deep Learning

Location: Hall C

Josh Tenenbaum (MIT)
Vikash K Mansinghka (MIT)

Recent successes in computer vision, natural language processing and other areas of artificial intelligence have been largely driven by methods for sophisticated pattern recognition — most prominently deep neural networks. But human intelligence is more than just pattern recognition. In particular, it depends on a suite of cognitive capacities for modeling the world: for making judgment calls in ambiguous situations, explaining and understanding what we see, imagining things we could see but haven't yet, solving problems and planning actions to make these things real, and building new models as we learn more about the world. We will talk about prospects for reverse-engineering these capacities at the heart of human intelligence, and using what we learn to make machines smarter in more human-like ways. We introduce basic concepts and techniques of Probabilistic programs, inference programming and program induction, which together with tools from deep learning and modern video game engines provide an approach to capturing many aspects of everyday intelligence.

Specific units in our tutorial will show how:

(1) Defining Probabilistic programs over algorithms and representations drawn from modern video game engines — graphics engines, physics engines, and planning engines — allows us to capture how people can perceive rich three-dimensional structure in visual scenes and objects, perceive and predict objects' motion based on their physical characteristics, and infer the mental states of other people from observing their actions.

(2) By formulating model learning as higher-order inference in these systems, we can construct ``program-learning programs". These programs can learn new concepts from just one or a few examples.

(3) It is possible to build Probabilistic programming systems in which scalable, general-purpose, efficient inference and model discovery algorithms can be easily and flexibly programmed by end users. These languages provide powerful tools for robotics, interactive data analysis, and scientific discovery.

Posner Lecture - 5:30 PM - Hall A

Energy Strategies to Decrease CO2 Emissions

The problem of climate change is very difficult to solve. On the one hand, fossil fuels are ubiquitous in human civilization: we get 16 trillion watts of power by burning fossil fuels. On the other hand, according to climate modeling, we have less than 30 years at current burn rates before we run out of carbon budget to keep the mean global temperature rise below 2 degrees C.

There are many different proposed strategies to combat climate change. This talk will attempt to clarify the confusion via economic modeling. First, I will give a tutorial about the energy system. Then, I will show a simple economic model which predicts the cost of and CO2 produced by electricity generation, given a number of assumptions. We will go through possible scenarios and see how we can reduce CO2 dramatically at least cost.

The biggest lesson from the economic model is that we need a "strong energy miracle": a zero-carbon 24x7 technology that can produce electricity cheaper than the isolated cost of burning the fossil fuel. Currently, there is no such technology. I'll talk about one technology that may become a strong energy miracle, and discuss progress towards making that a reality.

John Platt
(Google)

John Platt is best known for his work in machine learning: the SMO algorithm for support vector machines and calibrating the output of models. He was an early adopter of convolutional neural networks in the 1990s. However, John has worked in many different fields: data systems, computational geometry, object recognition, media UIs, analog computation, handwriting recognition, and applied math. He has discovered two asteroids, and won a Technical Academy Award in 2006 for his work in computer graphics.

John currently leads the Applied Science branch of Google Research, which works at the intersection between computer science and physical or biological science.

Monday Poster Session

#1 Learning Active Learning from Data
Ksenia Konyushkova, Raphael Sznitman, Pascal Fua

#2 Scalable Variational Inference for Dynamical Systems
Stefan Bauer, Nico S Gorbach, Joachim M Buhmann

#3 Active Learning from Peers
Keerthiram Murugesan, Jaime Carbonell

#4 Gradient Episodic Memory for Continuum Learning
David Lopez-Paz, Marc'Aurelio Ranzato

#5 Consistent Multitask Learning with Nonlinear Output Relations
Carlo Ciliberto, Alessandro Rudi, Lorenzo Rosasco, Massimiliano Pontil

#6 Joint distribution optimal transportation for domain adaptation
Nicolas Courty, Rémi Flamary, Amaury Habrard, Alain Rakotomamonjy

#7 Learning Multiple Tasks with Deep Relationship Networks
Mingsheng Long, Jianmin Wang, Philip S Yu

#8 Label Efficient Learning of Transferable Representations acrosss Domains and Tasks
Alan Luo, Yuliang Zou, Judy Hoffman, Li Fei-Fei

#9 Matching neural paths: transfer from recognition to correspondence search
Nikolay Savinov, Lubor Ladicky, Marc Pollefeys

#10 Do Deep Neural Networks Suffer from Crowding?
Anna Volokitin, Gemma Roig, Tomaso A Poggio

#11 SVCCA: Singular Vector Canonical Correlation Analysis for Deep Understanding and Improvement
Maithra Raghu, Justin Gilmer, Jason Yosinski, Jascha Sohl-Dickstein

#12 Neural Expectation Maximization
Klaus Greff, Sjoerd van Steenkiste, Jürgen Schmidhuber

#13 PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space
Charles Ruizhongtai Qi, Li Yi, Hao Su, Leonidas J Guibas

#14 Preserving Proximity and Global Ranking for Node Embedding
Yi-An Lai, Chin-Chi Hsu, Wen Hao Chen, Mi-Yen Yeh, Shou-De Lin

#15 Unsupervised Transformation Learning via Convex Relaxations
Tatsunori B Hashimoto, Percy Liang, John C Duchi

#16 Hunt For The Unique, Stable, Sparse And Fast Feature Learning On Graphs
Saurabh Verma, Zhi-Li Zhang

#17 Deep Subspace Clustering Network
Pan Ji, Tong Zhang, Hongdong Li, Mathieu Salzmann, Ian Reid

#18 Learning Graph Embeddings with Embedding Propagation
Alberto Garcia Duran, Mathias Niepert

#19 Unsupervised Sequence Classification using Sequential Output Statistics
Yu Liu, Jianshu Chen, Li Deng

#20 Context Selection for Embedding Models
Liping Liu, Francisco Ruiz, David Blei

#21 Probabilistic Rule Realization and Selection
Haizi Yu, Tianxi Li, Lav Varshney

#22 Trimmed Density Ratio Estimation
Song Liu, Akiko Takeda, Taiji Suzuki, Kenji Fukumizu

#23 A Minimax Optimal Algorithm for Crowdsourcing
Richard Combes, Thomas Bonald

#24 Introspective Classification with Convolutional Nets
Long Jin, Justin Lazarow, Zhuowen Tu

#25 Adaptive Classification for Prediction Under a Budget
Feng Nan, Venkatesh Saligrama

#26 Learning with Feature Evolvable Streams
Bo-Jian Hou, Lijun Zhang, Zhi-Hua Zhou

#27 Aggressive Sampling for Multi-class to Binary Reduction with Applications to Text Classification
Bikash Joshi, Massih-Reza Amini, Ioannis Partalas, Franck Iutzeler, Yury Maximov

#28 Adversarial Surrogate Losses for Ordinal Regression
Rizal Fathony, Mohammad Ali Bashiri, Brian Ziebart

#29 Formal Guarantees on the Robustness of a Classifier against Adversarial Manipulation
Matthias Hein, Maksym Andriushchenko

#30 Cost efficient gradient boosting
Sven Peter, Ferran Diego, Fred Hamprecht, Boaz Nadler

#31 A Highly Efficient Gradient Boosting Decision Tree
Guolin Ke, Qi Meng, Taifeng Wang, Wei Chen, Weidong Ma, Tie-Yan Liu

#32 Estimating Accuracy from Unlabeled Data: A Probabilistic Logic Approach
Emmanouil Platanios, Hoifung Poon, Tom M Mitchell, Eric J Horvitz

#33 Inferring Generative Model Structure with Static Analysis
Paroma Varma, Bryan He, Payal Bajaj, Nish Khandwala, Chris Ré

#34 Scalable Model Selection for Belief Networks
Zhao Song, Yusuke Muraoka, Ryohei Fujimaki, Lawrence Carin

#35 Time-dependent spatially varying graphical models, with application to brain fMRI data analysis
Kristjan Greenewald, Seyoung Park, Shuheng Zhou, Alexander Giessing

#36 A Bayesian Data Augmentation Approach for Learning Deep Models
Toan Tran, Trung Pham, Gustavo Carneiro, Lyle Palmer, Ian Reid

#37 Union of Intersections (UoI) for Interpretable Data Driven Discovery and Prediction
Kris Bouchard, Alejandro Bujan, Farbod Roosta-Khorasani, Shashanka Ubaru, Mr. Prabhat, Antoine Snijders, Jian-Hua Mao, Edward Chang, Michael W Mahoney, Sharmodeep Bhattacharya

#38 Deep Learning with Topological Signatures
Christoph Hofer, Roland Kwitt, Marc Niethammer, Andreas Uhl

#39 Practical Hash Functions for Similarity Estimation and Dimensionality Reduction
Søren Dahlgaard, Mathias Knudsen, Mikkel Thorup

#40 Maxing and Ranking with Few Assumptions
Venkatadheeraj Pichapati, Alon Orlitsky, Vaishakh Ravindrakumar, Moein Falahatgar, Yi Hao

#41 Kernel functions based on triplet comparisons
Matthäus Kleindessner, Ulrike von Luxburg

#42 Learning A Structured Optimal Bipartite Graph for Co-Clustering
Feiping Nie, Xiaoqian Wang, Heng Huang

#43 Multi-way Interacting Regression via Factorization Machines
Mikhail Yurochkin, Long Nguyen, nikolaos Vasiloglou

#44 Maximum Margin Interval Trees
Alexandre Drouin, Toby Hocking, Francois Laviolette

#45 Kernel Feature Selection via Conditional Covariance Minimization
Jianbo Chen, Mitchell Stern, Martin J Wainwright, Michael Jordan

#46 Improved Graph Laplacian via Geometric Self-Consistency
Dominique Joncas, Marina Meila, James McQueen

#47 Mixture-Rank Matrix Approximation for Collaborative Filtering
Dongsheng Li, Kehan Chen, Wei Liu, Tun Lu, Ning Gu, Stephen Chu

Monday Poster Session

#48 Predictive State Recurrent Neural Networks
Carlton Downey, Ahmed Hefny, Byron Boots, Geoffrey Gordon, Boyue Li

#49 Hierarchical Methods of Moments
Matteo Ruffini, Guillaume Rabusseau, Borja Balle

#50 Multitask Spectral Learning of Weighted Automata
Guillaume Rabusseau, Borja Balle, Joelle Pineau

#51 Generative Local Metric Learning for Kernel Regression
Yung-Kyun Noh, Masashi Sugiyama, Kee-Eung Kim, Frank Park, Daniel Lee

#52 Principles of Riemannian Geometry in Neural Networks
Michael Hauser, Asok Ray

#53 Subset Selection for Sequential Data
Ehsan Elhamifar

#54 On Quadratic Convergence of DC Proximal Newton Algorithm in Nonconvex Sparse Learning
Xingguo Li, Lin Yang, Jason Ge, Jarvis Haupt, Tong Zhang, Tuo Zhao

#55 Fast, Sample-Efficient Algorithms for Structured Phase Retrieval
Gauri Jagatap, Chinmay Hegde

#56 k-Support and Ordered Weighted Sparsity for Overlapping Groups: Hardness and Algorithms
Cong Han Lim, Stephen Wright

#57 Parametric Simplex Method for Sparse Learning
Haotian Pang, Tuo Zhao, Han Liu, Robert J Vanderbei

#58 Learned D-AMP: Principled Neural-network-based Compressive Image Recovery
Chris Metzler, Ali Mousavi, Richard Baraniuk

#59 FALKON: An Optimal Large Scale Kernel Method
Alessandro Rudi, Luigi Carratino, Lorenzo Rosasco

#60 Recursive Sampling for the Nystrom Method
Cameron Musco, Christopher Musco

#61 Efficient Approximation Algorithms for Strings Kernel Based Sequence Classification
Muhammad Farhan, Juvaria Tariq, Arif Zaman, Mudassir Shabbir, Imdad Khan

#62 Robust Hypothesis Test for Functional Effect with Gaussian Processes
Jeremiah Liu, Brent Coull

#63 Invariance and Stability of Deep Convolutional Representations
Alberto Bietti, Julien Mairal

#64 Testing and Learning on Distributions with Symmetric Noise Invariance
Law Ho Chung, Christopher Yau, Dino Sejdinovic

#65 An Empirical Study on The Properties of Random Bases for Kernel Methods
Maximilian Alber, Pieter-Jan Kindermans, Kristof Schütt, Klaus-Robert Müller, Fei Sha

#66 Max-Margin Invariant Features from Transformed Unlabelled Data
Dipan Pal, Ashwin Kannan, Gautam Arakalgud, Marios Savvides

#67 SafetyNets: Verifiable Execution of Deep Neural Networks on an Untrusted Cloud
Zahra Ghodsi, Tianyu Gu, Siddharth Garg

#68 Multi-output Polynomial Networks and Factorization Machines
Mathieu Blondel, Vlad Niculae, Takuma Otsuka, Naonori Ueda

#69 The Neural Hawkes Process: A Neurally Self-Modulating Multivariate Point Process
Hongyuan Mei, Jason Eisner

#70 Maximizing the Spread of Influence from Training Data
Eric Balkanski, Nicole Immorlica, Yaron Singer

#71 Inductive Representation Learning on Large Graphs
Will Hamilton, Rex Ying, Jure Leskovec

#72 A Meta-Learning Perspective on Cold-Start Recommendations for Items
Manasi Vartak, Hugo Larochelle, Arvind Thiagarajan

#73 DropoutNet: Addressing Cold Start in Recommender Systems
Maksims Volkovs, Guangwei Yu, Tomi Poutanen

#74 Federated Multi-Task Learning
Ginger Smith, Maziar Sanjabi, Chao-Kai Chiang, Ameet S Talwalkar

#75 Flexpoint: An Adaptive Numerical Format for Efficient Training of Deep Neural Networks
Urs Köster, Tristan Webb, Xin Wang, Marcel Nassar, Arjun K Bansal, William Constable, Oguz Elibol, Stewart Hall, Luke Hornof, Amir Khosrowshahi, Carey Kloss, Ruby J Pai, Naveen Rao

#76 Bayesian Inference of Individualized Treatment Effects using Multi-task Gaussian Processes
Ahmed M. Alaa, Mihaela van der Schaar

#77 Tomography of the London Underground: a Scalable Model for Origin-Destination Data
Nicolò Colombo, Ricardo Silva, Soong Moon Kang

#78 Matching on Balanced Nonlinear Representations for Treatment Effects Estimation
Sheng Li, Yun Fu

#79 MolecuLeNet: A continuous-filter convolutional neural network for modeling quantum interactions
Kristof Schütt, Pieter-Jan Kindermans, Huziel Enoc Sauceda Felix, Stefan Chmiela, Alexandre Tkatchenko, Klaus-Robert Müller

#80 Hiding Images in Plain Sight: Deep Steganography
Shumeet Baluja

#81 Universal Style Transfer via Feature Transforms
Yijun Li, Chen Fang, Jimei Yang, Zhaowen Wang, Xin Lu, Ming-Hsuan Yang

#82 Attend and Predict: Understanding Gene Regulation by Selective Attention on Chromatin
Ritambhara Singh, Jack Lanchantin, Yanjun Qi

#83 Unbounded cache model for online language modeling with open vocabulary
Edouard Grave, Moustapha Cisse, Armand Joulin

#84 Deconvolutional Paragraph Representation Learning
Yizhe Zhang, Dinghan Shen, Guoyin Wang, Zhe Gan, Ricardo Henao, Lawrence Carin

#85 Analyzing Hidden Representations in End-to-End Automatic Speech Recognition Systems
Yonatan Belinkov, Jim Glass

#86 Best of Both Worlds: Transferring Knowledge from Discriminative Learning to a Generative Visual Dialog Model
Jiasen Lu, Anitha Kannan, Jianwei Yang, Dhruv Batra, Devi Parikh

#87 Teaching Machines to Describe Images with Natural Language Feedback
huan ling, Sanja Fidler

#88 High-Order Attention Models for Visual Question Answering
Idan Schwartz, Alex Schwing, Tamir Hazan

#89 Visual Reference Resolution using Attention Memory for Visual Dialog
Paul Hongsuck Seo, Andreas Lehrmann, Bohyung Han, Leonid Sigal

#90 Semi-Supervised Learning for Optical Flow with Generative Adversarial Networks
Wei-Sheng Lai, Jia-Bin Huang, Ming-Hsuan Yang

#91 Associative Embedding: End-to-End Learning for Joint Detection and Grouping
Alejandro Newell, Zhiao Huang, Jia Deng

#92 Learning Deep Structured Multi-Scale Features using Attention-Gated CRFs for Contour Prediction
Dan Xu, Wanli Ouyang, Xavier Alameda-Pineda, Elisa Ricci, Xiaogang Wang, Nicu Sebe

#93 Incorporating Side Information by Adaptive Convolution
Di Kang, Debarun Dhar, Antoni Chan

#94 Learning a Multi-View Stereo Machine
Abhishek Kar, Jitendra Malik, Christian Häne

#95 Pose Guided Person Image Generation
Liqian Ma, Xu Jia, Qianru Sun, Bernt Schiele, Tinne Tuytelaars, Luc Van Gool

#96 Working hard to know your neighbor's margins: Local descriptor learning loss
Anastasiia Mishchuk, Dmytro Mishkin, Filip Radenovic, Jiri Matas

#97 Multimodal Image-to-Image Translation by Enforcing Bi-Cycle Consistency
Jun-Yan Zhu, Richard Zhang, Deepak Pathak, Prof. Darrell, Oliver Wang, Eli Shechtman, Alexei Efros

#98 Deep supervised discrete hashing
Qi Li, Zhenan Sun, Ran He, Tieniu Tan

#99 SVD-Softmax: Fast Softmax Approximation on Large Vocabulary Neural Networks
Kyuhong Shim, Minjae Lee, Iksoo Choi, Yoonho Boo, Wonyong Sung

#100 Hash Embeddings for Efficient Word Representations
Dan Tito Svenstrup, Jonas Hansen, Ole Winther

#101 A Regularized Framework for Sparse and Structured Neural Attention
Vlad Niculae, Mathieu Blondel

#102 Attentional Pooling for Action Recognition
Rohit Girdhar, Deva Ramanan

#103 Plan, Attend, Generate: Planning for Sequence-to-Sequence Models
Caglar Gulcehre, Francis Dutil, Adam Trischler, Yoshua Bengio

#104 Dilated Recurrent Neural Networks
Shiyu Chang, Yang Zhang, Wei Han, Mo Yu, Xiaoxiao Guo, Wei Tan, Xiaodong Cui, Michael Witbrock, Mark A Hasegawa-Johnson, Thomas Huang

#105 Thalamus Gated Recurrent Modules
Danijar Hafner, Alexander Irpan, James Davidson, Nicolas Heess

#106 Wasserstein Learning of Deep Generative Point Process Models
Benjamin XIAO, Mehrdad Farajtabar, Xiaojing Ye, Junchi Yan, Le Song, Hongyuan Zha

#107 Stabilizing Training of Generative Adversarial Networks through Regularization
Kevin Roth, Aurelien Lucchi, Sebastian Nowozin, Thomas Hofmann

#108 Neural Variational Inference and Learning in Undirected Graphical Models
Volodymyr Kuleshov, Stefano Ermon

#109 Adversarial Symmetric Variational Autoencoder
Yuchen Pu, Weiyao Wang, Ricardo Henao, Liqun Chen, Zhe Gan, Chunyuan Li, Lawrence Carin

#110 Diverse and Accurate Image Description Using a Variational Auto-Encoder with an Additive Gaussian Encoding Space
Liwei Wang, Alex Schwing, Svetlana Lazebnik

#111 Z-Forcing: Training Stochastic Recurrent Networks
Anirudh Goyal ALIAS PARTH GOYAL, Alessandro Sordoni, Marc-Alexandre Côté, Rosemary Ke, Yoshua Bengio

#112 One-Shot Imitation Learning
Yan Duan, Marcin Andrychowicz, Bradly Stadie, OpenAI Jonathan Ho, Jonas Schneider, Ilya Sutskever, Pieter Abbeel, Wojciech Zaremba

#113 Reconstruct & Crush Network
Erinc Merdivan, Mohammad Reza Loghmani, Matthieu Geist

#114 Fader Networks: Generating Image Variations by Sliding Attribute Values
Guillaume Lample, Neil Zeghidour, Nicolas Usunier, Antoine Bordes, Ludovic DENOYER, Marc'Aurelio Ranzato

#115 PredRNN: Recurrent Neural Networks for Video Prediction using Spatiotemporal LSTMs
Yunbo Wang, Mingsheng Long, Jianmin Wang, Philip S Yu

#116 Multi-agent Predictive Modeling with Attentional CommNets
Yedid Hoshen

#117 Real Time Image Saliency for Black Box Classifiers
Piotr Dabkowski, Yarin Gal

#118 Prototypical Networks for Few-shot Learning
Jake Snell, Kevin Swersky, Richard Zemel

#119 Few-Shot Learning Through an Information Retrieval Lens
Eleni Triantafillou, Richard Zemel, Raquel Urtasun

#120 The Reversible Residual Network: Backpropagation Without Storing Activations
Aidan N Gomez, Mengye Ren, Raquel Urtasun, Roger Grosse

#121 Gated Recurrent Convolution Neural Network for OCR
Jianfeng Wang, Xiaolin Hu

#122 Learning Efficient Object Detection Models with Knowledge Distillation
Guobin Chen, Wongun Choi, Xiang Yu, Tony Han, Manmohan Chandraker

#123 Active Bias: Training a More Accurate Neural Network by Emphasizing High Variance Samples
Haw-Shiuan Chang, Andrew McCallum, Erik Learned-Miller

#124 Decoupling "when to update" from "how to update"
Eran Malach, Shai Shalev-Shwartz

#125 Langevin Dynamics with Continuous Tempering for Training Deep Neural Networks
Lincoln Ye, Zhanxing Zhu, Rafal Mantiuk

#126 Differentiable Learning of Logical Rules for Knowledge Base Reasoning
Fan Yang, Zhilin Yang, William W Cohen

#127 Deliberation Networks: Sequence Generation Beyond One-Pass Decoding
Yingce Xia, Lijun Wu, Jianxin Lin, Fei Tian, Tao Qin, Tie-Yan Liu

#128 Neural Program Meta-Induction
Jacob Devlin, Rudy R Bunel, Rishabh Singh, Matthew Hausknecht, Pushmeet Kohli

#129 Saliency-based Sequential Image Attention with Multiset Prediction
Sean Welleck, Kyunghyun Cho, Zheng Zhang

#130 Protein Interface Prediction using Graph Convolutional Networks
Alex Fout, Basir Shariat, Jonathon Byrd, Asa Ben-Hur

#131 Dual-Agent GANs for Photorealistic and Identity Preserving Profile Face Synthesis
Jian Zhao, Lin Xiong, Panasonic Karlekar Jayashree, Jianshu Li, Fang Zhao, Zhecan Wang, Panasonic Sugiri Pranata, Panasonic Shengmei Shen, Jiashi Feng

#132 Toward Robustness against Label Noise in Training Deep Discriminative Neural Networks
Arash Vahdat

#133 Soft-to-Hard Vector Quantization for End-to-End Learning Compressible Representations
Eirikur Agustsson, Fabian Mentzer, Michael Tschannen, Lukas Cavigelli, Radu Timofte, Luca Benini, Luc V Gool

#134 Selective Classification for Deep Neural Networks
Yonatan Geifman, Ran El-Yaniv

#135 Deep Lattice Networks and Partial Monotonic Functions
Seungil You, David Ding, Kevin Canini, Jan Pfeifer, Maya Gupta

#136 Learning to Prune Deep Neural Networks via Layer-wise Optimal Brain Surgeon
Xin Dong, Shangyu Chen, Sinno Pan

#137 Bayesian Compression for Deep Learning
Christos Louizos, Karen Ullrich, Max Welling

#138 Lower bounds on the robustness to adversarial perturbations
Jonathan Peck, Yvan Saeys, Bart Goossens, Joris Roels

#139 Sobolev Training for Neural Networks
Wojciech M. Czarnecki, Simon Osindero, Max Jaderberg, Grzegorz Swirszcz, Razvan Pascanu

#140 Structured Bayesian Pruning via Log-Normal Multiplicative Noise
Kirill Neklyudov, Dmitry Molchanov, Arsenii Ashukha, Dmitry Vetrov

#141 Population Matching Discrepancy and Applications in Deep Learning
Jianfei Chen, Chongxuan LI, Yizhong Ru, Jun Zhu

#142 Investigating the learning dynamics of deep neural networks using random matrix theory
Jeffrey Pennington, Sam Schoenholz, Surya Ganguli

#143 Robust Imitation of Diverse Behaviors
Ziyu Wang, Josh Merel, Scott Reed, Nando de Freitas, Greg Wayne, Nicolas Heess

#144 Question Asking as Program Generation
Anselm Rothe, Brenden Lake, Todd Gureckis

#145 Variational Laws of Visual Attention for Dynamic Scenes
Dario Zanca, Marco Gori

#146 Flexible statistical inference for mechanistic models of neural dynamics
Jan-Matthis Lueckmann, Pedro J Goncalves, Giacomo Bassetto, Kaan Oecal, Marcel Nonnenmacher, Jakob H Macke

#147 Training recurrent networks to generate hypotheses about how the brain solves hard navigation problems
Ingmar Kanitscheider, Ila Fiete

#148 YASS: Yet Another Spike Sorter
Jin Hyung Lee, David E Carlson, Hooshmand Shokri Razaghi, Weichi Yao, Georges A Goetz, chichilnisky Chichilnisky, Espen Hagen, Gaute T. Einevoll, Liam Paninski

#149 Neural system identification for large populations separating "what" and "where"
David Klindt, Alexander Ecker, Thomas Euler, Matthias Bethge

#150 A simple model of recognition and recall memory
Nisheeth Srivastava, Edward Vul

#151 Gaussian process based nonlinear latent structure discovery in multivariate spike train data
Anqi Wu, Nicholas Roy, Stephen Keeley, Jonathan W Pillow

#152 Deep adversarial neural decoding
Yağmur Güçlütürk, Umut Güçlü, Katja Seeliger, Sander Bosch, Rob van Lier, Marcel A. J. van Gerven

#153 Cross-Spectral Factor Analysis
Neil Gallagher, Kyle Ulrich, Austin Talbot, Kafui Dzirasa, David E Carlson, Lawrence Carin

#154 Cognitive Impairment Prediction in Alzheimer's Disease with Regularized Modal Regression
Xiaoqian Wang, Hong Chen, Dinggang Shen, Heng Huang

#155 Stochastic Submodular Maximization: The Case of Coverage Functions
Mohammad Karimi, Mario Lucic, Hamed Hassani, Andreas Krause

#156 Gradient Methods for Submodular Maximization
Hamed Hassani, Mahdi Soltanolkotabi, Amin Karbasi

#157 Non-convex Finite-Sum Optimization Via SCSG Methods
Lihua Lei, Cheng Ju, Jianbo Chen, Michael Jordan

#158 Influence Maximization with ε-Almost Submodular Threshold Function
Qiang Li, Wei Chen, Institute of Computing Xiaoming Sun, Institute of Computing Jialin Zhang

#159 Subset Selection under Noise
Chao Qian, Jing-Cheng Shi, Yang Yu, Ke Tang, Zhi-Hua Zhou

#160 Polynomial time algorithms for dual volume sampling
Chengtao Li, Stefanie Jegelka, Suvrit Sra

#161 Lookahead Bayesian Optimization with Inequality Constraints
Remi Lam, Karen Willcox

#162 Non-monotone Continuous DR-submodular Maximization: Structure and Algorithms
An Bian, Joachim M Buhmann, Andreas Krause, Kfir Levy

#163 Solving (Almost) all Systems of Random Quadratic Equations
Gang Wang, Georgios Giannakis, Yousef Saad, Jie Chen

#164 Learning ReLUs via Gradient Descent
Mahdi Soltanolkotabi

#165 Stochastic Mirror Descent for Non-Convex Optimization
Zhengyuan Zhou, Panayotis Mertikopoulos, Nicholas Bambos, Stephen Boyd, Peter W Glynn

#166 Accelerated First-order Methods for Geodesically Convex Optimization on Riemannian Manifolds
Yuanyuan Liu, Fanhua Shang, James Cheng, Hong Cheng, Licheng Jiao

#167 On the Fine-Grained Complexity of Empirical Risk Minimization: Kernel Methods and Neural Networks
Arturs Backurs, Piotr Indyk, Ludwig Schmidt

#168 Large-Scale Quadratically Constrained Quadratic Program via Low-Discrepancy Sequences
Kinjal Basu, Ankan Saha, Shaunak Chatterjee

#169 A New Alternating Direction Method for Linear Programming
Sinong Wang, Ness Shroff

#170 Dykstra's Algorithm, ADMM, and Coordinate Descent: Connections, Insights, and Extensions
Ryan Tibshirani

#171 Smooth Primal-Dual Coordinate Descent Algorithms for Nonsmooth Convex Optimization
Ahmet Alacaoglu, Quoc Tran Dinh, Olivier Fercoq, Volkan Cevher

#172 First-Order Adaptive Sample Size Methods to Reduce Complexity of Empirical Risk Minimization
Aryan Mokhtari, Alejandro Ribeiro

#173 Accelerated consensus via Min-Sum Splitting
Patrick Rebeschini, Sekhar C Tatikonda

#174 Integration Methods and Optimization Algorithms
Damien Scieur, Vincent Roulet, Francis Bach, Alexandre d'Aspremont

#175 Efficient Use of Limited-Memory Resources to Accelerate Linear Learning
Celestine Dünner, Thomas Parnell, Martin Jaggi

#176 A Screening Rule for l1-Regularized Ising Model Estimation
Charles Kuang, Sinong Geng, David Page

#177 Uprooting and Rerooting Higher-order Graphical Models
Adrian Weller, Mark Rowland

#178 Concentration of Multilinear Functions of the Ising Model with Applications to Network Data
Constantinos Daskalakis, Nishanth Dikkala, Gautam Kamath

#179 Inference in Graphical Models via Semidefinite Programming Hierarchies
Murat A. Erdogdu, Yash Deshpande, Andrea Montanari

#180 Beyond normality: Learning sparse Probabilistic graphical models in the non-Gaussian setting
Rebecca Morrison, Ricardo Baptista, Youssef Marzouk

#181 Dynamic Importance Sampling for Anytime Bounds of the Partition Function
Qi Lou, Rina Dechter, Alexander Ihler

#182 Nonbacktracking Bounds on the Influence in Independent Cascade Models
Emmanuel Abbe, Sanjeev Kulkarni, Eun Jee Lee

#183 Rigorous Dynamics and Consistent Estimation in Arbitrarily Conditioned Linear Systems
Allie Fletcher, Sundeep Rangan, Moji Sahraee-Ardakan, Phil Schniter

#184 Learning Disentangled Representations with Semi-Supervised Deep Generative Models
Siddharth Narayanaswamy, T. Brooks Paige, Jan-Willem van de Meent, Alban Desmaison, Frank Wood, Noah Goodman, Pushmeet Kohli, Philip Torr

#185 Gauging Variational Inference
Sung-Soo Ahn, Michael Chertkov, Jinwoo Shin

#186 Variational Inference via χ Upper Bound Minimization
Adji Dieng, Dustin Tran, Rajesh Ranganath, John Paisley, David Blei

#187 Collapsed variational Bayes for Markov jump processes
Boqian Zhang, Jiangwei Pan, Vinayak A Rao

#188 Bayesian Dyadic Trees and Histograms for Regression
Stéphanie van der Pas, Veronika Rockova

#189 Differentially private Bayesian learning on distributed data
Mikko Heikkilä, Eemil Lagerspetz, Samuel Kaski, Kana Shimizu, Sasu Tarkoma, Antti Honkela

#190 Model-Powered Conditional Independence Test
Rajat Sen, Ananda Theertha Suresh, Karthikeyan Shanmugam, Alex Dimakis, Sanjay Shakkottai

#191 When Worlds Collide: Integrating Different Counterfactual Assumptions in Fairness
Chris Russell, Ricardo Silva, Matt Kusner, Joshua Loftus

#192 Q-LDA: Uncovering Latent Patterns in Text-based Sequential Decision Processes
Jianshu Chen, Chong Wang, Lin Xiao, Ji He, Lihong Li, Li Deng

#193 Probabilistic Models for Integration Error in the Assessment of Functional Cardiac Models
Chris Oates, Steven Niederer, Angela Lee, François-Xavier Briol, Mark Girolami

#194 Expectation Propagation for t-Exponential Family Using Q-Algebra
Futoshi Futami, Issei Sato, Masashi Sugiyama

#195 A Probabilistic Framework for Nonlinearities in Stochastic Neural Networks
Qinliang Su, xuejun Liao, Lawrence Carin

#196 Clone MCMC: Parallel High-Dimensional Gaussian Gibbs Sampling
Andrei-Cristian Barbos, Francois Caron, Jean-François Giovannelli, Arnaud Doucet

#197 Learning spatiotemporal piecewise-geodesic trajectories from longitudinal manifold-valued data
Stéphanie ALLASSONNIERE, Juliette Chevallier

#198 Scalable Levy Process Priors for Spectral Kernel Learning
Phillip A Jang, Andrew Loeb, Matthew Davidow, Andrew Wilson

#199 Inferring The Latent Structure of Human Decision-Making from Raw Visual Inputs
Yunzhu Li, Jiaming Song, Stefano Ermon

#200 Hybrid Reward Architecture for Reinforcement Learning
Harm Van Seijen, Laroche Laroche, Mehdi Fatemi, Joshua Romoff

#201 Shallow Updates for Deep Reinforcement Learning
Nir Levine, Tom Zahavy, Daniel J Mankowitz, Aviv Tamar, Shie Mannor

#202 Towards Generalization and Simplicity in Continuous Control
Aravind Rajeswaran, Kendall Lowrey, Emanuel Todorov, Sham Kakade

#203 Interpolated Policy Gradient: Merging On-Policy and Off-Policy Gradient Estimation for Deep Reinforcement Learning
Shixiang Gu, Timothy Lillicrap, Richard E Turner, Zoubin Ghahramani, Bernhard Schölkopf, Sergey Levine

#204 Scalable Planning with Tensorflow for Hybrid Nonlinear Domains
Ga Wu, Buser Say, Scott Sanner

#205 Task-based End-to-end Model Learning in Stochastic Optimization
Priya Donti, J. Zico Kolter, Brandon Amos

#206 Value Prediction Network
Junhyuk Oh, Satinder Singh, Honglak Lee

#207 Variable Importance Using Decision Trees
Arash Amini, Jalil Kazemitabar, Adam Bloniarz, Ameet S Talwalkar

#208 The Expressive Power of Neural Networks: A View from the Width
Zhou Lu, Hongming Pu, Feicheng Wang, Zhiqiang Hu, Liwei Wang

#209 SGD Learns the Conjugate Kernel Class of the Network
Amit Daniely

#210 Radon Machines: Effective Parallelisation for Machine Learning
Michael Kamp, Mario Boley, Olana Missura, Thomas Gärtner

#211 Noise-Tolerant Interactive Learning Using Pairwise Comparisons
Yichong Xu, Hongyang Zhang, Aarti Singh, Artur Dubrawski, Kyle Miller

#212 A PAC-Bayesian Analysis of Randomized Learning with Application to Stochastic Gradient Descent
Ben London

#213 Revisiting Perceptron: Efficient and Label-Optimal Learning of Halfspaces
Songbai Yan, Chicheng Zhang

#214 Sample and Computationally Efficient Learning Algorithms under S-Concave Distributions
Maria-Florina Balcan, Hongyang Zhang

#215 Nearest-Neighbor Sample Compression: Efficiency, Consistency, Infinite Dimensions
Aryeh Kontorovich, Sivan Sabato, Roi Weiss

#216 Learning Identifiable Gaussian Bayesian Networks in Polynomial Time and Sample Complexity
Asish Ghoshal, Jean Honorio

#217 From which world is your graph
Cheng Li, Varun Kanade, Felix MF Wong, Zhenming Liu

#218 Mean Field Residual Networks: On the Edge of Chaos
Ge Yang

#219 Learning from uncertain curves: The 2-Wasserstein metric for Gaussian processes
Anton Mallasto, Aasa Feragen

#220 On clustering network-valued data
Soumendu Sundar Mukherjee, Purnamrita Sarkar, Lizhen Lin

#221 On the Power of Truncated SVD for General High-rank Matrix Estimation Problems
Simon Du, Yining Wang, Aarti Singh

#222 AdaGAN: Boosting Generative Models
Ilya Tolstikhin, Sylvain Gelly, Olivier Bousquet, Carl-Johann SIMON-GABRIEL, Bernhard Schölkopf

#223 Discovering Potential Influence via Information Bottleneck
Weihao Gao, Sreeram Kannan, Hyeji Kim, Sewoong Oh, Pramod Viswanath

#224 Phase Transitions in the Pooled Data Problem
Jonathan Scarlett, Volkan Cevher

#225 Coded Distributed Computing for Inverse Problems
Yaoqing Yang, Pulkit Grover, Soummya Kar

#226 Query Complexity of Clustering with Side Information
Arya Mazumdar, Barna Saha

#227 Revisit Fuzzy Neural Network: Demystifying Batch Normalization and ReLU with Generalized Hamming Network
Lixin Fan

TUESDAY SESSIONS

7:30 - 9:00 AM	Coffee	
9:00 - 9:50 AM	**Invited Talk: Brendan Frey** *Why AI Will Make it Possible to Reprogram the Human Genome*	**Hall A**
9:50 - 10:10 AM	T**est Of Time Award: Ali Rahimi, Benjamin Recht** *Random Features for Large-Scale Kernel Machines*	
10:10 - 10:40 AM	Coffee break	
10:40 - 12:00 PM	**Parallel Tracks:** **Algorithms** **Optimization**	 **Hall A** **Hall C**
12:00 - 1:50 PM	Lunch on your own	
1:50 - 2:40 PM	**Invited Talk: Kate Crawford** *The Trouble with Bias*	**Hall A**
2:40 - 2:50 PM	Quick break	
2:50 - 3:50 PM	**Parallel Tracks:** **Algorithms, Optimization** **Theory**	 **Hall A** **Hall C**
3:50 - 4:20 PM	Coffee break	
4:20 - 6:00 PM	**Parallel Tracks:** **Deep Learning, Applications** **Algorithms**	 **Hall A** **Hall C**
6:00 - 7:00 PM	Light snack	
7:00 - 10:30 PM	Poster session and Demos	**Pacific Ballroom**

Tuesday Invited Talks

Why AI Will Make it Possible to Reprogram the Human Genome

Hall A, 9:00 - 9:50 AM

We have figured out how to write to the genome using DNA editing, but we don't know what the outcomes of genetic modifications will be. This is called the "genotype-phenotype gap". To close the gap, we need to reverse-engineer the genetic code, which is very hard because biology is too complicated and noisy for human interpretation. Machine learning and AI are needed. The data? Six billion letters per genome, hundreds of thousands of types of biomolecules, hundreds of cell types, over seven billion people on the planet. A new generation of "Bio-AI" researchers are poised to crack the problem, but we face extraordinary challenges. I'll discuss these challenges, focusing on which branches of AI and machine learning will have the most impact and why.

Brendan J Frey
(Vector Institute for AI & U. of Toronto)

Brendan Frey is Co-Founder and CEO of Deep Genomics, a Co-Founder of the Vector Institute for Artificial Intelligence, and a Professor of Engineering and Medicine at the University of Toronto. He is internationally recognized as a leader in machine learning and in genome biology and his group has published over a dozen papers on these topics in Science, Nature and Cell. His work on using deep learning to identify protein-DNA interactions was recently highlighted on the front cover Nature Biotechnology (2015), while his work on deep learning dates back to an early paper on what are now called variational autoencoders (Science 1995). He is a Fellow of the Royal Society of Canada, a Fellow of the Institute for Electrical and Electronic Engineers, and a Fellow of the American Association for the Advancement of Science. He has consulted for several industrial research and development laboratories in Canada, the United States and England, and has served on the Technical Advisory Board of Microsoft Research

The Trouble With Bias

Hall A, 1:50 - 2:40 PM

Computer scientists are increasingly concerned about the many ways that machine learning can reproduce and reinforce forms of bias. When ML systems are incorporated into core social institutions, like healthcare, criminal justice and education, issues of bias and discrimination can be extremely serious. But what can be done about it? Part of the trouble with bias in machine learning in high-stakes decision making is that it can be the result of one or many factors: the training data, the model, the system goals, and whether the system works less well for some populations, among several others. Given the difficulty of understanding how a machine learning system produced a particular result, bias is often discovered after a system has been producing unfair results in the wild. But there is another problem as well: the definition of bias changes significantly depending on your discipline, and there are exciting approaches from other fields that have not yet been included by computer science. This talk will look at the recent literature on bias in machine learning, consider how we can incorporate approaches from the social sciences, and offer new strategies to address bias.

Kate Crawford
(Microsoft Research)

Prof. Kate Crawford is leading researcher, academic and author who has spent 10 years studying the social impacts of large-scale data, machine learning and artificial intelligence. She is a Distinguished Research Professor at New York University, a Principal Researcher at Microsoft Research New York, and a Visiting Professor at the MIT Media Lab. In 2016, she co-chaired the White House symposium on the social and economic implications of artificial intelligence in the next decade. She is on the World Economic Forum Global Agenda Council on AI and Robotics, and she was recently appointed as a Richard von Weizsaecker Fellow. She has outlined in Nature a social systems approach to studying the impacts of machine learning systems, and she has also written about AI, ethics and discrimination for The New York Times and Harper's Magazine. She is the co-founder and co-director of the AI Now Institute at NYU, along with Meredith Whittaker, an interdisciplinary research institute that work across computer science, social science, law, and the humanities to better understand and address the social implications of artificial intelligence.

Student Paper Awards

A Linear-Time Kernel Goodness-of-Fit Test
Wittawat Jitkrittum, Wenkai Xu, Zoltan Szabo, Kenji Fukumizu, Arthur Gretton

Variance-based Regularization with Convex Objectives
Hongseok Namkoong, John Duchi

Safe and Nested Subgame Solving for Imperfect-Information Games
Noam Brown, Tuomas Sandholm

Track 1 - 10:40 am - 12:20 pm Algorithms

Location: Hall A

Diffusion Approximations for Online Principal Component Estimation and Global Convergence

Chris Junchi Li, Mengdi Wang, Tong Zhang

In this paper, we propose to adopt the diffusion approximation tools to study the dynamics of Oja's iteration which is an online stochastic gradient method for the principal component analysis. Oja's iteration maintains a running estimate of the true principal component from streaming data and enjoys less temporal and spatial complexities. We show that the Oja's iteration for the top eigenvector generates a continuous-state discrete-time Markov chain over the unit sphere. We characterize the Oja's iteration in three phases using diffusion approximation and weak convergence tools. Our three-phase analysis further provides a finite-sample error bound for the running estimate, which matches the minimax information lower bound for PCA under bounded noise.

Positive-Unlabeled Learning with Non-Negative Risk Estimator

Ryuichi Kiryo, Gang Niu, Marthinus C du Plessis,
Masashi Sugiyama

From only positive ~(P) and unlabeled ~(U) data, a binary classifier can be trained with PU learning, in which the state of the art is unbiased PU learning. However, if its model is very flexible, its empirical risk on training data will go negative and we will suffer from serious overfitting. In this paper, we propose a non-negative risk estimator for PU learning. When being minimized, it is more robust against overfitting and thus we are able to train very flexible models given limited P data. Moreover, we analyze the bias, consistency and mean-squared-error reduction of the proposed risk estimator and the estimation error of the corresponding risk minimizer. Experiments show that the proposed risk estimator successfully fixes the overfitting problem of its unbiased counterparts.

An Applied Algorithmic Foundation for Hierarchical Clustering

Joshua Wang (Stanford)
Benjamin Moseley (Washington U. in St Louis)

Hierarchical clustering is a data analysis method that has been used for decades. Despite its widespread use, there is a lack of an analytical foundation for the method. Having such a foundation would both support the methods currently used and guide future improvements. This paper gives an applied algorithmic foundation for hierarchical clustering. The goal of this paper is to give an analytic framework supporting observations seen in practice. This paper considers the dual of a problem framework for hierarchical clustering introduced by Dasgupta. The main results are that one of the most popular algorithms used in practice, average-linkage agglomerative clustering, has a small constant approximation ratio. Further, this paper establishes that using recursive *k*-means divisive clustering has a very poor lower bound on its approximation ratio, perhaps explaining why is it not as popular in practice. Motivated by the poor performance of *k*-means, we seek to find divisive algorithms that do perform well theoretically and this paper gives two constant approximation algorithms. This paper represents some of the first work giving a foundation for hierarchical clustering algorithms used in practice

SPOTLIGHTS

- **Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results**
  Antti Tarvainen, Harri Valpola
- **Communication-Efficient Stochastic Gradient Descent, with Applications to Neural Networks**
  Dan Alistarh, Demjan Grubic, Jerry Li, Ryota Tomioka, Milan Vojnovic
- **Inhomogoenous Hypergraph Clustering with Applications**
  Pan Li, Olgica Milenkovic
- **K-Medoids For K-Means Seeding**
  James Newling, François Fleuret
- **Online Learning with Transductive Regret**
  Scott Yang, Mehryar Mohri
- **Matrix Norm Estimation from a Few Entries**
  Sewoong Oh, Ashish Khetan
- **Semisupervised Clustering, AND-Queries and Locally Encodable Source Coding**
  Arya Mazumdar, Soumyabrata Pal

Track 2 - 10:40 am - 12:20 pm Optimization

Location: Hall C

On the Optimization Landscape of Tensor Decompositions

Rong Ge (Duke University)
Tengyu Ma (Facebook AI Research)

Non-convex optimization with local search heuristics has been widely used in machine learning, achieving many state-of-art results. It becomes increasingly important to understand why they can work for these NP-hard problems on typical data. The landscape of many objective functions in learning has been conjectured to have the geometric property that ``all local optima are (approximately) global optima", and thus they can be solved efficiently by local search algorithms. However, establishing such property can be very difficult. In this paper, we analyze the optimization landscape of the random over-complete tensor decomposition problem, which has many Applications in unsupervised leaning, especially in learning latent variable models. In practice, it can be efficiently solved by gradient ascent on a non-convex objective. We show that for any small constant $\epsilon > 0$, among the set of points with function values $(1+\epsilon)$-factor larger than the expectation of the function, all the local maxima are approximate global maxima. Previously, the best-known result only characterizes the geometry in small neighborhoods around the true components. Our result implies that even with an initialization that is barely better than the random guess, the gradient ascent algorithm is guaranteed to solve this problem. Our main technique uses Kac-Rice formula and random matrix theory. To our best knowledge, this is the first time when Kac-Rice formula is successfully applied to counting the number of local minima of a highly-structured random polynomial with dependent coefficients.

Robust Optimization for Non-Convex Objectives

Yaron Singer (Harvard University)
Robert S Chen (Harvard University)
Vasilis Syrgkanis (Microsoft Research)
Brendan Lucier (Microsoft Research)

We consider robust optimization problems, where the goal is to optimize in the worst case over a class of objective functions. We develop a reduction from robust improper optimization to Bayesian optimization: given an oracle that returns $\propto$-approximate solutions for distributions over objectives, we compute a distribution over solutions that is $\propto$-approximate in the worst case. We show that derandomizing this solution is NP-hard in general, but can be done for a broad class of statistical learning tasks. We apply our results to robust neural network training and submodular optimization. We evaluate our approach experimentally on a character classification task subject to adversarial distortion, and robust influence maximization on large networks.

Bayesian Optimization with Gradients

Jian Wu (AQR Capital Management)
Matthias Poloczek (Cornell University)
Andrew Wilson (Cornell University)
Peter Frazier (Cornell University)

Bayesian optimization has shown success in global optimization of expensive-to-evaluate multimodal objective functions. However, unlike most optimization methods, Bayesian optimization typically does not use derivative information. In this paper we show how Bayesian optimization can exploit derivative information to find good solutions with fewer objective function evaluations. In particular, we develop a novel Bayesian optimization algorithm, the derivative-enabled knowledge-gradient (dKG), which is one-step Bayes-optimal, asymptotically consistent, and provides greater one-step value of information than in the derivative-free setting. dKG accommodates noisy and incomplete derivative information, comes in both sequential and batch forms, and can optionally reduce the computational cost of inference through automatically selected retention of a single directional derivative. We also compute the dKG acquisition function and its gradient using a novel fast discretization-free technique. We show dKG provides state-of-the-art performance compared to a wide range of optimization procedures with and without gradients, on benchmarks including logistic regression, deep learning, kernel learning, and k-nearest neighbors.

SPOTLIGHTS

- **Gradient Descent Can Take Exponential Time to Escape Saddle Points**
  Simon Du, Chi Jin, Jason D Lee, Michael Jordan, Aarti Singh, Barnabas Poczos

- **Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration**
  Jason Altschuler, Jon Weed, Philippe Rigollet

- **Limitations on Variance-Reduction and Acceleration Schemes for Finite Sums Optimization**
  Yossi Arjevani

- **Implicit Regularization in Matrix Factorization**
  Suriya Gunasekar, Blake Woodworth, Srinadh Bhojanapalli, Behnam Neyshabur, Nati Srebro

- **Linear Convergence of a Frank-Wolfe Type Algorithm over Trace-Norm Balls**
  Zeyuan Allen-Zhu, Elad Hazan, Wei Hu, Yuanzhi Li

- **Acceleration and Averaging in Stochastic Descent Dynamics**
  Walid Krichene

- **When Cyclic Coordinate Descent Beats Randomized Coordinate Descent**
  Mert Gurbuzbalaban, Denizcan Vanli, Asuman Ozdaglar

Track 1 - 2:50 - 3:50 pm
Algorithms, Optimization

Location: Hall A

Streaming Weak Submodularity: Interpreting Neural Networks on the Fly

Ethan Elenberg, Alex Dimakis, Moran Feldman, Amin Karbasi

In many machine learning Applications, it is important to explain the predictions of a black-box classifier. For example, why does a deep neural network assign an image to a particular class? We cast interpretability of black-box classifiers as a combinatorial maximization problem and propose an efficient streaming algorithm to solve it subject to cardinality constraints. By extending ideas from Badanidiyuru et al. [2014], we provide a constant factor approximation guarantee for our algorithm in the case of random stream order and a weakly submodular objective function. This is the first such theoretical guarantee for this general class of functions, and we also show that no such algorithm exists for a worst case stream order. Our algorithm obtains similar explanations of Inception V3 predictions 10 times faster than the state-of-the-art LIME framework of Ribeiro et al. [2016]

A Unified Approach To Interpreting Model Predictions

Scott M Lundberg, Su-In Lee

Understanding why a model made a certain prediction is crucial in many Applications. However, with large modern datasets the best accuracy is often achieved by complex models even experts struggle to interpret, such as ensemble or deep learning models. This creates a tension between accuracy and interpretability. In response, a variety of methods have recently been proposed to help users interpret the predictions of complex models. Here, we present a unified framework for interpreting predictions, namely SHAP (SHapley Additive exPlanations), which assigns each feature an importance for a particular prediction. The key components of the SHAP framework are the identification of a class of additive feature importance measures and theoretical results that there is a unique solution in this class with a set of desired properties. This class unifies six existing methods, and several recent methods in this class do not have these desired properties. This means that our framework can inform the development of new methods for explaining prediction models. We demonstrate that several new methods we presented in this paper based on the SHAP framework show better computational performance and better consistency with human intuition than existing methods

SPOTLIGHTS

- **Differentiable Learning of Submodular Functions**
  Josip Djolonga, Andreas Krause
- **Generalized Linear Model Regression under Distance-to-set Penalties**
  Jason Xu, Eric Chi, Kenneth Lange
- **Decomposable Submodular Function Minimization: Discrete and Continuous**
  Alina Ene, Huy Nguyen, László A. Végh
- **Unbiased estimates for linear regression via volume sampling**
  Michal Derezinski, Manfred Warmuth
- **On Frank-Wolfe and Equilibrium Computation**
  Jacob D Abernethy, Jun-Kun Wang
- **On Separability of Loss Functions, and Revisiting Discriminative Vs Generative Models**
  Adarsh Prasad, Pradeep Ravikumar

Track 2 - 2:50 - 3:50 pm
Theory

Location: Hall C

Safe and Nested Subgame Solving for Imperfect-Information Games

Noam Brown, Tuomas Sandholm

Unlike perfect-information games, imperfect-information games cannot be solved by decomposing the game into subgames that are solved independently. Thus more computationally intensive equilibrium-finding techniques are used, and all decisions must consider the strategy of the game as a whole. While it is not possible to solve an imperfect-information game exactly through decomposition, it is possible to approximate solutions, or improve existing solutions, by solving disjoint subgames. This process is referred to as subgame solving. We introduce subgame solving techniques that outperform prior methods both in theory and practice. We also show how to adapt them, and past subgame-solving techniques, to respond to opponent actions that are outside the original action abstraction; this significantly outperforms the prior state-of-the-art approach, action translation. Finally, we show that subgame solving can be repeated as the game progresses down the tree, leading to significantly lower exploitability. We applied these techniques to develop the first AI to defeat top humans in heads-up no-limit Texas hold'em poker.

A Graph-theoretic Approach To Multitasking

Noga Alon, Daniel Reichman, Igor Shinkar, Tal Wagner, Sebastian Musslick, Tom Griffiths, Jonathan D Cohen, Biswadip dey, Kayhan Ozcimder

A key feature of neural network architectures is their ability to support the simultaneous interaction among large numbers of units in the learning and processing of representations. However, how the richness of such interactions trades off against the ability of a network to simultaneously carry out multiple independent processes -- a salient limitation in many domains of human cognition -- remains largely unexplored. In this paper we use a graph-theoretic analysis of network architecture to address this question, where tasks are represented as edges in a bipartite graph $G = (A \cup B, E)$. We define a new measure of multitasking capacity of such networks, based on the assumptions that tasks that \emph{need} to be multitasked rely on independent resources, i.e., form a matching, and that tasks **can** be performed without interference if they form an induced matching. Our main result is an inherent tradeoff between the multitasking capacity and the average degree of the network that holds \ emph{regardless of the network architecture}. These results are also extended to networks of depth greater than ***2***. On the positive side, we demonstrate that networks that are random-like (e.g., locally sparse) can have desirable multitasking properties. Our results shed light into the parallel-processing limitations of neural systems and provide insights that may be useful for the analysis and design of parallel architectures.

SPOTLIGHTS

- **Information-theoretic analysis of generalization capability of learning algorithms**
  Maxim Raginsky, Aolin Xu
- **Net-Trim: Convex Pruning of Deep Neural Networks with Performance Guarantee**
  Alireza Aghasi, Nam Nguyen, Justin Romberg
- **Clustering Billions of Reads for DNA Data Storage**
  Cyrus Rashtchian, Konstantin Makarychev, Luis Ceze, Karin Strauss, Sergey Yekhanin, Djordje Jevdjic, Miklos Racz, Siena Ang
- **On the Complexity of Learning Neural Networks**
  Le Song, Santosh Vempala, John Wilmes, Bo Xie
- **Multiplicative Weights Update with Constant Step-Size in Congestion Games: Convergence, Limit Cycles and Chaos**
  Gerasimos Palaiopanos, Ioannis Panageas, Georgios Piliouras
- **Estimating Mutual Information for Discrete-Continuous Mixtures**
  Weihao Gao, Sreeram Kannan, Sewoong Oh, Pramod Viswanath

Track 1 - 4:20 - 6:00 pm
Deep Learning, Applications

Location: Hall A

Unsupervised Object Learning From Dense Equivariant Image Labelling

James Thewlis, Andrea Vedaldi, Hakan Bilen

One of the key challenges of visual perception is to extract abstract models of 3D objects and object categories from visual measurements, which are affected by complex nuisance factors such as viewpoint, occlusion, motion, and deformations. Starting from the recent idea of viewpoint factorization, we propose a new approach that, given a large number of images of an object and no other supervision, can extract a dense object-centric coordinate frame. This coordinate frame is invariant to deformations of the images and comes with a dense equivariant labelling neural network that can map image pixels to their corresponding object coordinates. We demonstrate the applicability of this method to simple articulated objects and deformable objects such as human faces, learning embeddings from random synthetic transformations or optical flow correspondences, all without any manual supervision.

Interpretable and Globally Optimal Prediction for Textual Grounding using Image Concepts

Raymond Yeh, Jinjun Xiong, Wen-Mei Hwu, Minh Do, Alexander Schwing

Textual grounding is an important but challenging task for human-computer interaction, robotics and knowledge mining. Existing algorithms generally formulate the task as selection of the solution from a set of bounding box proposals obtained from deep net based systems. In this work, we demonstrate that we can cast the problem of textual grounding into a unified framework that permits efficient search over all possible bounding boxes. Hence, we able to consider significantly more proposals and, due to the unified formulation, our approach does not rely on a successful first stage. Beyond, we demonstrate that the trained parameters of our model can be used as word-embeddings which capture spatial-image relationships and provide interpretability. Lastly, our approach outperforms the current state-of-the-art methods on the Flickr 30k Entities and the ReferItGame dataset by 3.08 and 7.77 respectively.

Eigen-Distortions of Hierarchical Representations

Alexander Berardino, Valero Laparra, Johannes Ballé, Eero Simoncelli

We develop a method for comparing hierarchical image representations in terms of their ability to explain perceptual sensitivity in humans. Specifically, we utilize Fisher information to establish a model-derived prediction of local sensitivity to perturbations around a given natural image. For a given image, we compute the eigenvectors of the Fisher information matrix with largest and smallest eigenvalues, corresponding to the model-predicted most- and least-noticeable image distortions, respectively. For human subjects, we then measure the amount of each distortion that can be reliably detected when added to the image, and compare these thresholds to the predictions of the corresponding model. We use this method to test the ability of a variety of representations to mimic human perceptual sensitivity. We find that the early layers of VGG16, a deep neural network optimized for object recognition, provide a better match to human perception than later layers, and a better match than a 4-stage convolutional neural network (CNN) trained on a database of human ratings of distorted image quality. On the other hand, we find that simple models of early visual processing, incorporating one or more stages of local gain control, trained on the same database of distortion ratings, predict human sensitivity significantly better than both the CNN and all layers of VGG16.

SPOTLIGHTS

- **Towards Accurate Binary Convolutional Neural Network**
  Wei Pan, Xiaofan Lin, Cong Zhao
- **Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model**
  Xingjian Shi, Hao Wang, Zhihan Gao, Leonard Lausen, Dit-Yan Yeung, Wang-chun WOO, Wai-kin Wong
- **Poincaré Embeddings for Learning Hierarchical Representations**
  Maximillian Nickel, Douwe Kiela
- **Deep Hyperspherical Learning**
  Weiyang Liu, Yan-Ming Zhang, Xingguo Li, Zhiding Yu, Bo Dai, Tuo Zhao, Le Song
- **What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?**
  Alex Kendall, Yarin Gal
- **One-Sided Unsupervised Domain Mapping**
  Sagie Benaim, Lior Wolf
- **Deep Mean-Shift Priors for Image Restoration**
  Siavash Arjomand Bigdeli, Matthias Zwicker, Paolo Favaro, Meiguang Jin
- **Deep Voice 2: Multi-Speaker Neural Text-to-Speech**
  Andrew Gibiansky
- **Graph Matching via Multiplicative Update Algorithm**
  Bo Jiang, Jin Tang, Bin Luo
- **Dynamic Routing Between Capsules**
  Sara Sabour, Nicholas Frosst, Geoffrey E Hinton
- **Modulating early visual processing by language**
  Harm de Vries, Florian Strub, Jeremie Mary, Hugo Larochelle, Olivier Pietquin, Aaron C Courville

Track 2 - 4:20 am - 6:00 pm Algorithms

Location: Hall C

A Linear-Time Kernel Goodness-of-Fit Test

Wittawat Jitkrittum, Wenkai Xu, Zoltan Szabo, Kenji Fukumizu, Arthur Gretton

We propose a novel adaptive test of goodness-of-fit, with computational cost linear in the number of samples. We learn the test features that best indicate the differences between observed samples and a reference model, by minimizing the false negative rate. These features are constructed via Stein's method, meaning that it is not necessary to compute the normalising constant of the model. We analyse the asymptotic Bahadur efficiency of the new test, and prove that under a mean-shift alternative, our test always has greater relative efficiency than a previous linear-time kernel test, regardless of the choice of parameters for that test. In experiments, the performance of our method exceeds that of the earlier linear-time test, and matches or exceeds the power of a quadratic-time kernel test. In high dimensions and where model structure may be exploited, our goodness of fit test performs far better than a quadratic-time two-sample test based on the Maximum Mean Discrepancy, with samples drawn from the model.

Generalization Properties of Learning with Random Features

Alessandro Rudi, Lorenzo Rosasco

We study the generalization properties of ridge regression with random features in the statistical learning framework. We show for the first time that $O(1/\sqrt{n})$ learning bounds can be achieved with only $O(\sqrt{n}\, logn)$ random features rather than $O(n)$ as suggested by previous results. Further, we prove faster learning rates and show that they might require more random features, unless they are sampled according to a possibly problem dependent distribution. Our results shed light on the statistical computational trade-offs in large scale kernelized learning, showing the potential effectiveness of random features in reducing the computational complexity while keeping optimal generalization properties.

Communication-Efficient Distributed Learning of Discrete Distributions

Ilias Diakonikolas, Elena Grigorescu, Jerry Li, Abhiram Natarajan, Krzysztof Onak, Ludwig Schmidt

We initiate a systematic study of distribution learning (or density estimation) in the distributed model. In this problem the data drawn from an unknown distribution is partitioned across multiple machines. The machines must succinctly communicate with a referee so that in the end the referee can estimate the underlying distribution of the data. The problem is motivated by the pressing need to build communication-efficient protocols in various distributed systems, where power consumption or limited bandwidth impose stringent communication constraints. We give the first upper and lower bounds on the communication complexity of nonparametric density estimation of discrete probability distributions under both l1 and the l2 distances. Specifically, our results include the following: 1. In the case when the unknown distribution is arbitrary and each machine has only one sample, we show that any interactive protocol that learns the distribution must essentially communicate the entire sample. 2. In the case of structured distributions, such as k-histograms and monotone, we design distributed protocols that achieve better communication guarantees than the trivial ones, and show tight bounds in some regimes.

SPOTLIGHTS

- **Posterior sampling for Reinforcement learning: worst-case regret bounds**
  Shipra Agrawal, Randy Jia
- **Regret Analysis for Continuous Dueling Bandit**
  Wataru Kumagai
- **Minimal Exploration in Structured Stochastic Bandits**
  Stefan Magureanu, Richard Combes, Alexandre Proutiere
- **Fast Rates for Bandit Optimization with Upper-Confidence Frank-Wolfe**
  Quentin Berthet, Vianney Perchet
- **Diving into the shallows: a computational perspective on large-scale shallow learning**
  Siyuan Ma, Mikhail Belkin
- **Monte-Carlo Tree Search by Best Arm Identification**
  Emilie Kaufmann, Wouter Koolen
- **A framework for Multi-A(rmed)/B(andit) Testing with Online FDR Control**
  Fanny Yang, Aaditya Ramdas, Kevin Jamieson, Martin Wainwright
- **Parameter-Free Online Learning via Model Selection**
  Dylan J Foster, Satyen Kale, Mehryar Mohri, Karthik Sridharan
- **Bregman Divergence for Stochastic Variance Reduction: Saddle-Point and Adversarial Prediction**
  Zhan Shi, Xinhua Zhang, Yaoliang Yu
- **Gaussian Quadrature for Kernel Features**
  Tri Dao, Christopher M De Sa, Christopher Ré
- **Online Learning of Linear Dynamical Systems**
  Elad Hazan, Karan Singh, Cyril Zhang

Tuesday Poster Session

#1 Posterior sampling for Reinforcement learning: worst-case regret bounds
Shipra Agrawal, Randy Jia

#2 A framework for Multi-A(rmed)/B(andit) Testing with Online FDR Control
Fanny Yang, Aaditya Ramdas, Kevin Jamieson, Martin Wainwright

#3 Monte-Carlo Tree Search by Best Arm Identification
Emilie Kaufmann, Wouter Koolen

#4 Minimal Exploration in Structured Stochastic Bandits
Stefan Magureanu, Richard Combes, Alexandre Proutiere

#5 Regret Analysis for Continuous Dueling Bandit
Wataru Kumagai

#6 Elementary Symmetric Polynomials for Optimal Experimental Design
Zelda E. Mariet, Suvrit Sra

#7 Online Learning of Linear Dynamical Systems
Elad Hazan, Karan Singh, Cyril Zhang

#8 Efficient and Flexible Inference for Stochastic Systems
Stefan Bauer, Djordje Miladinovic, Nico S Gorbach, Joachim M Buhmann

#9 Group Sparse Additive Machine
Hong Chen, Xiaoqian Wang, Heng Huang

#10 Bregman Divergence for Stochastic Variance Reduction: Saddle-Point and Adversarial Prediction
Zhan Shi, Xinhua Zhang, Yaoliang Yu

#11 Online multiclass boosting
Young Jung, Jack Goetz, Ambuj Tewari

#12 Universal consistency and minimax rates for online Mondrian Forest
Jaouad Mourtada, Stéphane Gaïffas, Erwan Scornet

#13 Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results
Antti Tarvainen, Harri Valpola

#14 Learning from Complementary Labels
Takashi Ishida, Gang Niu, Weihua Hu, Masashi Sugiyama

#15 Positive-Unlabeled Learning with Non-Negative Risk Estimator
Ryuichi Kiryo, Gang Niu, Marthinus C du Plessis, Masashi Sugiyama

#16 Semisupervised Clustering, AND-Queries and Locally Encodable Source Coding
Arya Mazumdar, Soumyabrata Pal

#17 On Learning Errors of Structured Prediction with Approximate Inference
Yuanbin Wu

#18 On Optimal Generalizability in Parametric Learning
Ahmad Beirami, Meisam Razaviyayn, Shahin Shahrampour, Vahid Tarokh

#19 Multi-Objective Non-parametric Sequential Prediction
Guy Uziel, Ran El-Yaniv

#20 Fixed-Rank Approximation of a Positive-Semidefinite Matrix from Streaming Data
Joel A Tropp, Alp Yurtsever, Madeleine Udell, Volkan Cevher

#21 Communication-Efficient Stochastic Gradient Descent, with Applications to Neural Networks
Dan Alistarh, Demjan Grubic, Jerry Li, Ryota Tomioka, Milan Vojnovic

#22 Machine Learning with Adversaries: Byzantine Tolerant Gradient Descent
Peva Blanchard, El Mahdi El Mhamdi, Rachid Guerraoui, Julien Stainer

#23 Ranking Data with Continuous Labels through Oriented Recursive Partitions
Stéphan Clémençon, Mastane Achab

#24 Practical Data-Dependent Metric Compression with Provable Guarantees
Piotr Indyk, Ilya Razenshteyn, Tal Wagner

#25 Simple strategies for recovering inner products from coarsely quantized random projections
Ping Li, Martin Slawski

#26 Clustering Stable Instances of Euclidean k-means.
Aravindan Vijayaraghavan, Abhratanu Dutta, Alex Wang

#27 On Distributed Hierarchical Clustering
Mahsa Derakhshan, Soheil Behnezhad, Mohammadhossein Bateni, Vahab Mirrokni, MohammadTaghi Hajiaghayi, Silvio Lattanzi, Raimondas Kiveris

#28 Sparse k-Means Embedding
Weiwei Liu, Xiaobo Shen, Ivor Tsang

#29 K-Medoids For K-Means Seeding
James Newling, François Fleuret

#30 An Applied Algorithmic Foundation for Hierarchical Clustering
Joshua Wang, Benjamin Moseley

#31 Inhomogoenous Hypergraph Clustering with Applications
Pan Li, Olgica Milenkovic

#32 Subspace Clustering via Tangent Cones
Amin Jalali, Rebecca Willett

#33 Tensor Biclustering
Soheil Feizi, Hamid Javadi, David Tse

#34 A unified approach to interpreting model predictions
Scott M Lundberg, Su-In Lee

#35 Efficient Sublinear-Regret Algorithms for Online Sparse Linear Regression
Shinji Ito, Akihiro Yabe, Ken-Ichi Kawarabayashi, Naonori Kakimura, Takuro Fukunaga, Daisuke Hatano, Hanna Sumita

#36 Unbiased estimates for linear regression via volume sampling
Michal Derezinski, Manfred Warmuth

#37 On Separability of Loss Functions, and Revisiting Discriminative Vs Generative Models
Adarsh Prasad, Pradeep Ravikumar

#38 Generalized Linear Model Regression under Distance-to-set Penalties
Jason Xu, Eric Chi, Kenneth Lange

#39 Group Additive Structure Identification for Kernel Nonparametric Regression
Chao Pan, Michael Zhu

#40 Learning Overcomplete HMMs
Vatsal Sharan, Sham Kakade, Percy Liang, Gregory Valiant

#41 Matrix Norm Estimation from a Few Entries
Sewoong Oh, Ashish Khetan

#42 Optimal Shrinkage of Singular Values Under Random Data Contamination
Danny Barash, Matan Gavish

#43 A New Theory for Nonconvex Matrix Completion
Guangcan Liu, Xiaotong Yuan, Qingshan Liu

#44 Learning Low-Dimensional Metrics
Blake Mason, Lalit Jain, Robert Nowak

#45 Fast Alternating Minimization Algorithms for Dictionary Learning
Niladri Chatterji, Peter Bartlett

#46 Consistent Robust Regression
Kush Bhatia, Prateek Jain, Puru Kar

#47 Partial Hard Thresholding: A Towards Unified Analysis of Support Recovery
Jie Shen, Ping Li

#48 Minimax Estimation of Bandable Precision Matrices
Addison Hu, Sahand Negahban

#49 Diffusion Approximations for Online Principal Component Estimation and Global Convergence
Chris Junchi Li, Mengdi Wang, Tong Zhang

#50 Estimation of the covariance structure of heavy-tailed distributions
Xiaohan Wei, Stanislav Minsker

#51 Learning Koopman Invariant Subspaces for Dynamic Mode Decomposition
Naoya Takeishi, Yoshinobu Kawahara, Takehisa Yairi

#52 Stochastic Approximation for Canonical Correlation Analysis
Raman Arora, Teodor Vanislavov Marinov, Poorya Mianjy

#53 Diving into the shallows: a computational perspective on large-scale shallow learning
SIYUAN MA, Mikhail Belkin

#54 The Unreasonable Effectiveness of Structured Random Orthogonal Embeddings
Krzysztof M Choromanski, Mark Rowland, Adrian Weller

#55 Generalization Properties of Learning with Random Features
Alessandro Rudi, Lorenzo Rosasco

#56 Gaussian Quadrature for Kernel Features
Tri Dao, Christopher M De Sa, Chris Ré

#57 A Linear-Time Kernel Goodness-of-Fit Test
Wittawat Jitkrittum, Wenkai Xu, Zoltan Szabo, Kenji Fukumizu, Arthur Gretton

#58 Convergence rates of a partition based Bayesian multivariate density estimation method
Linxi Liu, Dangna Li, Wing Hung Wong

#59 The power of absolute discounting: all-dimensional distribution estimation
Moein Falahatgar, Mesrob Ohannessian, Alon Orlitsky, Venkatadheeraj Pichapati

#60 Optimally Learning Populations of Parameters
Kevin Tian, Weihao Kong, Gregory Valiant

#61 Communication-Efficient Distributed Learning of Discrete Distributions
Ilias Diakonikolas, Elena Grigorescu, Jerry Li, Abhiram Natarajan, Krzysztof Onak, Ludwig Schmidt

#62 Improved Dynamic Regret for Non-degeneracy Functions
Lijun Zhang, Tianbao Yang, Jinfeng Yi, Rong Jin, Zhi-Hua Zhou

#63 Parameter-Free Online Learning via Model Selection
Dylan J Foster, Satyen Kale, Mehryar Mohri, Karthik Sridharan

#64 Fast Rates for Bandit Optimization with Upper-Confidence Frank-Wolfe
Quentin Berthet, Vianney Perchet

#65 Online Learning with Transductive Regret
Scott Yang, Mehryar Mohri

#66 Multi-Armed Bandits with Metric Movement Costs
Tomer Koren, Roi Livni, Yishay Mansour

Tuesday Poster Session

#67 Differentially Private Empirical Risk Minimization Revisited: Faster and More General
Di Wang, Minwei Ye, Jinhui Xu

#68 Certified Defenses for Data Poisoning Attacks
Jacob Steinhardt, Pang Wei W Koh, Percy Liang

#69 Sparse Approximate Conic Hulls
Greg Van Buskirk, Ben Raichel, Nicholas Ruozzi

#70 On Tensor Train Rank Minimization : Statistical Efficiency and Scalable Algorithm
Masaaki Imaizumi, Takanori Maehara, Kohei Hayashi

#71 Sparse convolutional coding for neuronal assembly detection
Sven Peter, Elke Kirschbaum, Martin Both, Intramural Lee Campbell, Intramural Brandon Harvey, Intramural Conor Heins, Daniel Durstewitz, Ferran Diego, Fred Hamprecht

#72 Estimating High-dimensional Non-Gaussian Multiple Index Models via Stein's Lemma
Zhuoran Yang, krishnakumar balasubramanian, Princeton Zhaoran Wang, Han Liu

#73 Solid Harmonic Wavelet Scattering: Predicting Quantum Molecular Energy from Invariant Descriptors of 3D Electronic Densities
Michael Eickenberg, Georgios Exarchakis, Matthew Hirn, Stephane Mallat

#74 Clustering Billions of Reads for DNA Data Storage
Cyrus Rashtchian, Konstantin Makarychev, Luis Ceze, Karin Strauss, Sergey Yekhanin, Djordje Jevdjic, Miklos Racz, Siena Ang

#75 Deep Recurrent Neural Network-Based Identification of Precursor microRNAs
Seunghyun Park, Seonwoo Min, Hyun-Soo Choi, Sungroh Yoon

#76 Decoding with Value Networks for Neural Machine Translation
Di He, Hanqing Lu, Yingce Xia, Tao Qin, Liwei Wang, Tieyan Liu

#77 Towards the ImageNet-CNN of NLP: Pretraining Sentence Encoders with Machine Translation
Bryan McCann, James Bradbury, Caiming Xiong, Richard Socher

#78 Deep Voice 2: Multi-Speaker Neural Text-to-Speech
Andrew Gibiansky

#79 Modulating early visual processing by language
Harm de Vries, Florian Strub, Jeremie Mary, Hugo Larochelle, Olivier Pietquin, Aaron C Courville

#80 Multimodal Learning and Reasoning for Visual Question Answering
Ilija Ilievski, Jiashi Feng

#81 Learning to Model the Tail
Yu-Xiong Wang, Deva Ramanan, Martial Hebert

#82 Interpretable and Globally Optimal Prediction for Textual Grounding using Image Concepts
Raymond Yeh, Jinjun Xiong, Wen-Mei Hwu, Minh Do, Alex Schwing

#83 Multiscale Quantization for Fast Similarity Search
Xiang Wu, Ruiqi Guo, Ananda Theertha Suresh, Daniel Holtmann-Rice, David Simcha, Felix Yu, Sanjiv Kumar

#84 MaskRNN: Instance Level Video Object Segmentation
Yuan-Ting Hu, Jia-Bin Huang, Alex Schwing

#85 Flat2Sphere: Learning Spherical Convolution for Fast Features from 360° Imagery
Yu-Chuan Su, Kristen Grauman

#86 Deep Mean-Shift Priors for Image Restoration
Siavash Arjomand Bigdeli, Matthias Zwicker, Paolo Favaro, Meiguang Jin

#87 Pixels to Graphs by Associative Embedding
Alejandro Newell, Jia Deng

#88 3D Shape Reconstruction by Modeling 2.5D Sketch
Jiajun Wu, Yifan Wang, Tianfan Xue, Xingyuan Sun, Bill Freeman, Josh Tenenbaum

#89 Temporal Coherency based Criteria for Predicting Video Frames using Deep Multi-stage Generative Adversarial Networks
Prateep Bhattacharjee, S. Das

#90 Learning to Generalize Intrinsic Images with a Structured Disentangling Autoencoder
Michael Janner, Jiajun Wu, Tejas Kulkarni, Ilker Yildirim, Josh Tenenbaum

#91 Unsupervised object learning from dense equivariant image labelling
James Thewlis, Andrea Vedaldi, Hakan Bilen

#92 One-Sided Unsupervised Domain Mapping
Sagie Benaim, Lior Wolf

#93 Contrastive Learning for Image Captioning
Bo Dai, Dahua Lin

#94 Dynamic Routing Between Capsules
Sara Sabour, Nicholas Frosst, Geoffrey E Hinton

#95 What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?
Alex Kendall, Yarin Gal

#96 Efficient Optimization for Linear Dynamical Systems with Applications to Clustering and Sparse Coding
Wenbing Huang, Fuchun Sun, Tong Zhang, Junzhou Huang, Mehrtash Harandi

#97 Label Distribution Learning Forests
Wei Shen, KAI ZHAO, Yilu Guo, Alan Yuille

#98 Graph Matching via Multiplicative Update Algorithm
Bo Jiang, Jin Tang, Bin Luo

#99 Training Quantized Nets: A Deeper Understanding
Hao Li, Soham De, Zheng Xu, Christoph Studer, Hanan Samet, Tom Goldstein

#100 Inner-loop free ADMM using Auxiliary Deep Neural Networks
Kai Fan, Qi Wei, Katherine A Heller

#101 Towards Accurate Binary Convolutional Neural Network
Wei Pan, Xiaofan Lin, Cong Zhao

#102 Runtime Neural Pruning
Ji Lin, Yongming Rao, Jiwen Lu

#103 Structured Embedding Models for Grouped Data
Maja Rudolph, Francisco Ruiz, David Blei

#104 Poincaré Embeddings for Learning Hierarchical Representations
Maximillian Nickel, Douwe Kiela

#105 Language modeling with recurrent highway hypernetworks
Joseph Suarez

#106 Preventing Gradient Explosions in Gated Recurrent Units
Sekitoshi Kanai, Yasuhiro Fujiwara, Sotetsu Iwamura

#107 Wider and Deeper, Cheaper and Faster: Tensorized LSTMs for Sequence Learning
Zhen He, Shaobing Gao, Liang Xiao, David Barber

#108 Fast-Slow Recurrent Neural Networks
Asier Mujika, Florian Meier, Angelika Steger

#109 Cold-Start Reinforcement Learning with Softmax Policy Gradients
Nan Ding, Radu Soricut

#110 Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model
Xingjian Shi, Hao Wang, Zhihan Gao, Leonard Lausen, Dit-Yan Yeung, Wang-chun WOO, Wai-kin Wong

#111 Recurrent Ladder Networks
Isabeau Prémont-Schwarz, Alexander Ilin, Hotloo Hao, Antti Rasmus, Rinu Boney, Harri Valpola

#112 Predictive-State Decoders: Encoding the Future into Recurrent Networks
Arun Venkatraman, Nick Rhinehart, Wen Sun, Lerrel Pinto, Martial Hebert, Byron Boots, Kris Kitani, J. Bagnell

#113 QMDP-Net: Deep Learning for Planning under Partial Observability
Peter Karkus, David Hsu, Wee Sun Lee

#114 Filtering Variational Objectives
Chris Maddison, Dieterich Lawson, George Tucker, Mohammad Norouzi, Nicolas Heess, Andriy Mnih, Yee Teh, Arnaud Doucet

#115 Unsupervised Learning of Disentangled Latent Representations from Sequential Data
Wei-Ning Hsu, Yu Zhang, James Glass

#116 Neural Discrete Representation Learning
Aaron van den Oord, Oriol Vinyals, koray kavukcuoglu

#117 Variational Memory Addressing in Generative Models
Jörg Bornschein, Andriy Mnih, Daniel Zoran, Danilo Jimenez Rezende

#118 Cortical microcircuits as gated-recurrent neural networks
Rui Costa, Yannis Assael, Brendan Shillingford, Nando de Freitas, TIm Vogels

#119 Continual Learning with Deep Generative Replay
Hanul Shin, Jung Kwon Lee, Jaehong Kim, Jiwon Kim

#120 Hierarchical Attentive Recurrent Tracking
Adam Kosiorek, Alex Bewley, Ingmar Posner

#121 VAE Learning via Stein Variational Gradient Descent
Yuchen Pu, Zhe Gan, Ricardo Henao, Chunyuan Li, Shaobo Han, Lawrence Carin

#122 Learning to Inpaint for Image Compression
Mohammad Haris Baig, Vladlen Koltun, Lorenzo Torresani

#123 Visual Interaction Networks
Nick Watters, Daniel Zoran, Theophane Weber, Peter Battaglia, Razvan Pascanu, Andrea Tacchetti

#124 NeuralFDR: Learning Discovery Thresholds from Hypothesis Features
Martin J Zhang, Fei Xia, James Zou, David Tse

#125 Eigen-Distortions of Hierarchical Representations
Alexander Berardino, Valero Laparra, Johannes Ballé, Eero Simoncelli

#126 On-the-fly Operation Batching in Dynamic Computation Graphs
Graham Neubig, Yoav Goldberg, Chris Dyer

#127 Learning Affinity via Spatial Propagation Networks
Sifei Liu, Guangyu Zhong, Ming-Hsuan Yang, Shalini De Mello, Jan Kautz, Jinwei Gu

#128 Supervised Adversarial Domain Adaptation
Saeid Motiian, Quinn Jones, Gianfranco Doretto

#129 Deep Hyperspherical Learning
Weiyang Liu, Yan-Ming Zhang, Xingguo Li, Zhiding Yu, Bo Dai, Tuo Zhao, Le Song

#130 Riemannian approach to batch normalization
Minhyung Cho, Jaehyung Lee

#131 Backprop without Learning Rates Through Coin Betting
Francesco Orabona, Tatiana Tommasi

#132 On the Convergence of Block Coordinate Descent in Training DNNs with Tikhonov Regularization
Ziming Zhang, Matthew Brand

#133 Collaborative Deep Learning in Fixed Topology Networks
Zhanhong Jiang, Aditya Balu, Chinmay Hegde, Soumik Sarkar

#134 How regularization affects the critical points in linear networks
Amir Taghvaei, Jin W Kim, Prashant Mehta

#135 Predicting Organic Reaction Outcomes with Weisfeiler-Lehman Network
Wengong Jin, Connor Coley, Regina Barzilay, Tommi Jaakkola

#136 Predicting Scene Parsing and Motion Dynamics in the Future
Xiaojie Jin, Jiashi Feng, Huaxin Xiao, Yunpeng Chen, Shuicheng Yan, Xiaohui Shen, Jimei Yang, Zequn Jie, Li Ping

#137 Houdini: Democratizing Adversarial Examples
Moustapha Cisse, adiyoss Adi, Natalia Neverova, Yossi Keshet

#138 Geometric Matrix Completion with Recurrent Multi-Graph Neural Networks
Federico Monti, Michael Bronstein, Xavier Bresson

#139 Compression-aware Training of Deep Neural Networks
Jose Alvarez, Mathieu Salzmann

#140 Non-parametric Neural Networks
Andreas Lehrmann, Leonid Sigal

#141 GibbsNet: Iterative Adversarial Inference for Deep Graphical Models
Alex Lamb, devon Hjelm, Yaroslav Ganin, Joseph Paul Cohen, Aaron C Courville, Yoshua Bengio

#142 Exploring Generalization in Deep Learning
Behnam Neyshabur, Srinadh Bhojanapalli, Nati Srebro

#143 Regularizing Deep Neural Networks by Noise: Its Interpretation and Optimization
Hyeonwoo Noh, Tackgeun You, Jonghwan Mun, Bohyung Han

#144 Extracting low-dimensional dynamics from multiple large-scale neural population recordings by learning to predict correlations
Marcel Nonnenmacher, Srini C Turaga, Jakob H Macke

#145 Adaptive sampling for a population of neurons
Benjamin Cowley, Ryan Williamson, Katerina Clemens, Matthew Smith, Byron M Yu

#146 OnACID: Online Analysis of Calcium Imaging Data in Real Time
Andrea Giovannucci, Johannes Friedrich, Matt Kaufman, Anne Churchland, Dmitri Chklovskii, Liam Paninski, Eftychios Pnevmatikakis

#147 Detrended Partial Cross Correlation for Brain Connectivity Analysis
Jaime Ide, Fábio Cappabianco, Fabio Faria, Chiang-shan R Li

#148 Practical Bayesian Optimization for Model Fitting with Bayesian Adaptive Direct Search
Luigi Acerbi, Wei Ji

#149 An Error Detection and Correction Framework for Connectomics
Jonathan Zung, Ignacio Tartavull

#150 GP CaKe: Effective brain connectivity with causal kernels
Luca Ambrogioni, Max Hinne, Marcel Van Gerven, Eric Maris

#151 Learning Neural Representations of Human Cognition across Many fMRI Studies
Arthur Mensch, Julien Mairal, Danilo Bzdok, Bertrand Thirion, Gael Varoquaux

#152 Mapping distinct timescales of functional interactions among brain networks
Mali Sundaresan, Arshed Nabeel, Devarajan Sridharan

#153 Robust Estimation of Neural Signals in Calcium Imaging
Hakan Inan, Murat A. Erdogdu, Mark Schnitzer

#154 Learning the Morphology of Brain Signals Using Alpha-Stable Convolutional Sparse Coding
Mainak Jas, Tom Dupré la Tour, Umut Simsekli, Alex Gramfort

#155 Streaming Weak Submodularity: Interpreting Neural Networks on the Fly
Ethan Elenberg, Alex Dimakis, Moran Feldman, Amin Karbasi

#156 Decomposable Submodular Function Minimization: Discrete and Continuous
Alina Ene, Huy Nguyen, László A. Végh

#157 Differentiable Learning of Submodular Functions
Josip Djolonga, Andreas Krause

#158 Robust Optimization for Non-Convex Objectives
Yaron Singer, Robert S Chen, Vasilis Syrgkanis, Brendan Lucier

#159 On the Optimization Landscape of Tensor Decompositions
Rong Ge, Tengyu Ma

#160 Gradient Descent Can Take Exponential Time to Escape Saddle Points
Simon Du, Chi Jin, Jason D Lee, Michael Jordan, Aarti Singh, Barnabas Poczos

#161 Convolutional Phase Retrieval
Qing Qu, Yuqian Zhang, Yonina Eldar, John Wright

#162 Implicit Regularization in Matrix Factorization
Suriya Gunasekar, Blake Woodworth, Srinadh Bhojanapalli, Behnam Neyshabur, Nati Srebro

#163 Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration
Jason Altschuler, Jon Weed, Philippe Rigollet

#164 On Frank-Wolfe and Equilibrium Computation
Jacob D Abernethy, Jun-Kun Wang

#165 Greedy Algorithms for Cone Constrained Optimization with Convergence Guarantees
Francesco Locatello, Michael Tschannen, Gunnar Raetsch, Martin Jaggi

#166 When Cyclic Coordinate Descent Beats Randomized Coordinate Descent
Mert Gurbuzbalaban, Denizcan Vanli, Asuman Ozdaglar

#167 Linear Convergence of a Frank-Wolfe Type Algorithm over Trace-Norm Balls
Zeyuan Allen-Zhu, Elad Hazan, Wei Hu, Yuanzhi Li

#168 Adaptive Accelerated Gradient Converging Method under H\"{o}lderian Error Bound Condition
Mingrui Liu, Tianbao Yang

#169 Searching in the Dark: Practical SVRG Methods under Error Bound Conditions with Guarantee
Yi Xu, Qihang Lin, Tianbao Yang

#170 Geometric Descent Method for Convex Composite Minimization
Shixiang Chen, Shiqian Ma, Wei Liu

#171 Faster and Non-ergodic O(1/K) Stochastic Alternating Direction Method of Multipliers
Cong Fang, Feng Cheng, Zhouchen Lin

#172 Doubly Accelerated Stochastic Variance Reduced Dual Averaging Method for Regularized Empirical Risk Minimization
Tomoya Murata, Taiji Suzuki

#173 Limitations on Variance-Reduction and Acceleration Schemes for Finite Sums Optimization
Yossi Arjevani

#174 Nonlinear Acceleration of Stochastic Algorithms
Damien Scieur, Francis Bach, Alexandre d'Aspremont

#175 Acceleration and Averaging in Stochastic Descent Dynamics
Walid Krichene

#176 Multiscale Semi-Markov Dynamics for Intracortical Brain-Computer Interfaces
Daniel Milstein, Jason Pacheco, Leigh Hochberg, John D Simeral, Beata Jarosiewicz, Erik Sudderth

#177 EEG-GRAPH: A Factor Graph Based Model for Capturing Spatial, Temporal, and Observational Relationships in Electroencephalograms
Yogatheesan Varatharajah, Min Jin Chong, Krishnakant Saboo, Brent Berry, Benjamin Brinkmann, Gregory Worrell, Ravishankar Iyer

#178 Asynchronous Parallel Coordinate Minimization for MAP Inference
Ofer Meshi, Alex Schwing

#179 Speeding Up Latent Variable Gaussian Graphical Model Estimation via Nonconvex Optimization
Pan Xu, Jian Ma, Quanquan Gu

#180 The Expxorcist: Nonparametric Graphical Models Via Conditional Exponential Densities
Arun Suggala, Mladen Kolar, Pradeep Ravikumar

#181 Reducing Reparameterization Gradient Variance
Andrew Miller, Nick Foti, Alexander D'Amour, Ryan Adams

#182 Robust Conditional Probabilities
Yoav Wald, Amir Globerson

#183 Stein Variational Gradient Descent as Gradient Flow
Qiang Liu

#184 Parallel Streaming Wasserstein Barycenters
Matt Staib, Sebastian Claici, Justin M Solomon, Stefanie Jegelka

#185 AIDE: An algorithm for measuring the accuracy of Probabilistic inference algorithms
Marco Cusumano-Towner, Vikash K Mansinghka

#186 Deep Dynamic Poisson Factorization Model
Chengyue Gong, win-bin huang

#187 On the Model Shrinkage Effect of Gamma Process Edge Partition Models
Iku Ohama, Issei Sato, Takuya Kida, Hiroki Arimura

#188 Model evidence from nonequilibrium simulations
Michael Habeck

#189 A-NICE-MC: Adversarial Training for MCMC
Jiaming Song, Shengjia Zhao, Stefano Ermon

#190 Identification of Gaussian Process State Space Models
Stefanos Eleftheriadis, Tom Nicholson, Marc Deisenroth, James Hensman

#191 Streaming Sparse Gaussian Process Approximations
Thang D Bui, Cuong Nguyen, Richard E Turner

#192 Bayesian Optimization with Gradients
Jian Wu, Matthias Poloczek, Andrew Wilson, Peter Frazier

#193 Variational Inference for Gaussian Process Models with Linear Complexity
Ching-An Cheng, Byron Boots

#194 Efficient Modeling of Latent Information in Supervised Learning using Gaussian Processes
Zhenwen Dai, Mauricio A. Álvarez, Neil Lawrence

Tuesday Poster Session

#195 Non-Stationary Spectral Kernels
Sami Remes, Markus Heinonen, Samuel Kaski

#196 Scalable Log Determinants for Gaussian Process Kernel Learning
David Eriksson, Kun Dong, David Bindel, Andrew Wilson, Hannes Nickisch

#197 Spectral Mixture Kernels for Multi-Output Gaussian Processes
Gabriel Parra, Felipe Tobar

#198 Linearly constrained Gaussian processes
Carl Jidling, Niklas Wahlström, Adrian Wills, Thomas B Schön

#199 Hindsight Experience Replay
Marcin Andrychowicz, fjwolski Wolski, Alex Ray, Jonas Schneider, rfong Fong, Peter Welinder, Bob McGrew, Josh Tobin, OpenAI Pieter Abbeel, Wojciech Zaremba

#200 Log-normality and Skewness of Estimated State/Action Values in Reinforcement Learning
Liangpeng Zhang, Ke Tang, Xin Yao

#201 Finite sample analysis of the GTD Policy Evaluation Algorithms in Markov Setting
Yue Wang

#202 Inverse Filtering for Hidden Markov Models
Robert Mattila, Cristian Rojas, Vikram Krishnamurthy, Bo Wahlberg

#203 Safe Model-based Reinforcement Learning with Stability Guarantees
Felix Berkenkamp, Matteo Turchetta, Angela Schoellig, Andreas Krause

#204 Data-Efficient Reinforcement Learning in Continuous State-Action Gaussian-POMDPs
Rowan McAllister, Carl Edward Rasmussen

#205 Linear regression without correspondence
Daniel Hsu, Kevin Shi, Xiaorui Sun

#206 On the Complexity of Learning Neural Networks
Le Song, Santosh Vempala, John Wilmes, Bo Xie

#207 Near Optimal Sketching of Low-Rank Tensor Regression
Jarvis Haupt, Xingguo Li, David Woodruff

#208 Is Input Sparsity Time Possible for Kernel Low-Rank Approximation?
Cameron Musco, David Woodruff

#209 Higher-Order Total Variation Classes on Grids: Minimax Theory and Trend Filtering Methods
Veeranjaneyulu Sadhanala, Yu-Xiang Wang, James Sharpnack, Ryan Tibshirani

#210 Alternating Estimation for Structured High-Dimensional Multi-Response Models
Sheng Chen, Arindam Banerjee

#211 Adaptive Clustering through Semidefinite Programming
Martin Royer

#212 Compressing the Gram Matrix for Learning Neural Networks in Polynomial Time
Surbhi Goel, Adam Klivans

#213 Learning with Average Top-k Loss
Yanbo Fan, Siwei Lyu, Yiming Ying, Baogang Hu

#214 Hierarchical Clustering Beyond the Worst-Case
Vincent Cohen-Addad, Varun Kanade, Frederik Mallmann-Trenn

#215 Net-Trim: Convex Pruning of Deep Neural Networks with Performance Guarantee
Alireza Aghasi, Nam Nguyen, Justin Romberg

#216 A graph-theoretic approach to multitasking
Noga Alon, Daniel Reichman, Igor Shinkar, Tal Wagner, Sebastian Musslick, Tom Griffiths, Jonathan D Cohen, Biswadip dey, Kayhan Ozcimder

#217 Information-theoretic analysis of generalization capability of learning algorithms
Maxim Raginsky, Aolin Xu

#218 Independence clustering (without a matrix)
Daniil Ryabko

#219 Polynomial Codes: an Optimal Design for High-Dimensional Coded Matrix Multiplication
Qian Yu, Mohammad Maddah-Ali, Salman Avestimehr

#220 Estimating Mutual Information for Discrete-Continuous Mixtures
Weihao Gao, Sreeram Kannan, Sewoong Oh, Pramod Viswanath

#221 Best Response Regression
Omer Ben Porat, Moshe Tennenholtz

#222 Statistical Cost Sharing
Eric Balkanski, Umar Syed, Sergei Vassilvitskii

#223 A Sample Complexity Measure with Applications to Learning Optimal Auctions
Vasilis Syrgkanis

#224 Multiplicative Weights Update with Constant Step-Size in Congestion Games: Convergence, Limit Cycles and Chaos
Gerasimos Palaiopanos, Ioannis Panageas, Georgios Piliouras

#225 Efficiency Guarantees from Data
Darrell Hoy, Denis Nekipelov, Vasilis Syrgkanis

#226 Safe and Nested Subgame Solving for Imperfect-Information Games
Noam Brown, Tuomas Sandholm

See Page 9 For Specific Demo Locations

D1 A Deep Reinforcement Learning Chatbot

Iulian Vlad Serban, Chinnadhurai Sankar, Mathieu Germain, Saizheng Zhang, Zhouhan Lin, Sandeep Subramanian, Taesup Kim, Mike J Pieper, Sarath Chandar, Rosemary Ke, Sai Rajeswar Mudumba, Alexandre de Brébisson, Jose Sotelo, Dendi A Suhubdy, Vincent Michalski, Joelle Pineau, Yoshua Bengio

Dialogue systems and conversational agents - including chatbots, personal assistants and voice-control interfaces - are becoming ubiquitous in modern society. Examples include personal assistants on mobile devices, customer service assistants and technical support help, as well as online bots selling anything from fashion clothes, cosmetics to legal advice and self-help therapy. Nevertheless, building high-quality intelligent conversational agents remains a major challenge for the machine learning community. This demo shows the chatbot developed by the team at the Montreal Institute of Learning Algorithms (MILA), which also participated in the Amazon Alexa Prize competition held between 2016 - 2017. The system is a socialbot: a spoken conversational agent capable of conversing engagingly with humans on popular small talk topics. Between April and August, 2017, the system had over ten thousand conversations with real-world users in the Amazon Alexa Prize competition

D2 CTRL-Labs: Non-invasive Neural Interface

Patrick Kaifosh, Tudor Giurgica-Tiron, Alan Du, Adam Al-natsheh, Jeffrey Seely, Steven Demers

CTRL-Labs is developing a non-invasive neural interface for everyday use. This interactive demo will showcase a complete end-to-end neural control application. Users will be able to: (A) wear our non-invasive device prototype on the wrist (B) map their own choice of neuromotor control schemes to multiple continuous and discrete degrees of freedom (C) play an arcade-style computer game in real time based on their own choice of control schemes.

D3 TincyYolo: Smaller still, faster, and more efficient

Michaela Blott, Nicholas Fraser

Recent research demonstrated that even extreme reduced precision works well for convolutional neural networks used for object classification. We leveraged similar quantization techniques in combination with filter pruning to reduce the computational footprint of YOLO networks such that high performance implementations within power-constraint embedded compute environments can be achieved. The demo will consist of a small embedded platform at ~6Watts power consumption, directly connected to a USB camera and a display port. The compute is performed by a Xilinx Zynq Ultrascale+ device which consists of a quadcore ARM processor and a medium-sized FPGA fabric. The live camera video stream will be processed by the MPSOC device's ARM processors, NEON cores and a NN accelerator in the FPGA fabric in real-time and shown on a monitor, whereby the 20 object classes of Pascal VOC are live classified and indicated through bounding boxes. The run-time environment is fully integrated with DarkNet and demonstrated with dynamic off-loading and on- loading the accelerators. Users can directly interact with the demo through holding different types of objects in front of the camera to test the accuracy of the heavily quantized and pruned neural network. Furthermore, users can dynamically move layers from ARM processors and NEON to the FPGA fabric to experience the speed up and latency reduction of custom hardware accelerators. To the best of our knowledge, this is the first extreme reduced precision and pruned variant of YOLO demonstrated. While FPGA-based neural networks have started to emerge, this is a first which demonstrates high performance and reduced power for object recognition. Furthermore, our extensions of a DarkNet run-time that allows for dynamic on- and offloading on ARM, NEON and FPGA is novel.

D4 A cortical neural network simulator for kids

Michiro Negishi

Technology: An educational tool for learning, building, and testing cortical micro-circuit. The student can learn about the cortical tissue structure by reading descriptions and knowing default properties presented in kid-friendly manners, and then experiment by tuning simple parameters (e.g. dendric and axonal distributions by layers and cell types) and running predefined learning algorithms (Self Organizing Feature Maps, contrastive Hebbian).

Activity: The audience will define the cortical tissue micro-structure using a simple, kid- friendly interface and run the network using an OCR database. The operation of the neural network is displayed in semi-realistic 3D graphics. The audience can also draw digits online and run the network. The network parameters and test performance are recorded and the best performers are listed on a screen.

D5 Libratus: Beating Top Humans in No-Limit Poker

Noam Brown, Tuomas Sandholm

Heads-up no-limit Texas Hold'em is a primary benchmark challenge for AI. Due to the hidden information in poker, techniques used for games like chess and Go are ineffective. We present Libratus, the first AI to defeat top human professionals in no-limit poker. Libratus features three main components: pre-computing a solution to an abstraction of the game which provides a high-level blueprint for how the AI should play, a new nested subgame-solving algorithm which repeatedly calculates a more detailed strategy as play progresses, and a self-improving module which augments the pre-computed blueprint based on opponent behavior.

D6 Deep Robotic Learning using Visual Imagination and Meta-Learning

Chelsea Finn, Frederik Ebert, Tianhe Yu, Annie Xie, Sudeep Dasari, Pieter Abbeel, Sergey Levine

A key, unsolved challenge for learning with real robotic systems is the ability to acquire vision-based behaviors from raw RGB images that can generalize to new objects and new goals. We present two approaches to this goal that we plan to demonstrate: first, learning task-agnostic visual models for planning, which can generalize to new objects and goals, and second, learning to quickly adapt to new objects and environments using meta-imitation learning. In essence, these two approaches seek to generalize and dynamically adapt to new settings, respectively, as we discuss next.

D7 Conversational Speech Search on Massive Audio Datasets

Anthony Scodary, Wonkyum Lee, Nico Benitez, Samuel Kim

We present the Sift conversational search system. Our search system is designed to search through billions of minutes of long-form, conversational speech data. The core technology allows for complex searches that combine semantic and signal information, and a method for executing constraints on time, logical structure, and metadata.

D8 Symbol Grounding and Program Induction using Multi-modal instructions, Visual Cues and Eye Tracking.

Yordan Hristov, Emmanuel Kahembwe, Subramanian Ramamoorthy

As situated agents begin to cohabit with humans in semi-structured environments, the need arises to understand their instructions, conveyed to the agent via a combination of natural language utterances and physical actions. Understanding the instructions involves decoding the speaker's intended message from their signal, and this involves learning how to ground the symbols in the physical world. The latter can be ambiguous due to variability in the physical instantiations of concepts - different people might use turquoise, sky blue, light blue and blue while referring to the same color or small-sized building blocks for one person could be determined as medium-sized by another. Realistically, symbol grounding is a task which must cope with small datasets consisting of a particular user's contextual assignment of meaning to terms. We demonstrate a framework for inferring abstract plans and symbol groundings over human Demos of a task.

D9 Sensomind: Democratizing deep learning for the food industry

Michael Sass Hansen, Sebastian Brandes Kraaijenzank

Microsoft exhibits great use of technology at the Technology Centers around the globe. Sensomind is one of the solutions showcased at their Copenhagen center and shows how the cloud and the power of arti cial intelligence can be put to use with the purpose of increasing product quality and optimizing production processes for manufacturing companies of every type all over the world. At the Technology Center, visitors get to experience what it is like to be a modern plant manager in an Industry 4.0 world. The demo allows visitors to train a neural network that can detect aws in products of di erent kinds. At the stand, there are various types of fake plastic foods available for visitors to use when training their models. Using a simple and intuitive web interface, visitors can deploy their newly trained neural network into production and see it running live making predictions on products passing by on a conveyor belt at the stand. The predictions made on the products are being uploaded to Sensomind's solution in the cloud where the data are being visualized in an easy-to-use dashboard hosted in Power BI (https://powerbi.microsoft.com/). Power BI allows the visitor to dig into the data and make analyses on the predictions made. This enables the visitor in their function as a plant manager to get insights about the production and potentially identify pattern and causes for errors. At this stage, the data become very actionable, as the visitor can act upon the insights and resolve the issue causing the errors.

D10 Deep Neural Net implementations with FPGAs

Thomas Boser, Paolo Calafiura, Ian Johnson

With recent increases in the luminosity of Large Hadron Collider (LHC) collisions creating more tracking data an efficient track reconstruction solution has become necessary. As it currently stands during the level 1 trigger it is necessary to identify 50 million particle tracks per second with lower than 5 microsecond latency per track. This requires a low latency highly parallel implementation or a connectthedots track reconstruction algorithm. Current algorithms are implemented on ASIC chips or FPGAs and scale O(N2)or worse. It is projected that we'll experience a O(10x) resource shortage with current implementations.

Simultaneously deep learning has become a standard technique in computer vision. We explore the viability of a deep learning solution for track reconstruction. We have explored various implementations of DNNs applied to the tracking problem and have promising preliminary results. We've explored using CNNs, RNNs, LSTMs, and Deep Kalman Filters. Current popular deep learning libraries are all heavily reliant on Graphics Processing Units (GPUs) to shoulder the bulk of heavy computation. These libraries show incredible results with rapidly improving throughput. Unfortunately this cannot be applied for latency sensitive Applications such as our track reconstruction problem because GPUs cannot guarantee low latency.

WEDNESDAY SESSIONS

7:30 - 9:00 AM	Coffee	
9:00 - 9:50 AM	**Invited Talk: Lise Getoor** *The Unreasonable Effectiveness of Structure*	**Hall A**
9:50 - 10:20 AM	Coffee break	
10:20 - 12:00 PM	**Parallel Tracks:**	
	Theory, Probabilistic Methods	**Hall A**
	Deep Learning	**Hall C**
12:00 - 1:50 PM	Lunch on your own	
1:50 - 2:40 PM	**Invited Talk: Pieter Abbeel** *Deep Learning for Robotics*	**Hall A**
2:40 - 2:50 PM	Quick break	
2:50 - 3:50 PM	**Parallel Tracks:**	
	Reinforcement Learning, Deep Learning	**Hall A**
	Optimization	**Hall C**
3:50 - 4:20 PM	Coffee break	
4:20 - 6:00 PM	**Parallel Tracks:**	
	Reinforcement Learning, Algorithms, Applications	**Hall A**
	Probabilistic Methods, Applications	**Hall C**
6:00 - 7:00 PM	Light snack	
7:00 - 10:30 PM	Poster session and Demos	**Pacific Ballroom**

The Unreasonable Effectiveness of Structure

Hall A, 9:00 - 9:50 AM

Our ability to collect, manipulate, analyze, and act on vast amounts of data is having a profound impact on all aspects of society. Much of this data is heterogeneous in nature and interlinked in a myriad of complex ways. From information integration to scientific discovery to computational social science, we need machine learning methods that are able to exploit both the inherent uncertainty and the innate structure in a domain. Statistical relational learning (SRL) is a subfield that builds on principles from probability theory and statistics to address uncertainty while incorporating tools from knowledge representation and logic to represent structure. In this talk, I will give a brief introduction to SRL, present templates for common structured prediction problems, and describe modeling approaches that mix logic, Probabilistic inference and latent variables. I'll overview our recent work on Probabilistic soft logic (PSL), an SRL framework for large-scale collective, Probabilistic reasoning in relational domains. I'll close by highlighting emerging opportunities (and challenges!!) in realizing the effectiveness of data and structure for knowledge discovery.

Lise Getoor
(UC Santa Cruz)

Lise Getoor is a professor in the Computer Science Department at the University of California, Santa Cruz. Her research areas include machine learning, data integration and reasoning under uncertainty, with an emphasis on graph and network data. She has over 250 publications and extensive experience with machine learning and Probabilistic modeling methods for graph and network data. She is a Fellow of the Association for Artificial Intelligence, an elected board member of the International Machine Learning Society, serves on the board of the Computing Research Association (CRA), and was co-chair for ICML 2011. She is a recipient of an NSF Career Award and eleven best paper and best student paper awards. She received her PhD from Stanford University in 2001, her MS from UC Berkeley, and her BS from UC Santa Barbara, and was a professor in the Computer Science Department at the University of Maryland, College Park from 2001-2013.

Deep Learning for Robotics

Hall A, 1:50 - 2:40 PM

Computer scientists are increasingly concerned about the many ways that machine learning can reproduce and reinforce forms of bias. When ML systems are incorporated into core social institutions, like healthcare, criminal justice and education, issues of bias and discrimination can be extremely serious. But what can be done about it? Part of the trouble with bias in machine learning in high-stakes decision making is that it can be the result of one or many factors: the training data, the model, the system goals, and whether the system works less well for some populations, among several others. Given the difficulty of understanding how a machine learning system produced a particular result, bias is often discovered after a system has been producing unfair results in the wild. But there is another problem as well: the definition of bias changes significantly depending on your discipline, and there are exciting approaches from other fields that have not yet been included by computer science. This talk will look at the recent literature on bias in machine learning, consider how we can incorporate approaches from the social sciences, and offer new strategies to address bias.

Pieter Abbeel
(OpenAI / UC Berkeley / Gradescope)

Pieter Abbeel (Associate Professor at UC Berkeley, Research Scientist at OpenAI, Co-Founder Gradescope) works in machine learning and robotics, in particular his research is on making robots learn from people (apprenticeship learning) and how to make robots learn through their own trial and error (Reinforcement learning). His robots have learned advanced helicopter aerobatics, knot-tying, basic assembly, and organizing laundry. His group has pioneered deep Reinforcement learning for robotics, including learning visuomotor skills and simulated locomotion. He has won various awards, including best paper awards at ICML, NIPS and ICRA, the Sloan Fellowship, the Air Force Office of Scientific Research Young Investigator Program (AFOSR-YIP) award, the Office of Naval Research Young Investigator Program (ONR-YIP) award, the DARPA Young Faculty Award (DARPA-YFA), the National Science Foundation Faculty Early Career Development Program Award (NSF-CAREER), the Presidential Early Career Award for Scientists and Engineers (PECASE), the CRA-E Undergraduate Research Faculty Mentoring Award, the MIT TR35, the IEEE Robotics and Automation Society (RAS) Early Career Award, and the Dick Volz Best U.S. Ph.D. Thesis in Robotics and Automation Award.

Track 1 - 10:20 am - 12:00 pm
Theory, Probabilistic Methods

Location: Hall A

On Structured Prediction Theory with Calibrated Convex Surrogate Losses

Anton Osokin, Francis Bach, Simon Lacoste-Julien

We provide novel theoretical insights on structured prediction in the context of efficient convex surrogate loss minimization with consistency guarantees. For any task loss, we construct a convex surrogate that can be optimized via stochastic gradient descent and we prove tight bounds on the so-called "calibration function" relating the excess surrogate risk to the actual risk. In contrast to prior related work, we carefully monitor the effect of the exponential number of classes in the learning guarantees as well as on the optimization complexity. As an interesting consequence, we formalize the intuition that some task losses make learning harder than others, and that the classical 0-1 loss is ill-suited for structured prediction.

REBAR: Low-variance, unbiased gradient estimates for discrete latent variable models

George Tucker, Andriy Mnih, Chris J Maddison, Dieterich Lawson, Jascha Sohl-Dickstein

Learning in models with discrete latent variables is challenging due to high variance gradient estimators. Generally, approaches have relied on control variates to reduce the variance of the REINFORCE estimator. Recent work {**jang2016categorical, maddison2016concrete**} has taken a different approach, introducing a continuous relaxation of discrete variables to produce low-variance, but biased, gradient estimates. In this work, we combine the two approaches through a novel control variate that produces low-variance, \emph{unbiased} gradient estimates. Then, we introduce a novel continuous relaxation and show that the tightness of the relaxation can be adapted online, removing it as a hyperparameter. We show state-of-the-art variance reduction on several benchmark generative modeling tasks, generally leading to faster convergence to a better final log likelihood.

Variance-based Regularization with Convex Objectives

Hong Namkoong, John C Duchi

We develop an approach to risk minimization and stochastic optimization that provides a convex surrogate for variance, allowing near-optimal and computationally efficient trading between approximation and estimation error. Our approach builds off of techniques for distributionally robust optimization and Owen's empirical likelihood, and we provide a number of finite-sample and asymptotic results characterizing the theoretical performance of the estimator. In particular, we show that our procedure comes with certificates of optimality, achieving (in some scenarios) faster rates of convergence than empirical risk minimization by virtue of automatically balancing bias and variance. We give corroborating empirical evidence showing that in practice, the estimator indeed trades between variance and absolute performance on a training sample, improving out-of-sample (test) performance over standard empirical risk minimization for a number of classification problems.

More powerful and flexible rules for online FDR control with memory and weights

Aaditya Ramdas, Fanny Yang, Martin Wainwright, Michael Jordan

In the online multiple testing problem, p-values corresponding to different null hypotheses are presented one by one, and the decision of whether to reject a hypothesis must be made immediately, after which the next p-value is presented. Alpha-investing algorithms to control the false discovery rate were first formulated by Foster and Stine and have since been generalized and applied to various settings, varying from quality-preserving databases for science to multiple A/B tests for internet commerce. This paper improves the class of generalized alpha-investing algorithms (GAI) in four ways : (a) we show how to uniformly improve the power of the entire class of GAI procedures under independence by awarding more alpha-wealth for each rejection, giving a near win-win resolution to a dilemma raised by Javanmard and Montanari, (b) we demonstrate how to incorporate prior weights to indicate domain knowledge of which hypotheses are likely to be null or non-null, (c) we allow for differing penalties for false discoveries to indicate that some hypotheses may be more meaningful/important than others, (d) we define a new quantity called the decaying memory false discovery rate, or memfdr that may be more meaningful for Applications with an explicit time component, using a discount factor to incrementally forget past decisions and alleviate some potential problems that we describe and name "piggybacking"' and "alpha-death"'. Our GAI++ algorithms incorporate all four generalizations (a, b, c, d) simulatenously, and reduce to more powerful variants of earlier algorithms when the weights and decay are all set to unity.

SPOTLIGHTS

- **Submultiplicative Glivenko-Cantelli and Uniform Convergence of Revenues**
  Noga Alon, Moshe Babaioff, Yannai A. Gonczarowski, Yishay Mansour, Shay Moran, Amir Yehudayoff
- **Fast Black-box Variational Inference through Stochastic Trust-Region Optimization**
  Jeff Regier, Michael Jordan, Jon McAuliffe
- **A Universal Analysis of Large-Scale Regularized Least Squares Solutions**
  Ashkan Panahi, Babak Hassibi
- **A Disentangled Recognition and Nonlinear Dynamics Model for Unsupervised Learning**
  Marco Fraccaro, Simon Kamronn, Ulrich Paquet, Ole Winther
- **Accelerated Stochastic Greedy Coordinate Descent by Soft Thresholding Projection onto Simplex**
  Chaobing Song, Shaobo Cui, Shu-Tao Xia, Yong Jiang
- **Early stopping for kernel boosting algorithms: A general analysis with localized complexities**
  Yuting Wei, Fanny Yang, Martin Wainwright
- **Spectrally-normalized margin bounds for neural networks Soft Thresholding Projection onto Simplex**
  Matus Telgarsky, Peter Bartlett, Dylan J Foster
- **The Scaling Limit of High-Dimensional Online Independent Component Analysis**
  Chuang Wang, Yue Lu

Track 2 - 10:20 am - 12:00 pm
Deep Learning

Location: Hall C

TernGrad: Ternary Gradients to Reduce Communication in Distributed Deep Learning

Wei Wen, Cong Xu, Feng Yan, Chunpeng Wu, Yandan Wang, Yiran Chen, Helen Li

High network communication cost for synchronizing gradients and parameters is the well-known bottleneck of distributed training. In this work, we propose TernGrad that uses ternary gradients to accelerate distributed deep learning in data parallelism. Our approach requires only three numerical levels {-1,0,1} which can aggressively reduce the communication time. We mathematically prove the convergence of TernGrad under the assumption of a bound on gradients. Guided by the bound, we propose layer-wise ternarizing and gradient clipping to improve its convergence. Our experiments show that applying TernGrad on AlexNet doesn't incur any accuracy loss and can even improve accuracy. The accuracy loss of GoogLeNet induced by TernGrad is less than 2% on average. Finally, a performance model is proposed to study the scalability of TernGrad. Experiments show significant speed gains for various deep neural networks.

Train longer, generalize better: closing the generalization gap in large batch training of neural networks

Elad Hoffer, Itay Hubara, Daniel Soudry

Background: Deep learning models are typically trained using stochastic gradient descent or one of its variants. These methods update the weights using their gradient, estimated from a small fraction of the training data. It has been observed that when using large batch sizes there is a persistent degradation in generalization performance - known as the "generalization gap" phenomena. Identifying the origin of this gap and closing it had remained an open problem. Contributions: We examine the initial high learning rate training phase. We find that the weight distance from its initialization grows logarithmicaly with the number of weight updates. We therefore propose a "random walk on random landscape" statistical model which is known to exhibit similar "ultra-slow" diffusion behavior. Following this hypothesis we conducted experiments to show empirically that the "generalization gap" stems from the relatively small number of updates rather than the batch size, and can be completely eliminated by adapting the training regime used. We further investigate different techniques to train models in the large-batch regime and present a novel algorithm named "Ghost Batch Normalization" which enables significant decrease in the generalization gap without increasing the number of updates. To validate our findings we conduct several additional experiments on MNIST, CIFAR-10, CIFAR-100 and ImageNet. Finally, we reassess common practices and beliefs concerning training of deep models and suggest they may not be optimal to achieve good generalization.

End-to-end Differentiable Proving

Tim Rocktäschel, Sebastian Riedel

We introduce deep neural networks for end-to-end differentiable theorem proving that operate on dense vector representations of symbols. These neural networks are recursively constructed by following the backward chaining algorithm as used in Prolog. Specifically, we replace symbolic unification with a differentiable computation on vector representations of symbols using a radial basis function kernel, thereby combining symbolic reasoning with learning subsymbolic vector representations. The resulting neural network can be trained to infer facts from a given incomplete knowledge base using gradient descent. By doing so, it learns to (i) place representations of similar symbols in close proximity in a vector space, (ii) make use of such similarities to prove facts, (iii) induce logical rules, and (iv) it can use provided and induced logical rules for complex multi-hop reasoning. On four benchmark knowledge bases we demonstrate that this architecture outperforms ComplEx, a state-of-the-art neural link prediction model, while at the same time inducing interpretable function-free first-order logic rules.

Gradient descent GAN optimization is locally stable

Vaishnavh Nagarajan, J. Zico Kolter

Despite their growing prominence, optimization in generative adversarial networks (GANs) is still a poorly-understood topic. In this paper, we analyze the gradient descent" form of GAN optimization (i.e., the natural setting where we simultaneously take small gradient steps in both generator and discriminator parameters). We show that even though GAN optimization does \emph{not} correspond to a convex-concave game even for simple parameterizations, under proper conditions, equilibrium points of this optimization procedure are still \emph{locally asymptotically stable} for the traditional GAN formulation. On the other hand, we show that the recently-proposed Wasserstein GAN can have non-convergent limit cycles near equilibrium. Motivated by this stability analysis, we propose an additional regularization term for gradient descent GAN updates, which \emph{is} able to guarantee local stability for both the WGAN and for the traditional GAN, and which also shows practical promise in speeding up convergence and addressing mode collapse.

SPOTLIGHTS

- **f-GANs in an Information Geometric Nutshell**
  Richard Nock, Zac Cranko, Aditya K Menon, Lizhen Qu, Robert C Williamson
- **Unsupervised Image-to-Image Translation Networks**
  Ming-Yu Liu, Thomas Breuel, Jan Kautz
- **The Numerics of GANs**
  Lars Mescheder, Sebastian Nowozin, Andreas Geiger
- **Dual Discriminator Generative Adversarial Nets**
  Tu Nguyen, Trung Le, Hung Vu, Dinh Phung
- **Bayesian GANs**
  Yunus Saatci, Andrew Wilson
- **Approximation and Convergence Properties of Generative Adversarial Learning**
  Shuang Liu, Olivier Bousquet, Kamalika Chaudhuri
- **Dualing GANs**
  Yujia Li, Alex Schwing, Kuan-Chieh Wang, Richard Zemel
- **Generalizing GANs: A Turing Perspective**
  Roderich Gross, Yue Gu, Wei Li, Melvin Gauci

Track 1 - 2:50 - 3:50 pm Reinforcement Learning, Deep Learning

Location: Hall A

ELF: An Extensive, Lightweight and Flexible Research Platform for Real-time Strategy Games

Yuandong Tian, Qucheng Gong, Wendy Shang, Yuxin Wu, Larry Zitnick

In this paper, we propose ELF, an Extensive, Lightweight and Flexible platform for fundamental Reinforcement learning research. Using ELF, we implement a highly customizable real-time strategy (RTS) engine with three game environments (Mini-RTS, Capture the Flag and Tower Defense). Mini-RTS, as a miniature version of StarCraft, captures key game dynamics and runs at 165K frame- per-second (FPS) on a Macbook Pro notebook. When coupled with modern Reinforcement learning methods, the system can train a full-game bot against built-in AIs end-to-end in one day with 6 CPUs and 1 GPU. In addition, our platform is flexible in terms of environment-agent communication topologies, choices of RL methods, changes in game parameters, and can host existing C/C++-based game environments like ALE. Using ELF, we thoroughly explore training parameters and show that a network with Leaky ReLU and Batch Normalization coupled with long-horizon training and progressive curriculum beats the rule-based built-in AI more than 70% of the time in the full game of Mini-RTS. Strong performance is also achieved on the other two games. In game replays, we show our agents learn interesting strategies. ELF, along with its RL platform, will be open-sourced.

Imagination-Augmented Agents for Deep Reinforcement Learning

Seb Racanière, David Reichert, Theophane Weber, Oriol Vinyals, Daan Wierstra, Lars Buesing, Peter Battaglia, Razvan Pascanu, Yujia Li, Nicolas Heess, Arthur Guez, Danilo Jimenez Rezende, Adrià Puigdomènech Badia, David Silver

We introduce Imagination-Augmented Agents (I2As), a novel architecture for deep Reinforcement learning combining model-free and model-based aspects. In contrast to most existing model-based Reinforcement learning and planning methods, which prescribe how a model should be used to arrive at a policy, I2As learn to interpret predictions from a trained environment model to construct implicit plans in arbitrary ways, by using the predictions as additional context in deep policy networks. I2As show improved data efficiency, performance, and robustness to model misspecification compared to several strong baselines.

SPOTLIGHTS

- **Dual Path Networks**
  Yunpeng Chen, Jianan Li, Huaxin Xiao, Xiaojie Jin, Shuicheng Yan, Jiashi Feng
- **A simple neural network module for relational reasoning**
  Adam Santoro, David Raposo, David Barrett, Mateusz Malinowski, Razvan Pascanu, Peter Battaglia, Timothy Lillicrap
- **Second-order Optimization in Deep Reinforcement Learning using Kronecker-factored Approximation**
  Yuhuai Wu, Elman Mansimov, Roger Grosse, Shun Liao, Jimmy Ba
- **Attention is All you Need**
  Ashish Vaswani, Noam Shazeer, Niki Parmar, Llion Jones, Jakob Uszkoreit, Aidan N Gomez, Łukasz Kaiser
- **Learning Combinatorial Optimization Algorithms over Graphs**
  Elias Khalil, Hanjun Dai, Yuyu Zhang, Bistra Dilkina, Le Song
- **Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles**
  Balaji Lakshminarayanan, Alexander Pritzel, Charles Blundell

Track 2 - 2:50 - 3:50 pm
Optimization

Location: Hall C

The Marginal Value of Adaptive Gradient Methods in Machine Learning

Ashia C Wilson, Becca Roelofs, Mitchell Stern, Nati Srebro, Benjamin Recht

Adaptive optimization methods, which perform local optimization with a metric constructed from the history of iterates, are becoming increasingly popular for training deep neural networks. Examples include AdaGrad, RMSProp, and Adam. We show that for simple over-parameterized problems, adaptive methods often find drastically different solutions than vanilla stochastic gradient descent (SGD). We construct an illustrative binary classification problem where the data is linearly separable, SGD achieves zero test error, and AdaGrad and Adam attain test errors arbitrarily close to 1/2. We additionally study the empirical generalization capability of adaptive methods on several state-of-the-art deep learning models. We observe that the solutions found by adaptive methods generalize worse (often significantly worse) than SGD, even when these solutions have better training performance. These results suggest that practitioners should reconsider the use of adaptive methods to train neural networks.

Can Decentralized Algorithms Outperform Centralized Algorithms? A Case Study for Decentralized Parallel Stochastic Gradient Descent

Xiangru Lian, Ce Zhang, Huan Zhang, Cho-Jui Hsieh, Wei Zhang, Ji Liu

Most distributed machine learning systems nowadays, including TensorFlow and CNTK, are built in a centralized fashion. One bottleneck of centralized algorithms lies on high communication cost on the central node. Motivated by this, we ask, can decentralized algorithms be faster than its centralized counterpart? Although decentralized PSGD (D-PSGD) algorithms have been studied by the control community, existing analysis and theory do not show any advantage over centralized PSGD (C-PSGD) algorithms, simply assuming the application scenario where only the decentralized network is available. In this paper, we study a D-PSGD algorithm and provide the first theoretical analysis that indicates a regime in which decentralized algorithms might outperform centralized algorithms for distributed stochastic gradient descent. This is because D-PSGD has comparable total computational complexities to C-PSGD but requires much less communication cost on the busiest node. We further conduct an empirical study to validate our theoretical analysis across multiple frameworks (CNTK and Torch), different network configurations, and computation platforms up to 112 GPUs. On network configurations with low bandwidth or high latency, D-PSGD can be up to one order of magnitude faster than its well-optimized centralized counterparts.

SPOTLIGHTS

- **Breaking the Nonsmooth Barrier: A Scalable Parallel Method for Composite Optimization**
  Fabian Pedregosa, Rémi Leblond, Simon Lacoste-Julien
- **Stochastic Optimization with Variance Reduction for Infinite Datasets with Finite Sum Structure**
  Alberto Bietti, Julien Mairal
- **Process-constrained batch Bayesian optimisation**
  Pratibha Vellanki, Santu Rana, Sunil Gupta, David Rubin, Alessandra Sutti, Thomas Dorin, Murray Height, Paul Sanders, Svetha Venkatesh
- **Safe Adaptive Importance Sampling**
  Sebastian Stich, Anant Raj, Martin Jaggi
- **Beyond Worst-case: A Probabilistic Analysis of Affine Policies in Dynamic Optimization**
  Omar El Housni, Vineet Goyal
- **Straggler Mitigation in Distributed Optimization Through Data Encoding**
  Can Karakus, Yifan Sun, Suhas Diggavi, Wotao Yin

Track 1 - 4:20 - 6:00 pm
Reinforcement Learning, Algorithms, Applications

Location: Hall A

Off-policy evaluation for slate recommendation

Adith Swaminathan, Akshay Krishnamurthy, Alekh Agarwal, Miro Dudik, John Langford, Damien Jose, Imed Zitouni

This paper studies the evaluation of policies that recommend an ordered set of items (e.g., a ranking) based on some context---a common scenario in web search, ads, and recommendation. We build on techniques from combinatorial bandits to introduce a new practical estimator. A thorough empirical evaluation on real-world data reveals that our estimator is accurate in a variety of settings, including as a subroutine in a learning-to-rank task, where it achieves competitive performance. We derive conditions under which our estimator is unbiased---these conditions are weaker than prior heuristics for slate evaluation---and experimentally demonstrate a smaller bias than parametric approaches, even when these conditions are violated. Finally, our theory and experiments also show exponential savings in the amount of required data compared with general unbiased estimators.

Robust and Efficient Transfer Learning with Hidden Parameter Markov Decision Processes

Sam Daulton, Taylor Killian, Finale Doshi-Velez, George Konidaris

We introduce a new formulation of the Hidden Parameter Markov Decision Process (HiP-MDP), a framework for modeling families of related tasks using low-dimensional latent embeddings. We replace the original Gaussian Process-based model with a Bayesian Neural Network. Our new framework correctly models the joint uncertainty in the latent weights and the state space and has more scalable inference, thus expanding the scope the HiP-MDP to Applications with higher dimensions and more complex dynamics.

Inverse Reward Design

Dylan Hadfield-Menell, Smitha Milli, Stuart J Russell, Pieter Abbeel, Anca Dragan

Autonomous agents optimize the reward function we give them. What they don't know is how hard it is for us to design a reward function that actually captures what we want. When designing the reward, we might think of some specific scenarios (driving on clean roads), and make sure that the reward will lead to the right behavior in ***those*** scenarios. Inevitably, agents encounter ***new*** scenarios (snowy roads), and optimizing the reward can lead to undesired behavior (driving too fast). Our insight in this work is that reward functions are merely ***observations*** about what the designer ***actually*** wants, and that they should be interpreted in the context in which they were designed. We introduce ***Inverse Reward Design*** (IRD) as the problem of inferring the true reward based on the designed reward and the training MDP. We introduce approximate methods for solving IRD problems, and use their solution to plan risk-averse behavior in test MDPs. Empirical results suggest that this approach takes a step towards alleviating negative side effects and preventing reward hacking.

SPOTLIGHTS

- **Dynamic Safe Interruptibility for Decentralized Multi-Agent Reinforcement Learning**
  El Mahdi El Mhamdi, Rachid Guerraoui, Hadrien Hendrikx, Alexandre Maurer
- **Unifying PAC and Regret: Uniform PAC Bounds for Episodic Reinforcement Learning**
  Christoph Dann, Tor Lattimore, Emma Brunskill
- **Repeated Inverse Reinforcement Learning**
  Kareem Amin, Nan Jiang, Satinder Singh
- **Learning multiple visual domains with residual adapters**
  Sylvestre-Alvise Rebuffi, Hakan Bilen, Andrea Vedaldi
- **Natural value approximators: learning when to trust past estimates**
  Tom Schaul, Zhongwen Xu, Joseph Modayil, Hado van Hasselt, Andre Barreto, David Silver
- **EX2: Exploration with Exemplar Models for Deep Reinforcement Learning**
  Justin Fu, John Co-Reyes, Sergey Levine
- **Regret Minimization in MDPs with Options without Prior Knowledge**
  Ronan Fruit, Matteo Pirotta, Alessandro Lazaric, Emma Brunskill
- **Successor Features for Transfer in Reinforcement Learning**
  Andre Barreto, Will Dabney, Remi Munos, Jonathan Hunt, Tom Schaul, David Silver, Hado van Hasselt
- **Overcoming Catastrophic Forgetting by Incremental Moment Matching**
  Sang-Woo Lee, Jin-Hwa Kim, Jaehyun Jun, Jung-Woo Ha, Byoung-Tak Zhang
- **Fair Clustering Through Fairlets**
  Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, Sergei Vassilvitskii
- **Fitting Low-Rank Tensors in Constant Time**
  Kohei Hayashi, Yuichi Yoshida

Track 2 - 4:20 - 6:00 pm Probabilistic Methods, Applications

Location: Hall C

What-If Reasoning using Counterfactual Gaussian Processes

Peter Schulam, Suchi Saria

Answering "What if?" questions is important in many domains. For example, would a patient's disease progression slow down if I were to give them a dose of drug A? Ideally, we answer our question using an experiment, but this is not always possible (e.g., it may be unethical). As an alternative, we can use non-experimental data to learn models that make counterfactual predictions of what we would observe had we run an experiment. In this paper, we propose the counterfactual GP, a counterfactual model of continuous-time trajectories (time series) under sequences of actions taken in continuous-time. We develop our model within the potential outcomes framework of Neyman and Rubin. The counterfactual GP is trained using a joint maximum likelihood objective that adjusts for dependencies between observed actions and outcomes in the training data. We report two sets of experimental results using the counterfactual GP. The first shows that it can be used to learn the natural progression (i.e. untreated progression) of biomarker trajectories from observational data. In the second, we show how the CGP can be used for medical decision support by learning counterfactual models of renal health under different types of dialysis.

Convolutional Gaussian Processes

Mark van der Wilk, Carl Edward Rasmussen, James Hensman

We introduce a practical way of introducing convolutional structure into Gaussian processes, which makes them better suited to high-dimensional inputs like images than existing kernels. The main contribution of our work is the construction of an inter-domain inducing point approximation that is well-tailored to the convolutional kernel. This allows us to gain the generalisation benefit of a convolutional kernel, together with fast but accurate posterior inference. We investigate several variations of the convolutional kernel, and apply it to MNIST and CIFAR-10 that have been known to be challenging for Gaussian processes. We also show how the marginal likelihood can be used to find an optimal weighting between convolutional and RBF kernels to further improve performance. We hope this illustration of the usefulness of a marginal likelihood will help to automate discovering architectures in larger models.

Counterfactual Fairness

Matt Kusner, Joshua Loftus, Chris Russell, Ricardo Silva

Machine learning can impact people with legal or ethical consequences when it is used to automate decisions in areas such as insurance, lending, hiring, and predictive policing. In many of these scenarios, previous decisions have been made that are unfairly biased against certain subpopulations, for example those of a particular race, gender, or sexual orientation. Since this past data may be biased, machine learning predictors must account for this to avoid perpetuating or creating discriminatory practices. In this paper, we develop a framework for modeling fairness using tools from causal inference. Our definition of counterfactual fairness captures the intuition that a decision is fair towards an individual if it the same in (a) the actual world and (b) a counterfactual world where the individual belonged to a different demographic group. We demonstrate our framework on a real-world problem of fair prediction of success in law school.

SPOTLIGHTS

- **An Empirical Bayes Approach to Optimizing Machine Learning Algorithms**
  James McInerney
- **PASS-GLM: polynomial approximate sufficient statistics for scalable Bayesian GLM inference**
  Jonathan Huggins, Ryan Adams, Tamara Broderick
- **Multiresolution Kernel Approximation for Gaussian Process Regression**
  Yi Ding, Risi Kondor, Jon Eskreis-Winkler
- **Multi-Information Source Optimization**
  Matthias Poloczek, Jialei Wang, Peter Frazier
- **Doubly Stochastic Variational Inference for Deep Gaussian Processes**
  Hugh Salimbeni, Marc Deisenroth
- **Permutation-based Causal Inference Algorithms with Interventions**
  Yuhao Wang, Liam Solus, Karren Yang, Caroline Uhler
- **Gradients of Generative Models for Improved Discriminative Analysis of Tandem Mass Spectra**
  John T Halloran, David M Rocke
- **Style Transfer from Non-parallel Text by Cross-Alignment**
  Tianxiao Shen, Tao Lei, Regina Barzilay, Tommi Jaakkola
- **Premise Selection for Theorem Proving by Deep Graph Embedding**
  Mingzhe Wang, Yihe Tang, Jian Wang, Jia Deng
- **Deep Multi-task Gaussian Processes for Survival Analysis with Competing Risks**
  Ahmed M. Alaa, Mihaela van der Schaar
- **Unsupervised Learning of Disentangled Representations from Video**
  Emily Denton, vighnesh Birodkar

Wednesday Poster Session

#1 Deep Reinforcement Learning from Human Preferences
Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, Dario Amodei

#2 Multi-Modal Imitation Learning from Unstructured Demos using Generative Adversarial Nets
Karol Hausman, Yevgen Chebotar, Stefan Schaal, Gaurav Sukhatme, Joseph J Lim

#3 EX2: Exploration with Exemplar Models for Deep Reinforcement Learning
Justin Fu, John Co-Reyes, Sergey Levine

#4 #Exploration: A Study of Count-Based Exploration for Deep Reinforcement Learning
Haoran Tang, Pieter Abbeel, Davis Foote, Yan Duan, OpenAI Xi Chen, Rein Houthooft, Adam Stooke, Filip DeTurck

#5 Thinking Fast and Slow with Deep Learning and Tree Search
Thomas Anthony, Zheng Tian, David Barber

#6 Natural value approximators: learning when to trust past estimates
Tom Schaul, Zhongwen Xu, Joseph Modayil, Hado van Hasselt, Andre Barreto, David Silver

#7 Active Exploration for Learning Symbolic Representations
Garrett Andersen, George Konidaris

#8 State Aware Imitation Learning
Yannick Schroecker, Charles L Isbell

#9 Successor Features for Transfer in Reinforcement Learning
Andre Barreto, Will Dabney, Remi Munos, Jonathan Hunt, Tom Schaul, David Silver, Hado van Hasselt

#10 Bridging the Gap Between Value and Policy Based Reinforcement Learning
Ofir Nachum, Mohammad Norouzi, Kelvin Xu, Dale Schuurmans

#11 Using Options and Covariance Testing for Long Horizon Off-Policy Policy Evaluation
Daniel Guo, Philip S. Thomas, Emma Brunskill

#12 Compatible Reward Inverse Reinforcement Learning
Alberto Maria Metelli, Matteo Pirotta, Marcello Restelli

#13 Adaptive Batch Size for Safe Policy Gradients
Matteo Papini, Matteo Pirotta, Marcello Restelli

#14 Regret Minimization in MDPs with Options without Prior Knowledge
Ronan Fruit, Matteo Pirotta, Alessandro Lazaric, Emma Brunskill

#15 Is the Bellman residual a bad proxy?
Matthieu Geist, Bilal Piot, Olivier Pietquin

#16 Learning Unknown Markov Decision Processes: A Thompson Sampling Approach
Yi Ouyang, Mukul Gagrani, Ashutosh Nayyar, Rahul Jain

#17 Online Reinforcement Learning in Stochastic Games
Chen-Yu Wei, Yi-Te Hong, Chi-Jen Lu

#18 Reinforcement Learning under Model Mismatch
Aurko Roy, Huan Xu, Sebastian Pokutta

#19 Zap Q-Learning
Adithya M Devraj, Sean Meyn

#20 Ensemble Sampling
Xiuyuan Lu, Benjamin Van Roy

#21 Action Centered Contextual Bandits
Kristjan Greenewald, Ambuj Tewari, Susan Murphy, Predag Klasnja

#22 Conservative Contextual Linear Bandits
Abbas Kazerouni, Mohammad Ghavamzadeh, Yasin Abbasi, Benjamin Van Roy

#23 Rotting Bandits
Nir Levine, Koby Crammer, Shie Mannor

#24 Identifying Outlier Arms in Multi-Armed Bandit
Honglei Zhuang, Chi Wang, Yifan Wang

#25 Multi-Task Learning for Contextual Bandits
Aniket Anand Deshmukh, Urun Dogan, Clay Scott

#26 Boltzmann Exploration Done Right
Nicolò Cesa-Bianchi, Claudio Gentile, Gergely Neu, Gabor Lugosi

#27 Improving the Expected Improvement Algorithm
Chao Qin, Diego Klabjan, Daniel Russo

#28 A KL-LUCB algorithm for Large-Scale Crowdsourcing
Ervin Tanczos, Robert Nowak, Bob Mankoff

#29 Scalable Generalized Linear Bandits: Online Computation and Hashing
Kwang-Sung Jun, Aniruddha Bhargava, Robert Nowak, Rebecca Willett

#30 Bandits Dueling on Partially Ordered Sets
Julien Audiffren, Liva Ralaivola

#31 Position-based Multiple-play Multi-armed Bandit Problem with Unknown Position Bias
Junpei Komiyama, Junya Honda, Akiko Takeda

#32 Online Influence Maximization under Independent Cascade Model with Semi-Bandit Feedback
Zheng Wen, Branislav Kveton, Michal Valko, Sharan Vaswani

#33 A Scale Free Algorithm for Stochastic Bandits with Bounded Kurtosis
Tor Lattimore

#34 Adaptive Active Hypothesis Testing under Limited Information
Fabio Cecchi, Nidhi Hegde

#35 Near-Optimal Edge Evaluation in Explicit Generalized Binomial Graphs
Sanjiban Choudhury, Shervin Javdani, Siddhartha Srinivasa, Sebastian Scherer

#36 Robust and Efficient Transfer Learning with Hidden Parameter Markov Decision Processes
Sam Daulton, Taylor Killian, Finale Doshi-Velez, George Konidaris

#37 Overcoming Catastrophic Forgetting by Incremental Moment Matching
Sang-Woo Lee, Jin-Hwa Kim, Jaehyun Jun, Jung-Woo Ha, Byoung-Tak Zhang

#38 Hypothesis Transfer Learning via Transformation Functions
Simon Du, Jayanth Koushik, Aarti Singh, Barnabas Poczos

#39 Learning multiple visual domains with residual adapters
Sylvestre-Alvise Rebuffi, Hakan Bilen, Andrea Vedaldi

#40 Self-supervised Learning of Motion Capture
Hsiao-Yu Tung, Hsiao-Wei Tung, Ersin Yumer, Katerina Fragkiadaki

#41 Information Theoretic Properties of Markov Random Fields, and their Algorithmic Applications
Linus Hamilton, Frederic Koehler, Ankur Moitra

#42 Maximizing Subset Accuracy with Recurrent Neural Networks in Multi-label Classification
Jinseok Nam, Eneldo Loza Mencía, Hyunwoo J Kim, Johannes Fürnkranz

#43 Local Aggregative Games
Vikas Garg, Tommi Jaakkola

#44 An Empirical Bayes Approach to Optimizing Machine Learning Algorithms
James McInerney

#45 Learning Chordal Markov Networks via Branch and Bound
Kari Rantanen, Antti Hyttinen, Matti Järvisalo

#46 Optimal Sample Complexity of M-wise Data for Top-K Ranking
Minje Jang, Sunghyun Kim, Changho Suh, Sewoong Oh

#47 Translation Synchronization via Truncated Least Squares
Xiangru Huang, Zhenxiao Liang, Chandrajit Bajaj, Qixing Huang

#48 From Bayesian Sparsity to Gated Recurrent Nets
Hao He, Bo Xin, David Wipf

#49 Online Learning for Multivariate Hawkes Processes
Yingxiang Yang, Jalal Etesami, Niao He, Negar Kiyavash

#50 Efficient Second-Order Online Kernel Learning with Adaptive Embedding
Daniele Calandriello, Michal Valko, Alessandro Lazaric

#51 Online to Offline Conversions and Adaptive Minibatch Sizes
Kfir Levy

#52 Nonparametric Online Regression while Learning the Metric
Ilja Kuzborskij, Nicolò Cesa-Bianchi

#53 Stochastic and Adversarial Online Learning without Hyperparameters
Ashok Cutkosky, Kwabena A Boahen

#54 Affine-Invariant Online Optimization
Tomer Koren, Roi Livni

#55 Online Convex Optimization with Stochastic Constraints
Hao Yu, Michael Neely, Xiaohan Wei

#56 Online Learning with a Hint
Ofer Dekel, arthur flajolet, Nika Haghtalab, Patrick Jaillet

#57 Efficient Online Linear Optimization with Approximation Algorithms
Dan Garber

#58 Random Permutation Online Isotonic Regression
Wojciech Kotlowski, Wouter Koolen, Alan Malek

#59 Minimax Optimal Players for the Finite-Time 3-Expert Prediction Problem
Yasin Abbasi, Peter Bartlett, Victor Gabillon

#60 Online Learning of Optimal Bidding Strategy in Repeated Multi-Commodity Auctions
Sevi Baltaoglu, Lang Tong, Qing Zhao

#61 Online Prediction with Selfish Experts
Tim Roughgarden, Okke Schrijvers

#62 Real-Time Bidding with Side Information
arthur flajolet, Patrick Jaillet

#63 Improving Regret Bounds for Combinatorial Semi-Bandits with Probabilistically Triggered Arms and Its Applications
Qinshi Wang, Wei Chen

#64 A General Framework for Robust Interactive Learning
Ehsan Emamjomeh-Zadeh, David Kempe

#65 Practical Locally Private Heavy Hitters
kobbi nissim, Raef Bassily, Uri Stemmer, Abhradeep Guha Thakurta

#66 Deanonymization in the Bitcoin P2P Network
Giulia Fanti, Pramod Viswanath

#67 Accuracy First: Selecting a Differential Privacy Level for Accuracy Constrained ERM
Steven Wu, Bo Waggoner, Seth Neel, Aaron Roth, Katrina Ligett

#68 Renyi Differential Privacy Mechanisms for Posterior Sampling
Joseph Geumlek, Shuang Song, Kamalika Chaudhuri

#69 Collecting Telemetry Data Privately
Bolin Ding, Janardhan Kulkarni, Sergey Yekhanin

#70 Generating steganographic images via adversarial training
Jamie Hayes, George Danezis

#71 Fitting Low-Rank Tensors in Constant Time
Kohei Hayashi, Yuichi Yoshida

#72 Thy Friend is My Friend: Iterative Collaborative Filtering for Sparse Matrix Estimation
Christian Borgs, Jennifer Chayes, Christina Lee, Devavrat Shah

#73 Fair Clustering Through Fairlets
Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, Sergei Vassilvitskii

#74 On Fairness and Calibration
Geoff Pleiss, Manish Raghavan, Felix Wu, Jon Kleinberg, Kilian Weinberger

#75 Avoiding Discrimination through Causal Reasoning
Niki Kilbertus, Mateo Rojas Carulla, Giambattista Parascandolo, Moritz Hardt, Dominik Janzing, Bernhard Schölkopf

#76 Optimized Pre-Processing for Discrimination Prevention
Flavio Calmon, Dennis Wei, Karthi Natesan Ramamurthy, Bhanu Vinzamuri, Kush R Varshney

#77 Recycling for Fairness: Learning with Conditional Distribution Matching Constraints
Novi Quadrianto, Viktoriia Sharmanska

#78 From Parity to Preference: Learning with Cost-effective Notions of Fairness
Bilal Zafar, Isabel Valera, Manuel Rodriguez, Krishna Gummadi, Adrian Weller

#79 Beyond Parity: Fairness Objectives for Collaborative Filtering
Sirui Yao, Bert Huang

#80 Multi-view Matrix Factorization for Linear Dynamical System Estimation
Mahdi Karami, Martha White, Dale Schuurmans, Csaba Szepesvari

#81 Random Projection Filter Bank for Time Series Data
Amir-massoud Farahmand, Sepideh Pourazarm, Daniel Nikovski

#82 A Dirichlet Mixture Model of Hawkes Processes for Event Sequence Clustering
Hongteng Xu, Hongyuan Zha

#83 Predicting User Activity Level In Point Process Models With Mass Transport Equation
Yichen Wang, Xiaojing Ye, Hongyuan Zha, Le Song

#84 Off-policy evaluation for slate recommendation
Adith Swaminathan, Akshay Krishnamurthy, Alekh Agarwal, Miro Dudik, John Langford, Damien Jose, Imed Zitouni

#85 Expectation Propagation with Stochastic Kinetic Model in Complex Interaction Systems
Wen Dong, Le Fang, Fan Yang, Tong Guan, Chunming Qiao

#86 A multi-agent Reinforcement learning model of common-pool resource appropriation
Julien Pérolat, Joel Leibo, Vinicius Zambaldi, Charles Beattie, Karl Tuyls, Thore Graepel

#87 Balancing information exposure in social networks
Kiran Garimella, Aris Gionis, Nikos Parotsidis, Nikolaj Tatti

#88 Scalable Demand-Aware Recommendation
Jinfeng Yi, Cho-Jui Hsieh, Kush R Varshney, Lijun Zhang, Yao Li

#89 A Greedy Approach for Budgeted Maximum Inner Product Search
Hsiang-Fu Yu, Cho-Jui Hsieh, Qi Lei, Inderjit S Dhillon

#90 DPSCREEN: Dynamic Personalized Screening
Kartik Ahuja, William Zame, Mihaela van der Schaar

#91 Deep Multi-task Gaussian Processes for Survival Analysis with Competing Risks
Ahmed M. Alaa, Mihaela van der Schaar

#92 Premise Selection for Theorem Proving by Deep Graph Embedding
Mingzhe Wang, Yihe Tang, Jian Wang, Jia Deng

#93 Gradients of Generative Models for Improved Discriminative Analysis of Tandem Mass Spectra
John T Halloran, David M Rocke

#94 Style Transfer from Non-parallel Text by Cross-Alignment
Tianxiao Shen, Tao Lei, Regina Barzilay, Tommi Jaakkola

#95 Emergence of Language with Multi-agent Games: Learning to Communicate with Sequences of Symbols
Serhii Havrylov, Ivan Titov

#96 ELF: An Extensive, Lightweight and Flexible Research Platform for Real-time Strategy Games
Yuandong Tian, Qucheng Gong, Wendy Shang, Yuxin Wu, Larry Zitnick

#97 ExtremeWeather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events
Evan Racah, Christopher Beckham, Tegan Maharaj, Mr. Prabhat, Chris Pal

#98 Approximation and Convergence Properties of Generative Adversarial Learning
Shuang Liu, Olivier Bousquet, Kamalika Chaudhuri

#99 Gradient descent GAN optimization is locally stable
Vaishnavh Nagarajan, J. Zico Kolter

#100 f-GANs in an Information Geometric Nutshell
Richard Nock, Zac Cranko, Aditya K Menon, Lizhen Qu, Robert C Williamson

#101 The Numerics of GANs
Lars Mescheder, Sebastian Nowozin, Andreas Geiger

#102 Generalizing GANs: A Turing Perspective
Roderich Gross, Yue Gu, Wei Li, Melvin Gauci

#103 Dualing GANs
Yujia Li, Alex Schwing, Kuan-Chieh Wang, Richard Zemel

#104 Fisher GAN
Youssef Mroueh, Tom Sercu

#105 Learning to Pivot with Adversarial Networks
Gilles Louppe, Michael Kagan, Kyle Cranmer

#106 Improved Training of Wasserstein GANs
Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, Aaron C Courville

#107 MMD GAN: Towards Deeper Understanding of Moment Matching Network
Chun-Liang Li, Wei-Cheng Chang, Yu Cheng, Yiming Yang, Barnabas Poczos

#108 Two Time-Scale Update Rule for Generative Adversarial Nets
Hubert Ramsauer, Martin Heusel, Sepp Hochreiter, Bernhard Nessler, Thomas Unterthiner

#109 VEEGAN: Reducing Mode Collapse in GANs using Implicit Variational Learning
Akash Srivastava, Lazar Valkoz, Chris Russell, Michael U. Gutmann, Charles Sutton

#110 Improved Semi-supervised Learning with GANs using Manifold Invariances
Abhishek Kumar, Prasanna Sattigeri, Tom Fletcher

#111 Good Semi-supervised Learning That Requires a Bad GAN
Zihang Dai, Zhilin Yang, Fan Yang, William W Cohen, Ruslan Salakhutdinov

#112 Bayesian GANs
Yunus Saatci, Andrew Wilson

#113 Dual Discriminator Generative Adversarial Nets
Tu Nguyen, Trung Le, Hung Vu, Dinh Phung

#114 Towards Understanding Adversarial Learning for Joint Distribution Matching
Chunyuan Li, Hao Liu, Ricardo Henao, Liqun Chen, Yuchen Pu, Changyou Chen, Lawrence Carin

#115 Triple Generative Adversarial Nets
Chongxuan LI, Kun Xu, Jun Zhu, Bo Zhang

#116 Triangle Generative Adversarial Networks
Zhe Gan, Liqun Chen, Weiyao Wang, Yuchen Pu, Yizhe Zhang, Lawrence Carin

#117 Structured Generative Adversarial Networks
Hao Zhang, Zhijie Deng, Xiaodan Liang, Jun Zhu, Eric P Xing

#118 PixelGAN Autoencoders
Alireza Makhzani, Brendan J Frey

#119 Learning to Compose Domain-Specific Transformations for Data Augmentation
Alexander Ratner, Henry Ehrenberg, Zeshan Hussain, Jared Dunnmon, Chris Ré

#120 Unsupervised Image-to-Image Translation Networks
Ming-Yu Liu, Thomas Breuel, Jan Kautz

#121 Adversarial Invariant Feature Learning
Qizhe Xie, Zihang Dai, Yulun Du, Eduard Hovy, Graham Neubig

#122 Adversarial Ranking for Language Generation
Dianqi Li, Kevin Lin, Xiaodong He, Ming-ting Sun, Zhengyou Zhang

#123 Efficient Computation of Moments in Sum-Product Networks
Han Zhao, Geoffrey Gordon

#124 Attention is All you Need
Ashish Vaswani, Noam Shazeer, Niki Parmar, Llion Jones, Jakob Uszkoreit, Aidan N Gomez, Łukasz Kaiser

#125 Masked Autoregressive Flow for Density Estimation
George Papamakarios, Iain Murray, Theo Pavlakou

#126 Variational Walkback: Learning a Transition Operator as a Stochastic Recurrent Net
Anirudh Goyal ALIAS PARTH GOYAL, Rosemary Ke, Surya Ganguli, Yoshua Bengio

#127 TernGrad: Ternary Gradients to Reduce Communication in Distributed Deep Learning
Wei Wen, Cong Xu, Feng Yan, Chunpeng Wu, Yandan Wang, Yiran Chen, Helen Li

#128 End-to-end Differentiable Proving
Tim Rocktäschel, Sebastian Riedel

#129 A simple neural network module for relational reasoning
Adam Santoro, David Raposo, David Barrett, Mateusz Malinowski, Razvan Pascanu, Peter Battaglia, Timothy Lillicrap

#130 Dual Path Networks
Yunpeng Chen, Jianan Li, Huaxin Xiao, Xiaojie Jin, Shuicheng Yan, Jiashi Feng

#131 Spherical convolutions and their application in molecular modelling
Wouter Boomsma, Jes Frellsen

#132 Deep Sets
Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Ruslan Salakhutdinov, Alex Smola

#133 Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles
Balaji Lakshminarayanan, Alexander Pritzel, Charles Blundell

#134 Self-Normalizing Neural Networks
Günter Klambauer, Thomas Unterthiner, Andreas Mayr, Sepp Hochreiter

#135 Batch Renormalization: Towards Reducing Minibatch Dependence in Batch-Normalized Models
Sergey Ioffe

#136 Train longer, generalize better: closing the generalization gap in large batch training of neural networks
Elad Hoffer, Itay Hubara, Daniel Soudry

#137 Nonlinear random matrix theory for deep learning
Jeffrey Pennington, Pratik Worah

#138 DisTraL: Robust multitask Reinforcement learning
Yee Teh, Victor Bapst, Razvan Pascanu, Nicolas Heess, John Quan, James Kirkpatrick, Wojciech M. Czarnecki, Raia Hadsell

#139 Imagination-Augmented Agents for Deep Reinforcement Learning
Seb Racanière, David Reichert, Theophane Weber, Oriol Vinyals, Daan Wierstra, Lars Buesing, Peter Battaglia, Razvan Pascanu, Yujia Li, Nicolas Heess, Arthur Guez, Danilo Jimenez Rezende, Adrià Puigdomènech Badia, David Silver

#140 Second-order Optimization in Deep Reinforcement Learning using Kronecker-factored Approximation
Yuhuai Wu, Elman Mansimov, Roger Grosse, Shun Liao, Jimmy Ba

#141 Learning Combinatorial Optimization Algorithms over Graphs
Elias Khalil, Hanjun Dai, Yuyu Zhang, Bistra Dilkina, Le Song

#142 Targeting EEG/LFP Synchrony with Neural Nets
Yitong Li, David E Carlson, Lawrence Carin

#143 Toward Goal-Driven Neural Network Models for the Rodent Whisker-Trigeminal System
Chengxu Zhuang, Jonas Kubilius, Mitra JZ Hartmann, Daniel Yamins

#144 Fast amortized inference of neural activity from calcium imaging data with variational autoencoders
Artur Speiser, Jinyao Yan, Evan Archer, Lars Buesing, Srini C Turaga, Jakob H Macke

#145 Scene Physics Acquisition via Visual De-animation
Jiajun Wu, Erika Lu, Pushmeet Kohli, Bill Freeman, Josh Tenenbaum

#146 Shape and Material from Sound
zhoutong zhang, Qiujia Li, Zhengjia Huang, Jiajun Wu, Josh Tenenbaum, Bill Freeman

#147 Deep Networks for Decoding Natural Images from Retinal Signals
Nikhil Parthasarathy, Eleanor Batty, William Falcon, Thomas Rutten, Mohit Rajpal, chichilnisky Chichilnisky, Liam Paninski

#148 Quantifying how much sensory information in a neural code is relevant for behavior
Giuseppe Pica, Eugenio Piasini, Houman Safaai, Caroline Runyan, Christopher Harvey, Mathew Diamond, Christoph Kayser, Tommaso Fellin, Stefano Panzeri

#149 Model-based Bayesian inference of neural activity and connectivity from all-optical interrogation of a neural circuit
Laurence Aitchison, Lloyd Russell, Adam Packer, Jinyao Yan, Philippe Castonguay, Michael Hausser, Srini C Turaga

#150 Deep Hyperalignment
Muhammad Yousefnezhad, Daoqiang Zhang

#151 Tensor encoding and decomposition of brain connectomes with application to tractography evaluation
Cesar F Caiafa, Olaf Sporns, Andrew Saykin, Franco Pestilli

#152 Online Dynamic Programming
Holakou Rahmanian, Manfred Warmuth

#153 Unsupervised Learning of Disentangled Representations from Video
Emily Denton, vighnesh Birodkar

#154 Interactive Submodular Bandit
Lin Chen, Andreas Krause, Amin Karbasi

#155 Streaming Robust Submodular Maximization:A Partitioned Thresholding Approach
Slobodan Mitrovic, Ilija Bogunovic, Ashkan Norouzi-Fard, Jakub M Tarnawski, Volkan Cevher

#156 Minimizing a Submodular Function from Samples
Eric Balkanski, Yaron Singer

#157 Process-constrained batch Bayesian optimisation
Pratibha Vellanki, Santu Rana, Sunil Gupta, David Rubin, Alessandra Sutti, Thomas Dorin, Murray Height, Paul Sanders, Svetha Venkatesh

#158 The Marginal Value of Adaptive Gradient Methods in Machine Learning
Ashia C Wilson, Becca Roelofs, Mitchell Stern, Nati Srebro, Benjamin Recht

#159 Breaking the Nonsmooth Barrier: A Scalable Parallel Method for Composite Optimization
Fabian Pedregosa, Rémi Leblond, Simon Lacoste-Julien

#160 Beyond Worst-case: A Probabilistic Analysis of Affine Policies in Dynamic Optimization
Omar El Housni, Vineet Goyal

#161 Approximate Supermodularity Bounds for Experimental Design
Luiz Chamon Chamon, Alejandro Ribeiro

#162 On Blackbox Backpropagation and Jacobian Sensing
Krzysztof M Choromanski, Vikas Sindhwani

#163 Asynchronous Coordinate Descent under More Realistic Assumptions
Tao Sun, Robert Hannah, Wotao Yin

#164 Clustering with Noisy Queries
Arya Mazumdar, Barna Saha

#165 Approximation Algorithms for ℓ_0-Low Rank Approximation
Karl Bringmann, Pavel Kolev, David Woodruff

#166 Convergence Analysis of Two-layer Neural Networks with ReLU Activation
Yuanzhi Li, Yang Yuan

#167 Can Decentralized Algorithms Outperform Centralized Algorithms? A Case Study for Decentralized Parallel Stochastic Gradient Descent
Xiangru Lian, Ce Zhang, Huan Zhang, Cho-Jui Hsieh, Wei Zhang, Ji Liu

#168 Decomposition-Invariant Conditional Gradient for General Polytopes with Line Search
Mohammad Ali Bashiri, Xinhua Zhang

#169 Straggler Mitigation in Distributed Optimization Through Data Encoding
Can Karakus, Yifan Sun, Suhas Diggavi, Wotao Yin

#170 No More Fixed Penalty Parameter in ADMM: Faster Convergence with New Adaptive Penalization
Yi Xu, Mingrui Liu, Tianbao Yang, Qihang Lin

#171 Accelerated Stochastic Greedy Coordinate Descent by Soft Thresholding Projection onto Simplex
Chaobing Song, Shaobo Cui, Shu-Tao Xia, Yong Jiang

#172 Safe Adaptive Importance Sampling
Sebastian Stich, Anant Raj, Martin Jaggi

#173 Sharpness, Restart and Acceleration
Vincent Roulet, Alexandre d'Aspremont

#174 Stochastic Optimization with Variance Reduction for Infinite Datasets with Finite Sum Structure
Alberto Bietti, Julien Mairal

#175 Min-Max Propagation
Christopher Srinivasa, Inmar Givoni, Siamak Ravanbakhsh, Brendan J Frey

#176 A Disentangled Recognition and Nonlinear Dynamics Model for Unsupervised Learning
Marco Fraccaro, Simon Kamronn, Ulrich Paquet, Ole Winther

#177 Concrete Dropout
Yarin Gal, Jiri Hron, Alex Kendall

#178 REBAR: Low-variance, unbiased gradient estimates for discrete latent variable models
George Tucker, Andriy Mnih, Chris J Maddison, Dieterich Lawson, Jascha Sohl-Dickstein

#179 Hierarchical Implicit Models and Likelihood-Free Variational Inference
Dustin Tran, Rajesh Ranganath, David Blei

#180 Sticking the Landing: Simple, Lower-Variance Gradient Estimators for Variational Inference
Geoffrey Roeder, Yuhuai Wu, David Duvenaud

#181 Perturbative Black Box Variational Inference
Cheng Zhang, Robert Bamler, Manfred Opper, Stephan Mandt

#182 Fast Black-box Variational Inference through Stochastic Trust-Region Optimization
Jeff Regier, Michael Jordan, Jon McAuliffe

#183 Excess Risk Bounds for the Bayes Risk using Variational Inference in Latent Gaussian Models
Rishit Sheth, Roni Khardon

#184 Learning Causal Graphs with Latent Variables
Murat Kocaoglu, Karthikeyan Shanmugam, Elias Bareinboim

#185 Permutation-based Causal Inference Algorithms with Interventions
Yuhao Wang, Liam Solus, Karren Yang, Caroline Uhler

#186 Learning Causal Structures Using Regression Invariance
AmirEmad Ghassami, Saber Salehkaleybar, Negar Kiyavash, Kun Zhang

#187 Counterfactual Fairness
Matt Kusner, Joshua Loftus, Chris Russell, Ricardo Silva

#188 Causal Effect Inference with Deep Latent Variable Models
Christos Louizos, Uri Shalit, Joris M Mooij, David Sontag, Richard Zemel, Max Welling

#189 Conic Scan Coverage algorithm for nonparametric topic modeling
Mikhail Yurochkin, Aritra Guha, Long Nguyen

#190 Tractability in Structured Probability Spaces
Arthur Choi, Yujia Shen, Adnan Darwiche

#191 PASS-GLM: polynomial approximate sufficient statistics for scalable Bayesian GLM inference
Jonathan Huggins, Ryan Adams, Tamara Broderick

#192 Adaptive Bayesian Sampling with Monte Carlo EM
Anirban Roychowdhury, Srinivasan Parthasarathy

#193 What-If Reasoning using Counterfactual Gaussian Processes
Peter Schulam, Suchi Saria

#194 Multi-Information Source Optimization
Matthias Poloczek, Jialei Wang, Peter Frazier

#195 Doubly Stochastic Variational Inference for Deep Gaussian Processes
Hugh Salimbeni, Marc Deisenroth

#196 Convolutional Gaussian Processes
Mark van der Wilk, Carl Edward Rasmussen, James Hensman

#197 Multiresolution Kernel Approximation for Gaussian Process Regression
Yi Ding, Risi Kondor, Jon Eskreis-Winkler

#198 Unifying PAC and Regret: Uniform PAC Bounds for Episodic Reinforcement Learning
Christoph Dann, Tor Lattimore, Emma Brunskill

#199 Repeated Inverse Reinforcement Learning
Kareem Amin, Nan Jiang, Satinder Singh

#200 Inverse Reward Design
Dylan Hadfield-Menell, Smitha Milli, Stuart J Russell, Pieter Abbeel, Anca Dragan

#201 Utile Context Tree Weighting
Joao V Messias, Shimon Whiteson

#202 Policy Gradient With Value Function Approximation For Collective Multiagent Planning
Duc Nguyen, Akshat Kumar, Hoong Chuin Lau

#203 A Unified Game-Theoretic Approach to Multiagent Reinforcement Learning
Marc Lanctot, Vinicius Zambaldi, Audrunas Gruslys, Angeliki Lazaridou, karl Tuyls, Julien Perolat, David Silver, Thore Graepel

#204 Dynamic Safe Interruptibility for Decentralized Multi-Agent Reinforcement Learning
El Mahdi El Mhamdi, Rachid Guerraoui, Hadrien Hendrikx, Alexandre Maurer

#205 Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments
Ryan Lowe, YI WU, Aviv Tamar, Jean Harb, OpenAI Pieter Abbeel, Igor Mordatch

#206 Spectrally-normalized margin bounds for neural networks
Matus Telgarsky, Peter Bartlett, Dylan J Foster

#207 On Structured Prediction Theory with Calibrated Convex Surrogate Losses
Anton Osokin, Francis Bach, Simon Lacoste-Julien

#208 Collaborative PAC Learning
Avrim Blum, Nika Haghtalab, Ariel D Procaccia, IIIS Mingda Qiao

#209 Submultiplicative Glivenko-Cantelli and Uniform Convergence of Revenues
Noga Alon, Moshe Babaioff, Yannai A. Gonczarowski, Yishay Mansour, Shay Moran, Amir Yehudayoff

#210 Discriminative State Space Models
Vitaly Kuznetsov, Mehryar Mohri

#211 Delayed Mirror Descent in Continuous Games
Zhengyuan Zhou, Panayotis Mertikopoulos, Nicholas Bambos, Peter W Glynn, Claire Tomlin

#212 Variance-based Regularization with Convex Objectives
Hong Namkoong, John C Duchi

#213 Learning Mixture of Gaussians with Streaming Data
Aditi Raghunathan, Prateek Jain, Ravishankar Krishnawamy

#214 On the Consistency of Quick Shift
Heinrich Jiang

#215 Early stopping for kernel boosting algorithms: A general analysis with localized complexities
Yuting Wei, Fanny Yang, Martin Wainwright

#216 A Sharp Error Analysis for the Fused Lasso, with Implications to Broader Settings and Approximate Screening
Kevin Lin, James Sharpnack, Alessandro Rinaldo, Ryan Tibshirani

#217 The Scaling Limit of High-Dimensional Online Independent Component Analysis
Chuang Wang, Yue Lu

#218 A Universal Analysis of Large-Scale Regularized Least Squares Solutions
Ashkan Panahi, Babak Hassibi

#219 Statistical Convergence Analysis of Gradient EM on General Gaussian Mixture Models
Bowei Yan, Mingzhang Yin, Purnamrita Sarkar

#220 More powerful and flexible rules for online FDR control with memory and weights
Aaditya Ramdas, Fanny Yang, Martin Wainwright, Michael Jordan

#221 Learning with Bandit Feedback in Potential Games
Amélie Heliou, Johanne Cohen, Panayotis Mertikopoulos

#222 Fully Decentralized Policies for Multi-Agent Systems: An Information Theoretic Approach
Roel Dobbe, David Fridovich-Keil, Claire Tomlin

#223 Revenue Optimization with Approximate Bid Predictions
Andres Munoz, Sergei Vassilvitskii

#224 A Decomposition of Forecast Error in Prediction Markets
Miro Dudik, Sebastien Lahaie, Ryan M Rogers, Jenn Wortman Vaughan

#225 Dynamic Revenue Sharing
Santiago Balseiro, Max Lin, Vahab Mirrokni, Renato Leme, IIIS Song Zuo

#226 Multi-View Decision Processes
Christos Dimitrakakis, David Parkes, Goran Radanovic, Paul Tylkin

See Page 9 For Specific Demo Locations

D1 Humans Attributes Extraction And Search With A deep Learning Based Real-time Video Analysis System

Matthieu Ospici, Benoit Pelletier, Antoine Cecchi

Our demo is a real-time computer vision Demo with two main features. Firstly, attendees can visualize the detected person on a camera with a set of estimated attributes such as the clothes color or the gender. Then, a search engine enables the participants to request past detections by criteria or photo.

D2 MAgent: A Many-Agent Reinforcement Learning Research Platform for Artificial Collective Intelligence

Lianmin Zheng, Jiacheng Yang, Han Cai, Weinan Zhang, Jun Wang, Yong Yu

We introduce MAgent, a platform to support research and development of many-agent Reinforcement learning. Unlike previous research platforms on single or multi-agent Reinforcement learning, MAgent focuses on supporting the tasks and the Applications that require hundreds to millions of agents. Within the interactions among a population of agents, it enables not only the study of learning algorithms for agents' optimal polices, but more importantly, the observation and understanding of individual agent's behaviors and social phenomena emerging from the AI society. MAgent also provides flexible configurations and a description language for AI researchers to easily design their customized environment, agents, and rewards. In this demo, we present several environments designed on MAgent and show emerged collective intelligence. Visitors can also play interactive games provided by MAgent.

D3 Electronic Screen Protector with Efficient and Robust Mobile Vision

Hee Jung Ryu, Florian Schroff

Face authentication, in the context of privacy for phones, has been explored for some time. However, face recognition alone is not enough when you want to have private online conversations or watch a confidential video in a crowded space where there are many other people present. Each of them may or may not be looking at your private content displayed on your device, e.g. a smart phone. Because of the quick, robust, and accurate gaze detection mobile model we can now easily identify the face identity and gaze simultaneously in real time. Hence, the application, an electronic screen protector, can enable its enrolled users to continue reading private and confidential contents on your mobile device, while protecting their privacy from onlookers in a crowded space such as the subway or an elevator. We enable this by transfer learning from one mobile model to a different, but related task. Our final multihead mobile model is robust under varying lighting conditions and head poses. The runtime is 2ms per face for gaze detection, 47ms per face for face recognition, and 115ms per frame for face detection in average.

D4 3D Surface-to-Structure Translation with Deep Convolutional Networks

Takumi Moriya, Kazuyuki Saito

Our Demo shows a system that estimates internal body structures from 3D surface models using deep convolutional neural networks trained on CT (computed tomography) images of the human body. To take pictures of structures inside the body, we need to use a CT scanner or an MRI (Magnetic Resonance Imaging) scanner. However, assuming that the mutual information between outer shape of the body and its inner structure is not zero, we can obtain an approximate internal structure from a 3D surface model based on MRI and CT image database. This suggests that we could know where and what kind of disease a person is likely to have in his/her body simply by 3D scanning surface of the body. As a first prototype, we developed a system for estimating internal body structures from surface models based on Visible Human Project DICOM CT Datasets from the University of Iowa Magnetic Resonance Research Facility 1. The estimation process given a surface model is shown in Figure 1. The input surface model is not limited to the human body. For instance, our method enables us to create Stanford Armadillo that has internal structures of the human body.

D5 Sharkzor: Interactive Deep Learning for Image Triage, Sort and Summary

Nathan Hodas, Nathan Hilliard, Artem Yankov, Megan Pirrung, Courtney Corley

Sharkzor leverages multiple deep learning techniques to facilitate image identification and organization, exemplar based regression and few-shot learning. These algorithms and methods combined in a user-centric web UI, allowing users to triage large amounts of images, using n-shot learning to "find the needle in the haystack" Sharkzor captures users interactions with a 2D canvas of images. It tracks where users position images on the screen, and which groups they form. It then attempts to position all of the remaining images based on the user's mental model.To address the requirement of users being able to create arbitrary image-related mental models, we aren't able to use traditional multi-label classification techniques. This is because the user may be interested in clustering images into arbitrarily complex arrangements. To make a robust system that can adapt to user supplied groups, we leverage learning techniques requiring few training examples. We have developed our own few-shot learning techniques and exemplar-based regression to transform Sharkzor into an interactive deep learning platform with networks that require no retraining or weight tuning to adapt to each user's unique mental model of their task.

D6 Magenta and deeplearn.js: Real-time Control of DeepGenerative Music Models in the Browser

Curtis “Fjord” Hawthorne, Ian Simon, Adam Roberts, Jesse Engel, Daniel Smilkov, Nikhil Thorat, Douglas Eck

There has recently been increased interest in generating music using deep learning techniques, leading to remarkable improvements in the quality and expressiveness of sequence-based models. Beyond unconditional generation, we aim to explore the ability of the generative models to augment the creativity of musicians and novices alike. To be successful, both the model and the user interface must expose high-level and expressive controls that empower users to explore novel musical possibilities. Furthermore, the interface must be easy both for casual users to access and for professional users to integrate into existing creative workflows. This is key to new directions in adaptive feedback and training of models based on user preferences. To this end, we train state-of-the-art generative models with conditional controls for several musical domains — virtuosic piano performances, looping melodies and drum beats — and demonstrate user interfaces to control generation from these models in real time using only code running in a browser-based JavaScript environment via deeplearn.js.

D7 Matrix Calculus: The Power of Symbolic Differentiation

Soeren Laue, Matthias Mitterreiter, Joachim Giesen

Numerical optimization is a work horse of machine learning that often requires the deriva- tion and computation of gradients and Hessians. For learning problem that are modeled by some loss or likelihood function, the gradients and Hessians are typically derived man- ually, which is a time consuming and error prone process. Computing gradients (and Hessians) is also an integral part of deep learning frameworks that mostly employ au- tomatic differentiation, aka algorithmic differentiation (typically in reverse mode). At (www.MatrixCalculus.org) we provide a tool for symbolically computing gradients and Hessians that can be used in the classical setting of loss and likelihood functions, but also for deep learning.

D8 Babble Labble: Learning from Natural Language Explanations

Braden Hancock, Stephanie Wang, Paroma Varma, Percy Liang, Christopher Ré

We introduce Babble Labble, a system for converting natural language explanations into massive training sets with Probabilistic labels. In this demo, users will be shown unlabeled examples for a simple relation extraction task (identifying mentions of spouses in the news). For each example, instead of providing a label, users provide a sentence describing one reason why the given example should receive a certain label. These explanations are parsed into executable functions in real-time and applied to the unlabeled dataset. We use data programming to resolve conflicts between the functions and combine their weak labels into a single Probabilistic label per example. This large weakly labeled training set is then used to train a discriminative model that improves generalization as it includes features never mentioned in the small set of explanations. Using the explanations the user wrote, we calculate the final quality of the complete system, finding in most cases that one to two dozen explanations achieve the same quality as hundreds or thousands of labels.

D9 Interactive-Length Multi-Task Video Captioning with Cooperative Feedback

Han Guo, Ramakanth Pasunuru, Mohit Bansal

We present a fast and accurate demo system for our state-of-the-art multi-task video captioning model, with additional interactive-length paragraph generation and cooperative user feedback techniques. The task of automatic video captioning has various Applications such as assistance to a visually impaired person and improving the quality of online visual content search or retrieval. Our recent multi-task model uses auxiliary temporal video-to-video and logical premise-to-entailment generation tasks to achieve the best results on three popular community datasets. To address the lack of useful online demo systems for video captioning, we present a fast and interactive demo system of our state-of-the-art multi-task model, that allows users to upload any video file or YouTube link, with the additional novel aspect of generating multi-sentence, paragraph-style captions based on redundancy filtering (especially useful for real-world lengthy videos), where the user can ask for longer captions on the fly. Our demo system also allows for cooperative user feedback, where the user can click on a displayed alternative top-k beam option or rewrite corrections directly, providing us with valuable data for discriminative retraining.

D10 Fast-speed Intelligent Video Analytics using Deep Learning Algorithms on Low-power FPGA

Yi Shan, Song Yao, Song Han, Yu Wang

Deep learning algorithms, such as CNN (Convolutional Neural Network), could provide high 8 accuracy for great number of Applications including video analytics for surveillance and 9 automotive. Considering processing speed and energy efficiency, FPGA is a good hardware to construct customized CNN solution. In this demo session, we want to benefit from hardware technology, and show a fast speed 12 and accurate video analytics system using state-of-the-art deep learning algorithms running 13 on low power FPGA. This system could process 16 channels of continuous input video 14 with the resolution of 1080p. Two functionalities could be easily switched by just clicking a 15 button in this live demo: one for vehicle, non-motorized vehicle, and pedestrian detection, 16 tracking, and attributes analytics; and the other for face detection and recognition. The deep 17 learning algorithms used are SSD and densebox for two kinds of objects’ detection, which 18 have state-of-the-art accuracy. The FPGA used is Xilinx MPSoC ZU9, and the whole 19 board including this FPGA only cost about 50 Watts with Peak performance at 5.6 TOPS.

THURSDAY SESSIONS

7:30 - 9:00 AM	Coffee	
9:00 - 9:50 AM	**Invited talk: Yael Niv** *Learning State Representations*	**Hall A**
9:50 - 10:40 AM	**Invited Talk: Breiman Lecture** **Yee Whye Teh** *On Bayesian Deep Learning and Deep Bayesian Learning*	**Hall A**
10:40 - 11:10 AM	Coffee break	
11:10 - 12:30 PM	**Parallel Tracks:** **Neuroscience** **Deep Learning, Algorithms**	 **Hall A** **Hall C**
12:30 - 2:00 PM	Lunch on your own	
2:00 - 4:00 PM	**SYMPOSIA**	**Hall A, Hall C, Grand Ballroom, Beverly Theater**
4:00 - 4:30 PM	Coffee break	
4:30 - 6:30 PM	**SYMPOSIA**	**Hall A, Hall C, Grand Ballroom, Beverly Theater**
6:30 - 7:30 PM	Light dinner	
7:30 - 9:30 PM	**SYMPOSIA**	**Hall A, Hall C, Grand Ballroom, Beverly Theater**

Thursday Invited Talks

Learning State Representations

Hall A, 9:00 - 9:50 AM

On the face of it, most real-world world tasks are hopelessly complex from the point of view of Reinforcement learning mechanisms. In particular, due to the "curse of dimensionality", even the simple task of crossing the street should, in principle, take thousands of trials to learn to master. But we are better than that.. How does our brain do it? In this talk, I will argue that the hardest part of learning is not assigning values or learning policies, but rather deciding on the boundaries of similarity between experiences, which define the "states" that we learn about. I will show behavioral evidence that humans and animals are constantly engaged in this representation learning process, and suggest that in a not too far future, we may be able to read out these representations from the brain, and therefore find out how the brain has mastered this complex problem. I will formalize the problem of learning a state representation in terms of Bayesian inference with infinite capacity models, and suggest that an understanding of the computational problem of representation learning can lead to insights into the machine learning problem of transfer learning, and psychological/neuroscientific questions about the interplay between memory and learning.

Yael Niv
(Princeton University)

Yael Niv received her MA in psychobiology from Tel Aviv University and her PhD from the Hebrew University in Jerusalem, having conducted a major part of her thesis research at the Gatsby Computational Neuroscience Unit in UCL. After a short postdoc at Princeton she became faculty at the Psychology Department and the Princeton Neuroscience Institute. Her lab's research focuses on the neural and computational processes underlying Reinforcement learning and decision-making in humans and animals, with a particular focus on representation learning. She recently co-founded the Rutgers-Princeton Center for Computational Cognitive Neuropsychiatry, and is currently taking the research in her lab in the direction of computational psychiatry.

Breiman Lecture

On Bayesian Deep Learning and Deep Bayesian Learning

Hall A, 9:50 - 10:40 AM

Probabilistic and Bayesian reasoning is one of the principle theoretical pillars to our understanding of machine learning. Over the last two decades, it has inspired a whole range of successful machine learning methods and influenced the thinking of many researchers in the community. On the other hand, in the last few years the rise of deep learning has completely transformed the field and led to a string of phenomenal, era-defining, successes. In this talk I will explore the interface between these two perspectives on machine learning, and through a number of projects I have been involved in, explore questions like: How can Probabilistic thinking help us understand deep learning methods or lead us to interesting new methods? Conversely, how can deep learning technologies help us develop advanced Probabilistic methods?

Yee Whye Teh
(Princeton University)

I am a Professor of Statistical Machine Learning at the Department of Statistics, University of Oxford and a Research Scientist at DeepMind. I am also an Alan Turing Institute Fellow and a European Research Council Consolidator Fellow. I obtained my Ph.D. at the University of Toronto (working with Geoffrey Hinton), and did postdoctoral work at the University of California at Berkeley (with Michael Jordan) and National University of Singapore (as Lee Kuan Yew Postdoctoral Fellow). I was a Lecturer then a Reader at the Gatsby Computational Neuroscience Unit, UCL, and a tutorial fellow at University College Oxford, prior to my current appointment. I am interested in the statistical and computational foundations of intelligence, and works on scalable machine learning, Probabilistic models, Bayesian nonparametrics and deep learning. I was programme co-chair of ICML 2017 and AISTATS 2010.

Track 1 - 11:10 am - 12:30 pm Neuroscience

Location: Hall A

Toward Goal-Driven Neural Network Models for the Rodent Whisker-Trigeminal System

Chengxu Zhuang, Jonas Kubilius, Mitra JZ Hartmann, Daniel Yamins

In large part, rodents "see" the world through their whiskers, a powerful tactile sense enabled by a series of brain areas that form the whisker-trigeminal system. Raw sensory data arrives in the form of mechanical input to the exquisitely sensitive, actively-controllable whisker array, and is processed through a sequence of neural circuits, eventually arriving in cortical regions that communicate with decision making and memory areas. Although a long history of experimental studies has characterized many aspects of these processing stages, the computational operations of the whisker-trigeminal system remain largely unknown. In the present work, we take a goal-driven deep neural network (DNN) approach to modeling these computations. First, we construct a biophysically-realistic model of the rat whisker array. We then generate a large dataset of whisker sweeps across a wide variety of 3D objects in highly-varying poses, angles, and speeds. Next, we train DNNs from several distinct architectural families to solve a shape recognition task in this dataset. Each architectural family represents a structurally-distinct hypothesis for processing in the whisker-trigeminal system, corresponding to different ways in which spatial and temporal information can be integrated. We find that most networks perform poorly on the challenging shape recognition task, but that specific architectures from several families can achieve reasonable performance levels. Finally, we show that Representational Dissimilarity Matrices (RDMs), a tool for comparing population codes between neural systems, can separate these higher performing networks with data of a type that could plausibly be collected in a neurophysiological or imaging experiment. Our results are a proof-of-concept that DNN models of the whisker-trigeminal system are potentially within reach.

Model-based Bayesian Inference of Neural Activity And Connectivity From All-optical Interrogation Of A Neural Circuit

Laurence Aitchison, Lloyd Russell, Adam Packer, Jinyao Yan, Philippe Castonguay, Michael Hausser, Srinivas C Turaga

Population activity measurement by calcium imaging can be combined with cellular resolution optogenetic activity perturbations to enable the mapping of neural connectivity in vivo. This requires accurate inference of perturbed and unperturbed neural activity from calcium imaging measurements, which are noisy and indirect, and can also be contaminated by photostimulation artifacts. We have developed a new fully Bayesian approach to jointly inferring spiking activity and neural connectivity from in vivo all-optical perturbation experiments. In contrast to standard approaches that perform spike inference and analysis in two separate maximum-likelihood phases, our joint model is able to propagate uncertainty in spike inference to the inference of connectivity and vice versa. We use the framework of variational autoencoders to model spiking activity using discrete latent variables, low-dimensional latent common input, and sparse spike-and-slab generalized linear coupling between neurons. Additionally, we model two properties of the optogenetic perturbation: off-target photostimulation and photostimulation transients. Our joint model includes at least two sets of discrete random variables; to avoid the dramatic slowdown typically caused by being unable to differentiate such variables, we introduce two strategies that have not, to our knowledge, been used with variational autoencoders. Using this model, we were able to fit models on 30 minutes of data in just 10 minutes. We performed an all-optical circuit mapping experiment in primary visual cortex of the awake mouse, and use our approach to predict neural connectivity between excitatory neurons in layer 2/3. Predicted connectivity is sparse and consistent with known correlations with stimulus tuning, spontaneous correlation and distance.

Quantifying How Much Sensory Information In A Neural Code Is Relevant For Behavior

Giuseppe Pica, Eugenio Piasini, Houman Safaai, Caroline Runyan, Christopher Harvey, Mathew Diamond, Christoph Kayser, Tommaso Fellin, Stefano Panzeri

Determining how much of the sensory information carried by a neural code contributes to behavioral performance is key to understand sensory function and neural information flow. However, there are as yet no analytical tools to compute this information that lies at the intersection between sensory coding and behavioral readout. Here we develop a novel measure, termed the information-theoretic intersection information II(R), that quantifies how much sensory information carried by a neural response R is also used for behavior during perceptual discrimination tasks. Building on the Partial Information Decomposition framework, we define II(R) as the mutual information between the presented stimulus S and the consequent behavioral choice C that can be extracted from R. We compute III(R) in the analysis of two experimental cortical datasets, to show how this measure can be used to compare quantitatively the contributions of spike timing and spike rates to task performance, and to identify brain areas or neural populations that specifically transform sensory information into choice.

SPOTLIGHTS

- **Scene Physics Acquisition via Visual De-animation**
  Jiajun Wu, Erika Lu, Pushmeet Kohli, Bill Freeman, Josh Tenenbaum
- **Shape and Material from Sound**
  Zhoutong zhang, Qiujia Li, Zhengjia Huang, Jiajun Wu, Josh Tenenbaum, Bill Freeman
- **Deep Hyperalignment**
  Muhammad Yousefnezhad, Daoqiang Zhang
- **Fast amortized inference of neural activity from calcium imaging data with variational autoencoders**
  Artur Speiser, Jinyao Yan, Evan Archer, Lars Buesing, Srini C Turaga, Jakob H Macke
- **Tensor encoding and decomposition of brain connectomes with application to tractography evaluation**
  Cesar F Caiafa, Olaf Sporns, Andrew Saykin, Franco Pestilli
- **Targeting EEG/LFP Synchrony with Neural Nets**
  Yitong Li, David E Carlson, Lawrence Carin
- **Deep Networks for Decoding Natural Images from Retinal Signals**
  Nikhil Parthasarathy, Eleanor Batty, William Falcon, Thomas Rutten, Mohit Rajpal, EJ Chichilnisky, Liam Paninski

Track 2 - 11:10 am - 12:30 pm
Deep Learning, Algorithms

Location: Hall C

Masked Autoregressive Flow for Density Estimation

George Papamakarios, Iain Murray, Theo Pavlakou

Autoregressive models are among the best performing neural density estimators. We describe an approach for increasing the flexibility of an autoregressive model, based on modelling the random numbers that the model uses internally when generating data. By constructing a stack of autoregressive models, each modelling the random numbers of the next model in the stack, we obtain a type of normalizing flow suitable for density estimation, which we call Masked Autoregressive Flow. This type of flow is closely related to Inverse Autoregressive Flow and is a generalization of Real NVP. Masked Autoregressive Flow achieves state-of-the-art performance in a range of general-purpose density estimation tasks.

Deep Sets

Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Ruslan Salakhutdinov, Alex Smola

We study the problem of designing objective models for machine learning tasks defined on finite \emph{sets}. In contrast to the traditional approach of operating on fixed dimensional vectors, we consider objective functions defined on sets and are invariant to permutations. Such problems are widespread, ranging from the estimation of population statistics, to anomaly detection in piezometer data of embankment dams, to cosmology. Our main theorem characterizes the permutation invariant objective functions and provides a family of functions to which any permutation invariant objective function must belong. This family of functions has a special structure which enables us to design a deep network architecture that can operate on sets and which can be deployed on a variety of scenarios including both unsupervised and supervised learning tasks. We demonstrate the applicability of our method on population statistic estimation, point cloud classification, set expansion, and outlier detection.

From Bayesian Sparsity to Gated Recurrent Nets

Hao He, Bo Xin, David Wipf

The iterations of many first-order algorithms, when applied to minimizing common regularized regression functions, often resemble neural network layers with pre-specified weights. This observation has prompted the development of learning-based approaches that purport to replace these iterations with enhanced surrogates forged as DNN models from available training data. For example, important NP-hard sparse estimation problems have recently benefitted from this genre of upgrade, with simple feedforward or recurrent networks ousting proximal gradient-based iterations. Analogously, this paper demonstrates that more powerful Bayesian algorithms for promoting sparsity, which rely on complex multi-loop majorization-minimization techniques, mirror the structure of more sophisticated long short-term memory (LSTM) networks, or alternative gated feedback networks previously designed for sequence prediction. As part of this development, we examine the parallels between latent variable trajectories operating across multiple time-scales during optimization, and the activations within deep network structures designed to adaptively model such characteristic sequences. The resulting insights lead to a novel sparse estimation system that, when granted training data, can estimate optimal solutions efficiently in regimes where other algorithms fail, including practical direction-of-arrival (DOA) and 3D geometry recovery problems. The underlying principles we expose are also suggestive of a learning process for a richer class of multi-loop algorithms in other domains.

SPOTLIGHTS

- **Self-Normalizing Neural Networks**
  Günter Klambauer, Thomas Unterthiner, Andreas Mayr, Sepp Hochreiter
- **Batch Renormalization: Towards Reducing Minibatch Dependence in Batch-Normalized Models**
  Sergey Loffe
- **Nonlinear random matrix theory for deep learning**
  Jeffrey Pennington, Pratik Worah
- **Spherical convolutions and their application in molecular modelling**
  Wouter Boomsma, Jes Frellsen
- **Translation Synchronization via Truncated Least Squares**
  Xiangru Huang, Zhenxiao Liang, Chandrajit Bajaj, Qixing Huang
- **Self-supervised Learning of Motion Capture**
  Hsiao-Yu Tung, Hsiao-Wei Tung, Ersin Yumer, Katerina Fragkiadaki
- **Maximizing Subset Accuracy with Recurrent Neural Networks in Multi-label Classification**
  Jinseok Nam, Eneldo Loza Mencía, Hyunwoo J Kim, Johannes Fürnkranz

SYMPOSIUM

THURSDAY SESSIONS: 2:00 - 9:30 PM

Interpretable Machine Learning
Hall C

Andrew G Wilson (Cornell U.)
Jason Yosinski (Uber AI Labs)
Patrice Simard (Microsoft Research)
Rich Caruana (Microsoft Research)
William Herlands (Carnegie Mellon U.)

Complex machine learning models, such as deep neural networks, have recently achieved outstanding predictive performance in a wide range of Applications, including visual object recognition, speech perception, language modeling, and information retrieval. There has since been an explosion of interest in interpreting the representations learned by these models, with profound implications for research into explainable ML, causality, safe AI, social science, automatic scientific discovery, human computer interaction (HCI), crowdsourcing, machine teaching, and AI ethics. This symposium is designed to broadly engage the machine learning community on these topics -- tying together many threads which are deeply related but often considered in isolation.

Deep Reinforcement Learning
Hall A

Pieter Abbeel (Open AI, UC Berkeley, Gradescope)
Yan Duan (UC Berkeley)
David Silver (DeepMind)
Satinder Singh (U. Of Michigan)
Junhyuk Oh (U. Of Michigan)
Rein Houthooft (Ghent U., Open AI)

Although the theory of Reinforcement learning addresses an extremely general class of learning problems with a common mathematical formulation, its power has been limited by the need to develop task-specific feature representations. A paradigm shift is occurring as researchers figure out how to use deep neural networks as function approximators in Reinforcement learning algorithms; this line of work has yielded remarkable empirical results in recent years. This workshop will bring together researchers working at the intersection of deep learning and Reinforcement learning, and it will help researchers with expertise in one of these fields to learn about the other.

Metalearning
Grand Ballroom

Risto Miikkulainen (UT Austin)
Quoc V Le (Google)
Ken Stanley (Uber AI Labs, U. Of Central Florida)
Chrisantha Fernando (DeepMind)

Modern learning systems, such as the recent deep learning, Reinforcement learning, and Probabilistic inference architectures, have become increasingly complex, often beyond the human ability to comprehend them. Such complexity is important: The more complex these systems are, the more powerful they often are. A new research problem has therefore emerged: How can the complexity, i.e. the design, components, and hyperparameters, be configured automatically so that these systems perform as well as possible? This is the problem of metalearning. Several approaches have emerged, including those based on Bayesian optimization, gradient descent, Reinforcement learning, and evolutionary computation. The symposium presents an overview of these approaches, given by the researchers who developed them. Panel discussion compares the strengths of the different approaches and potential for future developments and Applications. The audience will thus obtain a practical understanding of how to use metalearning to improve the learning systems they are using, as well as opportunities for research on metalearning in the future.

Kinds Of Intelligence: Types, Tests and Meeting The Needs of Society
Beverly Theater

José Hernández-Orallo (U. Of Valencia)
Zoubin Ghahramani (Uber, U. Of Cambridge)
Tomaso A Poggio (MIT)
Adrian Weller (U. Of Cambridge)
Matthew Crosby (Imperial College Of London)

Existing research in machine learning and artificial intelligence has been constrained by a focus on specific tasks chosen either for their perceived importance in human intelligence, their expected practical impact, their suitability for testing and comparison, or simply by an accident of research trends. However, the intelligence landscape extends far beyond our current capabilities, with many unexplored dimensions that present themselves as new opportunities for research. This symposium explores this landscape across three main topics: a broader perspective of the possible types of intelligence beyond human intelligence, better measurements providing an improved understanding of research objectives and breakthroughs, and a more purposeful analysis of where progress should be made in this landscape in order to best benefit society.

FRIDAY WORKSHOPS

8:00 AM - 6:30 PM

Room S-4

- **Machine Learning for Molecules and Materials**

  Stefan Chmiela, José Miguel Hernández-Lobato, Kristof T. Schütt, Alan Aspuru-Guzik, Alexandre Tkatchenko, Bharath Ramsundar, Anatole von Lilienfeld, Matt Kusner, Koji Tsuda, Brooks Paige, Klaus-Robert Müller

Seaside Ballroom

- **Advances in Approximate Bayesian Inference**

  Francisco Ruiz, Stephan Mandt, Cheng Zhang, James McInerney, Dustin Tran, Tamara Broderick, Michalis Titsias, David Blei, Max Welling

Room 204

- **Transparent and interpretable Machine Learning in Safety Critical Environments**

  Alessandra Tosi, Alfredo Vellido, Mauricio A. Álvarez

Room 101-A

- **Learning in the Presence of Strategic Behavior**

  Nika Haghtalab, Yishay Mansour, Tim Roughgarden, Vasilis Syrgkanis, Jennifer Wortman Vaughan

Grand Ballroom B

- **Conversational AI - Today's Practice & Tomorrow's Potential**

  Alborz Geramifard, Jason Williams, Larry Heck, James Glass, Antoine Bordes, Steve Young, Gerald Tesauro

Room 102-C

- **6th Workshop on Automated Knowledge Base Construction**

  Jay Pujara, Danqi Chen, Bhavana Dalvi Mishra, Tim Rocktäschel

Room 102 A+B

- **Advances in Modeling and Learning Interactions from Complex Data**

  Gautam Dasarathy, Mladen Kolar, Richard Baraniuk

Room 101-B

- **Visually Grounded Interaction and Language**

  Florian Strub, Harm de Vries, Abhishek Das, Satwik Kottur, Stefan Lee, Mateusz Malinowski, Olivier Pietquin, Devi Parikh, Dhruv Batra, Aaron C Courville, Jeremie Mary

Room S-7

- **Machine Learning for the Developing World**

  Maria De-Arteaga, William Herlands

Grand Ballroom A

- **NIPS 2017 Time Series Workshop**

  Vitaly Kuznetsov, Oren Anava, Scott Yang, Azadeh Khaleghi

Hyatt Hotel, Regency Ballroom (A, B & C)

- **Extreme Classification: Multi-class & Multi-label Learning in Extremely Large Label Spaces**

  Manik Varma, Marius Kloft, Krzysztof Dembczynski

Room 201-B

- **Nearest Neighbors for Modern Applications with Massive Data: An Age-old Solution with New Challenges**

  George H Chen, Devavrat Shah, Christina Lee

Room 104-B

- **Acting and Interacting in the Real World: Challenges in Robot Learning**

  Ingmar Posner, Raia Hadsell, Martin Riedmiller, Markus Wulfmeier, Rohan Paul

Room 202

- **Machine Deception**

  Ian Goodfellow, Tim Hwang, Bryce Goodman, Mikel Rodriguez

Hall A

- **OPT 2017: Optimization for Machine Learning**

  Suvrit Sra, Sashank J. Reddi, Alekh Agarwal, Benjamin Recht

Hyatt Hotel, Regency Ballroom D+E+F+H

- **Learning on Distributions, Functions, Graphs and Groups**

  Florence d'Alché-Buc, Krikamol Muandet, Bharath Sriperumbudur, Zoltán Szabó

Hyatt Hotel, Shoreline

- **Machine Learning and Computer Security**

  Jacob Steinhardt, Nicolas Papernot, Bo Li, Chang Liu, Percy Liang, Dawn Song

Room S-5

- **Workshop on Worm's Neural Information Processing**

  Ramin Hasani, Manuel Zimmer, Stephen Larson, Radu Grosu

Room 1 04-C

- **Deep Learning for Physical Sciences**

  Atilim Gunes Baydin, Mr. Prabhat, Kyle Cranmer, Frank Wood

Room 104-A

- **Machine Learning for Health (ML4H) - What Parts of Healthcare are Ripe for Disruption by Machine Learning Right Now?**

  Andrew Beam, Madalina Fiterau, Peter Schulam, Jason Fries, Michael Hughes, Alex Wiltschko, Jasper Snoek, Natalia Antropova, Rajesh Ranganath, Bruno Jedynak, Tristan Naumann, Adrian Dalca, Adrian Dalca, Tim Althoff, SHUBHI ASTHANA, Prateek Tandon, Jaz Kandola, Alexander Ratner, David Kale, Uri Shalit, Marzyeh Ghassemi, Isaac S Kohane

Room 201-A

- **A Machine Learning for Audio Signal Processing (ML4)**

  Hendrik Purwins, Bob L. Sturm, Mark Plumbley

Room 103 A+B

- **Competition track**

  Sergio Escalera, Markus Weimer

Room 203

- **Discrete Structures in Machine Learning**

  Yaron Singer, Jeff A Bilmes, Andreas Krause, Stefanie Jegelka, Amin Karbasi

Hyatt Hotel, Seaview Ballroom

- **Machine Learning for Creativity and Design**

  Douglas Eck, David Ha, S. M. Ali Eslami, Sander Dieleman, Rebecca Fiebrink, Luba Elliott

Room S-1

- **ML Systems Workshop @ NIPS 2017**

  Aparna Lakshmiratan, Sarah Bird, Siddhartha Sen, Christopher Ré, Li Erran Li, Joseph Gonzalez, Daniel Crankshaw

Room 103-C

- **Synergies in Geometric Data Analysis**

  Marina Meila, Frederic Chazal

Hall C

- **From 'What If?' To 'What Next?' : Causal Inference and Machine Learning for Intelligent Decision Making**

  Alexander Volfovsky, Adith Swaminathan, Panagiotis Toulis, Nathan Kallus, Ricardo Silva, John S Shawe-Taylor, Thorsten Joachims, Lihong Li

SATURDAY WORKSHOPS

8:00 AM - 6:30 PM

Room 102 A+B

- **Machine Learning on the Phone & other Consumer Devices**

  Hrishikesh Aradhye · Joaquin Quinonero Candela · Rohit Prasad

Room 101-B

- **Deep Learning at Supercomputer Scale**

  Erich Elsen · Danijar Hafner · Zak Stone · Brennan Saeta

Seaside Ballroom

- **Teaching Machines, Robots, and Humans**

  Maya Cakmak · Anna Rafferty · Adish Singla · Xiaojin Zhu · Sandra Zilles

Room 201-A

- **2017 NIPS Workshop on Machine Learning for Intelligent Transportation Systems**

  Li Erran Li · Anca Dragan · Juan Carlos Niebles · Silvio Savarese

Grand Ballroom A

- **Hierarchical Reinforcement Learning**

  Andrew G Barto · Doina Precup · Shie Mannor · Tom Schaul · Roy Fox · Carlos Florensa Campo

Hyatt Hotel, Regency Ballroom D+E+F+H

- **Workshop on Meta-Learning**

  Roberto Calandra · Frank Hutter · Hugo Larochelle · Sergey Levine

Room 104-B

- **Machine Learning in Computational Biology**

  James Zou · Anshul Kundaje · Gerald Quon · Nicolo Fusi · Sara Mostafavi

Room 101-A

- **(Almost) 50 shades of Bayesian Learning: PAC-Bayesian trends and insights**

  Benjamin Guedj · Pascal Germain · Francis Bach

Room104-A

- **Cognitively Informed Artificial Intelligence: Insights From Natural Intelligence**

  Michael Mozer · Brenden Lake · Angela J Yu

Room S-7

- **Bayesian Optimization for Science and Engineering**

  Ruben Martinez-Cantin · José Miguel Hernández-Lobato · Javier Gonzalez

Room 103-C

- **Workshop on Prioritising Online Content**

  John S Shawe-Taylor · Massimiliano Pontil · Nicolò Cesa-Bianchi · Emine Yilmaz · Chris Watkins · Sebastian Riedel · Marko Grobelnik

Hyatt Hotel, Shoreline

- **Collaborate & Communicate: An Exploration and Practical Skills Workshop That Builds On The Experience of AIML Experts Who Are Both Successful Collaborators and Great Communicators**

  Katherine Gorman

Grand Ballroom B

- **Learning with Limited Labeled Data: Weak Supervision and Beyond**

  Isabelle Augenstein · Stephen Bach · Eugene Belilovsky · Matthew Blaschko · Christoph Lampert · Edouard Oyallon · Emmanouil Antonios Platanios · Alexander Ratner · Christopher Ré

Room S-4

- **Emergent Communication Workshop**

  Jakob Foerster · Igor Mordatch · Angeliki Lazaridou · Kyunghyun Cho · Douwe Kiela · Pieter Abbeel

Hall A

- **Deep Learning: Bridging Theory and Practice**

  Sanjeev Arora · Maithra Raghu · Ruslan Salakhutdinov · Ludwig Schmidt · Oriol Vinyals

Room 203

- **Learning Disentangled Features: from Perception to Control**

  Emily Denton · Siddharth Narayanaswamy · Tejas Kulkarni · Honglak Lee · Diane Bouchacourt · Josh Tenenbaum · David Pfau

Hall C

- **Bayesian Deep Learning**

  Yarin Gal · José Miguel Hernández-Lobato · Christos Louizos · Andrew G Wilson · Diederik P. (Durk) Kingma · Zoubin Ghahramani · Kevin P Murphy · Max Welling

Room 104-C

- **The Future of Gradient-Based Machine Learning Software and Techniques**

  Alex Wiltschko · Bart van Merriënboer · Pascal Lamblin

Hyatt Hotel, Regency Ballroom A+B+C

- **Interpreting, Explaining and Visualizing Deep Learning - Now what?**

  Klaus-Robert Müller · Andrea Vedaldi · Lars K Hansen · Wojciech Samek · Grégoire Montavon

Hyatt Hotel, Seaview Ballroom

- **Optimal Transport and Machine Learning**

  Olivier Bousquet · Marco Cuturi · Gabriel Peyré · Fei Sha · Justin Solomon

Room 204

- **BigNeuro 2017: Analyzing brain data from nano to macroscale**

  Eva Dyer · Gregory Kiar · William Gray Roncal · · Konrad P Koerding · Joshua T Vogelstein

Room 201-B

- **Aligned Artificial Intelligence**

  Dylan Hadfield-Menell · Jacob Steinhardt · David Duvenaud · David Krueger · Anca Dragan

Room 102-C

- **Synergies in Geometric Data Analysis (2nd day)**

  Marina Meila · Frederic Chazal

RoomS-1

- **Machine Learning Challenges as a Research Tool**

  Isabelle Guyon · Evelyne Viegas · Sergio Escalera · Jacob D Abernethy

Room 103 A+B

- **Medical Imaging meets NIPS**

  Ben Glocker · Ender Konukoglu · Hervé Lombaert · Kanwal Bhatia

Room 203

- **NIPS Highlights (MLTrain), Learn How to code a paper with state of the art frameworks**

  Alexandros Dimakis · Nikolaos Vasiloglou · Guy Van den Broeck · Alexander

REVIEWERS

A. Rupam Mahmood
Aaditya Ramdas
Aamir Ahmad
Aapo Hyvarinen (SAC)
Aaron Schein
Aaron Defazio
Aaron Courville (SAC)
Aasa Feragen
Abbas Kazerouni
Abdel-rahman Mohamed
Abdeslam Boularias
Abhishek Das
Abhishek Kumar
Abhradeep Guha Thakurta
Abir De
Abou-Moustafa Karim
Abram Friesen
Achim Rettinger
Adam Lelkes
Adam Kalai
Adam Pocock
Adam Scibior
Adam White
Adam Charles
Adam Lerer
Adams Wei Yu
Adish Singla
Adith Swaminathan
Aditya Bhaskara
Aditya Menon
Aditya Gilra
Adji Dieng
Adler Perotte
Adrian Weller
Adriana Romero
Afshin Rostamizadeh
Agata Lapedriza
Aharon Bar-Hillel
Ahmad Shabbar Kazmi
Ahmed El Alaoui
Aika Terada
Aishwarya Agrawal
Akshay Krishnamurthy
Alain Rakotomamonjy
Alan Malek
Alan Stocker
Albert Berahas
Alberto Albiol
Alberto Bietti
Alborz Geramifard
Aldo Faisal
Alec Radford
Aleix Martinez
Alejandra Quiros-Ramirez
Alekh Agarwal
Aleksandrs Slivkins
Alessandra Tosi
Alessandro Rudi
Alex Graves
Alex Wein
Alex Gittens
Alex Thiery
Alex Alemi
Alex Peysakhovich
Alex Beutel
Alexander Ecker
Alexander Schwing
Alexander Ihler
Alexander Rush
Alexander Kirillov
Alexandra Carpentier
Alexandre Allauzen
Alexandre Gramfort
Alexandre Passos
Alexandre Lacoste
Alexandre Defossez
Alexandre Proutiere
Alexandros Dimakis
Alexandros Kalousis
Alexandros Beskos
Alexandru Niculescu-Mizil
Alexey Pozdnukhov
Alexey Dosovitskiy
Alexis Conneau
Alfredo Vellido
Ali Cemgil
Ali Zarezade
Alina Beygelzimer (SAC)
Alireza Makhzani
Allison Chaney
Alon Cohen
Alon Gonen
Alona Fyshe
Alp Kucukelbir
Amadou Ba
Amar Shah
Amaury Habrard
Ambedkar Dukkipati
Ambuj Tewari
Ameesh Makadia
Ameet Talwalkar
Amelia Perry
Amir Salavati
Amir Globerson
Amir Sani
Amir Saffari
Amir-massoud Farahmand
Amit Sharma
Amit Deshpande
Amjad Mahayri
Amr Ahmed
Ananda Suresh
Anastasia Pentina
Anastasia Podosinnikova
Anastasios Zouzias
Anastasios Kyrillidis
Anastasis Georgoulas
Andre Altmann
Andre Marquand
Andre Barreto
Andrea Albarelli
Andrea Torsello
Andrea Locatelli
Andrea Vedaldi
Andrea Frome
Andrea Passerini
Andreas Damianou
Andreas Lehrmann
Andreas Ruttor
Andreas Veit
Andreas Stuhlmüller
Andreas Vlachos
Andreas Geiger
Andreea Gane
Andrei Rusu
Andrej Risteski
Andrej Gisbrecht
Andres Munoz
Andres Masegosa Arrendondo
Andres Alvarez
Andrew Miller
Andrew Lan
Andrew Owens
Andrew Rabinovich
Andrew McCallum
Andrew Maas
Andrew Saxe
Andrew Holbrook
Andrew Wilson
Andrew Dai
Andrew Cotter
Andriy Mnih
Anelia Angelova
Ang Li
Angela Yu
Angela Schoellig
Anh Nguyen
Anima Anandkumar
Anirban Roychowdhury
Anitha Kannan
Ankan Saha
Ankit Singh Rawat
Ankur Parikh
Anna Goldenberg
Annalisa Barla
Anne Auger
Anoop Cherian
Anoop Korattikara
Anqi Liu
Anru Zhang
Anshul Kundaje
Anshumali Shrivastava
Antoine Bordes (SAC)
Anton Osokin
Antoni Chan
Antonio Artes
Antonio Torralba
Anup Rao
Aphinyanaphongs Yindalon
Aram Galstyan
Arash Vahdat
Aravindan Vijayaraghavan
Ari Pakman
Ariadna Quattoni
Arjun Jain
Arjun Chandrasekaran
Armand Joulin
Arnak Dalalyan
Aron Culotta
Arthur Mensch
Arthur Guez
Arthur Tenenhaus
Arthur Szlam
Arthur Choi
Arto Klami
Arun Narayanan
Arya Mazumdar
Aryeh Kontorovich
Asaf Weinstein
Ashesh Jain
Ashia Wilson
Ashish Vaswani
Ashok Cutkosky
Ashwin Kalyan
Ashwinkumar Badanidiyuru
Asim Kadav
Asja Fischer
Asli Celikyilmaz
Aswin Sankaranarayanan
Ata Kaban
Atil Iscen
Atilim Gunes Baydin
Atsuyoshi Nakamura
Audrunas Gruslys
Aurelie Lozano
Aurelien Lucchi
Aurko Roy
Aurélien Bellet
Austin Benson
Aviv Tamar
Avrim Blum
Awais Athar
Ayan Acharya
Ayan Chakrabarti
Aysegul Dundar
Babis Tsourakakis
Badrul Sarwar
Balaji Krishnapuram
Balaji Lakshminarayanan
Balaraman Ravindran
Balazs Csaji
Balazs Szorenyi
Balu Sivan
Bamdev Mishra
Bang Vu
Barbara Engelhardt
Barbara Hammer
Barbara Plank
Barlas Oguz
Barna Saha
Barnabas Poczos
Barret Zoph
Bart Selman
Basura Fernando
Been Kim
Behnam Neyshabur
Behrooz Ghorbani
Behrouz Behmardi
Behzad Golshan
Ben Calderhead
Ben Glocker
Ben Poole
Benigno Uria
Benjamin Rubinstein
Benjamin Rosman
Benjamin Marlin
Benjamin Van Roy
Benjamin Guedj
Bernard Ghanem
Bernard Ng
Bernardino Romera-Paredes
Bernardo Pires
Bert Huang
Bert De Brabandere
Bertrand Thirion
Beyza Ermis
Bharath Hariharan
Bharath Sriperumbudur
Bhuvana Ramabhadran
Bilal Piot
Bill Triggs
Bin Dong
Bin Shen
Bing Xu
Bing Xiang
Bingni Brunton
Biwei Huang
Bjoern Jensen
Bjoern Schuller
Bjoern Menze
Bjoern Andres
Bo Chen
Bo Liu
Bo Dai
Bo Li
Bo Waggoner
Bo Zhao
Bo Xie
Boaz Nadler
Bob Price
Bobby Jaros
Bohyung Han
Bojan Pepik
Boqing Gong
Borja Balle
Bowei Yan
Bowen Xu
Bowen Zhou
Brahim Chaib-Draa
Brandon Malone
Branislav Kveton
Brendan O'Donoghue
Brian Patton
Brian Milch
Brian Ziebart
Brian Kingsbury
Brian Mcwilliams
Brian Mcfee
Brian Kulis
Brian Roark
Brijnesh Jain
Brooks Paige
Brunel Victor-Emmanuel
Bruno Olshausen
Bruno Ribeiro
Bryan Tripp
Bryan Russell
Bryant Chen
Byron Wallace
Çağlar Gülçehre
Caiming Xiong
Camille Couprie
Can Le
Canyi Lu
Carl Henrik Ek
Carl-Johann Simon-Gabriel
Carlo Ciliberto
Carlos Stein
Carlos Fernandez-Granda
Carlotta Domeniconi
Carmen Vidaurre
Cassio De Campos
Catalin Ionescu
Cecile Capponi
Cedric Archambeau
Chandrashekar Lakshm Narayanan
Chang Xu
Changhe Yuan
Changho Suh
Changyou Chen
Chansoo Lee
Chao Chen
Chao Qian
Chaohui Wang
Chaouki Regoui
Charalampos Mavroforakis
Charles Bouveyron
Charles Blundell
Charles-Alban Deledalle
Charlie Tang
Chelsea Finn
Chen Yanover
Cheng Tang
Cheng Zhang
Cheng Soon Ong
Chengtao Li
Chenxin Ma
Chi Jin
Chicheng Zhang
Chien-Ju Ho
Ching-pei Lee
Chinmay Hegde
Chiranjib Bhattacharya
Chiyuan Zhang
Chloé-Agathe Azencott
Cho-Jui Hsieh
Chong Wang
Chongxuan Li
Chris Lucas
Chris Eliasmith
Chris Russell
Chris Oates
Chris Pal
Chris Williams
Chris Junchi Li
Christian Szegedy
Christian Osendorfer
Christian Igel
Christian Steinruecken
Christian Gagne
Christian Wolf
Christian Naesseth
Christina Lioma
Christof Angermueller
Christoph Lippert
Christoph Dann
Christoph Lampert
Christoph Lassner
Christophe Giraud-Carrier
Christophe Devleeschouwer
Christophe Denis
Christopher Frantz
Christopher Yau
Christopher Musco
Christopher Maddison
Christos Boutsidis
Christos Dimitrakakis
Chuan-Yung Tsai
Chun-Nam Yu
Chung-chieh Shan
Chunhua Shen
Cicero Nogueira dos Santos
Claire Vernade
Claire Monteleoni (SAC)
Claudio Gentile (SAC)
Clement Levrard
Clément Calauzènes
Colin Raffel
Constantin Rothkopf
Constantine Caramanis
Cordelia Schmid
Corinna Cortes (SAC)
Cosmin Paduraru
Cristian Sminchisescu
Cristina Savin
Cristobal Guzmàn
Csaba Szepesvari
Cuong Nguyen
Cynthia Rudin
Cyril Goutte
D. Sculley
Daan Christiaens
Dacheng Tao
Dahua Lin
Daichi Mochihashi
Dale Schuurmans (SAC)
Damian Roqueiro
Dan Rosenbaum
Dan Holtmann-Rice
Dan Sheldon
Dan Lizotte
Dan Feldman
Dan Ventura
Dan Yamins
Dan Goldwasser
Dan Stowell
Dan Garber
Dani Yogatama
Daniel Povey
Daniel Hendrycks
Daniel Russo
Daniel Freeman
Daniel Hernandez Lobato
Daniel Rudolf
Daniel Vainsencher
Daniel Soudry
Daniel Hsu
Daniel Neil
Daniel Cremers
Daniel Lowd
Daniel Neill
Danilo Rezende
Danny Tarlow
Danqi Chen
Daoqiang Zhang
Daozheng Chen
Dario Garcia-Garcia
Dave Moore
David Kale
David Wingate
David Jensen
David Sussillo
David Leslie
David Warde-Farley
David Forsyth
David Belanger
David Knowles
David Duverle
David Andrzejewski
David Carlson
David Rosenberg
David Ginsbourger
David Greenberg
David Mimno
David Pfau
David Krueger
David Pal
David Silver
David Wipf
David Balduzzi
David Mcallester
David Eigen
David Gleich
David Hofmeyr
David Duvenaud
David Woodruff
David Sontag
David Salinas
David Kelley
Davide Chicco
Dawen Liang
De-Chuan Zhan
Dean Bodenham
Debadeepta Dey
Debarghya Ghoshdastidar
Debing Zhang
Deepak Venugopal
Demian Battaglia
Deng Cai
Deniz Erdogmus
Dennis DeCoste
Devon Hjelm
Deyu Meng
Dhruv Mahajan
Dhruv Batra
Di He
Diane Bouchacourt
Diego Marcheggiani
Dilan Gorur
Dilip Krishnan
Dimitri Palaz
Dimitri Yatsenko
Dimitrios Milios
Dimitris Papailiopoulos
Ding-Xuan Zhou
Dinh Phung
Dino Oglic
Dit-Yan Yeung
Djalel Benbouzid
Djallel Bouneffouf
Dmitri Chklovskii
Dmitry Pechyony
Dmitry P. Vetrov
Dmitry Malioutov
Doina Precup
Dominik Janzig
Dominique Chu
Dongryeol Lee
Dongwoo Kim
Dotan Di Castro
Doug Downey
Dougal Sutherland
Dougal Maclaurin
Douwe Kiela
Dragomir Anguelov
Dumitru Erhan
Dustin Tran
Dylan Foster
Dzmitry Bahdanau
Edith Cohen
Edo Liberty
Edouard Oyallon
Edouard Pauwels
Edouard Grave
Eduardo F. Morales
Edward Grefenstette
Edwin Bonilla
Efstratios Gavves
Ehsan Elhamifar
Ehsan Adeli
Ehsan Variani
Eiji Takimoto
Elad Mezuman
Elad Ganmor
Elad Eban
Elad Hazan
Elena Zheleva
Eleni Vasilaki
Elias Bareinboim
Elias Khalil
Elisa Ricci
Ellen Vitercik
Elliot Crowley
Emanuele Olivetti
Emile Richard
Emile Contal
Emilie Morvant
Emilija Perkovic
Emilio Parisotto
Emily Denton
Emily Pitler
Emma Brunskill
Emmanuel Bengio
Emmanuel Abbe
Ender Konukoglu
Eran Mukamel
Eric Xing (SAC)
Eric Nalisnick
Eric Moulines
Eric Balkanski
Eric Cosatto
Erik Kruus
Erik Mcdermott
Erik Talvitie
Erik Linstead
Erik-Jan van Kampen
Ernesto De Vito
Ersin Yumer
Ethan Fang
Ethan Elenberg
Etienne Roquain
Eugene Belilovsky
Eugenio Culurciello
EunJee Lee
Eunho Yang
Eva Dyer
Evangelos Theodorou
Evgeny Burnaev
Evimaria Terzi
Eyke Hullermeier
Fabian Gieseke
Fabian Wauthier
Fabian Sinz
Fabian Theis
Fabian Pedregosa
Fabio A. Gonzalez
Fabio Anselmi
Fabio Cozman
Fabio Massimo Zanzotto
Fabrizio Costa
Fahad Shah
Fang Zhao
Fang Han
Fangpo Wang
Farhad Pourkamali-Anaraki
Federica Bogo
Fei Tian
Fei Sha
Feiping Nie
Felipe Llinares López
Felix Biessmann
Felix Hieber
Felix Reinhart
Felix Xinnan Yu
Feng Ruan
Ferenc Huszar
Fernando Perez-Cruz
Finale Doshi-Velez
Finn Kuusisto
Firas Hamze
Florence d'Alché-Buc
Florent Krzakala
Florian Stimberg
Florian Metze
Florian Yger
Florin Popescu
Francesc Moreno-Noguer
Francesco Casale
Francesco Orabona
Francis Bach (SAC)
Francisco J. R. Ruiz
Franco Pestilli
Francois Schnitzler
Francois Laviolette
Frank Wood
Frank Broz
Frank Dondelinger
Frank Hutter
Frank-Michael Schleif
Frans de Ruiter
Frans Oliehoek
François Fleuret
Frederic Koriche
Fredrik Johansson
Furong Huang
Fuxin Li
Gabriel Synnaeve
Gabriel Dulac-Arnold
Gael Varoquaux
Gal Chechik
Gal Dalal
Gamaleldin Elsayed
Gang Niu
Gao Huang
Gareth Peters
Gary Cottrell
Gaurav Sharma
Gautam Dasarathy
Gavin Whitaker
Gavin Taylor
Gavin C. Cawley
Gediminas Luksys
Gemma Roig
Geoffrey Roeder
Geoffrey Irving
Georg Martius
George Deligiannidis
George Toderici
George Azzopardi
George Papandreou
George Philipp
George Chen
George Michailidis
Georgia Gkioxari
Georgios Theocharous
Gerald Tesauro
Gerald Quon
Gerard Pons-Moll
Gergely Neu
Gerhard Neumann
Gilad Lerman
Gilles Blanchard
Gilles Louppe
Gilles Gasso
Giorgio Patrini
Giovanni Zappella
Giuseppe Jurman
Glenn Fung
Golan Pundak
Gonzalo Martinez-Munoz
Gonzalo Mena
Graham Neubig
Graham Taylor
Greg Shakhnarovich
Greg Mori
Greg Wayne
Gregoire Montavon
Gregory Rogez
Gregory Kahn
Grigory Yaroslavtsev
Grzegorz Swirszcz
Guang Cheng
Guang-Tong Zhou
Guangcan Liu
Guido Montufar Cuartas
Guillaume Wisniewski
Guillaume Alain
Guillaume Desjardins
Guillaume Bouchard
Guillaume Obozinski
Gungor Polatkan
Gunnar Raetsch (SAC)
Guo-Jun Qi
Guoguo Chen
Gustau Camps-Valls
Guy Van den Broeck
Gwenaelle Mabon
Gyorgy Turan
Hachem Kadri
Hado Van Hasselt
Hagen Soltau
Hai Leong Chieu
Haimonti Dutta
Haipeng Luo
Haiping Lu
Haiqin Yang
Hakan Grahn
Hamdi Dibeklioglu
Hamed Hassani
Hamed Valizadegan
Hamid Rabiee
Hang Su
Hang Li
Hanie Sedghi
Hanjun Dai
Hanli Wang
Hannes Nickisch
Hans Kersting
Hanxiao Liu
Harish Ramaswamy
Harm Van Seijen
Harm De Vries
Hartmut Neven
Hassan Ashtiani
He He
Heeyoul Choi
Heiko Strathmann

REVIEWERS

Heng Luo
Heng Huang
Hengshuai Yao
Herke van Hoof
Herve Lombaert
Hesam Sagha
Heung-Il Suk
Hexiang Hu
Hideitsu Hino
Hidekazu Oiwa
Himabindu Lakkaraju
Hiroto Saigo
Hoang Le
Hoda Heidari
Holakou Rahmanian
Holger Schwenk
Honglak Lee
Hongseok Yang
Hongyang Zhang
Hongyi Zhang
Hongyu Su
Hongyu Guo
Hongyuan Wang
Hongzhou Lin
Hossein Azari
Hossein Mobahi
Hsiang-Fu Yu
Hsuan-Tien Lin
Huahua Wang
Huan Wang
Hugh Salimbeni
Hugo Jair Escalante
Hugo Van hamme
Hugo Larochelle (SAC)
Hui Lin
Hung Bui
Huy Nguyen
Hyokun Yun
Hyun Oh Song
Iain Murray
Ian Goodfellow
Ian Osband
Ian Stevenson
Ian En-Hsu Yen
Iasonas Kokkinos
Ichigaku Takigawa
Ichiro Takeuchi
Igor Melnyk
Igor Mordatch
Il Park
Ilias Diakonikolas
Ilja Kuzborskij
Ilker Birbil
Ilya Shpitser
Ilya Loshchilov
Inderjit Dhillon (SAC)
Inmar Givoni
Ioannis Tsamardinos
Ioannis Mitliagkas
Ion Muslea
Irena Koprinska
Irene Rodriguez-Lujan
Irene Waldspurger
Irina Rish (SAC)
Irina Higgins
Irwin King
Isabel Valera
Issei Sato
Itamar Katz
Ivan Titov
Ivo Danihelka
Ivor Tsang
Izhak Shafran
Jaakko Peltonen
Jaan Altosaar
Jack Poulson
Jackie Chi Kit Cheung
Jackson Gorham
Jacob Abernethy
Jacob Steinhardt
Jacob Andreas
Jacob Eisenstein
Jaegul Choo
Jaesik Choi
Jake Gardner
Jakob Macke
Jakob Foerster
Jakob Verbeek
Jakub Marecek
Jalal Fadili
Jamal Atif
James Davidson
James Atwood
James McInerney
James Hays
James Bergstra
James Cussens
James Hensman
James Martens
James Saunderson
James Kwok
James Wright
James Sharpnack
Jamie Morgenstern
Jamie Kiros
Jan Ramon
Jan Stühmer
Jan Gasthaus
Jan Wegner
Jan Chorowski
Jan-Peter Calliess
Jan-Willem van de Meent
Janardhan Rao Doppa
Jane Wang
Jarvis Haupt
Jascha Sohl-Dickstein
Jason Pazis
Jason Corso
Jason Weston
Jason Eisner
Jason Yosinski
Jason Lee
Jason Rock
Jason Pacheco
Jasper Snoek
Javad Azimi
Javier Romero
Javier González
Jay Pujara
Jayadev Acharya
Jayashree Kalpathy-Cramer
Jayesh Gupta
Jean Honorio
Jean-Baptiste Tristan
Jean-Marc Odobez
Jean-Philippe Vert
Jeff Bilmes (SAC)
Jeff Donahue
Jeff Schneider
Jeffrey Regier
Jeffrey Pennington
Jenna Wiens
Jennifer Wortman Vaughan
Jennifer Listgarten
Jennifer Dy
Jennifer Neville (SAC)
Jennifer Healey
Jens Kober
Jeremiah Deng
Jeremie Mary
Jeremy Heng
Jeremy Jancsary
Jeremy Weiss
Jerome Revaud
Jerry Zhu (SAC)
Jes Frellsen
Jesse Krijthe
Jesse Engel
Jesse Davis
Jessica Hamrick
Jessica Hwang
Jesus Malo
Jesus Cid-Sueiro
Ji He
Ji Liu
Jia Deng
Jiajun Wu
Jian Peng
Jian Li
Jianbo Shi
Jianfei Chen
Jianhui Chen
Jianshu Chen
Jianwei Yang
Jianxin Wu
Jiasen Lu
Jiasen Yang
Jiashi Feng
Jiayu Zhou
Jie Liu
Jie Wang
Jie Shen
Jie Chen
Jihun Hamm
Jilles Vreeken
Jim Huang
Jimei Yang
Jimeng Sun
Jimmy Ba
Jimmy Vandel
Jinbo Bi
Jing Lei
Jingdong Wang
Jingrui He
Jingu Kim
Jingyi Yu
Jinwoo Shin
Jinyan Guan
Jivko Sinapov
Jo-Anne Ting
Joachim Buhmann
Joan Bruna
Joao Carreira
Joaquin Vanschoren
Joe Austerweil
Joelle Pineau
Joey Tianyi Zhou
Johan Suykens
Johan Ugander
Johanni Brea
John Agosta
John Wright
John Dickerson
John Lee
John Platt
John Hershey
John Schulman
John Blitzer
John Langford (SAC)
John Duchi
John Smith
John Paisley
Johnathan Rosenblatt
Jon Shlens
Jonas Richiardi
Jonathan LeRoux
Jonathan Hunt
Jonathan Scarlett
Jonathan Masci
Jonathan Huang
Jonathan Tompson
Jonathan May
Jonathan Huggins
Jonathan Robert Ullman
Jonathan Pillow
Jonathan Berant
Joni Pajarinen
Joost van Amersfoort
Jorg Bornschein
Joris Pelemans
Jorn Jacobsen
Joschka Boedecker
Jose Enrique Munoz De Cot
Jose Oramas M.
Jose Martin-Guerrero
Jose Alvarez
Jose Miguel Hernandez-Lobato
Josef Sivic
Joseph Salmon
Joseph Modayil
Joseph Lim
Joshua Loftus
Joshua Griffin
Josiah Hanna
Josip Djolonga
Joyce Whang
Juan Caicedo
Juan Carlos Niebles
Juergen Schmidhuber
Juergen Gall
Juho Lee
Juhyun Park
Julia Vogt
Julian Ibarz
Julian McAuley
Julie Nutini
Julien Chiquet
Julien Mairal
Jun Zhu
Jun Suzuki
Jun Sakuma
Jun Xu
Jun Liu
Jun Zhu
Jun-Kun Wang
Jun-ichiro Hirayama
Junhong Lin
Junhyuk Oh
Junier Oliva
Junpei Komiyama
Junwei Han
Junya Honda
Junyang Qian
Junyoung Chung
Junyuan Xie
Junzhou Huang
Jure Leskovec
Jurgen Van Gael
Justin Johnson
Justin Bayer
Justin Solomon
Justin Domke
Justin Thaler
Jyotishka Datta
Jyrki Kivinen
Kacper Chwialkowski
Kai Zhang
Kai Fan
Kai Puolamaki
Kai-Wei Chang
Kamalika Chaudhuri
Kamyar Azizzadenesheli
Kangwook Lee
Kannan Achan
Karen Sachs
Karen Simonyan
Karl Stratos
Karol Gregor
Karsten Borgwardt
Karthik Mohan
Karthik Sridharan
Karthikeyan Ramamurthy
Karthikeyan Shanmugam
Kashif Javed
Kashif Zafar
Kate Saenko
Kate Jeong Lee
Katerina Fragkiadaki
Katerina Gkirtzou
Katherine Heller (SAC)
Katherine Kinnaird
Katja Hofmann
Kaushik Sinha
Kazuho Watanabe
Ke Jiang
Ke Sun
Kee-Eung Kim
Keerthiram Murugesan
Keisuke Yamazaki
Keith Hall
Kelvin Xu
Ken Caluwaerts
Kenji Fukumizu
Kenneth Latimer
Kevin Swersky
Kevin Roth
Kevin McGuinness
Kevin Small
Kevin Shih
Kevin Smith
Kevin Jamieson
Keyan Ghazi-Zahedi
Keyon Vafa
Kfir Y. Levy
Khoat Than
Kian Ming Adam Chai
Kihyuk Sohn
Kimon Fountoulakis
Kinjal Basu
Kirthevasan Kandasamy
Kishore Konda
Klaus-Robert Muller
Koby Crammer
Kohei Hayashi
Kohei Hatano
Koji Tsuda
Konstantinos Rematas
Konstantinos Bousmalis
Koray Kavukcuoglu
Kosta Derpanis
Kota Yamaguchi
Kriste Krstovski
Kristen Grauman (SAC)
Kristiaan Pelckmans
Krzysztof Dembczynski
Krzysztof Choromanski
Kshipra Bhawalkar
Kuang-chih Lee
Kui Tang
Kun Zhang
Kunal Talwar
Kush R. Varshney
Kwang In Kim
Kwang-Sung Jun
Kyle Cranmer
Kyunghyun Cho
L.A. Prashanth
Laetitia Chapel
Laetitia Papaxanthos
Laiwan Chan
Lam Nguyen
Lamiae Azizi
Lan Du
Lars Buesing
Lars Kai Hansen
Lars Mescheder
Laura Igual
Laura Sevilla
Laura Dietz
Laurens van der Maaten
Laurent Sifre
Lavi Shpigelman
Le Song
Lea Duncker
Lei Han
Lenka Zdeborova
Leo Boytsov
Leon Gatys
Leon Bottou
Leonel Rozo
Leonid Sigal
Leonid Razoumov
Leslie Pack Kaelbling
Lester Mackey
Li Cheng
Li Deng
Li Wang
Li Yao
Li Zhang
Li-Lun Wang
Liang Huang
Liang Zhang
Liang Lu
Liang-Chieh Chen
Liangliang Cao
Lihong Li
Lijun Zhang
Limin Li
Ling Chen
Lingpeng Kong
Lingxi Xie
Linhong Zhu
Linli Xu
Liping Liu
Liva Ralaivola
Liwei Wang
Lixin Duan
Locatello Francesco
Lorenzo Orecchia
Lorenzo Rosasco
Loris Bazzani
Louis Wehenkel
Lu Cheng
Luca Martino
Luca Rossi
Lucas Drumond
Ludovic Denoyer
Ludovic Arnold
Ludwig Schmidt
Luigi Malagò
Lukas Balles
Lukasik Michal
Lukasz Kaiser
Luke Vilnis
Luke Dickens
Lyle Muller
M. Pawan Kumar
Madalina Fiterau
Madeleine Udell
Magnus Rattray
Mahdi Fard
Mahesan Niranjan
Mahito Sugiyama
Maithra Raghu
Maja Rudolph
Makoto Yamada
Mandar Dixit
Maneesh Singh
Manohar Paluri
Manon Kok
Manuel Gomez Rodriguez
Manuel Lopes
Manzil Zaheer
Manzil Zaher
Marc Lanctot
Marc Deisenroth
Marc Lelarge
Marc'Aurelio Ranzato (SAC)
Marcel Van Gerven
Marcelo Fiori
Marco Lorenzi
Marco Loog
Marco Pedersoli
Marco Cuturi
Marco Cristani
Marco Baroni
Marco Duarte
Marco Tulio Ribeiro
Marcus Brubaker
Marcus Rohrbach
Marcus Liwicki
Marek Petrik
Marek Grzes
Maren Mahsereci
Margareta Ackerman
Margarita Kotti
Maria Lomeli
Maria-Florina Balcan (SAC)
Mariano Rivera Meraz
Marina Meila (SAC)
Mario Figueiredo
Mario Boley
Mario Marchand
Mario Fritz
Marius Leordeanu
Mark Schmidt
Mark Davenport
Mark Kozdoba
Mark Herbster
Markus Heinonen
Martha White
Martin Takac
Martin Slawski
Martin Renqiang Min
Martin Jaggi
Martin Boldt
Martin Arjovsky
Martin Zinkevich
Maruan Al-Shedivat
Marwa El Halabi
Maryam Esmaeili
Maryam Fazel
Marzyeh Ghassemi
Masaaki Imaizumi
Masayuki Karasuyama
Masrour Zoghi
Massih-Reza Amini
Massimo Rosa
Massimo Piccardi
Matan Gavish
Matey Neykov
Mathias Niepert
Mathias Berglund
Mathieu Sinn
Mathieu Aubry
Mathieu Blondel
Mathieu Salzmann
Mathieu Lerasle
Matias Quiroz
Matsumoto Yuji
Matt Johnson
Matt Hoffman
Matt Kusner
Matthew Hauknecht
Matthew Blaschko
Matthew Golub
Matthew Riemer
Matthew Walter
Matthias Hein
Matthias Seeger
Matthieu Cord
Matthieu Vignes
Matthieu Geist
Matthäus Kleindessner
Matti Vihola
Mattias Villani
Mauricio Villegas
Mauricio Alvarez
Maurilio Gutzeit
Maurizio Filippone
Max Jaderberg
Maxim Panov
Maxim Raginsky
Maxime Sangnier
Maxime Bouton
Maxime Berar
Maximilliam Igl
Maximillian Nickel
Maxwell Libbrecht
Maya Gupta
Meelis Kull
Meg Mitchell
Meghana Kshirsagar
Mehdi Mirza
Mehmet Gönen
Mehrdad Farajtabar
Mehrdad Mahdavi
Mehreen Saeed
Meisam Razaviyayn
Melanie Schmidt
Melanie Pradier
Melih Kandemir
Melissa Smith
Mengdi Wang
Mengye Ren
Mert Pilanci
Mesrob I. Ohannessian
Miao Xu
Miao Liu
Miaomiao Zhang
Micha Andriluka
Michael Moller
Michael Mathieu
Michael Mahoney
Michael Houle
Michael Cogswell
Michael Eickenberg
Michael Hirsch
Michael Kaisers
Michael James
Michael Sapienza
Michael Wick
Michael Stark
Michael Zhang
Michael Oliver
Michael Cohen
Michael Hughes
Michael Mandel
Michael Andersen
Michael Hofmann
Michael Collins
Michael Littman
Michael Mahoney
Michael Maire
Michael Katehakis
Michael Kemmler
Michael Mozer
Michael Brueckner
Michal Drozdzal
Michal Derezinski
Michal Valko
Michal Januszewski
Michalis K. Titsias
Michel Besserve
Michele Sebag
Miguel Lazaro-Gredilla
Miguel Carreira-Perpinan
Mihaela Rosca
Mihajlo Grbovic
Mihalis Nicolaou
Mihir Jain
Mijung Park
Mikael Henaff
Mikaela Keller
Mike Schuster
Mike Hawrylycz
Mikhail Yurochkin
Mikio AOI
Miklos Racz
Min Sun
Min Lin
Min Ye
Min-Ling Zhang
Ming Li
Ming Yan
Ming-Hsuan Yang
Ming-Wei Chang
Ming-Yu Liu
Mingjun Zhong
Mingming Gong
Mingsheng Long
Mingyi Hong
Mingyuan Zhou
Mingzhang Yin
Minhua Chen
Minje Kim
Minjie Xu
Minmin Chen
Miro Dudik
Misha Denil
Misha Chertkov
Mitesh Khapra
Moez Draief
Mohamed Amer
Mohamed Tamaazousti
Mohammad Mahdian
Mohammad Azar
Mohammad Rohban
Mohammad Emtiyaz Khan
Mohammad Rastegari
Mohammad Ghavamzadeh
Mohammad Bahadori
Mohammad Saberian
Mohammad Norouzi
Mohit Iyyer
Moritz Hardt
Moritz Grosse-Wentrup
Motoaki Kawanabe
Motonobu Kanagawa
Motonobu Kanagawa
Moustapha Cisse
Mu Li
Mukund Narasimhan
Murat Kocaoglu
Murat Dundar
Mykola Pechenizkiy
Myunghwan Kim
Nadav Cohen
Naejin Kong
Nagarajan Natarajan
Nakul Verma
Nal Kalchbrenner
Nan Feng
Nan Jiang
Nan Du
Nan Li
Nan Ye
Nan Ding
Naoki Abe
Narges Razavian
Natalia Ponomareva
Natalia Neverova
Nate Kushman
Nathan Kallus
Navdeep Jaitly
Ndapandula Nakashole
Neela Sawant
Negar Kiyavash
Neil Zeghidour
Nelly Pustelnik
Nematollah (Kayhan) Batmangheli
Nevena Lazic
Nevin Zhang
Ngoc Mai Tran
Nhat Ho
Niall Twomey
Niao He
Nic Mayoraz
Nicholas Andrews
Nicholas Ruozzi
Nicholas Carlini
Nicholas Foti
Nicolas Ballas
Nicolas Le Roux
Nicolas Chapados
Nicolas Flammarion
Nicolas Durrande
Nicolas Courty
Nicolas Dobigeon
Nicolas Papernot
Nicolas Usunier
Nicolas Thome
Nicole Mücke
Nicolo Fusi
Nicolo Cesa-Bianchi
Nicolo Colombo
Nika Haghtalab
Nikhil Devanur
Nikhil Rao
Niklas Lavesson
Niklas Wahlström
Nikos Karampatziakis
Nikunj Oza
Nils Kriege
Ning Zhang
Nir Rosenfeld
Niru Maheswaranathan
Nishant Mehta
Nitish Srivastava
Nitish Shirish Keskar
Noah Smith
Novi Quadrianto
Nuno Vasconcelos
Odalric Maillard
Ofer Dekel
Ofer Meshi
Ognjen Rudovic
Oleksandr Zadorozhnyi
Oliver Stegle
Oliver Obst
Olivier Fercoq
Olivier Pietquin
Olivier Delalleau
Olivier Sigaud
Olivier Bousquet
Omar Besbes
Omer Levy
Omid Madani
Omkar Parkhi
Omri Barak
Onaiza Maqbool
Onkar Dalal
Onno Zoeter
Onur Hamsici
Or Zuk
Oriol Vinyals
Oscar Dalmau
Oswin Krause

REVIEWERS

Ozan İrsoy
Ozgur Simsek
Ozlem Aslan
Pablo Sprechmann
Pablo G. Moreno
Padhraic Smyth (SAC)
Pallika Kanani
Pan Weike
Panagiotis Toulis
Pantelis Analytis
Paolo Frasconi
Parameswaran Raman
Paramveer Dhillon
Parikshit Ram
Paris Smaragdis
Partha Talukdar
Pascal Frossard
Pascal Germain
Pascal Denis
Pascal Vincent
Pascal Poupart
Patrick Haffner
Patrick Flaherty
Patrick Forré
Patrick Loiseau
Patrick Jähnichen
Patrick Gallinari
Patrick Jordan
Paul Smolensky
Paul Honeine
Paul Rubenstein
Paul Weng
Paul Mineiro
Paul Valiant
Paul Vernaza
Paul Wohlhart
Paul Swoboda
Pauline Tan
Pavel Tokmakov
Pavel Laskov
Pawel Swietojanski
Pedro Pinheiro
Pedro Ortega
Pedro Goncalves
Peilin Zhao
Pekka Parviainen
Pekka Marttinen
Peng Zhang
Peng Sun
Perry Groot
Peter Gehler
Peter Krafft
Peter Jacko
Peter Orbanz
Peter Stone (SAC)
Peter Flach
Peter Sadowski
Peter Richtarik
Peter Battaglia
Petr Motlicek
Philemon Brakel
Philip Bachman
Philip Thomas
Philipp Kraehenbuehl
Philipp Hennig
Philipp Geiger
Philipp Berens
Philipp Thomann
Philippe Preux
Phillip Isola
Phong Le
Pierre Geurts
Pierre Savalle
Pierre Gaillard
Pierre Bellec
Pierre Barbillon
Pierre Alquier
Pierre-Luc Bacon
Pieter-Jan Kindermans
Pin-Yu Chen
Pinar Yanardag
Ping Liu
Ping Li
Ping Li
Pinghua Gong
Piotr Mirowski
Piotr Bojanowski
Piyush Rai
Po-Wei Wang
Prajit Ramachandran
Pramod Viswanath
Pranjal Awasthi
Prateek Jain
Pratik Jawanpuria
Preethi Jyothi
Prem Gopalan
Prithiraj Sen
Pulkit Agrawal
Purnamrita Sarkar
Purushottam Kar
Qi Mao
Qiang Wu
Qiang Yang
Qianqian Xu
Qiaozhu Mei
Qibin Zhao
Qichao Que
Qihang Lin
Qin Zhang
Qing Sun
Qinqing Zheng
Qixing Huang
Quan Wang
Quan Zhang
Quan Sun
Quang Nguyen
Quanming Yao
Quanquan Gu
Quentin Paris
Quentin Berthet
Quoc Tran-Dinh
Quynh Nguyen
Rachel Tappenden
Radu Timofte
Rafael Frongillo
Ragesh Jaiswal
Raghu Bollapragada
Rahul Krishnan
Raia Hadsell
Rajat Monga
Rajesh Ranganath
Rajhans Samdani
Rajiv Khanna
Rakesh Ranjan
Ramakrishna Vedantam
Raman Arora
Ramin Raziperchikolaei
Ramprasaath Selvaraju
Ramya Vinayak
Ran Gilad Bachrach
Rana Forsati
Raphael Feraud
Raquel Urtasun (SAC)
Rashish Tandon
Rasmus Bonnevie
Rasool Fakoor
Rasul Tutunov
Raul Santos-Rodriguez
Ravi Ganti
Raviv Raich
Razvan Pascanu
Rebecca C. Steorts
Rebecca Willett
Reinhold Scherer
Remi Leblond
Remi Munos
Remi Bardenet
Remi Flamary
Remi Gilleron
Rene Vidal
Renjie Liao
Reshad Hosseini
Reza Babanezhad
Reza Sahraeian
Reza Shokri
Ricardo Silva
Ricardo Henao
Richard Wilkinson
Richard Nock
Richard Combes
Richard Socher
Ricky Loynd
Rina Panigrahy
Rishabh Singh
Rita Cucchiara
Ritwik Mitra
Rob Zinkov
Rob Cornish
Robert Legenstein
Robert Guetig
Robert Tillman
Robert D. Kleinberg
Robert Durrant
Robert Gens
Robert West
Robert Busa-Fekete
Robert Vandermeulen
Rocky Duan
Rodolphe Jenatton
Roger Frigola
Roger Grosse
Rohit Babbar
Rohit Prabhavalkar
Roi Reichart
Roi Weiss
Roi Livni
Roland Kwitt
Romain Herault
Roman Garnett
Romaric Gaudel
Romer Rosales
Ronan Collobert
Rong Ge
Rongrong Ji
Roozbeh Mottaghi
Rosanne Liu
Ross Goroshin
Roy Fox
Roy Frostig
Roy Adams
Ru He
Ruben Martinez-Cantin
Rui Ponte Costa
Ruichu Cai
Ruiqi Guo
Ruitong Huang
Ruiwen Zhang
Ruoyu Sun
Ruslan Salakhutdinov
Russ Greiner
Ruth Urner
Ryan Giordano
Ryan Turner
Ryan Adams
Ryan Rifkin
Ryohei Fujimaki
Ryota Tomioka
Sadid A. Hasan
Saehoon Kim
Saeid Haghighatshoar
Sahand Negahban
Sahil Garg
Sainbayar Sukhbaatar
Saketha Nath
Sally Goldman
Sam Livingstone
Sam Gershman
Sam Bowman
Saman Muthukumarana
Sameer Singh
Samet Oymak
Samira Kahou
Samuel Kaski
Samuel Vaiter
Samy Bengio
Sandeep Subramanian
Sander Bohte
Sander Dieleman
Sandhya Prabhakaran
Sandipan Roy
Sandor Szedmak
Sanja Fidler
Sanjiv Kumar
Sanjoy Dasgupta
Sanmi Koyejo
Santosh Vempala
Sara Mostafavi
Sara Magliacane
Sarah Bird
Sarath Chandar
Satoshi Hara
Satwik Kottur
Saurabh Gupta
Scott Yang
Scott Reed
Scott Linderman
Scott Sanner
Se-Young Yun
Sean Gerrish
Sebastian Trimpe
Sebastian Tschiatschek
Sebastian Bubeck
Sebastian Nowozin
Sebastian Seung
Sebastien Giguere
Sebastien Bratieres
Sebastien Jean
Sekhar Tatikonda
Sepideh Mahabadi
Seppo Virtanen
Serena Yeung
Sergei Vassilvitskii
Sergey Ioffe
Sergey Karayev
Sergey Levine
Sergey Bartunov
Sergio Escalera
Sergio Guadarrama
Sergios Theodoridis
Seth Flaxman
Seunghak Lee
Seunghoon Hong
Seungil You
Seungjin Choi
Sewoong Oh
Seyoung Kim
Shai Ben-David
Shakir Mohamed
Shalini Ghosh
Shan Yang
Shandian Zhe
Shang-Tse Chen
Shankar Vembu
Shanshan Wu
Shao-Yuan Li
Shaobo Han
Shaogao Lv
Shaojun Wang
Shaoting Zhang
Sharan Vaswani
Sharmodeep Bhattacharya
Shashank Srivastava
Shashank Singh
Sheeraz Ahmad
Sheng-Jun Huang
Shenghua Gao
Shengjie Wang
Shenlong Wang
Sherjil Ozair
Shiau Hong Lim
Shie Mannor
Shihao Ji
Shiming Ge
Shinichi Nakajima
Shinji Watanabe
Shipeng Yu
Shiqi Zhang
Shiqian Ma
Shivaram Kalyanakrishnan
Shixiang Gu
Shou-de Lin
Shuai Li
Shuang Li
Shuiwang Ji
Shusen Wang
Si Wu
Si Si
Siamak Ravanbakhsh
Siamak Dadaneh
Sida Wang
Siddharth Gopal
Siddharth Narayanaswamy
Sigurður Freyr Hafstein
Silvia Pintea
Silvia Chiappa
Silvio Lattanzi
Simon Lacoste-Julien
Sinan Yildirim
Sinead Williamson
Sinno Jialin Pan
Sira Ferradans
Siva Hari
Sivan Sabato
Siyu Tang
Slim Essid
Slobodan Vucetic
Smita Krishnaswamy
Snell Jake
Sohan Seth
Sokol Koco
Song Liu
Songbai Yan
Sotirios Chatzis
Soumen Chakrabarti
Soumik Sarkar
Soumith Chintala
Soumya Ghosh
Soumya Ray
Souvik Ghosh
Sreeram Kannan
Sridhar Mahadevan
Srinadh Bhojanapalli
Srinivas Turaga
Srinivasan Sriram
Sriraam Natarajan
Sriram Sankararaman
Stefan Roth
Stefan Sommer
Stefan Riezler
Stefan Wager
Stefan Lee
Stefan Depeweg
Stefan Schaal
Stefanie Jegelka
Stefano Ermon
Stefanos Poulis
Stefanos Zafeiriou
Steffen Grunewalder
Stepan Tulyakov
Stephan Zheng
Stephan Clemencon
Stephan Gouws
Stephan Mandt
Stephane Robin
Stephane Canu
Stephen Becker
Stephen Bach
Stephen Tyree
Stephen Gould
Steven Hoi
Steven Rennie
Steven Li
Stratis Ioannidis
Su-In Lee
Subhransu Maji
Sudheendra Vijayanarasimhan
Sudipto Guha
Sujith Ravi
Suman Ravuri
Sumio Watanabe
Sung ju Hwang
Sungjin Ahn
Suresh Venkatasubramanian
Suriya Gunasekar
Surya Ganguli
Sushant Sachdeva
Suvrit Sra
Sven Dähne
Svetlana Lazebnik
Syama Rangapuram
Sylvain Lamprier
Sylvain Chevallier
SÀren Hauberg
Sébastien Loustau
Sébastien Destercke
Taco Cohen
Tae Hyun Kim
Taehwan Kim
Taiji Suzuki
Takafumi Kanamori
Takanori Maehara
Takashi Takenouchi
Takayuki Osogami
Tal El-Hay
Tamara Broderick
Tameem Adel
Tamir Hazan
Tanguy Urvoy
Tanmingkui Tan
Tanya Schmah
Tao Yang
Tao Qin
Tara Sainath
Tatiana Tommasi
Teemu Roos
Tejas Kukarni
Teng Zhang
Tengyu Ma
Terrence Chen
Tetsu Matsukawa
Thang Luong
Thang Bui
Theofanis Karaletsos
Theophane Weber
Thierry Artieres
Thijs van Ommen
Thodoris Rekatsinas
Thomas Steinke
Thomas Gaertner
Thomas Hofmann
Thomas Oberlin
Thomas Schön
Thomas Liebig
Thomas Serre
Thomas Colthurst
Thomas Deselaers
Tian Tian
Tianbao Yang
Tianqi Chen
Tianyi Zhou
Tiberio Caetano
Tieyan Liu
Tim Vieira
Tim Rocktäschel
Timo Bolkart
Timothee Lacroix
Timothy Arthur Mann
Timothy Hospedales
Tin Kam Ho
Tinne Tuytelaars
Tobias Glasmachers
Tobias Domhan
Toby Hocking
Todd Hester
Todd Gureckis
Tolga Bolukbasi
Tom Diethe
Tom Schaul
Tom Claassen
Tom Finley
Tom Gunter
Tom Goldstein
Tom Nickson
Tom Erez
Tom Rainforth
Tomas Kazmar
Tomas Mikolov
Tomer Koren
Tommaso Mannucci
Tomoharu Iwata
Toon Goedeme
Tor Lattimore
Toshiyuki Tanaka
Trapeznikov Kirill
Travis Dick
Trevor Campbell
Tristan Mary-Huard
Truyen Tran
Tuan Anh Le
Tuo Zhao
Tuomas Virtanen
Ulf Brefeld
Ulrich Paquet
Umar Syed
Umut Simsekli
Uri Shalit
Urs Koster
Urun Dogan
Utkarsh Upadhyay
Uwe Schmidt
Vaishak Belle
Valentin Flunkert
Vamsi Ithapu
Varun Jog
Varun Jampani
Varun Kanade
Vasil Denchev
Vasilis Syrgkanis
Venkatesh Saligrama
Veronika Cheplygina
Vianney Perchet
Vicenç Gómez
Victor Gabillon
Vien Ngo
Vijay Peddinti
Vijay Janakiraman
Vijay Mahadevan
Vikas Sindhwani
Vikas Singh
Vikas Garg
Viktoriia Sharmanska
Vinay Jayaram
Vinay Namboodiri
Vinayak Rao
Vincent Lepetit
Vincent Guigue
Vincent Dumoulin
Vineet Chaoji
Vinod Nair
Viorica Patraucean
Viren Jain
Virginia de Sa
Virginia Smith
Visvanathan Ramesh
Vitaly Feldman
Vitaly Kuznetsov
Vittorio Murino
Viveck Cadambe
Vivek Singh
Vivek Srikumar
Vlad Mnih
Vladan Radosavljevic
Vladimir Pavlovic
Vladimir Kolmogorov
Vladimir Jojic
Vladlen Koltun
Vojtech Franc
Volkan Cevher
Volker Roth
Volker Tresp
Voot Tangkaratt
Vrancx Peter
Vu Thang
Walid Krichene
Wan Li
Wang Ling
Wanli Ma
Wee Sun Lee
Wei Chen
Wei Sun
Wei Wu
Wei Wang
Wei Gao
Wei Shi
Wei Liu
Weicong Ding
Weiwei Cheng
Weixin Li
Weiyu Xu
Wen Huang
Wenjie Pei
Wenliang Zhong
Wenlin Chen
Wenruo Bai
Wenwu Wang
Wenxin Zhou
Wenzhao Lian
Wieland Brendel
Wilker Aziz
Will Dabney
Willem Waegeman
William Fithian
William Bishop
Willie Neiswanger
Wittawat Jitkrittum
Wojciech Czarnecki
Wojciech Kotlowski
Wojciech Samek
Wray Buntine
Wu-Jun Li
Xaq Pitkow
Xavier Carreras
Xavier Boix
Xavier Giro-i-Nieto
Xavier Alameda-Pineda
Xi Li
Xi (Peter) Chen
Xiang Zhang
Xiang Wu
Xiang Bai
Xiangfeng Wang
Xiangju Qin
Xiangliang Zhang
Xiangnan Kong
Xiangyu Wang
Xiao Li
Xiao He
Xiaodan Zhu
Xiaodong Cui
Xiaodong He
Xiaojun Chang
Xiaoli Fern
Xiaoning Qian
Xiaotong Yuan
Xiaoxiao Guo
Xin Geng
Xin Guo
Xin Jiang
Xinchen Yan
Xinggang Wang
Xinghao Pan
Xingyuan Fang
Xinhua Zhang
Xinwei Deng
Xinxing Xu
Xiuyuan Lu
Xu Jia
Xu Hu
Xue-Xin Wei
Y-Lan Boureau
Yacine Jernite
Yair Weiss
Yakir Reshef
Yale Chang
Yali Wang
Yan Liu
Yan Xu
Yanan Sui
Yang Wang
Yang Yu
Yangfeng Ji
Yangqing Jia
Yangqiu Song
Yanina Shkel
Yaniv Ovadia
Yanjun Qi
Yanlong Huang
Yann Guermeur
Yann Lepoittevin
Yann Traonmilin
Yann Ollivier
Yann Dauphin
Yanwei Fu
Yanxia Zhang
Yanxun Xu
Yao Xie
Yao-Hung Tsai
Yaoliang Yu
Yarin Gal
Yaron Singer
Yaroslav Ganin
Yash Goyal
Yasin Abbasi Yadkori
Yasuhisa Fujii
Yejin Choi
Yevgeny Seldin
Yichen Wang
Yifan Wu
Yifan Sun
Yifeng Li
Yin Zheng
Yin Cheng Ng
Yin-Tat Lee
Yingyu Liang
Yingzhen Li
Yining Wang
Yinlam Chow
Yishai Shimoni
Yishay Mansour
Yishu Miao
Yisong Yue
Yixin Wang
Yizhe Zhang
Yoav Artzi
Yoni Halpern
Yoram Singer
Yoseph Barash
Yoshihiro Yamanishi
Yoshikazu Terada
Yoshinobu Kawahara
Yossi Arjevani
Youhei Akimoto
Young Hwan Chang
Youssef Mroueh
Yu Xiang
Yu Cheng
Yu Zhang
Yu-Feng Li
Yu-Xiang Wang
Yuandong Tian
Yuanming Shi
Yuanqing Lin
Yuchen Zhang
Yue Zhu
Yuening Hu
Yuh-Jye Lee
Yuhong Guo
Yuhuai Wu
Yujia Li
Yulai Cong
Yun Yang
Yung-Kyun Noh
Yunlong Feng
Yunlong Jiao
Yunus Saatci
Yunwen Lei
Yuri Grinberg
Yuta Umezu
Yutaka Sasaki
Yutian Chen
Yuting Zhang
Yuval Tassa
Yuxin Chen
Yuxin Peng
Yuyang Wang
Yuyi Wang
Yves Grandvalet
Zachary Lipton
Zaid Harchaoui
Zbigniew Wojna
Zechao Li
Zelda Mariet
Zeljko Agic
Zenglin Xu
Zeynep Akata
Zeyuan Allen-Zhu
Zhang Yu
Zhanyu Ma
Zhaohan Guo
Zhaoran Wang
Zhaowei Cai
Zhaoxiang Zhang
Zhe Hu
Zheng Wen
Zheng Wang
Zhengming Xing
Zhenguo Li
Zhenjie Zhang
Zhenwen Dai
Zhi Geng
Zhi-Hua Zhou
Zhiqiang Xu
Zhirong Yang
Zhitang Chen
Zhiwei Deng
Zhiwei Steven Wu
Zhongang Qi
Zhouchen Lin
Zhuowen Tu
Zichao Yang
Zico Kolter
Zihang Dai
Ziming Zhang
Ziyu Wang
Ziyu Guan
Zoltan Szabo
Zuzana Petrickova
hirotaka Hachiya
krishnakumar Balasubramanian
xuejun Liao

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX