Schedule: Workshop on Machine Learning and Applications at the University of Maryland
March 11, 2019

William E. Kirwan Hall, Math 3206
4176 Campus Drive
College Park, MD 20742

Lunch will be served in the Math Rotunda

8:30-9 a.m.	Breakfast
9-9:20 a.m.	Welcome
9:20-10:40 a.m.	Talks: Session I
9:20-9:40 a.m.	Max Leiserson
9:40-10 a.m.	Pratyush Tiwary
10-10:20 a.m.	Mark Fuge
10:20-10:40 a.m.	Joseph JaJa
10:40-11 a.m.	Break
11 a.m.-12 p.m.	Keynote Speaker: Kilian Weinberger, Cornell University
12-1 p.m.	Lunch (Lunch will be served in the Math Rotunda)
1-2:20 p.m.	Talks: Session II
1-1:20 p.m.	Philip Resnik
1:20-1:40 p.m.	Vanessa Frias-Martinez
1:40-2 p.m.	Naomi Feldman
2-2:20 p.m.	Jen Golbeck
2:20-2:40 p.m.	Break
2:40-4 p.m.	Talks: Session III
2:40-3 p.m.	Soheil Feizi
3-3:20 p.m.	Tudor Dumitraș
3:20-3:40 p.m.	Radu Balan
3:40-4 p.m.	Furong Huang
4-5 p.m.	Panel: "Machine Learning Outside the University"

"A Multifactorial Model of Response to Immunotherapy in Bladder Cancer Integrates Clinical, Tumor, and Immune Features"

Speaker: Max Leiserson, assistant professor in the Department of Computer Science with an appointment in the University of Maryland Institute for Advanced Computer Studies

Abstract: Checkpoint inhibitor (CPI) immunotherapy has proven to be one of the most successful new cancer therapies in decades. This success is exemplified by patients with

previously untreatable late-stage cancer showing durable clinical responses. However, in some cancer types only ~20-40% of patients respond to CPIs. Due to the cost and side effects of these drugs, discovering molecular/clinical biomarkers associated with response has become a priority.

We will present a multifactorial model of immune response to checkpoint inhibitor immunotherapy. While the most common biomarkers for response to CPIs are single molecular features (e.g. the number of non-synonymous mutations in the tumor), our model integrates data from the tumor, immune system, and the patient's clinical history.

We trained and tested this model on a small cohort of bladder cancers. Importantly, we find that each class of features (clinical, immune, and tumor) are required in order to make accurate predictions in this cohort. Further, we find that the predicted immune response is associated with improved progression-free survival, and in this cohort could have reduced the number of treated patients that did not respond to CPI by nearly 50% compared to standard biomarkers.

"Past-future Information Bottleneck Framework for Accelerated, Accurate Sampling of Molecular Systems"

Speaker: Pratyush Tiwary, assistant professor in the Department of Chemistry and Biochemistry

Abstract: The ability to rapidly learn from high-dimensional data to make reliable bets about future outcomes is crucial in many contexts. This could be a fly avoiding predators, or the retina processing gigabytes of data almost instantaneously to guide complex human actions. In this work, we draw parallels between such tasks, and the efficient sampling of complex biomolecules with hundreds of thousands of atoms.

For this we use the Predictive Information Bottleneck (PIB) framework developed and used for the first two classes of problems, and re-formulate it for the sampling of biomolecular structure and dynamics, especially when plagued with rare events [1-3]. Our method considers a given biomolecular trajectory expressed in terms of order parameters or basis functions, and uses a deep neural network to learn the minimally complex yet most predictive aspects of this trajectory, viz the PIB. This information is used to perform iterative rounds of biased simulations that enhance the sampling along the PIB to gradually improve its accuracy, directly obtaining associated thermodynamic and kinetic information. We demonstrate the method on different test-pieces, including benzene dissociation from the protein lysozyme, where we calculate the dissociation pathway and timescales slower than milliseconds.

Finally, by performing an analysis of residues contributing to the PIB, we predict the critical mutations in the system which would be most impactful on the stability of the crucial but ephemeral transition state. We believe this work marks a big step forward in the use of predictive artificial intelligence ideas for the sampling of molecules.

References:
[1] Wang, Ribeiro and Tiwary bioRxiv 507822; doi: https://doi.org/10.1101/507822
[2] Ribeiro, Bravo, Wang and Tiwary J. Chem. Phys. 2018, 149, 072301
[3] Ribeiro and Tiwary J. Chem. Theor. Comp. 2019, 15, 708

"Learning to Computationally Design Engineered Systems"

Speaker: Mark Fuge, assistant professor in the Department of Mechanical Engineering and a member of the Maryland Robotics Center and the Human-Computer Interaction Lab

Abstract: This talk will highlight our recent work on combining three scientific fields that may, at first glance, appear tangential to one another—machine learning, engineering systems design, and multi-physics simulation—in order to automatically design and explore new complex engineered systems ranging in size from atomic scale condensed matter physics/chemistry to systems the size of aircraft and cargo ships and even up to globally distributed communities of thousands of people collaborating together.

As we briefly walk through each of these size scales and see various physical phenomena emerge—e.g., how materials pack into 3-D space, how modifications to a surface of an airfoil affect fuel efficiency, how engineers collaborate together on global scales, etc.—we also see how each possesses unique mathematical structures that occur in or can be leveraged by machine learning—e.g., differential geometry, manifold subspaces, submodular functions, etc. I will discuss how those structures help us understand how humans design engineered systems and how that knowledge can help us better teach computers how to emulate that process. As a byproduct, this connection also points the way to advances in fundamental problems in computer science, like diverse bipartite matching and diverse ranking or clustering positive semi-definite matrices, among other topics.

Overall, this talk's central argument is that machine learning and engineering systems have a powerful symbiotic relationship: (1) that machine learning is not only a valuable tool for helping understand real-world engineering systems but also (2) that engineering itself helps shed light on phenomena with interesting and unique mathematical structure that can inform new algorithmic advances within machine learning not well covered by existing machine learning benchmarks. Both are trying, in complementary ways, to uncover fundamental representations of nature, of complexity, and human behavior. I argue (and attempt to briefly show) how the two fields are better together than they are apart, and that their merger provides a unique window into understanding fundamental questions of nature and human-kind in ways that are not possible otherwise.

“Brain Dynamics and Temporal Trajectories Using Reservoir Computing”

Speakers: Joseph JaJa, professor and interim chair of the Department of Electrical and Computer Engineering with an appointment in the University of Maryland Institute for Advanced Computer Studies

Abstract: Human functional Magnetic Resonance Imaging (fMRI) data are acquired while participants engage in diverse perceptual, motor, cognitive, and emotional tasks. Although data are acquired temporally, they are most often treated in a quasi-static manner. Yet, a fuller understanding of the mechanisms that support mental functions necessitates the characterization of dynamic properties.

Here, we describe an approach employing a class of recurrent neural networks called reservoir computing, and show the feasibility and potential of using it for the analysis of temporal properties of brain data. We show that reservoirs can be used effectively both for condition classification and for characterizing lower-dimensional trajectories of temporal data. Classification accuracy was approximately 90% for short clips of social interactions and around 70% for clips extracted from movie segments. Data representations with 12 or fewer dimensions (from an original space with over 300) attained classification accuracy within 5% of the full data. We hypothesize that such low-dimensional trajectories may provide signatures that can be associated with tasks and/or mental states.

The approach was applied across participants (that is, training in one set of participants, and testing in a separate group), showing that representations generalized well to unseen participants. Taken together, we believe the present approach provides a promising framework.

Keynote: “Deep Learning with Dense Connectivity”

Speaker: Kilian Q. Weinberger, associate professor in the Department of Computer Science at Cornell University and a member of the Cornell Center of Data Science for Improved Decision-Making

Abstract: Although half a decade has passed since Frank Rosenblatt’s original work on multi-layer perceptrons, modern artificial neural networks are still surprisingly similar to his original ideas.

In this talk, we will question some of their most fundamental design aspects and shed some lights on the inner workings of neural networks. Guided by the quest to reduce and understand the role of parameter redundancy, we will arrive at a novel connectivity pattern, which we refer to as DenseNets. DenseNets depart from the traditional layer-to-layer connectivity and are to-date the most efficient and accurate neural network architectures on a variety of tasks.

Finally, we will investigate the question why deep neural networks are so well suited for natural images, and provide evidence that they may linearize the underlying sub-manifold into an Euclidean feature space.

"Machine Learning for Mental Health"

Speaker: Philip Resnik, professor in the Department of Linguistics with an appointment in the University of Maryland Institute for Advanced Computer Studies

Abstract: Mental disorders are among one of the most significant global health problems we face, affecting approximately 450 million people worldwide—300 million people suffering depression, 60 million with bipolar affective disorder, 23 million people with schizophrenia and other psychoses, to name a few.

The cost of mental illness is staggering. In purely economic terms, the World Economic forum estimated the 2010 direct and indirect global cost of mental health conditions to be $2.5T. That's "T" for trillion. To calibrate, consider that in 2010, the total global expenditure for health—all of healthcare spending worldwide—was $6.5T.

Meanwhile, the human cost is incalculable. In the U.S. alone, well over a million people per year are seen in an emergency room following a suicide attempt, and more than 120 million people live in official "mental health care health professional shortage areas," mental healthcare deserts where they have a hard time finding a provider, even if they recognize that they need one. According to the National Alliance on Mental Illness, only 63% of adults with a serious mental illness received mental health services in the past year.

Language is a window into mental state. Indeed, what clinicians do when they are assessing and treating patients is largely a language analysis task, a search for relevant signal. So, what might language technology be able to do to help? In this talk I will provide a high level overview of mental health as an application area for natural language processing and machine learning, looking both at the promise and the challenges for these approaches.

"Characterization of Internal Migrant Behavior in the Immediate Post-Migration Period using Cell Phone Traces"

Speaker: Vanessa Frias-Martinez, assistant professor in the iSchool with an appointment in the University of Maryland Institute for Advanced Computer Studies

Abstract: Internal migrations have been studied using two types of approaches: macro-level and micro-level analyses. Macro-level studies are typically carried out using a combination of various survey and census datasets to model large-scale behaviors, however these models fail to provide more nuanced information about the physical or social status of the migrants. Micro approaches, which successfully use interviews and diaries to provide a window into more individual behaviors, could benefit from methods to identify novel or under-studied behaviors that should be addressed in the migration research agenda.

In this paper, we present a framework that uses information extracted from cell phone metadata to reveal internal migration behaviors that could guide or complement the research agenda of micro-level migration researchers working to understand the physical, social and psychological decision processes behind migration experiences. The proposed framework allows us to carry out micro-level analyses of internal migration with a focus on immediate post-migration behaviors and the role of pre-migration activities from two perspectives: spatial behaviors and social ties.

Ultimately, we expect our analyses to inform migration researchers of pre- and post-migration behaviors that would benefit from further qualitative analysis.

"How to Optimize Speech Perception for a Native Language"

Speaker: Naomi Feldman, associate professor in the Department of Linguistics with an appointment in the University of Maryland Institute for Advanced Computer Studies

Abstract: Speech technology for low-resource languages often requires training data from multiple high-resource languages. Yet human infants learn language in an unsupervised setting, without already knowing other languages. A large body of experimental research suggests that they do so by creating a native language perceptual "filter" that optimizes their speech perception for a particular language. This work uses tools from speech technology to create models of what the human native language perceptual filter might look like.

This is joint work with Thomas Schatz, Sharon Goldwater, and Emmanuel Dupoux.

Talk title and abstract from Jen Golbeck TBA

"Normalized Wasserstein Distance for Mixture Distributions with Applications in Adversarial Learning and Domain Adaptation"

Speaker: Soheil Feizi, assistant professor in the Department of Computer Science with an appointment in the University of Maryland Institute for Advanced Computer Studies

Abstract: Understanding proper distance measures between distributions is at the core of several learning tasks such as generative models, domain adaptation, clustering, etc.

In this talk, we focus on mixture distributions that arise naturally in several application domains where the data contains different sub-populations. For mixture distributions, established distance measures such as the Wasserstein distance do not take into account imbalanced mixture proportions. Thus, even if two mixture distributions have identical mixture components but different mixture proportions, the Wasserstein distance between them will be large. This often leads to undesired results in distance-based learning methods for mixture distributions.

In this work, we resolve this issue by introducing Normalized Wasserstein distance. The key idea is to introduce mixture proportions as optimization variables, effectively normalizing mixture proportions in the Wasserstein formulation. Using the proposed normalized Wasserstein distance, instead of the vanilla one, leads to significant gains working with mixture distributions with imbalanced mixture proportions. We demonstrate effectiveness of the proposed distance in GANs, domain adaptation, adversarial clustering and hypothesis testing over mixture of Gaussians, MNIST, CIFAR-10, CelebA and VISDA datasets.

"Research Challenges for the Security of Machine Learning"

Speaker: Tudor Dumitraș, assistant professor in the Department of Electrical and Computer Engineering with an appointment in the University of Maryland Institute for Advanced Computer Studies

Abstract: There is an emerging arms race in the field of adversarial machine learning. Recent results suggest that machine learning systems are vulnerable to a wide range of attacks; meanwhile, there are no systematic defenses.

In this talk, I will discuss what a similar arms race, which unfolded over the past 30 years in the field of malware detection, can teach us about the security of machine learning. I will use the lessons learned from malware research as a vehicle for highlighting both challenges and promising research directions for using machine learning under adversarial pressure.

"Discrete Optimization using Graph Deep Learning"

Speaker: Radu Balan, professor in the Department of Mathematics and a member of the Center for Scientific Computation and Mathematical Modeling (CSCAMM)

Abstract: This talk will present recent results using Graph Convolutive Networks to solve discrete assignment problems. Specifically the linear and quadratic assignment problems are analyzed. It is shown a well-designed network produces exact solutions for some cases. Preliminary results are obtained in the general case.

"Guaranteed Scalable Learning of Latent Tree Models"

Speaker: Furong Huang, assistant professor in the Department of Computer Science with an appointment in the University of Maryland Institute for Advanced Computer Studies

Abstract: We present an integrated approach for structure and parameter estimation in latent tree graphical models. Our overall approach follows a "divide-and-conquer" strategy that learns models over small groups of variables and iteratively merges onto a global solution. The structure learning involves combinatorial operations such as minimum spanning tree construction and local recursive grouping; the parameter learning is based on the method of moments and on tensor decompositions.

Our method is guaranteed to correctly recover the unknown tree structure and the model parameters with low sample complexity for the class of linear multivariate latent tree models which includes discrete and Gaussian distributions, and Gaussian mixtures. Our bulk asynchronous parallel algorithm is implemented in parallel and the parallel computation complexity increases only logarithmically with the number of variables and linearly with dimensionality of each variable.