New Ideas in Machine Learning Workshop
October 10, 2022
Abstracs

8:30AM – 9:00AM – Registration and Introduction

9:00AM – 9:45AM - Terry Rockafellar, University of Washington (rtr@uw.edu)

Title: Coherent Measures of Risk from Wasserstein Distances

Abstract: An important topic now in machine learning is distributional robustness as a way guarding against the training set being too special. The training set corresponds to an empirical distribution of data, but there is a perceived need do worst-case analysis with respect to some choice of nearby distributions. For this, so-called distances in the Wasserstein sense have become very popular. Of course, maximizing expectations over a set of alternative distributions is exactly the prescription for a coherent measure of risk, although researchers in machine learning have little familiarity with that theory.
In this talk, the measures of risk that arise from Wasserstein distances will be explained in more detail along with troubles about fitting them into the fundamental quadrangle of risk with its potential connections to generalized statistics and regression.

9:45AM – 10:30AM – Aaron Kim, Stony Brook University, NY (aaron.kim@stonybrook.edu)

Title: Option Pricing with Artificial Neural Network on the Tempered Stable GARCH Model

Abstract: This talk focuses on developing fast computational algorithms for option pricing and hedging using the artificial neural network. We first construct a stochastic volatility model with non-Gaussian infinitely divisible innovations. The model includes features such as fat-tailed probability distribution driven processes, variability clustering effects, and stochastic skewness & kurtosis. The option price on the model can be obtained by the Monte-Carlo simulation method because of the complexity of the model. However, the MCS method is not analytical and contains uncertainty errors. Also, it takes too much time to apply high frequency trading activity. To overcome this problem, we use the artificial neural network. We construct and train a multilayer perceptron network using large amounts of randomly generated data with different parameter sets in the MCS method. After training, it can be used to calibrate the model for market option prices. It can also be used for sensitivity analysis which is very difficult if we use the MCS method.

10:30AM-11:00AM – Break

11:00AM – 11:45AM – Douglas Johnston, Farmingdale State College, NY (douglas.johnston@farmingdale.edu)

Title: Bayesian Forecasting of Dynamic Extreme Quantiles

Abstract: In this talk, we discuss a novel Bayesian solution to forecasting extreme quantile levels that are dynamic in nature. This is an important problem in many fields of study including climatology, structural engineering, and finance. We utilize results from extreme value theory to provide the backdrop for developing a state-space model for the unknown parameters of the observed time-series. To solve for the requisite probability densities, we derive a Rao-Blackwellized particle filter and, most importantly, a computationally efficient, recursive solution. Using the filter, the predictive distribution of future observations, conditioned on the past data, is forecast at each time-step, and used to compute extreme quantile levels. We illustrate the improvement in forecasting ability, versus traditional methods, using simulations and apply our technique to the S&P 500.

11:45AM – 12:30PM – Pawel Polak, Stony Brook University, NY (pawel.polak@stonybrook.edu)

Title: CP-PINNs: Changepoints Detection in PDEs using Physics Informed Neural Networks with Total-Variation Penalty

Abstract: We consider the inverse problem for the Partial Differential Equations (PDEs) such that the parameters of the dependency structure can exhibit random changepoints over time. This can arise, for example, when the physical system is either under malicious attack (e.g., hacker attacks on power grids and internet networks) or subject to extreme external conditions (e.g., weather conditions impacting electricity grids or large market movements impacting valuations of derivative contracts). For that purpose, we employ Physics Informed Neural Networks (PINNs) -- universal approximators that can incorporate prior information from any physical law described by a system of PDEs. This prior knowledge acts in the training of the neural network as a regularization that limits the space of admissible solutions and increases the correctness of the function approximation. We show that when the true data generating process exhibits changepoints in the PDE dynamics, this regularization can lead to a complete miss-calibration and a failure of the model. Therefore, we propose an extension of PINNs using a Total-Variation penalty which accommodates (multiple) changepoints in the PDE dynamics. These changepoints can occur at random locations over time, and they are estimated together with the solutions. We propose an additional refinement algorithm that combines changepoints detection with a reduced dynamic programming method that is feasible for the computationally intensive PINNs methods, and we demonstrate the benefits of the proposed model empirically using examples of different equations with changes in the parameters. In case of no changepoints in the data, the proposed model reduces to the original PINNs model. In the presence of changepoints, it leads to improvements in parameter estimation, better model fitting, and a lower training error compared to the original PINNs model.

12:30PM-1:30PM – Lunch

1:30PM – 2:15PM – Stan Uryasev, Stony Brook University, NY (stanislav.uryasev@stonybrook.edu)
Cheng Peng, Stony Brook University, NY (cheng.peng@stonybrook.edu)

Title: Factor Model of Mixtures and Linear Regression

Abstract: This paper considers the problem of estimating conditional distribution conditioned on observing some factors. The quantile function of the conditional distribution is modeled by a mixture (linear combination) of some basis quantile functions. The weight of each basis quantile function is a nonlinear function of the factors. We use the spline function as a primary example. The spline function is adaptive to the data, and can be estimated with the model in one shot. The model calibration is formulated as a linear regression problem similar to quantile regression. It can be efficiently solved by convex and linear programming. Various types of constraints, such as cardinality of the factors in the model, and penalties can be included in the optimization problem. The quantile function has a flexible shape in both tail and body, and has an analytic expression. The calibration method can focus on the tail of the distribution by assigning higher weights to the tail confidence levels. We prove the asymptotic normality of the estimator in a special case. We also prove the equivalence of our calibration method to minimization of Continuous Probability Ranked Score (CRPS). The proposed approach is generalized to conditional distributions defined by CVaR (also known as Expected Shortfall, superquantile), expectile and other functions. The considered approach is based on Risk Quadrangle theory. We conducted numerical experiments and demonstrated the efficiency of the approach.

2:15PM – 3:00PM – Alejandro Lopez Lira, University of Florida (alejandro.lopez-lira@warrington.ufl.edu)

Title: Do Common Factors Really Explain the Cross-Section of Stock Returns

Abstract: We document challenges to the notion of a trade-off between systematic risk and expected returns when analyzing the empirical ability of stock characteristics to predict excess returns. First, we measure individual stocks' exposures to all common latent factors using an efficient high-dimensional method. These latent factors appear to earn negligible risk premia despite explaining virtually all of the common time-series variation in stock returns. Next, we use machine learning methods to construct out-of-sample forecasts of stock returns based on a wide range of characteristics. A zero-cost beta-neutral portfolio that exploits this predictability but hedges all undiversifiable risk delivers a Sharpe ratio above one with no correlation with any systematic factor, thus rejecting the central prediction of the arbitrage pricing theory.

3:00PM – 3:30PM – Break

3:30PM – 4:15PM – George Michailidis, University of Florida (gmichail@ufl.edu)

Title: Distributed Adaptive Gradient Methods for Online Optimization

Abstract: Adaptive gradient based optimization methods (Adam, Adagrad, RMSProp) are widely used in solving large scale machine learning problems including training deep learning neural networks. A number of schemes have been proposed in the literature aiming at parallelizing them, based on communications of peripheral nodes with a central node, or amongst themselves. In this presentation, we briefly review centralized adaptive gradient based algorithms and then discuss distributed variants. We discuss their convergence properties in both stochastic and deterministic settings. The algorithms are illustrated on applications, including training of deep neural networks.

4:15PM – 5:00PM – Arnaud Marsiglietti, University of Florida (a.marsiglietti@ufl.edu)

Title: Moments, concentration, and entropy of log-concave distributions

Abstract: In this talk I will present a simple mechanism, combining log-concavity and majorization in the convex order to derive moments, concentration, and entropy inequalities for random variables that are log-concave with respect to a reference measure.

5:00PM – 5:45PM – Yongpei Guan, University of Florida (guan@ise.ufl.edu)

Title: TBD

Abstract: TBD

5:30PM – 7:30PM - Reception

New Ideas in Machine Learning Workshop
October 11, 2022
Speakers

8:30AM – 9:00AM - Registration and Introduction

9:00AM – 9:45AM - Farid AitSahlia, University of Florida (farid.aitsahlia@warrington.ufl.edu)

Title: Differential and reinforcement learning in derivatives markets

Abstract: In this talk I introduce an approach that combines the speed of execution associated with differential learning together with the flexibility of reinforcement learning to price and hedge American options. With the reliance of differential machine learning on automatic adjoint differentiation, hedge parameters are computed path by path for the prices determined via simulation-driven reinforcement learning.

9:45AM – 10:30AM - Boyan Lazarov, Lawrence Livermore National Laboratory (lazarov2@llnl.gov)

Title: Risk-averse topology optimization

Abstract: This work focuses on incorporating manufacturing and operational uncertainties in topology-optimized additive manufacturing designs. Topology optimization is an iterative design process aiming to distribute material in a design domain by optimizing an objective function and fulfilling a set of constraints. The material distribution is represented by a density field taking values one at subdomains occupied with solid material and zero in the void regions. The optimization process is gradient-based, which requires relaxing the material distribution to take values between zero and one. The existence of a solution to the optimization process is ensured through filtering, and to minimize the so-called gray transition areas, we utilize smooth Heaviside projections. The Heaviside projection is often used to model geometric uncertainties in production.

Even with projection, the optimization formulation without any amendments results in designs with intermediate densities. The grey regions are often penalized by explicit penalization terms in the objective or by special material interpolation schemes. Another alternative that does not necessitate artificial penalization to the optimization is to require the design to behave robustly with respect to a uniform erosion or dilation of the solid boundary. Such an idealized model is relatively cheap and provides length scale in the optimized designs. It is based on the so-called min/max optimization formulation, where the worst-case scenario drives the optimization process. However, the model can model only uniform erosion/dilation of the design boundary. Thus, more realistic applications require the introduction of spatially varying uncertain thresholds.

Finding the worst-case scenario for these more complex models is far from trivial and significantly increases the computational cost. Therefore, the min/max formulation is often replaced with an objective easier to compute. Common choices are the weighted sum of expected value and variance or standard deviation or formulations based on the so-called First/second-order reliability methods (FORM/SORM). The main issue with the first approach is that the objective does not provide one to one map to a probability measure. The second set of approaches (FORM/SORM) relies on the behavior of the design around a single point, the so-called most probable failure point, in the stochastic space, and therefore, the design is tuned for the specified set of parameters that prevents the optimization algorithm from finding a better solution. The issue is especially pronounced in situations with multiple most probable failure points with similar probabilities of failure. The alleviate the issues mentioned above, we propose to modify the currently used optimization formulations with risk-averse measures used on a daily bases in the finance industry. These modifications require the introduction of new optimization algorithms and new solution strategies. Hence, we will present a detailed comparison between different formulations and their effect on the optimized results.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

10:30AM – 11:00AM – Break

11:00AM- 11:45AM - Artem Prokhorov, The University of Sydney (artem.prokhorov@sydney.edu.au)
Robert James, The University of Sydney (robert.james@sydney.edu.au)
Henry Leung, The University of Sydney (henry.leung@sydney.edu.au)

Title: A Machine Learning Attack on Illegal Trading

Abstract: We design an adaptive framework for the detection of illegal trading behavior. Its key component is an extension of a pattern recognition tool, originating from the field of signal processing and adapted to modern electronic systems of securities trading. The new method combines the flexibility of dynamic time warping with contemporary approaches from extreme value theory to explore large-scale transaction data and accurately identify illegal trading patterns. Importantly, our method does not need access to any confirmed illegal transactions for training. We use a high-frequency order book dataset provided by an international investment firm to show that the method achieves remarkable improvements over alternative approaches in the identification of suspected illegal insider trading cases.

11:45AM – 12:30PM – Kevin Maritato, Stony Brook University, NY (kevin.maritato@stonybrook.edu)

Title: Derivatives of Buffered CDF

Abstract: This paper considers the concept of Reduced Cumulative Probability Distribution Function (rCDF), which is also known as a Super CDF (inverse of Conditional Value at Risk (CVaR)). rCDF equals one minus Buffered Probability of Exceedance (bPOE). We introduce the concept of Reduced Probability Distribution Function (rPDF), as the derivative of rCDF. Then we describe three means of calculating the rPDF for a distribution, depending on what is known about the distribution. For functions with a closed-form formula for bPOE, we derive closed-form formulae for rPDF. We also examine using these formulae for parameter estimation with the maximum likelihood method. We then describe a method for deriving a formula for rPDF based on a numerical bPOE calculation, when there is a closed-form formula for CVaR. Finally, we give a method for numerically calculating rPDF for an empirical distribution, and compare the results of this method to those given by the other methods for known distributions.