Yufeng Zheng

yufeng.zheng@mail.utoronto.ca

EDUCATION

University of Toronto — **Ontario, Canada**
Ph.D. Candidate, Operations Management and Statistics

University of California, Berkeley — **California, United States**
MSc, Industrial Engineering & Operation Research

Shanghai Jiao Tong University — **Shanghai, China**
Bachelor of Engineering in Industrial Engineering & Operation Research

PUBLICATIONS

Robust Situational Reinforcement Learning in Face of Context Disturbances
Microsoft Research Asia
- Developed a framework called Robust Situational Markov Decision Process (RS-MDP) to address challenges in reinforcement learning when dealing with real-world tasks where context transitions can be contaminated, causing performance degradation.
- Explicitly captured possible deviations of context transitions, providing a more accurate representation of the problem compared to existing methods.
- Introduced a softmin smoothed robust Bellman operator to learn robust Q-values, allowing scalability to large context spaces.
- Demonstrated the effectiveness of the RS-MDP framework by applying it to the Soft Actor-Critic (SAC) algorithm in robot control and inventory management tasks, showcasing superior generalization and robustness compared to existing robust RL algorithms.
- Outlet: ICML, 2023. [Link]

Mind Your Step: Continuous Conditional GANs with Generator Regularization
UC Berkeley
- We suspect that the mode collapse issue of Conditional GAN may be caused by the generator not having enough smoothness. The smoothness means that the generator should output similar conditional distribution given similar conditions.
- We propose a simple generator regularization term on the GAN generator loss in the form of a Lipschitz penalty to promote smoothness.
- The regularization term will leverage the neighbor information and push the generator to generate samples with similar conditional distributions for neighboring conditions.
- Analyze the effect of the proposed regularization term and demonstrate its robust performance on a range of synthetic tasks as well as real-world conditional time series generation tasks.
- Outlet: NeurIPS SyntheticData4ML workshop, 2022. [Link]

An Adaptive Deep RL Method for Non-Stationary Environments with Piecewise Stable Context
Microsoft Research Asia
- Address the Reinforcement Learning (RL) problem with a piecewise stable unobservable environment context, such as changing terrains in robotic tasks and fluctuated bandwidth in congestion control.
- Existing works on adaptation to unknown environment contexts either assume the contexts are the same for the whole episode or assume the context variables are Markovian, which makes these methods unfit for our setting.
- Propose a Segmented Context Belief Augmented Deep (SeCBAD) RL method; with the help of the change point detection technique, our method can jointly infer the belief distribution over latent context and perform more accurate belief context inference.
- Empirically demonstrate that SeCBAD can accurately infer context segment length and outperform existing methods on a toy grid world environment, Mujoco tasks with piecewise-stable context, and a real-world bandwidth control task for real-time communications.
- Outlet: NeurIPS, 2022. [Link]

LOLA: LLM-Assisted Online Learning Algorithm for Content Experiments
University of Washington
- Developed LOLA, a framework that combines Large Language Models (LLMs) with adaptive experimentation to optimize content delivery for media firms.
- Analyzed Upworthy's dataset (17,681 A/B headline tests), achieving 82-84% accuracy in identifying engaging headlines with embedding-based and fine-tuned LLMs.
- Integrated LLMs with the Upper Confidence Bound algorithm in LOLA, enabling adaptive traffic allocation to maximize user engagement.
- Demonstrated LOLA's scalability and superior performance over standard A/B testing in experimental scenarios with limited traffic.
- Manuscript Under Major Revision for Marketing Science.

A Doubly Stochastic Simulator with Applications in Arrivals Modeling and Simulation
UC Berkeley
- Propose a new framework named DS-WGAN that integrates the doubly stochastic (DS) structure and the Wasserstein generative adversarial networks (WGAN) to model, estimate, and simulate a broad class of arrival processes with general non-stationary and random arrival rates.
- Prove consistency and convergence rate for the estimator solved by the DS-WGAN framework under a non-parametric smoothness condition.
- Regarding computational efficiency and tractability, we address a challenge in gradient evaluation and model estimation arising from the discontinuity in the simulator.
- Implement numerical experiments with synthetic and real-world data sets to demonstrate the performance of DS-WGAN. Measure performance from both statistical and operational perspectives.
- Outlet: Operation Research. [Informs]

Demand Prediction, Predictive Shipping, and Product Allocation for Large-scale E-commerce
UC Berkeley
- Explore the data set containing information about logistics, price, and order of the real business scene of nearly 130GB.
- Predict the sale of products by incorporating features like historical sales, prices, and page views. In addition to the price of an item itself, the price of other correlated items also affects sales, so we build a network to illustrate the competitive and complementary relationship among products and boost prediction accuracy through representation learning.
- Find out that the products are stored in warehouses too dense. By theoretical analysis and numerical experiment, we conclude that increasing the density of products stored in warehouses can considerably reduce transportation costs.
- Finalist: M&SOM Data-driven Research Challenge, 2018. [SSRN]

WORKING PROJECTS

Formal Math Theorem Proving via LLM and Reinforcement Learning
University of Toronto
- Use reinforcement learning to fine-tune LLMs, enhancing their capability in writing Lean code for formal theorem proving.

Enhancing Discrete Event Simulation with Advanced Language Model Techniques
UC Berkeley
- Develop a method for coding discrete event simulations using LLMs, leveraging prompt engineering, Retrieval-Augmented Generation

Revenue Management and Pricing for Machine Learning Products
University of Toronto
- Develop an algorithm to dynamically optimize ML product pricing and retraining, maximizing profits by balancing performance gains and cost savings

Analyzing Equity in Private Tutoring Markets Post-Regulation
University of Toronto
- Uses mean field game theory to analyze exam competition and assess how regulations affect fairness and access in tutoring markets, optimizing for educational equity.

INTERNS

Microsoft Research Asia
- Develop a Pretrained Video Foundation Model for Embodied Agents, focusing on generalizability and few shot learning in new real world environments.
- Implement an AI-based Control System for HVAC (Heating, Ventilation, and Air Conditioning) using Reinforcement Learning to optimize energy efficiency and operational performance.
- Execute a multi-agent Reinforcement Learning algorithm for multi-SKU inventory management, optimizing supply chain operations.

Shun Feng Express (Group) Co., Ltd
- Design and develop an Algorithm for Large-Scale Truck Scheduling, improving logistics efficiency and reducing operational costs.

COURSES AND SKILLS

Core Courses:
University of Toronto: Online and Data-Driven Decision Making (RSM3091H), Advanced Topics in Linear and Nonlinear Optimization (RSM3090H), Machine Learning for Mathematical Optimization (MIE1666H), Automated Reasoning with Machine Learning (CSC2547H)
UC Berkeley: Mathematical Programming (IDENG 262), Applied Stochastic Process (IDENG 263), Supply Chain and Logistics Management (IDENG 253), Introduction to Production Planning and Logistics Models (IDENG 250), Network Flows and Graphs (IDENG 266), Statistical Learning Theory (COMPSCI 281), Deep Reinforcement Learning (COMPSCI 285), Introduction to Machine Learning (COMPSCI 289A)
SJTU**:** Calculus, Probability and Statistics, Linear Algebra, C++ Programming, Fundamentals of Industrial Engineering, Microeconomics, Introduction to Operations Research, Logistics and Supply Chain, Production Plan and Control, Stochastic Model
Programming: PyTorch, TensorFlow, MATLAB, Gurobi, CPLEX, C++