YUDAN TANG

tangyd@uchicago.edu | tangyd.github.io | (+1) 773-406-3294 | (+86) 13967111566

EDUCATION

University of Chicago Chicago, IL, USA
M.Sc. in Statistics (Admission Scholarship) 2022 - 2024 (Expected)

· Cumulative GPA: 4.0/4.0
· Coursework: Mathematical Statistics, Transfer Learning, Distribution Theory, Causal Inference

University of Hong Kong Hong Kong SAR, China
B.Sc. in Statistics & Computer Science 2018 - 2022

· Cumulative GPA: 3.82/4.30 (Top 3%); Major GPA: 3.85/4.3 (Statistics), 4.02/4.3 (Computer Science)
· Coursework: Statistical Learning, Time Series, Linux, Artificial Intelligence, Algorithm Design

WORKING EXPERIENCE

Applied Scientist Intern, Amazon June 2023 - Sep 2023

· Incoming Applied Scientist intern in EOP Econ group at Amazon, Seattle

AI Intern, Institute of Computing Innovation, Zhejiang University May 2022 - Aug 2022

· Reproduced and modified Reinforcement Learning algorithms like Proximal Policy Optimization (PPO) and Deep Learning algorithms like Deep Factorization-Machine (Deepfm)
· Implemented new algorithms to extract effective feature set for Machine Learning with 5% improvement on various metrics like F1 and Accuracy on average. Related techniques include PPO and Neural Architecture Searching (NAS), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM)
· Written algorithms in AI apps to increase functionality on mobile phones and labelled images for AI training
· Applied the Energy function instead of the SoftMax to train and evaluate Machine Learning models, and achieve higher discrimination between In-Distribution (ID) and Out-Of-Distribution (OOD) samples

Quantitative Research Intern, Kubid Research Dec 2020 - Jan 2021

· Ruled out the noises collected in the daily stock return, suggested three different models, and constructed the Bayesian posterior distribution for the parameters to predict the change point. All those three models got incorporated into the existing frameworks in the company
· Applied *CmdStan* (a language based on C++ and R) to perform the Markov Chain Monte Carlo (MCMC) simulation and derive the posterior distribution for the parameters

RESEARCH EXPERIENCE

Transfer Learning (Master Thesis) 2023 – 2024 (Expected)

· Transfer Learning theory and its application

Semi-Supervised Learning (Best Undergraduate Research of 2022) Jun 2021 - Apr 2022

· Used Mathematical Optimization theoretically proved that combining semi-supervised learning methods with non-parametric techniques (i.e., kernel regression) can reduce the asymptotic Mean integrated square error (MISE) from $O(n^{-\frac{4}{5}})$ to $O(n^{-\frac{16}{19}})$. Compared to other semi-supervised learning methods, it has far less restrictions on the data set
· Discovered a by-product that artificially modifying a supervised dataset to a semi-labeled dataset with this semi-supervised method can further improve the prediction performance
· Reproduced the experiments in the past papers and through R and compared the results of our models. The new method has the MSE at most 40% smaller than the traditional supervised learning method
· Employed this idea to the Projection Estimator and Reproducing Kernel Hilbert Space (RKHS)

Approximate Bayesian Computation for High Dynamical Models Jan 2021 - Jun 2021

· Applied Approximate Bayesian Computational algorithms with the auxiliary likelihood function to better estimate the parameters in the stochastic models, especially for high dynamic models
· Proposed a more efficient algorithm by adding weights on the datapoint with minimum summary distance
· Performed numerical analysis using both simulation and real-world data set (SP500 Index in the latest 80 years) Improved time complexity, root mean square error (RMSE) and the variance of parameters in stochastic models

SKILLS

· Programming Language: Python, C/C++, Java, R, MySQL, VBA, Latex
· Frameworks: Pytorch, tensorflow, Scikit-learn, OpenCV, CmdStan
· Language: Mandarin (native), English (fluent), Cantonese (intermediate)
· Hobbies: Chess (National Level Two Chess Player), Badminton, Guzheng