Jinpu Zhou

zhoujinpu1@gmail.com | (+1) 248-214-6673 | github.com/annshuu

Education

Ph.D. candidate in Applied Mathematics *Sep. 2019 - Present*
Louisiana State University

- **Specialization**: Probability and Statistics
- **Research Focus**: Bayesian machine learning in time series with noise. **Advisor**: Arnab Ganguly
- Relevant Coursework: Convex and Stochastic Optimization, Statistical Learning, Stochastic Analysis, Bayesian Analysis
- GPA: 4.0/4.0. Expected graduation date: May 2024

Master in Computer Science *Jan. 2022 - Present*
Georgia Institute of Technology

- **Specialization**: Machine Learning
- Relevant Coursework: Deep Learning, Reinforcement Learning, Graph Networks, Big Data for Healthcare, Fairness in AI/ML
- GPA: 4.0/4.0. Expected graduation date: December 2023

Bachelor in Physics *Sep. 2014 Jul. 2018*
Beijing Normal University

- GPA: 3.3/4.0. Top of the elite class. 2015 Academic Excellence Scholarship
- Dissertation: A Study of the Planetary Orbits in the Solar System based on Machine Learning

Skills

- **Machine Learning**: machine learning (kNN, Naïve Bayes, Linear/Logistic Regression, Ridge, LASSO, elastic net, SVM, Kernel methods, CART, Bagging, Boosting, PCA), deep learning (RNN, CNN, LSTM, GAN, YOLO, U-net, ResNet, NST), reinforcement learning (SARSA, Q-learning, DQN, QMIX, COMA, PPO), graph networks (GNN, GCN), natural language processing (BERT, SentenceTransformers)
- **Mathematics**: convex and nonconvex optimization, graph theory, probability theory, stochastic analysis, markov chain monte carlo, filtering algorithms
- **Statistics**: A/B testing, statistical inference, bayesian analysis, experimental design, causal inference
- **Language and Framework**: Python (numpy, matplotlib, scipy, statsmodels, pandas, seaborn, ggplot; pytorch, tensorflow, keras, scikit-learn), MATLAB, R; Java, C/C++, Scala; SQL, Spark, Hadoop; D3, Tableau; Docker, Linux, High-Performance Computing, AWS, GCP

Experiences

Department of Mathematics, Louisiana State University Baton Rouge, LA
Graduate Teaching and Research Assistant *Sep. 2020 - Present*

- **Optimization in a Function Space**: • Proved a generalized representer theorem for the optimization problem in a Hilbert space which gives the foundation for linear functional regression. • Characterized the existence and uniqueness conditions for the solution of the optimization problem.
- **Nonparametric Learning Problems Related to Time Series with Noise**: • Applied nonparametric Bayesian techniques, kernel methods and Gibbs sampling to estimate the function components in SDEs (a general model for time series with noise). • Conducted various asymptotic analyses of the system including convergence of the estimators through central limit type theorems and large deviation techniques.
- **Sequential Monte Carlo Method for Noisy and Sparse Data**: Improved Sequential Monte Carlo method using bridge processes to recover the signal from noisy and sparse data.
- **Instructor of Calculus and Collage Algebra and TA for Real Analysis I, II and III**

Taboola Inc. Los Angeles, CA
Data Scientist Intern *May 2022 Aug. 2022*

- **Crawling Issues Detection**: • Compared different sentence embedding models and the similarity scores to define a method to classify bad crawled body text. • Achieved 80% of accuracy in bad crawled text detection. [NLP: Sentence Embedding; Python]
- **Improving the quality and speed of the NLU model**: • Improved the data preprocessing and model inference pipeline. Successfully reduced processing time by 20% to the previous version. • Helped with the production issues. [NLP; Kubernetes, Airflow, MLflow; Python, Java, SQL, Hadoop]

State Key Laboratory of Cognitive Neuroscience and Learning Beijing, China
Data Science Research Assistant *May 2015 Aug. 2019*

- **Data Mining of Brain Intrinsic Optical Signal Imaging**: •Implemented DNN and CNN to brain intrinsic optical signal images to get the functional brain mapping in the V1 area related to different eyes and cues. • The built model marked the receptive fields with higher signal-to-noise ratios than the traditional method. [CV: DNN, CNN; Matlab]
- **Nonlinear Dynamic in Working Memory Model**: • Constructed a spiking network with different types of interneurons and pyramidal neurons to simulate the brains short-term memory. • Successfully generated the model that holds simultaneously and sequentially inputs and the observed high and low-frequency oscillations match the experimental data. Studied the influential factors on the oscillation of different frequencies. [HPC; Matlab, Python]

Projects

Department of Computer Science, Georgia Institute of Technology Remote
Deep Learning Projects *Jan. 2022 Apr. 2022*

- **Fraud Detection on the Bitcoin Transitions Network**: • Applied Graph Convolutional Networks (GCN) to Bitcoin transactions data to classify illicit transitions. • Achieved 95% accuracy in fraud detection. [GCN; Python, Pytorch]
- **Multi-agent Deep Reinforcement Learning in Google Football**: • Applied multi-agent reinforcement learning algorithms and value function approximation with deep neural networks to train a team of agents to play 3 vs. 3 football games with built-in rule-based AI. • Implemented COMA, QMIX and PPO with centralized critic in Google Football environment to improve the winning rate. Successfully improved winning rate by 40% to the baseline models. [RL: COMA, QMIX, PPO; Python, Pytorch; Docker; AWS, GCP]
- **Predicting Hospital Readmission using Clinical Notes**: • Applied BERT and GNN to integrate clinical notes information and patient network topological structure to predict 30-day hospital readmission using MIMIC-III dataset. • Successfully reproduced the published result. [GNN, NLP: BERT; Python, Pytorch, Scala, Spark, SQL; Docker]

Kaggle Competitions Remote
Deep Learning Projects *Mar. 2022 Oct. 2022*

- **Google Universal Image Embedding**: • Applied EfficientNet, ConvNet, CLIP and fine-tuned the model to provide an embedding of dimension 64 for any input image. • Ranked in top 3% and won a silver medal in the competition. [CV: EfficientNet, ConvNet, CLIP; Python: Tensorflow]
- **HuBMAP and HPA - Hacking the Human Body**: • Applied U-net, DeepLab and CoAT to identify and segment functional tissue units across kidney, prostate, large intestine, spleen and lung. • Ranked in top 10% and won a Bronze medal in the competition. [CV: U-net, DeepLab, CoAT; Python, Pytorch]
- **HM Personalized Fashion Recommendations**: • Conducted feature engineering and applied GRU and LSTM to develop product recommendations based on data from previous transactions. • Ranked in top 25%. [REC: GRU, LSTM; Python, Tensorflow, RecBole]

Publications

- A. Ganguly, R. Mitra and J. Zhou. *Infinite-dimensional optimization and Bayesian nonparametric learning of stochastic differential equations.* Journal of Machine Learning Research, 24(159), 139. Retrieved from http://jmlr.org/papers/v24/22-0582.html.
- Y. Zhang, J. Zhou, D. Wang and X. Wang. *Gamma and Beta Band Oscillation in Working Memory given Sequential or Concurrent Multiple Items: A Spiking Network Model.* eNeuro (in revision).
- A. Ganguly, R. Mitra and J. Zhou. *Nonparametric learning of SDE from sparse, partial and noisy data.* Submitted.
- J. Zhou, A. Ganguly and H. Gietz. *Sparse estimation of drift functions of SDEs using Normal-GIG priors.* Submitted.

Conferences and Presentations

- Y. Zhang, J. Zhou, D. Wang, X. Wang. (2018). *Neural Circuit Maintains Simultaneously or sequentially presented multiple items in working memory.* Program No. 426.12. Society for Neuroscience 2018.
- A. Ganguly, R. Mitra and J. Zhou. (2022). *Bayesian nonparametric learning of stochastic differential equations.* Program G31. Society for Industrial and Applied Mathematics TX-LA 2022.
- A. Ganguly, R. Mitra and J. Zhou. (2023). *Bayesian nonparametric learning of stochastic differential equations.* Part of CP2 Machine Learning. SIAM Conference on Dynamical Systems 2023.
- A. Ganguly, R. Mitra and J. Zhou. (2023). *Bayesian nonparametric learning of stochastic differential equations.* Program 2693, 2023 Joint Statistical Meetings.

Awards

- Sliver Medal in Kaggle Google Universal Image Embedding competition (Top 3% of 1022) 2022
- Bronze Medal in Kaggle HuBMAP and HPA competition (Top 10% of 1175) 2022
- Dissertation Year Fellowship 2023
- A. K. and Shirley Barton Superior Graduate Student Scholarship 2022
- Academic Excellence Scholarship 2015

Activities

Vice President of LSU SIAM Student Chapter Baton Rouge, LA
Organized events, conducted and delivered workshops reaching hundreds of students. *2021-2022*

Volunteer of Capital Volunteers Association Beijing, China
Organized events and requested sponsorships, conducted regular community services. *2017-2019*

Last updated: September 2023