Yunfan Zhao

Email: yunfanzhao19@outlook.com | Phone: 443-240-2948 | Address: 11 Peabody Terr, Cambridge, MA

Education

Columbia University   New York, NY

Ph.D. in Operations Research, School of Engineering and Applied Science   September 2019 - May 2023

Johns Hopkins University   Baltimore, MD

Combined B.S. / M.S. in Applied Mathematics and Statistics   September 2014 - December 2018

Foundation Model / LLM Research Experience

Harvard University, *postdoc in CS*   July 2023 - present

Foundation models with fewer trainable parameters - scalable neural network kernels (SNNK). ICLR 2024

- Proposed a method to approximate feed forward layers with Random Fourier Feature Maps.

- Linearized the pooler layer in **Vision Transformer** with SNNK, and finetuned on ImageNet, Cifar10, and Cifar100.

- Replaced adapter blocks in in **Adapter Transformers** with SNNK, and finetuned on GLUE text datasets.

- Replaced Transformer layer's MLP block in **BERT and ViT** with SNNK, and finetuned on text and image datasets.

- Our method achieves SOTA performance with **3x fewer parameters** and dramatically **reduces the storage to 1/1000**.

Foundation models as automated planners for bandit tasks. IJCAI 2024 (part (ii) submitted to NeurIPS)

Part i): - Designed a foundation model for bandits with general zero-shot ability. Implemented the model in **Pytorch** and pretrained on Harvard cluster and google Cloud Compute, requiring **8x fewer finetuning samples** than SOTA.

Part ii): - Introduced a reward proposal loop that enhances LLM-generated reward functions using **feedback from bandit simulations** and iteratively refines reward design (**LLM reflection**), achieving 20% performance increase over baselines without feedback.

- Evaluated the performance of our system using **Gemini Pro**, demonstrating near human-level policy tuning to achieve human specified outcomes using only language prompts as input, outperforming SOTA by 63%.

Internship Experience

Experian DataLabs, *Data Scientist Intern*   May 2022 - August 2022

- Designed a credit rating model for DeFi; collected data for over $10 billion transactions from on-chain crypto lending protocols.

- Generated 100+ features from data with **PySpark SQL;** perform extensive feature selection and correlation analysis using **SHAP**.

- Fine-tuned an **XGBoost** model; improved model **AUC** by 14% and achieve SOTA performance.

Skills

Programming Languages: Python, SQL, Matlab

Machine Learning Packages: PyTorch, Tensorflow, HuggingFace Transformers, PySpark, Scikit-learn, MultiProcessing

Selected Machine Learning Research Papers (* indicates equal contribution)

1. *Zhao, Y., *Behari, N., et al. Towards a Pretrained Model for Restless Bandits via Multi-arm Generalization. *IJCAI 2024.*

2. *Sehanobish, A., *Choromanski, K., *Zhao, Y., et al. Scalable Neural Network Kernels. *ICLR, 2024.*

3. *Verma, S., *Zhao, Y., et al. Group Fairness in Predict-Then-Optimize Settings for Restless Bandits. *UAI 2024.*

4. *Zhao, Y., *Pan, Q., *Choromanski, K., Jain, D., Sindhwani, V. Implicit Two-Tower Policies. *ICLR 2024 PML4LRS workshop*

5. *Choromanski, K., *Sehanobish, A., *Lin H., *Zhao, Y., et al. Efficient Graph Field Integrators Meet Point Clouds. *ICML, 2023.*

6. Yuan, G., Zhao, Y., Kpotufe, S. Regimes of No Gain in Multi-class Active Learning. *Journal of Machine Learning Research, 2024.*

7. *Elmachtoub, A. N., *Gupta, V., *Zhao, Y. Balanced Off-Policy Evaluation for Personalized Pricing. *AISTATS, 2023.*

8. *Kpotufe, S., *Yuan, G., *Zhao, Y. Nuances in Margin Conditions Determine Gains in Active Learning. *AISTATS, 2022.*