Konghao (Shelton) Zhao

zhaok220@wfu.edu | https://sheltonzhaok.github.io/ | https://github.com/SheltonZhaoK

Education

Ph.D in Computer Science **August 2024 – Present**
University of Southern California Los Angeles, California
Advisor: Ruishan Liu, Research Interest: AI in Healthcare

B.S. in Computer Science and Mathematics **August 2020 – May 2024**
Wake Forest University, *Magna Cum Laude* with honors in Computer Science, GPA: 3.94 Winston Salem, North Carolina
Advisor: Natalia Khuri, Thesis: Evolutionary Multi-Objective Coreset Selection for Data-Centric AI

Skills

Research: large language model (LLM), natural language processing, interpretable deep neural networks, algorithmic instability and fairness, evolutionary algorithms for cluster analysis and coreset selection, single-cell RNA sequencing (scRNA-seq) data analysis

Programming: Python, C, R, Java, CUDA, Unix Shell, MATLAB, SQL

Toolkits: Pytorch, Tensorflow, Scikit-learn, Huggingface, OpenCV, Ray, Gym, Captum, PyCUDA, Python Multi-processing

Research Experience

Laboratory for Machine Learning, Health and Biomedicine, University of Southern California **August 2024 – Present**
Ph.D Student advised by Dr. Ruishan Liu *Los Angeles, CA*

- Retrieved structural tables from unstructured patient and trial corpus using LLM and named entity recognition
- Extracted feature embeddings and developed three feature alignment methods for patient-trial matching using Siamese BERT networks and cosine similarity
- Developed a reasoning approach for patient-trial matching based on criteria using an ensemble of LLMs with a ML model trained on LLM-extracted features to predict eligibility scores

DataMine Research Group, Wake Forest University **February 2021 – May 2024**
Undergraduate Researcher advised by Dr. Natalia Khuri *Winston Salem, NC*

- Created a metric and developed a system to access algorithmic and data bias in integration methods for heterogeneous dataset
- Implemented a systematic and parallel approach to evaluate the algorithmic instability of ML algorithms
- Co-developed an open-source R package for the metamorphic evaluation of the robustness of data analysis tools in scRNA-seq data
- Designed, implemented, and optimized two multi-objective genetic algorithms for clustering and coreset selection of high-dimensional, noisy single-cell RNA sequencing data

Advanced Agent-Robotics Technology Lab, Carnegie Mellon University **June 2023 – October 2023**
Robotics Institute Summer Scholar (RISS) advised by Dr. Katia Sycara *Pittsburgh, PA*

- Improved the performance and interpretability of Concept Bottleneck Models (CBMs) by mitigating the information leakage (IL)
- Implemented three novel latent space disentangling approaches with a semi-supervised CBM to address IL with CUB 200 dataset
- Benchmarked the performance and interpretability by test-time intervention through extensive empirical experiments

Publications

1. **K. Zhao**, S. Bhandari, N.P. Whitener, J. M. Grayson and N. Khuri, "An Ensemble Machine Learning Approach for Benchmarking and Selection of scRNA-seq Integration Methods", ACM-BCB, September 2023 [**Top 10% regular paper with oral presentation**]
2. **K. Zhao**, J. M. Grayson and N. Khuri, "Multi-Objective Genetic Algorithm for Cluster Analysis of Single-Cell Transcriptomes", J. Pers. Med., January 2023 [**Monthly cover**]
3. N. Khuri, S. Bhandari, E. Murillo Burford, N.P. Whitener and **K. Zhao**, "Multi-target integration and annotation of single-cell RNA-sequencing data", ACM-BCB, August 2022 [**Short paper**]
4. S. Bhandari, N. P. Whitener, **K. Zhao** and N. Khuri, "An evolutionary approach to data valuation", ACM-BCB, August 2022 [**Regular paper**]
5. **K.Zhao**, Y. Wang and N.Khuri, "Multi-Objective Coreset Selection for Data-Centric QSAR Modeling" [**In Submission**]
6. N.P. Whitener*, **K. Zhao***, J. M. Grayson and N. Khuri, "scrnabench: A Package for Metamorphic Benchmarking of scRNA-seq Data Analysis Methods" [**In Submission**]
7. R. Zabounidis, I. Oguntola, **K. Zhao**, J. Campbell, S. Stepputtis, K. P. Sycara, "Benchmarking and Enhancing Disentanglement in Concept-Residual Models" [**In Submission**]

Awards and Scholarship

- 2024 John W. Sawyer Award, WFU CS Department, Awardee — April 17, 2024
- CRA Outstanding Undergraduate Researcher Award, Computing Research Association, Honorable Mention — December 20, 2023
- Student HPC cluster competition (IndySCC), SC23, First Place in US — November 12, 2023
- RISS 2023 Summer Scholarship, Carnegie Mellon Robotics Institute, 2023 summer research — June 1, 2023
- Wake Forest Research Fellowship, Wake Forest URECA Center, 2022 summer research — May 1, 2022