

ZHECHENG SHENG

516 Delaware Street SE, Minneapolis, MN, 55455

📞 919-519-7176 ✉ sheng136@umn.edu 🌐 github.com/zcsheng95 🌐 zcsheng95.github.io

RESEARCH INTEREST

Trustworthy machine learning My current research interests focus on developing causal methods to ensure fairness and robustness of machine learning models and their application in Biomedicines. This involves addressing the presence of sensitive attributes or confounding shifts that may interfere with model performances. Besides, I am interested in designing methods for probing Large Language Models(LLM) to assess the faithfulness of language generation and adapting it for downstream domain-specific tasks.

EDUCATION

University of Minnesota, Twin Cities Minneapolis, MN

Aug. 2021 – Aug. 2025(Est.)

Ph.D. in Health Informatics & Data Science

M.Sc. in Data Science, GPA: 3.99/4.00

Advisor: Serguei Pakhomov

Duke University Durham, NC

Aug. 2017 – May. 2019

M.Sc. in Biostatistics, GPA: 3.82/4.00

Advisor: Benjamin Goldstein

SELECTED PUBLICATION

- Sheng, Z.***, Zhang, T.* , Jiang, C.* , Kang, D. (2024). BBScore: A Brownian Bridge Based Metric for Assessing Text Coherence. *In Proceedings of the 38th Annual AAAI Conference on Artificial Intelligence (AAAI)*, pages 14937-14945. Vancouver, Canada. (* equal contributions)
- Li, C., **Sheng, Z.**, Cohen, T., & Pakhomov, S. V. S. (2024). Too Big to Fail: Larger Language Models are Disproportionately Resilient to Induction of Dementia-related Linguistic Anomalies. *In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*. Bangkok, Thailand
- Sheng, Z.**, Finzel, R., Lucke, M., Dufresne, S., Gini, M., & Pakhomov, S. (2023). A Dialogue System for Assessing Activities of Daily Living: Improving Consistency with Grounded Knowledge. *In Proceedings of the Third DialDoc Workshop on Document-grounded Dialogue and Conversational Question Answering*, pages 68–79. Toronto, Canada. Association for Computational Linguistics.
- Ding, X., **Sheng, Z.**, Hur, B., Chen, F., Pakhomov, S., & Cohen, T. (2023). Enhancing Robustness of Foundation Model Representations under Provenance-related Distribution Shifts. *In NeurIPS 2023 Workshop on Distribution Shifts: New Frontiers with Foundation Models*. New Orleans, LA, USA.
- Sheng, Z.**, Finzel, R., Gaydhani, A., Lucke, M., Dufresne, S., Gini, M., & Pakhomov, S. (2023). A Chatbot for Activities of Daily Living Assessment Standardization. *In AMIA Annual Symposium*. New Orleans, LA, USA.
- Sheng, Z.**, Bollig, E., Granick, J., Zhang, R., & Beaudoin, A. (2022). Canine Parvovirus Diagnosis Classification Utilizing Veterinary Free-Text Notes. *In 2022 IEEE 10th International Conference on Healthcare Informatics (ICHI)* (pp. 614-615). Rochester, MN, USA: IEEE.
- Ding, X., **Sheng, Z.**, Yetişgen, M., Pakhomov, S., & Cohen, T. (2023). Backdoor Adjustment of Confounding by Provenance for Robust Text Classification of Multi-institutional Clinical Notes. *In AMIA Annual Symposium*. New Orleans, LA, USA.

RESEARCH EXPERIENCE

Deconfound Deep Transformers Networks

Sep. 2022 – Present

Mentor: Serguei Pakhomov, University of Minnesota Twin Cities

- Investigated language model's causal reasoning ability in dementia classification results and located weights within the deep network that associated with target labels.
- Proposed model weight filtering strategies that effectively reduce group disparities and improve fairness.
- Developed an probabilistic evaluation framework to quantify distribution shift within subgroups.
- Discovered BERT families captures gender specific-tokens during fine-tuning which resulting in prediction bias at model deployment in dementia detection.

- Discovered different subgroup training distributions can result in performance shift at test time.

Dialogue System for Activity of Daily Living Assessments

July. 2022 – Present

Mentor: Serguei Pakhomov, University of Minnesota Twin Cities

- Implemented transformer models like RoBERTa for text classification and improved classification accuracy by 10%.
- Fine-tuned Large Language Models such as LLaMA and Vicuna with clinical assessment notes for domain adaptation through parameter efficient methods(LoRA) and improved the usability of the system.
- Built a knowledge base for the dialogue system which reduces hallucinations from large language models and made the system more factual grounded.

Assess Text Coherence with Brownian Bridge

Jan. 2023 – Present

Mentor: Dongyeop Kang, University of Minnesota Twin Cities

- Introduced a likelihood-based evaluation score motivated by Brownian Bridge for measuring global text coherence.
- Demonstrated the proposed score can improve classification performance from previous SOTA in global artificial tasks without end-to-end training.
- Proposed an likelihood objective for language encoder training that is able to capture latent dependencies.
- Elaborated the proposed score can be used to compare text with different length both empirically and theoretically.
- Showcased the score can be adapted for diverse downstream tasks, such as distinguishing between human and AI-generated text and detecting different LLM generated text from mixed corpus.

An Unified Machine Learning Fairness Notion via Sparsity

May. 2023 – Present

Mentor: Enmao Diao, Duke University

- Proposed a novel machine learning fairness notion based on distributional difference and sparsity.
- Unified fairness evaluation measurements for both classification and regression problem.
- Extended the fairness notion for multiple sensitive group and multiple class scenarios.
- Implemented augmented ADMM in optimization for linear models with designated fairness constraint and provided theoretical guarantees.

WORK EXPERIENCE

Amazon, Inc. *Seattle, WA*

Jun. 2024 – Present

Applied Scientist Intern, Mentor: Kyle Willett

- Developed a Generative AI application using LLM agent framework through AWS Bedrock to orchestrate causal effect estimation workstream for third party seller FBA fee changes.
- Leveraged tool calling ability of LLM to implement chained and recursive tool invocation actions.
- Proposed evaluation frameworks for assessing causal reasoning ability and factual consistency of LLM outputs.

Duke Cancer Institute *Durham, NC*

Aug. 2019 – Jul. 2021

Bioinformatician I, Mentor: Kouros Owzar

- Designed and implemented an S4 R package for assessing sequencing and mapping quality of RNA-Seq data.
- Operated different bioinformatics pipelines through dockerized containers in remote HPC clusters.
- Delivered reproducible and well-documented reports and code.

SERVICES & AWARDS

- Reviewer Award for the AMIA Informatics Summit 2024
- Served as a reviewer of AMIA Informatics Summit 2024
- Served as a reviewer of EMNLP 2023

SKILLS & COURSEWORK

Programming Languages: Python, R, SAS, Shell, SQL, Java

Software: Pytorch, Huggingface, Scikit-learn, Pandas, Tidyverse, AWS

Relevant Coursework: Data Structure and Algorithms, Database Management System, Advanced Machine Learning, Artificial Intelligence, Statistical Programming with Big Data, Generalized Linear Model, Causal Inference, Computational Causal Analytics, Natural Language Processing with Deep Learning