# Zhiyi Sun

Tel: (734) 596-5591        Email: zhiyisun@umich.edu

## EDUCATION BACKGROUND

**University of Michigan, Ann Arbor, MI**                                                08/2022-05/2024
- Master of Science in Biostatistics (Conferred in May 2024)        GPA: 3.5/4.0

**University of Wisconsin-Madison, Madison, WI**                                    09/2018-05/2021
- Bachelor of Science in Statistics & Mathematics (Conferred in May 2021)        GPA: 3.7/4.0
- Honor: Dean's List (Fall 2018 & Fall 2019)

**Xiangnan University, Chenzhou, China**                                                09/2016-06/2018
- Major: Clinical Medicine        GPA: 3.8/4.0

## RESEARCH EXPERIENCE

***Research Assistant*, University of Michigan, Ann Arbor**                        07/2024-ongoing

*Advisor: Dr. Zach Landis-Lewis, Assistant Professor, Department of Learning Health Science, University of Michigan Medical School*
- Design and apply prompt engineering to evaluate large language models (LLMs) for authoring precision feedback messages and metadata for healthcare professionals in the development of a precision feedback system/knowledge base.
- Develop evaluation criteria for qualitative and quantitative assessing the correctness, consistency, and acceptability of precision feedback messages.
- Explore the application of Retrieval-Augmented Generation (RAG) in generating personalized and evidence-based feedback for healthcare professionals.

***Research Assistant*, University of Michigan-Dearborn**                        01/2023-ongoing

*Advisor: Dr. Francine Dolins, Associate Professor of Psychology, Department of Behavioral Science, University of Michigan-Dearborn*
- Proposed a proposal of statistical experimental design that arranged the testing trials and calculated the sample size to ensure significantly statistical power for the research on cognitive/behavioral foundations of social minds.
- Employ VR GIS data to investigate modes of thinking and to analyze and visualize foraging behavior data through change-points analysis, subgroup analysis, learning analytics, PCA, t-SNE, LDA, semi-supervised LFDA (SELF), and linear mixed effects models.
- Perform statistical analysis and visualization of the bonobo's fruit preference and emotional behavior data and explore the potential of LLMs for interpreting results regarding spatial cognition and memory.

***Research Assistant*, Research on Intestinal Polyps, Diarrhea and IBD**        05/2021-ongoing

*Advisor: Dr. Hongmei Zhao, Professor of Pediatrics, Department of Gastroenterology and Nutrition, The Affiliated Children's Hospital of Xiangya School of Medicine, Central South University (Hunan Children's Hospital), China*
- Identified risk factors and made prediction for secondary intussusception in children with intestinal polyps, constructing and validating a nomogram (e.g., logistic, lasso, machine learning models, etc.).
- Investigated multidisciplinary treatment outcomes for children with refractory gastrointestinal polyps, especially in Peutz-Jeghers Syndrome (PJS), and identified the genotype-phenotype correlation.
- Explored the potential risk factors of diarrhea among population under 5 years old and its correlations with other diseases, based on global data from 1990 to 2019.
- Studied on inflammatory bowel disease, investigating risk factors for the development of C.difficile infection in pediatric IBD patients and quantifying causality and potential co-morbid mechanisms with other diseases.
- Explore NLP for clinical notes and adapting deep learning methods, i.e., Vision Transformers, to colon polyp image detection.

***Research Projects*, University of Michigan, Ann Arbor**                        06/2023-05/2024
- Modeled risk prediction in health services and outcomes research with EHRs.
- Replicated the Bayesian Knowledge Tracing algorithm using Python and applied it to test performance in medical

education.

- Investigated the Micro-Randomized Trial design in mobile health and replicated the HeartSteps intervention study with longitudinal data analysis to estimate time-varying causal excursion effects.

### *Research Projects*, Global Suicide Rates Insight and Prediction                    07/2021-11/2021
*Advisor: Dr. Patrick Rebeschini, Associate Professor of Statistics, Department of Statistics, University of Oxford, UK*

- Analyzed and predicted the global suicide trend to provide a reference for the future suicide prevention and treatment.
- Conducted exploratory data analysis, statistical inference and hypothesis tests, modeling (i.e., linear regression, KNN, decision tree, random forest, MLPs, and XGBoost) and prediction, with a Shiny app for presentation.

### *Research Assistant*, Research of Image Recognition Based on Deep Learning        05/2020-08/2020
*Advisor: Dr. Pengzhi Chu, Lecturer, Department of Computer Science and Engineering, Shanghai Jiao Tong University, China*

- Adopted YOLOv3 and Fast/Faster R-CNN in TensorFlow and PyTorch respectively to train the dataset of waste classification images with an accuracy of 75%, the final model was used as a reference for urban pollution solutions.
- Adopted Faster R-CNN in TensorFlow to recognize the traffic lights, meeting the standard of autonomous vehicles, with an accuracy of over 99%.
- Collaborated with interdisciplinary medical team, adopting CNN, R-CNN, Fast/Faster R-CNN to the research of medical imaging (i.e., MRI and fMRI).


## PUBLICATIONS & CONFERENCE PRESENTATIONS
**Journals**

1. Liu, Y., **Sun, Z.**, Guo, Y., Liu, C., Tian, S., & Dong, W. (2023). Construction and validation of a nomogram of risk factors and cancer-specific survival prognosis for combined lymphatic metastases in patients with early-onset colorectal cancer. *International Journal of Colorectal Disease*, *38*(1), 128.
2. Liu, Y., Li, J., Tian, S., Lan, Q., **Sun, Z.**, Liu, C., & Dong, W. (2024). Identification and validation of hub genes expressed in ulcerative colitis with metabolic dysfunction-associated steatotic liver disease. *Frontiers in Immunology*, *15*, 1357632.
3. Deng, Y., Li, C., Huang, L., Xiong, P., Li, Y., Liu, Y., ..., **Sun, Z.**, ... & Zhao, H. (2024). Single-cell landscape of the cellular microenvironment in three different colonic polyp subtypes in children. *Clinical and Translational Medicine*, *14*(1), e1535.
4. Zhao, H., Zhou, J., Yuan, L., **Sun, Z.**, Liu, Y., Zhao, X., & Ye, F. (2024). Exploring the Alleviating Effects of Bifidobacterium Metabolite Lactic Acid on Non-Alcoholic Steatohepatitis Through the Gut-Liver Axis. *Frontiers in Microbiology*.


**Journals Under Review**

5. **Sun, Z.**, Liu, Y., Wang, Z., Zhou, J., Li, C., Luo, Y., Ouyang, H., Liu, L., Zhang, W., Jiang, N., Duan, J., Zhan, M., Liu, C., You, J., & Zhao, H. (2024). Enhancing pediatric inflammatory bowel disease care with large language models: personalized strategies for nutrition, psychology, and education. *JAMA pediatrics*.
6. **Sun, Z.**, Zhou, J., Liu, Y., Wang, Z., You, J., & Zhao, H. (2024). Nutritional Status and Influencing Factors in Pediatric Inflammatory Bowel Disease: Leveraging Large Language Models for Subtype Differentiation, Data Augmentation, and Clinical Insights. *Clinical Gastroenterology and Hepatology*.
7. **Sun, Z.**, Zhou, J., Liu, Y., Wang, Z., You, J., & Zhao, H. (2024). Integrating large language models to translate machine learning findings from historical NHANES IBD data into current actionable dietary insights and practices. *Journal of Crohn's and Colitis*.
8. **Sun, Z.**, Li, C., Zhou, J., Liu, Y., Wang, Z., Luo, Y., Ouyang, H., Liu, L., Zhang, W., Jiang, N., Duan, J., Zhan, M., Liu, C., You, J., Li, Y., & Zhao, H. (2024). Risk factors for secondary intussusception in children with intestinal polyps: a retrospective cohort study based on machine learning of 2669 children with intestinal polyps. *JAMA network open*.


**Conferences**

1. Liu, Y., **Sun, Z.**, & Wang, Z. (2024). Association between Abscess, Intestinal Obstruction, and Inflammatory Bowel Disease: Evidence from A Mendelian Randomization Study. *The 15th ACM Conference on Bioinformatics,*

*Computational Biology, and Health Informatics (ACM BCB)*, Accepted as Poster, Shenzhen, China, November 22-25, 2024.

2. **Sun, Z.**, Liu, Y., Wang, Z., & Zhao, H. (2024). Enhancing Pediatric IBD Management with Large Language Models: Personalized Nutritional Advice, Psychological Support, and Educational Tools. *IEEE EMBS BHI 2024 Conference*, Poster presentation, Huston, TX, USA, November 10-13, 2024.

3. **Sun, Z.**, Cao, Y., Shi, G., Flynn, A., & Landis-Lewis, Z. (2024). Human-Guided Iterative Prompt Engineering for Precision Feedback Message Authoring Using LLMs. *2024 ADSA Annual Meeting*, Poster presentation, Ann Arbor, MI, USA, October 29-31, 2024.

4. **Sun, Z.**, Cao, Y., Shi, G., Flynn, A., & Landis-Lewis, Z. (2024). Iterative Prompt Engineering Using LLMs for Authoring Precision Feedback Messages. *2024 MCBK Global Meeting*, Lighting Talk (Virtual), October 22-23, 2024.

5. Reviewer, *2025 ASPPH Annual Meeting for Academic Public Health*, Association of Schools and Programs of Public Health (ASPPH), October 2024.

6. **Sun, Z.** (2022, May). Global suicide rates insight and prediction. In *2nd International Conference on Applied Mathematics, Modelling, and Intelligent Computing (CAMMIC 2022)* (Vol. 12259, pp. 1396-1408). SPIE.

**Upcoming Presentations**

7. **Sun, Z.**, Liu, Y., Wang, Z., Zhou, J., & Zhao, H. (2025). Preventing Cardiovascular Disease in Pediatric IBD: A Health Management Approach with Large Language Models. Abstract submitted to the *2025 SER Mid-Year Meeting*, Virtual, February 10-14, 2025.

8. **Sun, Z.**, Wang, Z., Liu, Y., Shen, J., & Zhao, H. (2025). Utilizing LSTM and Large Language Models with Global Burden of Disease Data for Occupational Injury and Noise Risk Forecasting: An Example in the U.S. and Michigan. Paper submitted to the *Workshop on Large Language Models and Generative AI for Health at AAAI 2025*, Philadelphia, PA, USA, March 4, 2025.

9. **Sun, Z.**, Liu, Y., Wang, Z., & Zhao, H. (2025). Using Large Language Models for Child-Friendly Narratives in Pediatric IBD Education. *46th Annual Meeting & Scientific Sessions of the Society of Behavioral Medicine (SBM 2025)*, Accepted as Poster, San Francisco, CA, USA, March 26-29, 2025.

10. **Sun, Z.**, Liu, Y., Wang, Z., Zhou, J., & Zhao, H. (2025). Optimizing Cost-Effective and Personalized Drug Therapy for Pediatric IBD Using Large Language Models. *AMCP 2025 Annual Meeting*, Accepted as a highlighted poster in the moderated poster tour, Houston, TX, USA, March 31-April 3, 2025.

11. **Sun, Z.**, Liu, Y., Wang, Z., Ma, Z., & Zhao, H. (2025). Differential Diagnosis of Pediatric Inflammatory Bowel Disease Using Large Language Models. *2025 Translational Science (ACTS TS25)*, Accepted as Poster, Washington, DC, USA, April 14-17, 2025.

12. **Sun, Z.**, Liu, Y., Wang, Z., Zhou, J., Ma, Z., Duan, B., & Zhao, H. (2025). Supporting Healthcare Planning and Resource Management for Gastritis and Duodenitis Using Large Language Models. *AMIA 2025 Clinical Informatics Conference*, Accepted as Poster, Anaheim, CA, USA, May 20-22, 2025.

13. **Sun, Z.**, Liu, Y., Wang, Z., Zhou, J., & Zhao, H. (2025). Utilizing Diffusion Models to Generate Polyp Images for Familial Adenomatous Polyposis and Peutz-Jeghers Syndrome: Enhancing Clinician and Patient Understanding. Abstract submitted to the *SIIM25 Annual Meeting*, Portland, OR, USA, May 21-23, 2025.

## WORK EXPERIENCE

*Assistant Analyst*, **Healthcare Group of D&R Center, GF Securities**, Shanghai, China          03/2022-07/2022
- Conducted systematic research and evaluation of biomedicine industry.
- Tracked and analyzed the latest developments in the biomedical market and listed companies to predict future trends and publish reports, especially in the fields of COVID-19 rapid antigen test and self-test, disposable medical devices, surgical robots, assisted reproductive technology, aesthetic medicine, and ophthalmology market.
- Desk research, focused on emerging trends in medical devices (stocks).

## COMPUTER SKILLS

**Programming Languages**: Python, R, SAS, SQL, Java, C++, Stata, MATLAB, Julia, JavaScript, Shell scripting (Bash/Zsh)
**Software/Tools**: R Shiny, Django, Tableau, Power BI, Git, SPSS, TensorFlow, PyTorch, Spark, PLINK, Cytoscape, LaTeX, Excel, HTML, CSS