# Yixuan(Jay) Duan

| yduan54@wisc.edu |

## Education

**University of Wisconsin-Madison(GPA: 3.83/4.0)** Madison, WI

*Bachelor of Science in Computer Science* *Expected May 2027*

- **Relevant Coursework:** Data Structures and Algorithms(A), Deep Learning(A), Molecules to Life and Science(A), Artificial Intelligence(A)
- **Honors and Awards:** 2026 UW-Madison Dean's List, 2026 UW-Madison Undergraduate Summer Award
- **Skills:** Linux, Python, R, Java, SQL
- **Interests:** Biological Foundation Model, AI for Precise Health, Interpretable AI, Multi-omics Integration

## Publications

1. **Yixuan Duan**, Wei Qiu "Discovering Interpretable Features in Cardiac Multi-Modality Foundation Models via Sparse Autoencoders." *Manuscript in preparation*, 2026.

2. **Yixuan Duan**, Wei Qiu "InterBench: An Interpretable Benchmark for Cardiopulmonary Foundation Models." *Manuscript in preparation*, 2026.

## Research Experience

**Discovering interpretable features in Foundation model via SAE** Apr. 2026 – Present

*Advisor: Professor Wei Qiu, Department of Electrical and Computer Engineering* *Rice University*

- Pretrained Foundation Model to encode raw ECG and PPG biomedical signals, into hidden activations and embeddings
- Applied SAE to decompose the Foundation Model's internal representations into sparse, interpretable latent features.
- Analyzed top-activating samples and linked SAE features to waveform patterns, clinical reports, ICD phenotypes, and disease-related concepts

**Disentangle Aging Foundation Model** Feb. 2026 – Apr. 2026

*Advisor: Professor Wei Qiu, Department of Electrical and Computer Engineering* *Rice University*

- Used bulk RNA-seq and scRNA-seq foundation models backbones as encoders, and designed an ACE-style disentanglement strategy to separate aging-related transcriptomic variation from background factors such as tissue, sex, and batch.
- Extracted transcriptomic aging signals from large-scale bulk RNA-seq and scRNA-seq using chronological age, then adapts the learned embedding to predict epigenetic age using limited paired RNA-seq and DNA methylation data.
- Designed the final model outputs and evaluation strategy, including transcriptomic aging embeddings, epigenetic age prediction, age acceleration prediction, cross-tissue/cohort generalization, and biological interpretation of aging-related genes.

**Graph-based Integration Model of scATAC-seq and DNA-Methylation** Jan. 2026 – Present

*Advisor: Professor Lei Hou, Department of Medicine, Biomedical Genetics Section* *Boston University*

- Constructed a microglia multi-omics dataset by processing scATAC-seq and snm3C-seq data, generating a cell × peak matrix and a cell × CpG methylation matrix.
- Built a CpG feature space linked to regulatory elements by identifying CpG sites overlapping ATAC peaks and extracting methylation coverage/counts for microglia cells across 40 donors.
- Designed a TF–CRE–CpG graph-based framework to use DNA methylation embeddings for predicting differentially methylated CpGs and inferring dysregulated CRE modules and TF regulatory programs.

**Modeling to Predict the Strength of Glycan and Protein Binding** Jun. 2025 – Aug. 2025

*Advisor: Dr. Chao Wang, Institute of Systems and Physical Biology* *Shenzhen Bay Laboratory*

- Construct a database that can be queried for binding strengths and sites of different glycan and proteins.
- Encode the structure and sequence information of glycan and proteins then choose different deep learning model architectures to compare R Square.
- Train a model that can accurately predict the binding strength and binding site of glycan and proteins when inputting the structure and sequence information of different glycan and proteins.