

# Syed Asad Rizvi

(832)-643-9462 | syed.rizvi@yale.edu | New Haven, CT

Homepage: <https://syedarizvi.com> | GitHub: <https://github.com/SyedA5688>

Google Scholar: <https://scholar.google.com/citations?user=2rhnnZ4AAAAI>

## RESEARCH INTERESTS

---

My primary research interests lie at the intersection of Graph Neural Networks and Foundational Models, particularly in applications to large-scale biological data. I am deeply passionate about solving real-world scientific problems using AI solutions.

## EDUCATION

---

Graduate School of Arts and Sciences, Yale University, CT

**Ph.D. in Computer Science**

**August 2023 - Present**

Relevant Coursework: Deep Learning on Graph-Structured Data, AI Foundation Models, Artificial Intelligence

College of Natural Science and Mathematics, University of Houston, TX

**Bachelor of Science in Computer Science**

**August 2019 - December 2022**

Cumulative GPA: 3.92/4.0

## RESEARCH EXPERIENCE

---

**Yale University, New Haven, Connecticut**

**June 2022 - Present**

*Graduate Research Assistant, advised by Prof. David van Dijk*

- Research Topic: Large Language Models, Foundation Models, Graph Neural Networks
- **Foundation-Model Informed Message Passing (FIMP) for Graph Neural Networks**
  - Proposed a framework for adapting arbitrary transformer-based Foundation Models into message-passing operators on graphs.
  - Improved performance on graph-based self-supervised tasks in biological application datasets utilizing SOTA foundation models for brain activity recordings and single-cell transcriptomics.
- **Cell2Sentence: Teaching Large Language Models the Language of Biology**
  - Co-led the development of an LLM-based framework for single-cell analysis in natural language.
  - Collected and processed a dataset of over 36 million single cell transcriptomic profiles.
  - Formulated combinatorial label prediction for single-cell and bulk RNA-seq data as a sequence-to-sequence generation task for LLMs.
- **BrainLM: A Foundation Model for Brain Activity Recordings**
  - Co-developed a transformer-based foundation model for fMRI brain activity recordings.
  - Scaled model capacity to over 650 million parameters and optimized training compute requirements using the Huggingface library.

**Rice University, Houston, TX**

**August 2022 - March 2023**

*Undergraduate Research Intern, advised by Prof. Xia Hu and Prof. Xiaoqian Jiang*

- Research topic: Contrastive learning, Vision-Language Models, Interpretability
- Proposed a finetuning framework based on cross-modality contrastive learning to increase the interpretability of vision-language models on radiology data.

**Houston Methodist, Houston, TX**

**December 2021 - August 2022**

*Undergraduate Research Intern, advised by Prof. Vittorio Cristini and Prof. Prashant Dogra*

- Research topic: Spatiotemporal modeling, Graph Neural Networks
- Developed a spatiotemporal Graph Neural Network architecture for COVID-19 infection forecasting, focusing on modeling dynamic international flight patterns.
- Formulated a perturbation-based explainability framework for infection spread between regions.

**University of Houston, Houston, TX**

**September 2020 – May 2022**

*Undergraduate Research Intern, advised by Dr. Hien van Nguyen*

- Research topic: Convolutional Neural Networks, Image Data Augmentation
- Proposed an efficient semi-supervised Convolutional Neural Network architecture for high-resolution medical image generation and segmentation.

## **PUBLICATIONS**

---

### **BrainLM: A Foundation Model for Brain Activity Recordings**

International Conference on Learning Representations (ICLR) 2024.

Josue O. Caro\*, Antonio H. O. Fonseca\*, **Syed A Rizvi\***, Matteo Rosati\*, Christopher Averill, James L. Cross, Prateek Mittal, Emanuele Zappala, Daniel Levine, Rahul M. Dhodapkar, Insu Han, Amin Karbasi, Chadi G. Abdallah, David van Dijk

### **Local Contrastive Learning for Medical Image Recognition**

American Medical Informatics Association (AMIA) Annual Symposium, 2023.

**Syed A. Rizvi**, Ruixiang Tang, Xiaoqian Jiang, Xiaotian Ma, Xia Hu

### **Histopathology DatasetGAN: Synthesizing Large-Resolution Histopathology Datasets**

IEEE Signal Processing in Medicine and Biology (SPMB) Symposium, 2022.

**Syed A. Rizvi**, Pietro A. Cicalese, Surya V. Seshan, Savino Sciascia, Jan U. Becker, Hien van Nguyen

### **MorphSet: Improving Renal Histopathology Case Assessment Through Learned Prognostic Vectors**

Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2021.

Pietro A. Cicalese, **Syed A. Rizvi**, Victor Wang, Sai Patibandla, Pengyu Yuan, Samira Zare, Katherina Moos, Ibrahim Batal, Marian C. Groningen, Candice Roufousse, Jan Ulrich Becker, Chandra Mohan, Hien van Nguyen

## **PREPRINTS**

---

### **Cell2Sentence: Teaching Large Language Models the Language of Biology**

bioRxiv preprint, 2023.

Daniel Levine\*, **Syed Asad Rizvi\***, Sacha Lévy\*, Nazreen Pallikkavaliyaveetil, Ruiming Wu, Insu Han, Zihe Zheng, Antonio Henrique de Oliveira Fonseca, Xingyu Chen, Sina Ghadermarzi, Amin Karbasi, Rahul M Dhodapkar, David van Dijk

### **AMPNet: Attention as Message Passing for Graph Neural Networks**

arXiv preprint, 2022.

**Syed A. Rizvi**, Nhi Nguyen, Haoran Lyu, Benjamin Christensen, Josue Ortega Caro, Antonio HO Fonseca, Emanuele Zappala, Maryam Bagherian, Christopher Averill, Chadi G Abdallah, Rex Ying, Maria Brbic, Rahul Madhav Dhodapkar, David van Dijk

### **Deep Learning-Derived Optimal Aviation Strategies to Control Pandemics**

arXiv preprint, 2022.

**Syed A. Rizvi**, Akash Awasthi, Maria J. Peláez, Zhihui Wang, Vittorio Cristini, Hien van Nguyen, Prashant Dogra

## **INDUSTRY EXPERIENCE**

---

### **Amazon, Austin, TX**

**May 2022 – August 2022**

*Software Development Engineer Intern*

- Developed a launcher application for starting customer screen sharing sessions on Amazon devices using the Spring framework in Java.

### **Phillips 66, Houston, TX**

**May 2021 – August 2021**

*IT Intern (Natural Language Processing)*

- Trained entity recognition models on land exchange contract documents.

Last Update: February 23<sup>rd</sup>, 2024

Syed Asad Rizvi

- Identified 6 domain-specific entities with 87% overall model precision.
- Deployed models to AzureML Cloud Platform and developed an automated Azure Function App to run preprocessing and inference on new contract documents within 12 seconds.

**Taipei Medical University, Taipei, Taiwan**

**March 2021 – April 2021**

*Data Analyst Intern*

- Performed on data processing, correlation analysis, and visualization on wearable device data measurements taken from 18 Taiwanese patients.

**PRESENTATIONS**

---

- Attention as Message-Passing for Graph Neural Networks, Yale University
- Mayday Screen Sharing Intern Presentation, Amazon
- COVID-19 Infection Forecasting with Explainable Spatiotemporal GNNs, University of Houston
- Histopathology DatasetGAN Oral Abstract, IEEE SPMB 2022 Virtual Conference
- MorphSet Project Oral Abstract, 2021 AI in Nephropathology Workshop in Amsterdam
- Natural Language Processing and Entity Recognition Models, Phillips 66

**AWARDS AND RECOGNITIONS**

---

- Provost’s Undergraduate Research Scholarship (\$1000), University of Houston, Spring 2022
- Dean’s Distinguished Scholar’s List, University of Houston
- First prize in the 2021 HP & AWS Bot-a-thon
- Third prize (\$3000) in the 2020 AWS & NVIDIA Environmental Hackathon

**INDEPENDENT PROJECTS**

---

**AWS Lex Bot Generation Pipeline**

**January 2021**

- 1<sup>st</sup> place finish among 20+ teams at the 2021 HP & AWS Bot-a-thon competition.
- Wrote chatbot configuration files and led presentation development.

**Autoencoder Anomaly Detection**

**August 2020**

- 3<sup>rd</sup> place finish in the AWS & NVIDIA Environmental Hackathon (\$3000 award).
- Trained an unsupervised autoencoder machine learning model on environmental sensor data taken from Amazon's Seattle Sphere conservatories.

**NutrientView Mobile App**

**July 2020**

- Nutrient logging mobile app utilizing image recognition services to track consumed meals.
- Integrated an Azure Q&A chatbot to provide interactive feedback about different nutrients.
- Developed using React Native, IBM Watson image recognition, Azure bot service, Firebase, and the Edamam Nutrition Analysis API.

**SKILLS**

---

Programming Languages: Python, C++, Java, R, MATLAB, SQL, JavaScript  
 Libraries: Pytorch, Pytorch Geometric, Tensorflow, Scikit-learn, Pandas, Numpy  
 Tools: Parallel programming, Distributed training, HPC job scheduling, Git

**ACTIVITIES**

---

**Management Information Systems Student Organization**

**January 2020 – December 2020**

*Professional Development Committee Member*

- Worked with teams of 20+ committee members to perform 60+ resume reviews per semester.
- Delivered a presentation on IT candidate profile development at the MISSO professional development workshop to over 80 students.