

Shaan Pakala

Email: shaan.pakala@gmail.com

Website: <https://shaanpakala.github.io/>

[Google Scholar](#), [LinkedIn](#) & [GitHub](#)

About

I am a 1st year Computer Science Ph.D. student at the University of California, Riverside, where I work on machine learning research problems with Professor [Vagelis Papalexakis](#). Generally, I am interested in developing machine learning methods for interdisciplinary scientific applications in domains such as material design, physics, and healthcare. The majority of my work involves learning from multidimensional data through the use of tensor decomposition. I have also interned and collaborated with scientists at Lawrence Livermore National Laboratory.

Education

Ph.D. in Computer Science

Sept. 2025 – Present

University of California, Riverside

- Advisor: Prof. Vagelis Papalexakis

B.S. in Data Science & Engineering

Sept. 2021 – June 2025

University of California, Riverside

- 3.9 upper division GPA
- Chancellor's Honor List

Research Experience

Graduate Research Assistant

June 2025 – Present

University of California, Riverside

- Worked on using tensor decomposition to efficiently train image generation models [3]
- Worked on surrogate modeling in material design with Lawrence Livermore National Laboratory [4]

Graduate Research Intern

June 2025 – Present

Lawrence Livermore National Laboratory

- Worked in the predictive healthcare group with Dr. Braden Soper & Dr. Priyadip Ray, in collaboration with clinicians/neuroscientists from Stanford & The University of Tokyo

Undergrad Research Assistant

June 2024 – June 2025

University of California, Riverside

- Worked with Professors Papalexakis, Tsotras, and Chen on surrogate modeling to efficiently design optimal data science pipelines (hyperparameter optimization, SQL query cardinality estimation) [1]
- Also worked on using machine learning based surrogate modeling for material design [2]

Undergrad Research Assistant

March 2024 – Dec. 2024

University of California, Riverside

- Worked with Professor Lonardi on using machine learning for accelerated protein sequence analysis

Papers

Main Conference

- [1] [Shaan Pakala](#), B. Graw, D. Ahn, T. Dinh, M. T. Mahin, V. Tsotras, J. Chen, E. Papalexakis, “Automating Data Science Pipelines with Tensor Completion,” *IEEE International Conference on Big Data* (2024). **Received Student Travel Award.** [\[Link\]](#) [\[PDF\]](#) [\[Code\]](#)

Workshop

- [2] [Shaan Pakala](#), D. Ahn, E. Papalexakis, “Tensor Completion for Surrogate Modeling of Material Property Prediction,” *AAAI Bridge on Knowledge-Guided Machine Learning* (2025). [\[PDF\]](#)
- [3] P. Goulart*, [Shaan Pakala](#)*, E. Papalexakis, “Efficiently Generating Multidimensional Calorimeter Data with Tensor Decomposition Parameterization,” *ICCV Workshop on Representation Learning with Very Limited Resources* (2025). [\[PDF\]](#) [\[Code\]](#)
- [4] [Shaan Pakala](#), A. Gongora, B. Giera, E. Papalexakis, “Surrogate Modeling for the Design of Optimal Lattice Structures using Tensor Completion,” *NeurIPS Workshop on AI for Accelerated Materials Design* (2025). [\[Code\]](#)

* denotes equal contribution

Awards

Undergraduate Research Spotlight <i>University of California, Riverside</i>	2025
Student Travel Award <i>IEEE International Conference on Big Data</i>	2024
NSF REU Fellowship <i>University of California, Riverside</i>	2024
Chancellor’s Honor List <i>University of California, Riverside</i>	2023 – 2024

Other Experience

Computer Science Grader <i>University of California, Riverside</i>	March 2024 – June 2024
Data Science Challenge <i>Lawrence Livermore National Laboratory</i>	July 2023
Data Science Camp Mentor <i>Spotline, Inc.</i>	July 2022 – Sept. 2022