Shaan Pakala

https://shaanpakala.github.io/

Email: shaan.pakala@gmail.com   LinkedIn & GitHub

I am an incoming Computer Science Ph.D. student at the University of California, Riverside. I will be continuing my ongoing work with Professor Vagelis Papalexakis, where we look into interdisciplinary research applications of tensor decomposition, mainly for the surrogate modeling of various combinatorial problems. This summer I will intern at the Lawrence Livermore National Laboratory Data Science Institute.

Education

Ph.D. in Computer Science (incoming)   To begin Sept. 2025
University of California, Riverside

· Advisor: Professor Vagelis Papalexakis

B.S. in Data Science & Engineering   Sept. 2021 – June 2025
University of California, Riverside

· Upper Division GPA: 3.93

Research Experience

Data Science Graduate Research Intern (incoming)   Summer 2025
Lawrence Livermore National Laboratory

· Selected for 2025 Data Science Summer Institute (DSSI) for graduate students

Data Science Undergraduate Researcher   Sept. 2024 – Present
University of California, Riverside

· Led research work on using tensor completion to predict material properties (e.g. energy, band gap)
· Presented preliminary work [2] at AAAI 2025 Bridge on Knowledge-Guided Machine Learning
· Extended work into full paper [3] (currently under review at *Scientific Reports*)

NSF REU Undergraduate Research Intern   Summer 2024
University of California, Riverside

· Led team of 3 undergraduates in research project, in collaboration with UCR Ph.D. students
· Presented full conference paper [1] at IEEE International Conference on Big Data 2024
· Modeled hyperparameter tuning, neural network architecture search, and SQL query cardinality estimation as tensor completion problems to predict their optimal configurations
· Developed task-specific tensor completion algorithm to cut parameters without losing performance

Bioinformatics Undergraduate Researcher   March 2024 – Dec. 2024
University of California, Riverside

· Worked on bioinformatics research problems using machine learning, for protein sequence analysis
· Developed machine learning model to classify protein subsequences as antimicrobial peptides
· Conducted literature reviews, and experimented with data processing techniques and ESM (LLM)

Papers

Conference

[1] <u>Shaan Pakala</u>, 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]

Poster

[2] <u>Shaan Pakala</u>, D. Ahn, E. Papalexakis, "Tensor Completion for Surrogate Modeling of Material Property Prediction," *AAAI 2025 Bridge on Knowledge-Guided Machine Learning.* [PDF] [Venue]

Under review

[3] <u>Shaan Pakala</u>, E. Papalexakis, "Accelerating Material Design with Tensor Completion," Under review in *Scientific Reports.*

Awards

Dean's Distinguished Award 2025
· To fully fund my first year of Ph.D. studies at UC Riverside

IEEE BigData Travel Award 2024
· Travel award ($800) to present my research at IEEE BigData in Washington, D.C.

Chancellor's Honor List 2023 – 2024
· For academic achievement in coursework at UC Riverside

Other Experience

Computer Science Grader Spring 2024
University of California, Riverside

· Grader for upper division Data Analysis Methods (CS 105 at UCR)
· Facilitated lab and project demos, as well as graded quizzes and reports

Data Science Challenge July 2023
Lawrence Livermore National Laboratory

· Participated in the Data Science Challenge, to develop data-driven approaches to cardiology problems
· Used electrocardiogram time-series data to create machine learning disease classification tool, as well as displaying 3D activity map of heart (electrical activity of 75 locations in the heart over 500ms)

References

Dr. Evangelos E. Papalexakis
Associate Professor
Department of Computer Science & Engineering
University of California, Riverside
https://www.cs.ucr.edu/~epapalex/

Additional references available upon request.