Saurabh Sivakumar

sausiva@ucdavis.edu | linkedin.com/in/saurabhsivakumar | github.com/saurabhsivakumar

Education

University of California — Davis, CA
Ph.D. in Chemical Engineering — *Sep 2021 – Present*

Carnegie Mellon University — Pittsburgh, PA
M.S in Chemical Engineering — *Aug 2019 – Dec 2020*

National Institute of Technology — Tiruchirappalli, India
B.Tech in Chemical Engineering, Minor in Economics — *Jul 2015 – May 2019*

Interests

- Computational Catalysis and Molecular modelling
- Data science applications in Engineering
- Deep Learning and Active Learning
- High-performance computing and workflows

Skills

- **Software/Frameworks:** VASP, QuantumEspresso, LAMMPS, COMSOL, GAMS, OVITO, ChemCAD
- **Programming/Markup Languages:** Python (**Packages:** ASE, Pytorch, scikit-learn, Numpy, Scipy, RDKit, MDanalysis, Seaborn, Matplotlib, Plotly, Pandas, WandB, Deap), R, Bash, LaTeX, C/C++, Markdown

Research Experience

Graduate Student Researcher — Sep 2021 – Present
University of California — *Davis, CA*

- Advised by Dr. Ambarish Kulkarni, focusing on identifying and solving problems related to surface and electro-catalysis through molecular simulations and machine learning
- Developed machine-learned potentials (MLPs) through active learning and molecular dynamics to predict surface phenomena like diffusion and desorption of adsorbates at ab-initio accuracy for a fraction of the compute cost
- Leading efforts to use machine-learned potentials and datasets for complex reaction networks for use with micro-kinetic modeling by interfacing with national labs and developing open source workflows
- Working on identifying transition states for surface catalytic reactions using a combination of MLPs + DFT + Pynta (an open source package)

Graduate Researcher & Research Assistant — Jan 2020 – Jul 2021
Carnegie Mellon University — *Pittsburgh, PA*

- Focused on applying machine learning and mathematical modelling to the field of atomic scale simulations on projects with the Ulissi Group in the department of Chemical Engineering
- Developed and applied an Active learning framework for accelerating Nudged Elastic Band calculations
- Built and tested an open source program to identify lowest energy nanoclusters on a potential energy surface using a modified genetic algorithm
- Implemented a genetic algorithm with inbuilt active learning methods (Neural networks and Gaussian processes) to reduce the computational time required for identifying the optimal shape of nanoclusters
- Assisted the team working on surface segregation with Deep Reinforcement learning with dataset training and bayesian parameter optimization

Undergraduate Thesis — Jan 2019 – May 2019
National Institute of Technology — *Tiruchirappalli, India*

- Designed a chemical plant and associated equipment for manufacturing Trichloroethylene with a teammate, under the guidance of Dr. Saravanan
- Formulated a theoretical design with focus on cost analysis and safety using ChemCAD

Research Intern — May 2018 – Aug 2018
Singapore University of Technology and Design — *Singapore*

- Guided by Dr. Arief Budiman, worked with his group on projects related to optimizing lightweight solar PV modules with a polycarbonate substitute

Publications

- Enabling robust offline active learning for machine learning potentials using simple physics-based priors
  - Muhammed Shuaibi, **Saurabh Sivakumar**, Rui Qi Chen and Zachary W Ulissi
  - [Paper Link]
- Cluster-MLP: An Active Learning Genetic Algorithm Framework for Accelerated Discovery of Global Minimum Configurations of Pure and Alloyed Nanoclusters
  - Rajesh K. Raju, **Saurabh Sivakumar**, Xiaoxiao Wang and Zachary W Ulissi
  - [Paper Link]
- Studying the diffusion of adsorbates on Ag (111) using machine learning accelerated molecular dynamics
  - **Saurabh Sivakumar** and Ambarish Kulkarni
  - Manuscript in preparation

Selected Projects

Classification of Musk Dataset from UC Irvine Mar 2022 – June 2022

- Implemented a neural network, k-NN, decision tree, and logistic regression classifiers with a prediction accuracy of at least 96% across all the classifiers.
- Studied the effect of varying kernel types for SVM and different optimizers for NN and performed feature extraction with PCA
- Presented results including decision boundaries and errors for the classifiers

Linear regression with the Abalone dataset from UC Irvine Sep 2021 – Dec 2021

- Found the best linear regression model to predict the age of abalone using a four-step methodology using R
- Performed Box-cox transformation, then using a Greedy search strategy, a stepwise regression algorithm was implemented to find the best model containing the interaction terms according to AIC and BIC criterion. Ridge regression was also performed on the dataset due to high multicollinearity and the two models were compared and their metrics were found to be similar

Analysis of the COVID-19 dataset Jan 2022 – March 2022

- Explored the COVID-19 dataset curated by the New York Times and showcased the effect of the COVID-19 pandemic across the USA.
- Visualized the number of cases/mortalities to each county and state in the USA using plotly and Chloropleth plots. Subsequently, the trends in certain states and counties were analyzed by plotting the time series over 2 years with a weekly moving average to gain insights such as the effect of vaccination rates.

Optimal Scheduling of Copper concentrate operations under uncertainty Feb 2020 – May 2020

- Performed optimization for an industrial scheduling problem (Mixed Integer Nonlinear problem) for a copper plant using code written in GAMS
- Corroborated the results of the publication on which the project was based and presented a case for further improvement (recommended adding further uncertainty to certain parameters)

Conferences & Posters

- AIChE Annual Meeting 2020 (Virtual). November 20 2020
  - An Active Learning Framework for Accelerating Saddle Point Searches
  - [Poster Link]
- Sandia & CERCAS Symposiums 2022
  - Developing an integrated experiment theory approach to provide new insights into heterogeneous catalysis using atomically dispersed materials
  - [Poster Link]

Other Experiences & Awards

- ACCESS compute grant worth 750k credits (500k core hrs) (Oct 2021 - Present)
- Multiple NERSC compute grants totalling over 20k CPU and GPU hrs (Nov 2021 - Present)
- Member, AIChE (Oct 2019 - Present)
- Member, ACS (Feb 2021 - Present)
- Teaching Assistant, Chemical Reaction Engineering, 2023
- Head, Media Relations – Pragyan (NITT's Technical Organization)