Raviteja Vangara

rvangara.phd@gmail.com ravitejavangara github.com/rvangara ravitejavangara.com
rvangara San Diego, USA 505-488-1420

Summary

Experienced computational scientist with over 8 years of programming expertise and 5 years of dedicated research in machine learning. Seeking roles to apply my expertise in AI/ML and engineering to address complex challenges and drive innovation. Proven ability to collaborate effectively in multidisciplinary teams, lead projects, and communicate my work to technical teams and non-technical stakeholders.

Experience

- **University of California, San Diego** — San Diego, CA
  Post-Doctoral Research Scientist — *Sep 2021 - current*
  - **Machine learning for Cancer Genetics**:
    - Engaged as a key team member in the groundbreaking **Mutographs** research team, collaborating with international scientists to advance our understanding of cancer genetics and mutational patterns.
    - Developed **SigProfilerExtractor**, a widely used tool in the cancer community. Designed and implemented the tool using PyTorch, enabling seamless parallel computing on both GPU and CPU platforms.
    - Utilized Non-Negative Matrix Factorization (**NMF**) in combination with diverse optimization techniques, such as **MU** (Multiplicative Update) and **HALS** (Hierarchical Alternating Least Squares), alongside customized clustering and stability analysis.
    - Outperforming other existing tools by elucidating **20%** and **50%** more true-positive signatures while yielding **5-fold** less false-positive signatures.
    - Designed and implemented **SigProfilerAssignment**, a tool for the attribution of mutational signatures in known cancer signature catalogs for the somatic variant profiles obtained from single or multiple cancer patients, which leverages **NNLS** using **forward** and **reverse stagewise** algorithms for better convergence.
    - Developed **Deep learning** methods, specifically a weakly supervised **convolutional neural network** architecture that uses **multiple instance learning** (MIL) for predicting various biomarkers from Tumor and Normal digital H&E slides.
    - Employed a **multi-resolution decision system** designed to emulate the standard diagnostic protocol employed by pathologists which conducts an initial prediction at low magnification (5x magnification) and autonomously identifies regions of interest (ROI) for a subsequent high-resolution prediction within those ROIs (20x magnification).
    - Employed **Bayesian dropout** during inference of a tissue slide to calculate confidence intervals for the final model predictions.

- **Los Alamos National Laboratory** — Los Alamos, NM
  Post-Doctoral Research Scientist — *Mar 2020 - Aug 2021*
  - **Tensor Networks**: Robust Unsupervised Machine Learning for Big-Data Analytics. Developed various non-negative matrix and tensor factorization techniques using chemical graph theory.
  - **Supervised learning on TF-DNA binding models**: Predicting noncoding variants that affect TF-DNA binding using Large-scale simulations and computations to extract dynamical and structural properties of DNA, from publicly available genomic datasets and development of supervised machine learning tool with multimodal deep neural nets fusion to predict mutations affecting TF-binding which help in formulating biological hypotheses, facilitating the development of targeted therapeutics.

  Graduate Research Scientist — *Apr 2018 - Mar 2020*
  - **Scalable Machine Learning**: Design & Development of Scalable machine learning developer for high performance computing systems involving shared and distributed memory applications on LANL supercomputers

- **University of New Mexico** — Albuquerque, NM
  Graduate Research Scientist -Ph.D. — *Dec 2015 - Apr 2018*
  - **Computational Chemistry**:
    - Developing density functional models for charged interfaces involving electrolyte solutions;
    - Built a python framework for classical density functional theory coupled with surface charge regulation; model treats solvent explicitly and accounts surface charge basing on thermodynamic chemical equilibrium.
    - Observed new physical insights and the steps for a molecular theory to address important features like the role of non-columbic interactions - ionic solvation and surface-ion interactions, on the actual electrostatics of the system i.e. the electric double layer.
    - Published six research papers in reputed scientific peer reviewed journals

Education

- **University of New Mexico** — Albuquerque, NM, USA
  Ph.D. - Engineering [With Distinction]; GPA: 4.04/4.0 — *2019*
- **University of New Mexico** — Albuquerque, NM, USA
  MS - Chemical Engineering; GPA: 4.04/4.0 — *2017*
- **National Institute of Technology Warangal** — Warangal, India
  B.Tech - Chemical Engineering; GPA: 7.33/10.0 — *2015*

Skills Summary

- **Languages**: Python, Bash, Matlab, R, Julia,C,C++
- **Frameworks**: PyTorch, TensorFlow, Docker, EC2,scikit learn, OpenCV, HuggingFace
- **Machine Learning**: Matrix and Tensor Factorization, Dimension reduction techniques, Linear and Non linear models in classification and regression, Clustering, graphical models
- **Deep Learning**: Natural Language Processing, Computer Vision, Generative AI, GANs, VAEs, Diffusion Models, Transformers

Selected Publications

- Nebgen, Benjamin, **Raviteja Vangara**, Miguel A. Hombrados-Herrera, Svetlana Kuksova, and Boian Alexandrov. "A neural network for determination of latent dimensionality in Nonnegative Matrix Factorization." Machine Learning: Science and Technology (2020).
- Islam, SM Ashiqul, Marcos Díaz-Gay, Yang Wu, Mark Barnes, **Raviteja Vangara**, Erik N. Bergstrom, Yudou He et al. "Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor." Cell Genomics 2, no. 11 (2022): 100179.
- Bhattarai, Manish, Ben Nebgen, Erik Skau, Maksim Eren, Gopinath Chennupati, Raviteja Vangara, Hristo Djidjev, John Patchett, Jim Ahrens, and Boian Alexandrov. "pydnmfk: Python distributed non negative matrix factorization." (2021).
- **Raviteja Vangara**, Erik Skau , Gopinath Chennupati , Hirsto Djidjev , Thomas Tierney , James Smith , Manish Bhattarai , Valentin Stanev , Boian Alexandrov " Semantic Nonnegative Matrix Factorization with Automatic Model Determination for Topic Modeling" in 19th IEEE International Conference On Machine Learning And Applications.
- **Raviteja Vangara**, Manish Bhattarai, Erik Skau, Gopinath Chennupati, Hristo Djidjev, Tom Tierney, James P. Smith, Valentin G. Stanev, and Boian S. Alexandrov. "Finding the number of latent topics with semantic non-negative matrix factorization." IEEE Access 9 (2021): 117217-117231
- Chennupati, Gopinath, **Raviteja Vangara**, Erik Skau, Hristo Djidjev, and Boian Alexandrov. "Distributed non-negative matrix factorization with determination of the number of latent features." The Journal of Supercomputing (2020): 1-31.
- Manish Bhattarai, Gopinath Chennupati, Erik Skau, **Raviteja Vangara**, Hristo Djidjev and Boian Alexandrov "Distributed Non-Negative Tensor Train Decomposition", 2020 IEEE High Performance Extreme Computing.
- **Raviteja Vangara**, Kim Ø. Rasmussen, Dimiter N. Petsev, Golan Bel, and Boian S. Alexandrov. "Identification of anomalous diffusion sources by unsupervised learning." Physical Review Research 2, no. 2 (2020): 023248.
- Kazi Lutful Kabir, Gopinath Chennupati, **Raviteja Vangara**, Hristo Djidjev, Boian Alexandrov, and Amarda Shehu. "Decoy Selection in Protein Structure Determination via Symmetric Non-negative Matrix Factorization", In 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).
- Manzini, Gianmarco, Erik Skau, Duc P. Truong, and **Raviteja Vangara** "Nonnegative tensor-train low-rank approximations of the Smoluchowski coagulation equation." In Large-Scale Scientific Computing: 13th International Conference, LSSC 2021, Sozopol, Bulgaria, June 7–11, 2021, Revised Selected Papers, pp. 342-350. Cham: Springer International Publishing, 2022.

Patents/ Invention Disclosure

- Integral transforms through non-linear kernel: Joint disclosure filing between LANL and UNM. *-Patent pending*

Honors and Awards

- Part of winning team of prestigious 2021 R&D 100 Award [termed as Oscar of Inventions] for the team "Smart Tensors", Los Alamos National Laboratory ,

  Media/posts

  - https://www.rdworldonline.com/rd-100-winner-of-the-day-smarttensors-ai-platform/
  - https://collaboration.lanl.gov/smarttensors/

- Lead PI for Institutional Computing proposal at LANL to perform large scale high-performance computing calculations of non-linear breathing dynamics of DNA, awarding 4.5 million core hours on LANL HPC clusters.
- Milton Levy Award of Corrosion Science for poster presented in 2017, DOD Allied Nations Technical Corrosion Conference.
- Awarded International Amigo Scholarship for graduate studies at UNM
- Awarded Prime Minister (of India) Merit Scholarship (PMSS) for my under-graduation studies at NIT Warangal, India.