

Resume

✉ nr57@rice.edu | 📧 narein97 | 🐦 @rao_narein | 🎓 Narein Rao

Education

RICE University

PHD IN COMPUTER SCIENCE

Advisor: Dr. Luay Nakhleh

Houston, Texas

Aug. 2023 - Present

Indian Institute of Technology (IIT), Kanpur

MS BY RESEARCH IN COMPUTER SCIENCE AND ENGINEERING

CPI: 9.0/10.0

Kanpur, India

Sept. 2020 - Jul. 2023

MVJ College of Engineering (MVJCE), Bangalore

B.TECH IN COMPUTER SCIENCE AND ENGINEERING

CGPA: 8.04/10.0

Bangalore, India

June. 2015 - June. 2019

Research Interests

Scalable and efficient computational techniques involving aspects of probabilistic machine learning, statistics, and algorithms for analyzing genomic data.

Publications

TransLIST: A Transformer-Based Linguistically Informed Sanskrit Tokenizer

JIVNESH SANDHAN, RATHIN SINGHA, NAREIN RAO, SUVENDU SAMANTA, LAXMIDHAR BEHERA, PAWAN GOYAL

EMNLP22 (Findings) - DOI: arXiv:2210.11753

Charting spatial ligand-target activity using Renoir

NAREIN RAO, TANUSH KUMAR, RHEA PAI, ARCHITA MISHRA, FLORENT GINHOUX, JERRY CHAN, ANKUR SHARMA, HAMIM ZAFAR

(Under review - Nature Communications) biorxiv - DOI: doi.org/10.1101/2023.04.14.536833

Intratumor Heterogeneity Through the Lens of Gene Regulatory Networks

NAREIN RAO, NICOLAE SAPOVAL, HAMIM ZAFAR, LUAY NAKHLEH

RECOMB CCB 2024 - biorxiv - DOI: doi.org/10.1101/2025.04.01.646625

Research Experience

Intratumor Heterogeneity Through the Lens of Gene Regulatory Networks

RICE University

SUPERVISOR : DR. LUAY NAKHLEH

July. 2024 - Present

- Reliably infer differential network patterns between clonal subpopulations offering complementary insight to differential gene analyses.
- Analyses indicate that a substantial proportion of GRN edges can be attributed to clone-specific activity

Charting spatial ligand-target activity

IIT Kanpur

GRADUATE THESIS | SUPERVISOR : DR. HAMIM ZAFAR

Oct. 2021 - Oct. 2022

- Developed a novel approach to chart ligand target activity across spatial transcriptomic data.
- Allows inference of spatially resolved ligand-target interaction domains, colocalized celltype interactions and signalling pathways.

Gene Regulatory Networks for Spatial Transcriptomic Data

IIT Kanpur

RESEARCH ASSISTANT | SUPERVISOR: DR. HAMIM ZAFAR

Mar. 2021 - July 2021

- Developed an algorithm to extend gene regulatory network inference for spatial transcriptomic data.
- Discovering gene regulatory networks for domains defined by spatial gene expression and celltype distribution.

Automatic hyperparameter tuning for celltype deconvolution

IIT Kanpur

RESEARCH ASSISTANT | SUPERVISOR: DR. HAMIM ZAFAR

Nov. 2022 - March 2023

- Analyzed the impact of various distance metrics on cell type clustering and meta-cell inference in single-cell RNA sequencing datasets to evaluate clustering robustness and biological relevance.
- Developed an automated parameter search algorithm to identify optimal configurations for clustering and meta-cell inference tools, improving performance of downstream analyses.

Sanskrit Word Segmentation

IIT Kanpur

INDEPENDENT COLLABORATION

Jun. 2021 - Jan. 2022

- Developed a Transformer based Linguistically Informed Sanskrit Tokenizer capable of tackling Sandhi phenomenon.
- Outperformed the current state of the art system by an average 7.2 points absolute gain in terms of perfect match (PM) metric

Academic / Research Projects

Predicting drug resistance in Mycobacterium Tuberculosis

Course: Computational Genomics

MENTOR: DR. HAMIM ZAFAR

Oct. 2020 - Nov. 2020

- Developed statistical models to predict the resistance of Mycobacterium tuberculosis (MTB) towards several first and second line drugs commonly used for treating tuberculosis.
- Showcased an average predictive accuracy of 92% across 10 first and second line drugs.

Analysis of degree of contribution of mutations in Mycobacterium Tuberculosis

Undergraduate Thesis

MENTOR: DR. MANJU KHANNA

Sept. 2018 - May. 2019

- Developed an algorithm to understand the relation between mutations and drug susceptibility exhibited by mycobacterium tuberculosis
- The implementation is based on "Machine learning for classifying tuberculosis drug-resistance from DNA sequencing data" by Yang Yang et al.

Comparison of GAN and VAE in continual learning

Course: Probabilistic Modeling and Inference

MENTOR: DR. PIYUSH RAI

Mar. 2021 - May. 2021

- Provided a measure of realism for the images generated by GAN/VAE
- Found the relation between no. of tasks and forgetfulness in GAN/VAE
- Generated images over a domain perceptually-distant from the domain the models have been trained over

Study on the effect of Covid-19 lockdown on Air Quality in India

Course: Data Mining

MENTOR: DR. ARNAB BHATTACHARYA

Sept. 2020 - Nov. 2020

- Integrated Air Quality Index (AQI) and air borne disease data from multiple data sources.
- Performed time series predictions, statistical and probabilistic analysis to gain further insights between AQI, air borne diseases and number of Covid-19 cases.

Bluetooth attendance system

Research Project

INDEPENDENT PROJECT

Feb. 2018 - Apr. 2018

- Developed a multi-agent based bluetooth attendance system (Proxy) using JADE framework with a user friendly android application.
- Proxy is a bluetooth-based attendance system that employs smart phones and (optionally) bluetooth tags to speed up attendance calls and automate student registrations and provide log reports.

Multi agent system for power regulation

Research Project

MENTOR: DR. MANJU KHANNA

Feb. 2018 - Apr. 2018

- A multi-agent system that regulated power supplied by wind and solar energy sources was simulated and developed as a prototype. The simulation was executed over existing data sources.
- Anylogic simulations were used to examine the behaviour of the agents in the environment, and a prototype system was developed using the JADE framework.

Notable Achievements

- Cleared GATE 2020 entrance with an overall standing within the top 0.6 % of total participating students.
- Proxy (bluetooth attendance system) gained press attention from five publications, including some of India's most prominent news organisations (Times of India, 2018).

Relevant Coursework

- | | | |
|--|-------------------------------------|-----------------------------------|
| • Computational Genomics | • Data Mining | • Programming and Data Structures |
| • Probabilistic Modeling and Inference | • Big data analytics | • Unix and shell programming |
| • Introduction to Machine Learning | • Design and analysis of algorithms | |

Technical Skills

Programming	Python, R, C, Java, Bash, Latex
DevOps	Docker, Git, Firebird
Cloud Platforms	AWS, Microsoft Azure
Tools	PyTorch, Seurat, Scanpy, Tableau, JADE

Teaching Experience

Programming for Data Science (COMP614)

TEACHING ASSISTANT

RICE University

Aug. 2024 - Dec. 2024

- Core responsibilities included teaching classes, holding office hours to address questions and clarify concepts, contributing to the design of tests and assignments, and grading student work.

Graduate Tools and Models - Data Science (COMP543)

TEACHING ASSISTANT

RICE University

Jan. 2024 - May 2024

- Core responsibilities holding office hours to address questions and clarify concepts and grading student work.

Fundamentals of Computing (ESC101)

TEACHING ASSISTANT

IIT Kanpur

Jun. 2021 - Jun. 2022

- Core responsibilities included conducting first-year undergrad labs, quizzes and grading lab solutions.

Discrete Mathematics

TUTOR

Freelance

Sep. 2022 - Present

- Held one-to-one tutoring sessions for students which involved lectures, assignments and quizzes.

Operating Systems

TUTOR

Freelance

Jun. 2022 - Jul. 2022

- Held one-to-one tutoring sessions for students which involved lectures and assignments.