

# Rohan Sonthalia

[sonthaliarohan@gmail.com](mailto:sonthaliarohan@gmail.com) | +91 8334977754

[linkedin.com/in/rohan-sonthalia](https://linkedin.com/in/rohan-sonthalia)

[github.com/rohan-sonthalia](https://github.com/rohan-sonthalia)



## Education

- **University of Wisconsin-Madison**, Visiting International Student, Computer Science  
GPA: 3.91/4.00 Sept 2023 - Dec 2023
- **SRM Institute of Science and Technology**, Bachelor of Technology, Computer Science  
GPA: 9.9/10.0 May 2025
- **Don Bosco School Park Circus**, High School Diploma  
Percentage: 94.5% Mar 2021

## Research Experience

- **Research Assistant**, Solis-Lemus Lab, Wisconsin Institute of Discovery, University of Wisconsin-Madison  
Jul 2023 - Sept 2024
  - Extended the BioKlustering project to analyze phenotype-genotype correlations using genomic data.
  - Utilized transformers to embed DNA sequences into vectors and implemented neural networks for phenotype predictions.
  - Conducted research on semi-supervised learning for nonlinear genomic data.
  - Delivered a talk at the **ML+X Forum** highlighting advancements in leveraging transformer models for genomic applications.
  - Attended the yearly Wisconsin Institute for Discovery Symposium along with several other gatherings of researchers within the university.
- **MITACS Globalink Research Intern**, Dumeaux Lab, University of Western Ontario  
Jun 2024 - Aug 2024
  - Conducted Deep Archetypal Analysis on gut microbiome data to explore nonlinear combinations of Gene Expression Programs.
  - Curated our dataset by combining data from 38 different studies from various research laboratories across the globe.
  - Preprocessed compositional data through log-ratio transformations and compared multiple deepAA models.
  - Compared results of archetypal analysis generated by running several open source repositories on our data.

## Work Experience

- **Software Engineer Intern**, Tune AI (formerly NimbleBox.ai), Chennai, India  
Jan 2024 - May 2024
  - Curated Automated API Testing suites for web servers using Golang.
  - Implemented Structured Language Generation for serving as well as improving the output of Large Language Models.
  - Finetuned Stable-Diffusion-XL base along with a refiner on it's FLAX pipeline on a publicly curated dataset for coherent image generation on our platform.
  - Leveraged Golang packages such as Viper, to setup configurations for server applications.
- **Software Engineer Intern**, NimbleBox.ai, Chennai, India  
Jul 2022 - Aug 2023
  - Fine-tuned LLMs such as MPT-7B and LLaMa-7B using Parameter Efficient Fine-tuning and Lower Rank Adaptation (LoRA) for serving on our GenAI platform.
  - Built end-to-end API testing suites using Postman for our MLOps platform.
  - Developed Infrastructure as Code (IAC) templates on Pulumi to launch VMs and wrote DockerFiles for Azure CLI authentication.
  - Created an Internal Notification CronJob to alert our Engineering Team when bots were to create accounts on our platform, in order to prevent an imminent DDoS attack.
  - Developed a filter to remove bots and malicious users from our platform

## Publications and Presentations

- Presented research findings at the **ML+X Forum**, Wisconsin Institute of Discovery.  
*Title: Clustering of Mycoviruses using Transformers and Deep Learning.*

## Projects

- **BioKlustering: Web App for Phenotype Predictions from Genotype**  
Developed three DNA sequence embedding methods using transformers, autoencoders, and kmer-cants. Deployed an interactive web app for real-time DNA sequence analysis.
- **Automated Online Video KYC Tool**  
Designed a tool using OpenCV and gTTS libraries for automating KYC processes with a regional-language UI for accessibility.
- **AutoPreprocessor Library**  
Created a Python library to automate data preprocessing workflows using Scikit-learn, Pandas, and Numpy.

## Awards and Honors

- Performance Based Scholarship, SRM Institute of Science and Technology 2023
- National Stage Qualification, ABU Robocon 2022.

## Coursework and Technical Skills

- **Relevant Coursework:** Artificial Neural Networks, Image Processing, Linear Algebra, Data Structures and Algorithms.
- **Programming Languages:** Python, Java, C++, R, Golang, MATLAB, JavaScript.
- **Tools and Frameworks:** TensorFlow, PyTorch, Scikit-learn, OpenCV, Docker, Kubernetes, Flask, SQLAlchemy.

## Leadership and Extracurriculars

- **Technical Lead**, Gen-Y, SRMIST Feb 2024 - Dec 2024
  - Developed a Sales Analysis Dashboard for a fashion conglomerate based in India.
  - Worked on a project that automated the generation of technical articles through web-scraping
- **Technical Team Member**, SRM Team Robocon, SRMIST Oct 2021 - Jan 2023
  - Worked on the problem statement of ABU Robocon 2022, and qualified for the National Stage of the same.
  - Trained a group of 21 students on robotics and microprocessors for a duration of 3 months.
- **Secretary**, Science Group, Don Bosco School May 2019 - Mar 2020
  - Organized intra-school scientific exhibitions for over 300 students.