Sindhura Kommu

✉ sindhura@vt.edu   in https://www.linkedin.com/in/sindhura1997/

○ https://github.com/sindhura-cs   🔗 https://sindhura-cs.github.io/   ◆ Google Scholar

Education

Virginia Tech, *M.S. in Computer Science - Thesis, GPA: 4.0*   2023 Jan – present
**Teaching Assistant**: Machine Learning , Spring 2023, Prof. Lifu Huang
Teaching Assistant: Intro to AI, Fall 2023, Prof. Peter Ghavami
Teaching Assistant: Intro to Programming in Python, Spring 2024, Prof. Lewis John
Graduate Research Assistant: Summer 2024, Prof. Xuan Wang

Indian Institute Of Technology Bombay (IIT Bombay),   2014 Jul – 2018 Jul
*B.Tech, Computer Science & Engineering* ⇱
Teaching Assistant: Computer Programming and Utilization, Spring 2018, Prof. Krishna Subramaniam

Research Experience

Graduate Student Researcher, Virginia Tech,   2023 Apr – present | Blacksburg, US
Advisor: Dr. Xuan Wang
Conducting pioneering research on **large language models for biological sequences**, with a focus on
developing **foundational models** for predictive and **explainable AI** in **single-cell multi-modality biology**.
Currently, leading efforts on one of the world's largest **proteomics** datasets to predict **heart disease
progression**. Funded by NSF ⇱ , CCI ⇱ , Amazon and Sanghani Center for AI and Data Analytics ⇱ .

Publications

**Sindhura Kommu, Yizhi Wang, Yue Wang, and Xuan Wang, "Gene Regulatory Network Inference from
Pre-trained Single-Cell Transcriptomics Transformer with Joint Graph Learning",** *The 41st International
Conference on Machine Learning AI for Science Workshop (ICML'24 AI4Science), July 26th, Vienna, Austria* ⇱

***Sindhura Kommu, *Zhenyu Bi, *Sajib Acharjee Dip, *Daniel Hajialigol, *Hanwen Liu, *Meng Lu, Xuan
Wang, "AI for Biomedicine in the Era of Large Language Models",**
The 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2024, Submitted ⇱

Academic Projects

Model Reasoning and Hallucinations in Medical LLMs, *Trustworthiness of LLMs, CS6804, Spring 2024*
Investigated **hallucinations in medical large language models (LLMs)** by assessing their **reasoning
abilities** using the newly developed Med-HALT benchmark.

Medical Image Segmentation using text enhanced Vision Transformers,
Learning-based Computer Vision, CS5864, Fall 2023
Enhanced medical image segmentation by **training vision transformers with text annotations**, improving
performance with **multimodal data fusion. Code** ⇱

Text-to-Image Synthesis using CLIP-guided GANs, *Advanced Machine Learning, CS5824, Spring 2023*
Implemented **CLIP guided StyleGAN** for **interactive fashion design tool** using **PyTorch**. The model is used for **generating fashion images using textual descriptions**. *Code* ⧉

Visual Question Answering, *Intro to AI, CS5804, Spring 2023*
Performed a comparative study between the **CNN+LSTM** approach and **Hierarchical Question-Image Co-Attention** approach for Visual Query Response. *Code* ⧉

Professional Experience

Fractal Analytics ⧉

Senior Data Scientist - AI@Scale   2020 Feb – 2022 Dec │ London, UK
- **Data Science Lab:** Utilized big data analytics and machine learning to extract valuable insights from **financial data**, providing customized solutions for internal stakeholders. **Led a team of engineers** to enhance the lab's machine learning capabilities, fostering growth and development.
- **AI for Threat Detection:** Developed a real-time, **data-driven surveillance system** for IdeaForge, a leading drone manufacturer, leveraging **GPU-powered UAV drones** for rapid threat detection.
- Co-authored a white paper on building **scalable** machine learning workflows using **MLOps** practices on cloud platforms such as **Vertex AI** and **Cloud Composer**.

Data Scientist - Image & Video Analytics   2018 Jun – 2020 Jan │ Mumbai, India
- Developed and deployed **AI models for image** classification, detection, and segmentation. Achieved >90% accuracy on complex real-world datasets using **YOLOv3** and other state-of-the-art techniques.
- **Video Analysis:** Automated **action detection** (dishwashing, etc.) in videos, reducing size by 98%. Shorter videos helped P&G understand consumer interaction, leading to smarter home care products.

Samsung R&D Institute, *Software Intern*   2017 May – 2017 Jul │ Bangalore, India
- Built **usability extension** in Tizen Browser for mirroring and hooking website regions, reducing scrolling.

Scholastic Achievements

- Awarded "**Advancing Inclusion Scholarship**" to attend **Grace Hopper Celebration (GHC) 2024.**
- Recived **departmental funding guarantee** for entire master's thesis program, encompassing **100% tuition waiver** and living stipend.
- Served as a Reviewer for: IEEE BigData'23, TheWebConf'24, Bioinformatics'24, Nature Methods'24.
- Secured **99.17 percentile** in **IIT JEE (Advanced)** with an ALL India Rank of 1231 out of **1.5 lakh** candidates.

Certifications

Structuring Machine Learning Projects ⧉
Coursera, deeplearning.ai

Convolutional Neural Networks ⧉
Coursera, deeplearning.ai

Deep learning A-Z, Hand-On Artificial Neural Networks ⧉
Udemy

Neural Networks and Deep Learning ⧉
Coursera, deeplearning.ai

Improving Deep Neural Networks: Hyperparam Tuning, Regularization and Optimization ⧉
Coursera

Deep learning, Nanodegree Program
⧉
Udacity

Sequences, Time Series and Prediction
⧉
Coursera, deeplearning.ai

The Science of Well-Being ⧉
Coursera, Yale University