Varun Ullanat

617-669-8069 ● varappu.ram@gmail.com ● GitHub ● LinkedIn
Biomedical AI researcher and prospective PhD student

EDUCATION

Harvard Medical School, Boston, MA
Master of Biomedical Informatics, GPA 3.969/4 **2022 – 2024**

RV College of Engineering, Bengaluru, India
Bachelor of Engineering, Biotechnology, GPA 3.98/4 **2017 – 2021**

PROFESSIONAL EXPERIENCE

Massachusetts Institute of Technology, Cambridge, USA **March 2024 – November 2024**
Research Assistant, Bonnie Berger Group

- Developed a novel protein language model (PLM) focused on modeling protein-protein interactions (PPIs). Trained it in a data-distributed, multi-GPU manner on an academic compute cluster for 500,000+ iteration steps over 5 months. Conducted rigorous benchmarking of the model with existing PLMs over 6 general PPI prediction tasks to achieve a 10% performance increase. Designed and conducted two case studies applying our model to realistic research scenarios: predicting the effects of potential cancer-causing mutations on PPIs and estimating the cross-neutralization effects of SARS-CoV-2 antibodies on Omicron variants.
- Independently drafted the 2500-word manuscript, made 6 high quality figures and was responsible for submitting it to conferences and journals.

Dana-Farber Cancer Institute, Boston, USA **May 2023 – Present**
AI/ML Engineer, Department of Informatics & Analytics

- Led a team of 3 engineers in enhancing PathML, a widely used open-source computational pathology software. Directed the addition of 20,000 lines of code, including the addition of 4 new functionalities for enhanced representation of pathological slides and increased test coverage by 10%.
- Created an AI pipeline for extracting key cancer-related data elements from raw, unstructured EHR notes of 20,000 Dana-Farber patients. Leveraged a HIPAA-compliant version of GPT-4 and NLP techniques such as RAG and in-context learning to achieve an average accuracy of 90% compared to manual human extractions.
- Developed Dana-Farber's first AI-powered chatbot application to allow 13,000 providers and support staff to ask interactive questions about key institutional policy documents.

Massachusetts Eye and Ear, Boston, USA **May 2023 – September 2023**
Summer Intern, Harvard Ophthalmology AI Lab

- Developed custom statistical scripts in R to analyse the results of a large registry-based retrospective cohort study containing data from 2000+ endophthalmitis patients. Performed rigorous statistical analyses to assess the impact of pars plana vitrectomy (PPV) on long-term clinical outcomes, using linear regression and hypothesis testing.

Harvard Medical School, Boston, USA **January 2023 – March 2024**
Master's thesis student, Zitnik Lab

- Contributed to a multimodal drug-drug interaction (DDI) project in collaboration with AstraZeneca. Developed and validated 4 modality-specific encoders; designed a novel technique for using textual inputs for DDI prediction, resulting in a model with a zero-shot AUPRC of over 0.60.
- Developed a new paradigm of latent graph learning with applications to protein function prediction, resulting in a model which contented with state-of-the-art methods on protein-ligand binding affinity prediction tasks.

National Centre for Biological Sciences, Bengaluru, India **September 2021 – June 2022**
Research Assistant, Integrative Structure Biology Lab

- Developed an integrative protein structure analysis package in Python using classical machine learning principles that is currently being added to PDB-dev.
- Created the architecture of a protein-protein interaction model for predicting interactions of intrinsically disordered proteins.

Indian Institute of Science, Bengaluru, India **October 2020 – September 2021**
Computational Intern, Cancer Systems Biology Lab

- Performed computational analysis of 5+ regulatory networks implicated in epithelial-mesenchymal plasticity of cancer, leveraging Python, R, and shell scripting. Collaborated with other researchers to develop a novel theory that "teams" of regulatory nodes are active in cancer, constraining phenotypic outcomes and cell-fate decisions.

SELECTED ACADEMIC WORK

Cultural insights from restaurant reviews using long-context LLMs **September 2023 — December 2023**
Harvard Medical School, Boston, USA

- Fine-tuned a pretrained Large Language Model (LLM) for multi-document summarization; evaluated the performance using four common NLP metrics.
- Performed additional fine-tuning on restaurant reviews and summaries to produce a model capable of multi-document summarization on groups of restaurant reviews. Built a simple interactive application using React and FastAPI.

Classification and generation of live cell microscopy images **January 2023 — May 2023**
Harvard Medical School, Boston, USA

- Developed a convolutional neural network (CNN) classifier for distinguishing diverse cell types from a live cell microscopy dataset. Conducted explainability analysis to identify key differentiating features between cell types.
- Trained a denoising diffusion probabilistic model (DDPM) with a U-Net backbone to generate synthetic live cell microscopy images to augment sparse datasets, supporting the development of automated microscopy analysis workflows.

Transcriptomic Analysis of the Brain Regions of Symptomatic and Asymptomatic Alzheimer's Patients **May 2020**
RV College of Engineering, Bengaluru, India

- Analyzed microarray data from the brains of symptomatic and asymptomatic Alzheimer's patients, using a variety of gene expression analysis techniques including differential gene expression (DGE), overrepresentation analysis (ORA), and gene set enrichment analysis (GSEA).

TEACHING

Teaching Assistant, Biomedical Artificial Intelligence **January 2024 — May 2024**
Harvard Medical School, Boston, MA

- Designed a total of 14 weekly quizzes based on important papers in the field of biomedical AI over the course of the entire semester.
- Held weekly office hours and graded assignments weekly for 35+ students.

Teaching Assistant, Statistical Inference in Biology **October 2021 — January 2022**
National Centre for Biological Sciences, Bengaluru, India

- Held weekly office hours in addition to designing and grading homework for 20+ students.

PUBLICATIONS

Peer-reviewed

1) Ross, C., Ghauri, S., Gilbert, J.B., Hu, D., **Ullanat, V.,** Gong, D., Greenberg, P.B., Eliott, D., Elze, T., Lorch, A. and Miller, J.W., 2024. Intravitreal Antibiotics versus Early Vitrectomy plus Intravitreal Antibiotics for Post-Injection Endophthalmitis: an IRIS® Registry (Intelligent Research in Sight). *Ophthalmology Retina*.
2) Omar, M., **Ullanat, V**., Loda, M., Marchionni, L. and Umeton, R., 2024. ChatGPT for digital pathology research. *The Lancet Digital Health*, *6*(8), pp.e595-e600.
3) Hari, K., Harlapur, P., Gopalan, A., **Ullanat, V.,** Duddu, A.S. and Jolly, M.K., 2022. Emergent properties of coupled bistable switches. *Journal of Biosciences*, *47*(4), p.81.
4) Hari, K., **Ullanat, V**., Balasubramanian, A., Gopalan, A. and Jolly, M.K., 2022. Landscape of epithelial–mesenchymal plasticity as an emergent property of coordinated teams in regulatory networks. *Elife*, *11*, p.e76535.
5) **Ullanat, V**., Kasukurthi, N. and Viswanath, S., 2022. PrISM: precision for integrative structural models. *Bioinformatics*, *38*(15), pp.3837-3839.

6) DSouza, G.C., Sheriff, R.S., **Ullanat, V**., Shrikrishna, A., Joshi, A.V., Hiremath, L. and Entoori, K., 2021. Fungal biodegradation of low-density polyethylene using consortium of Aspergillus species under controlled conditions. *Heliyon*, *7*(5).

Conference

1) Ghauri, S., Ross, C., **Ullanat, V**., Hu, D., Gilbert, J., Gong, D., Greenberg, P.B., Eliott, D., Elze, T., Lorch, A.C. and Miller, J.W., 2024. Early Vitrectomy for Post-Injection Endophthalmitis: an IRIS® Registry Analysis. *Investigative Ophthalmology & Visual Science*, *65*(7), pp.3826-3826.
2) **Ullanat, V.,** Balamurali, V. and Rao, A., 2021, March. A Novel Residual 3-D Convolutional Network for Alzheimer's disease diagnosis based on raw MRI scans. In *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)* (pp. 82-87). IEEE.
3) **Ullanat, V.,** 2020, November. Variational autoencoder as a generative tool to produce de-novo lead compounds for biological targets. In *2020 14th international conference on innovations in information Technology (IIT)* (pp. 102-107). IEEE.

In preparation

1) **Ullanat V.**, Jing B., Sledzieski S., Berger B., 2024. Learning the language of protein-protein interactions with ESM-Multimer. (Under review at *Nature Machine Intelligence*).
2) Huang Y., Su X., **Ullanat V.**, Liang I., Clegg L., Olabode D., Ho N., John B., Gibbs M., and Zitnik M., 2024. A unified multimodal model for predicting drug combination effects. (Under review at *Nature Medicine*).
3) Majila K., **Ullanat V.**, Viswanath S. A deep learning method for predicting interactions for intrinsically disordered regions of proteins. (Under review at *Nature Communications*).
4) Omar, M., Fanelli, G.N., Socciarelli, F., **Ullanat, V.,** Puchala, S.R., Wen, J., Chowdhury, A., Valencia, I., Scatena, C., Marchionni, L. and Umeton, R., 2024. Multiplex Imaging Analysis in Pathology: a Comprehensive Review on Analytical Approaches and Digital Toolkits. *arXiv preprint arXiv:2411.00948*. (Under review at *Modern Pathology*).

SERVICE & OUTREACH

Reviewer **September 2024**
Machine Learning in Structural Biology @ NeurIPS 2024

Reviewer **June 2024**
Accessible and Efficient Foundation Models for Biological Discovery @ ICLR 2024

Reviewer **February 2023**
Conference on Intelligent Systems for Molecular Biology (ISMB)

POSTERS & PRESENTATIONS

Machine Learning in Structural Biology (NeurIPS 2024), Poster Session **December 2024**
Learning the language of protein-protein interactions with ESM-Multimer Vancouver, Canada

DBMI Science Day, Harvard Medical School, Poster Session **September 2023**
Translational drug pharmacology learning to predict safety and synergy in realistic settings Boston, USA
Learning protein graphs for predicting protein functions

Annual Talks, National Centre for Biological Sciences, Poster Session **January 2022**
Annotating precision for models of large protein assemblies Bengaluru, India

Festival of Genomics and Biodata, Young Scientist Session, Poster Session **January 2021**
Transcriptomic analysis of the brains of symptomatic and asymptomatic Alzheimer's patients London, UK (Virtual)

Quantitative Systems Pharmacology Summit, Poster Session **January 2022**
An Artificial Intelligence approach to drug discovery Virtual

Neuromatch 3.0, Interactive Session **November 2020**
Transcriptomic analysis of the brains of symptomatic and asymptomatic Alzheimer's patients Virtual

Individualizing Medicine Virtual Conference by Mayo Clinic, Poster Session **November 2020**
Transcriptomic Analysis of the Entorhinal Regions of Symptomatic and Asymptomatic Alzheimer's Patients Virtual

International Conference on Drug Discovery, Development and Lead Optimization, *Poster Session* **November 2020**
Variational autoencoders as a generative tool to produce de-novo lead compounds for biological targets Virtual

VOLUNTEERING

Mentor, Alexander Twilight Academy **September 2024 — October 2024**
Duties: Conduct mock interviews for students Boston, USA

Ed-Support Mentor, Make A Difference, a Non-profit organisation **October 2021 — June 2022**
Duties: Managing a team of tutors, scheduling, volunteer sourcing Bengaluru, India

Transition Readiness Volunteer, Make A Difference, a Non-profit organisation **October 2020 — May 2022**
Duties: Teaching senior year Physics and Chemistry Bengaluru, India

Technical Team Lead, National Internship Platform, RV College of Engineering **December 2019 — January 2020**
Duties: Website development, event management Bengaluru, India

Senior Associate, Entrepreneurship Cell, RV College of Engineering **August 2018 — August 2019**
Duties: Moderator for multiple large-scale events, volunteer sourcing, event management Bengaluru, India