Ananthan Nambiar

nambiar4@illinois.edu

Academic Background

Institution	Discipline	Awards	Graduation
University of Illinois at Urbana-Champaign	Ph.D. Bioengineering (CGPA: 4.0) Focusing on computational and systems biology.	Mavis Future Faculty Fellow 2022/23	2024
Reed College, Portland	B.A. Computer Science (CGPA: 3.7) With coursework in deep learning, statistics and numerical analysis.	Commendation for Excellence in Scholarship (2016/17, 2017/18, 2018/19)	2019

Select Research Experience

Institute for Genomic Biology, Graduate Research Assistant, August 2019 - present
Machine Learning for Biological Sequences. I lead a two-pronged project studying NLP inspired deep learning models for both DNA and protein sequences.
- To build an interpretable neural network that is **the first sequence-based model to predict condition specific gene expression**.
- To predict protein binding with cost efficient transformer neural networks (**170 times more compute efficient than past models**).

Inari Agriculture, Machine Learning Research Intern, Summer 2021, Summer 2022
Machine Learning for Variant Effect Prediction. I implemented zero-shot deep learning models to predict targets for gene editing in crops. These predictions were estimated to be **20% more accurate than traditional methods**.
Machine Learning for Protein Generation. Investigated the use of graph neural networks (GNN) for protein generation.

Reed College, Research Assistant, May 2018 – August 2019, Summer 2020
NLP and Computational Social Science. I combined ideas from natural language processing, economics, time-series analysis and theoretical ecology to develop data driven approaches that
- **Measure and predict the evolution of technology** based on patent records.
- **Quantify the diversity of the US economy and predict the sales of large US corporations** by mining information from publicly available financial documents.

Delhi University, Research Assistant, Summer 2019
Network Biology. I studied the controllability of genetic regulatory networks of bacteria, focusing on E. coli. I focused on the target controllability of functional classes of genes to understand the relationship between the controllability of a class and its function.

Papers

A. Nambiar, V. Dubinkina, S. Liu* and S. Maslov. FUN-PROSE: A Deep Learning Approach to Predict Condition Specific Gene Expression in Fungi. (In peer-review.)

A. Nambiar, J. M. Forsyth*, S. Liu* and S. Maslov. DR-BERT: A Protein Language Model to Annotate Disordered Regions. (In peer-review.)

S. Bhogale, V. Dubinkina, P-H. Hsieh, P. Dibaeinia, **A. Nambiar**, S. Maslov, Y. Yoshikuni, S. Sinha. A Transcriptomic Atlas of Low pH Stress Response in Multiple I.orientalis Strains. (In peer-review.)

A. Nambiar, C. Pan, V. Rana, Mahdi Cheraghchi, João Ribeiro, Sergei Maslov and Olgica Milenkovic. Semi-Quantitative Group Testing for Efficient and Accurate qPCR Screening of Pathogens with a Wide Range of Loads. (In peer-review)

N. Packard, N. Gigliotti, **A. Nambiar**, T. Janssen and M. Bedau. An Evolving Classification for Forecasting Technology. (Paper in preparation.)

A. Nambiar, T. Janssen, J. McCaull, M. Bedau. Dropping diversity of products of large US firms: Models and measures. *PLOS ONE* 17(3): e0264330. 2022.

A. Nambiar, M. Heflin*, S. Liu*, S. Maslov, M. Hopkins and A. Ritz. Transforming the Language of Life: Transformer Neural Networks for Protein Prediction Tasks. In Proceedings of *ACM-BCB'20: 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics,* September 2020.

M. Bedau, N. Gigliotti, T. Janssen, A. Kosik, **A. Nambiar** and N. Packard. Open-ended Technological Innovation. *Artificial Life,* 25 (1). 2019.

Underline: First authors, *: Mentored undergraduate student

Select Conference Presentations

A. Nambiar, V. Dubinkina, S. Liu and S. Maslov. Transcribing the Language of Life: An Interpretable Deep Neural Network to Predict Condition Specific Gene Regulation. Poster presentation at *ISMB/ECCB'21: 29th Conference on Intelligent Systems for Molecular Biology and the 20th European Conference on Computational Biology*, July 2021.

A. Nambiar, M. Hopkins and A. Ritz. Computing the Language of Life: NLP Approaches to Feature Extraction for Protein Family Classification. Poster presentation at *ISMB/ECCB'19: 27th Conference on Intelligent Systems for Molecular Biology and the 18th European Conference on Computational Biology*, July 2019.

M. Bedau, N. Gigliotti, T. Janssen, A. Kosik, **A. Nambiar** and N. Packard. Detecting the On-going Emergence of Technological Innovations. Oral presentation at *ALIFE'18: The 2018 Conference of Artificial Life – Workshop on Open-ended Evolution*, July 2018.

A. Nambiar. Controllability of Functional Classes in the Genetic Regulatory Network of *E. coli*. Poster presentation at *SIAMAN'18: 2018 SIAM Annual Meeting*, July 2018.

Students Supervised

Nisha Janmashtami (Bioe: UIUC '25)	Neural networks for microbial genomics	2023-present
Malcolm Forsyth (CS: UIUC '24)	Predicting systems level properties of proteins	2021-present
Simon Liu (CS: UIUC '22)	Transformer networks for embedding amino acid sequences	2019-2022
Maeve Heflin (CS: UIUC '22)	Unsupervised learning to on protein vector embeddings	2019/20

Chaired Conferences

The Special Session on **Representation Learning in Biology** at ISMB/ECCB '21. Co-chaired with Christian Dallago (TUM), Peter Koo (CSHL) and Ali Madani (Salesforce). **Sponsored by Microsoft Research, Inari Agriculture and Dyno Therapeutics.**

External Talks

Institute for New Economic Thinking, University of Oxford	Presented on **Using NLP to Obtain Vector Representations of Financial Entities**. Discussed various NLP techniques used for document embeddings and how they can be adapted to obtain representations of financial entities.
Google Developer Students Club, University of Malaya	Presented on **Linking the Language of Life**. I gave a talk on developing deep learning models for protein characterization tasks including protein family classification and protein-protein interaction prediction.

Pacific Northwest Quantitative Biology Symposium	Presented on **The Controllability of the Genetic Regulatory Network of *E. coli*.** Discussed controllability of networks, structural controllability and how the low controllability of the genetic regulatory network of E. coli can provide biological insight.	
St. Stephen's College Feynman Club Colloquium	Presented on **Analysing the Presence of Terrorist Organisations on Social Media**. Discussed sentiment analysis, scale-free networks and robustness of networks in this context.	

Grants Offered

Grant	Detail(s)	Year
Google Research Credits	To study the use of transfer learning for protein feature embedding	2020
Reed Opportunity Grant	To study the use of machine learning to predict protein functions.	2019
Reed Opportunity Fellowship & Summer Experience Award	To study the effect of semi-synchronous updating on homeostasis in Boolean networks. Chose not to accept.	2018
Reed Student Opportunity Subsidy	To study genetic regulation in bacterial cells.	2017

Outreach

Program	Detail(s)	Year
Reed College Science Outreach	As a **Lead Instructor**, I led several teams of Reed students whose duties include in-class teaching at Portland public schools, guiding the students through science projects and teaching them safety in the lab.	2015 - 2017
SMKWM After School Science	I taught **Biology, Chemistry, Physics, Additional Mathematics and ICT** to students from underprivileged backgrounds and were unable to afford tutoring	2013 - 2014

Other Activities

Vice-President of the UIUC Graduate Biomedical Engineering Society (BMES).

Courses Taught/TA-ed/Tutored

BIOE 310 (UIUC): Computational Tools for Biological Data	2021-present
PHYS 467 (Reed): Computational Methods for Physics	2019
MATH 201 (Reed): Linear Algebra	2017-2019

Sample Course Comments

"Ananthan was very responsive to student questions. He cared about teaching."

"Ananthan was always so willing to help and created slides that were logical. He was happy to help outside of class and was very approachable."

"[Ananthan] was very helpful. He understood what concepts confused students and always explained things thoroughly. He helped guide students in their studying and did extra office hours."

"I think you're a great instructor and I would suggest you give more lectures in biostats."

Skills

Programming Languages	Proficient: Python, Wolfram Familiar with: Standard ML, C, Go, R, x86 assembly
Select Frameworks	PyTorch, Gephi, NetworkX, PyOpenGL, SQLite, Scikit-Learn, Gensim, Google Cloud Platform, Amazon Web Service, Docker