

Ananthan Nambiar

nambiar4@illinois.edu

Academic Background

| <u>Institution</u> | <u>Discipline</u> | <u>Recognitions</u> | <u>Graduation</u> |
|--|---|---|-------------------|
| University of Illinois at Urbana-Champaign | Ph.D. Bioengineering (CGPA: 4.0) Focusing on computational and systems biology. | | 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

| <u>Research Area</u> | <u>Period</u> |
|--|--|
| Machine Learning for Gene Expression Prediction Supervised by Sergei Maslov (UIUC). Studying interpretable deep learning models that can be used to understand gene regulation in addition to predicting condition specific gene expression | August 2019 - present |
| Machine Learning for Proteomics Supervised by Mark Hopkins (Reed), Anna Ritz (Reed) and Sergei Maslov (UIUC). Using deep learning to predict protein families and protein-protein interactions. Focus is on treating protein sequences as a language and adapting ideas from natural language processing. | August 2018 - present |
| Machine Learning for Variant Effect Prediction Supervised by Karl Kremling (Inari Agriculture). Utilized zero-shot deep learning models to predict targets for gene editing in crops. | Summer 2021 |
| NLP and Computational Social Science Supervised by Norman Packard (Daptics) and Mark Bedau (Reed). Combined ideas from natural language processing, economics, and theoretical ecology to develop data driven approaches that <ul style="list-style-type: none">• 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. | May 2018 - August 2019, Summer 2020 |
| Network Biology Supervised by Sanjay Jain (Delhi U. & Santa Fe Inst.). 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. | Summer 2017 |

Select Papers and Conferences

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. Under peer-review. (Preprint on arxiv.org/abs/2110.08367)

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. (Journal paper in preparation.)

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.

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. Open-ended Technological Innovation. *Artificial Life*, 25 (1). 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

| | | |
|----------------------------|---|--------------|
| Malcolm Forsyth (UIUC '24) | Predicting systems level properties from amino acid sequences | 2021-present |
| Simon Liu (UIUC '22) | Transformer networks for embedding amino acid sequences | 2019-present |
| Maeve Heflin (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 <i>E. coli</i> . Discussed controllability of networks, structural controllability and how the low controllability of the genetic regulatory network of <i>E. coli</i> 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

| <u>Grant</u> | <u>Detail(s)</u> | <u>Year</u> |
|-------------------------|---|-------------|
| 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 |

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 |