Anupama Jha anupamaj@seas.upenn.edu

Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA 19104

Education

Doctor of Philosophy in Computer and Information Science
University of Pennsylvania, Philadelphia, PA, USA *August 2014 to present, GPA: 3.93/4.0*

Master of Science in Informatics
Technische Universität München, Munich, Bavaria, Germany *2012-2014, Grade: 1.5, Passed with Distinction*

Bachelor of Technology in Information Technology
Guru Gobind Singh Indraprastha University, New Delhi, India *2006–2010, CPI: 85.25%, First Class with Distinction*

Research and Professional Experience

Graduate Research Assistant at Biociphers Lab, University of Pennsylvania,
Advisor: Dr. Yoseph Barash *August 2014-Present*

My research projects have involved predicting splicing differences between tissues, regulatory networks between RNA-binding proteins with deep learning and developing EIG, an interpretation method for splicing code and other deep learning models based on genomic data. In parallel, I have contributed to research towards reliable identification and quantification of splicing events from RNA-Seq data and understanding the role of RNA-binding proteins in post-transcriptional regulation.

Master Thesis Student at Model-Based Systems and Qualitative Reasoning Group, TU Munich
Advisor: Dr. Peter Struss *November 2013–August 2014*

Conceptualization and implementation of a generic tool for selection of learning goals for a knowledge-based machine learning system, application of the tool for the fitness training domain.

Graduate Research Assistant for Human Brain Project, Neurorobotics
with Dr. Florian Röhrbein *November 2013–March 2014*

Surveyed on the availability of robot simulation tools and game engines for the neurorobotics platform of the Human Brain Project.

Graduate Research Assistant at fortiss GmbH, Munich,
Supervisor: Christian Vögele *October 2013–April 2014*

Project: PARO-Performance analysis for role-based applications. Created a tool for the performance modeling of an application and analysis of potential low performance areas during the design phase.

Graduate Research Assistant at Rostlab, TU Munich,
Supervisor: Dr. Tobias Hamp *March 2013–September 2013*

Project: Prediction of Interaction sites in Proteins using only sequences. Developed neural network models for prediction of interaction hot spots in proteins using amino acid sequences and evolutionary information.

Systems Engineer at Infosys Ltd., India. *June 2010–September 2012*

Trainings Undertaken at Infosys: Android, ASP.NET, ADO.NET, C#, C, UNIX, Oracle, SQL
Certifications at Infosys: STAR, Geographical Information Systems, PERL
Responsibilities included development of UI and back end for GPS enabled embedded navigation device with Android for TomTom Int. BV. (Eindhoven, Netherlands)

Teaching Experience

GCB 537: Advanced Computational Biology
TA with Dr. Yoseph Barash, UPenn *Spring 2016, Spring 2017, Spring 2018*

Teaching Assistant for the Ph.D. level course with three components: statistical data analysis and machine learning techniques for computational biology, discussion on current topics in genomics and computational biology, and hands on experience in data analysis, coding and evaluation of computational biology tools/algorithms.

Deep Learning Reading Group
Co-organizer with Dr. Yoseph Barash, UPenn *Summer 2016, Spring 2018*

Co-organizer of reading groups to cover deep learning book and interpretation methods for deep learning models.

Awards and Honors

Selected student abstract for oral presentation, SAGES 2019, June 2019, Philadelphia, Pennsylvania, USA.
GHC 18 Scholarship, Grace Hopper Celebration 2018, September 2018, Houston, Texas, USA.
Travel Fellowship, ISMB/ECCB 2017, July 2017, Prague, Czech Republic.
Best Poster Award, Symposium on Advances in Genomics, Epidemiology, and Statistics, June 2017, University of Pennsylvania.
Best Poster Award, RNA Biology & Cancer 2017 Symposium, May 2017, University of Pennsylvania.

Publications

Anupama Jha, Matthew R. Gazzara, Caleb Radens, Paul Jewell, and Yoseph Barash. RBP-Pokedex: Prediction of RBP knockdown effect via DNN experiment modeling. *In Preparation* (2020).

Xinjun Ji, **Anupama Jha**, Jesse Humenik, Louis R. Ghanem, Andrew Kromer , Christopher Duncan-Lewis, Elizabeth Traxler, Mitchell J. Weiss, Yoseph Barash, Stephen A, Liebhaber RNA binding proteins PCBP1 and PCBP2 are critical determinants of murine erythropoiesis. *In review, Blood* (2020).

Barry Slaff, Caleb Matthew Radens, Paul Jewell, **Anupama Jha**, Nicholas Lahens, Gregory R. Grant, Andrei Thomas-Tikhonenko, Kristen W. Lynch, Yoseph Barash MOCCASIN: A method for correcting for known and unknown confounders in RNA splicing analysis. *In Review, Nature Communications* (2020).

William P. Bone, Katherine M. Siewert, **Anupama Jha**, Derek Klarin, Scott M. Damrauer, the VA Million Veteran Project, Kyong-Mi Chang, Philip S. Tsao, Themistocles L. Assimes, Marylyn D. Ritchie, Benjamin F. Voight Multi-trait association studies discover pleiotropic loci between Alzheimers disease and cardiometabolic traits. *Submitted* (2020).

Anupama Jha, Joseph K Aicher, Matthew R Gazzara, Deependra Singh, and Yoseph Barash. Enhanced Integrated Gradients: improving interpretability of deep learning models using splicing codes as a case study. *Genome Biology* (2020).

Anupama Jha, Matthew R. Gazzara, and Yoseph Barash. Integrative Deep Models for Alternative Splicing. *Bioinformatics* (2017).

Matthew R. Gazzara, Michael J. Mallory, Renat Roytenberg, John Lindberg, **Anupama Jha**, Kristen W. Lynch, and Yoseph Barash. Ancient antagonism between CELF and RBFOX families tunes mRNA splicing outcomes. *Genome Research* (2017).

Pinaki Chakraborty, Shweta Taneja, Nandita Anand, **Anupama Jha**, Diksha Malik, and Ankit Nayar. An Optimizing Compiler for Turing Machine Description Language, *Journal of Computer Sciences, ICFAI University Press, Hyderabad, India* (2011).

Talks

RBP-Pokedex: Prediction of RBP knockdown effect via DNN experiment modeling *July 2020*
Talk at the iRNA COSI of ISMB 2020.

Improving interpretation of deep learning models: splicing codes as a case study *June 2019, July 2019*
Talk at the MLCSB COSI of ISMB/ECCB 2019.
Selected student abstract for oral talk at SAGES 2019

Integrative Deep Models for Alternative Splicing *April 2017, July 2017, October 2017*
Talk at the iRNA COSI of ISMB/ECCB 2017.
Talk at PRiML Group at Upenn
Talk at Penn RNA Group

GCB 537 Guest Lecture: Support Vector Machine *April 2016-2019*
Guest lecture on Support Vector Machine for Ph.D. level course at UPenn.

Panel: Is Graduate School for me? *March 2019*
Panel at CAPWIC 2019 to discuss entry requirements and graduate school experience.

Guest Lecture: AI and Computational Biology *March 2018-2019*
Guest lecture with Dr. Yoseph Barash at the WICS High School Day for Girls at UPenn.

Posters

Anupama Jha, Joseph K. Aicher, Mathew R. Gazzara, Deependra Singh and Yoseph Barash. Improving interpretability of deep learning models: splicing codes as a case study. *ISMB 2019*.

Anupama Jha, Matthew R. Gazzara and Yoseph Barash. Assembling the Building Blocks for a Unifed Splicing Code. *ISMB 2018*.

Anupama Jha, Matthew R. Gazzara, and Yoseph Barash. Deep Learning Models for Alternative Splicing. Symposium on Advances in Genomics, Epidemiology, and Statistics, June 2017, University of Pennsylvania. Best Poster Award

Anupama Jha, Matthew R. Gazzara, and Yoseph Barash. Integrative Deep Models for Alternative Splicing. RNA Biology & Cancer 2017 Symposium, May 2017, University of Pennsylvania. Best Poster Award

Anupama Jha, Matthew Gazzara, Jorge Vaquero-Garcia and Yoseph Barash. Predictive modeling framework for splice factor knockdown experiments. *ISMB/ECCB 2016*.

Skills

Languages, systems, and tools
Proficient: Python, Tensorflow, High Performance Computing, UNIX
Competent: R, LaTeX
Past work experience: Perl, MATLAB, Java, Android, C#, .NET Framework, C++, Oracle, SQL

Relevant graduate courses taken
Machine Learning, Bayesian Statistics, Mathematical Statistics, Deep Learning, Advanced Computational Biology, RNA World, High-throughput Datasets for Biologists, Interpretation of Deep Learning Models, Adversarial and Secure Machine Learning, Computational Linguistics.

Service and outreach

Reviewer
Co-reviewer with Prof. Yoseph Barash for NeurIPS 2015-2020, ISMB/ECCB 2016-2019, ICLR 2018-2020, ICML 2019-2020, PLOS Computational Biology, Bioinformatics, Nucleic Acids Research. Independent reviewer for TEAMC-2018, Nature Scientific Reports.

Project Mentor
Mentored undergraduate students at University of Pennsylvania in their final year project in Fall 2017 and Spring 2018.

References are available on request.