PRATIK DUTTA

Postdoctoral Research Associate

House No:-194, Belmuri Station Road, Beside Paschimbanga Gramin Bank, Belmuri, Hooghly, West Bengal, 712302
(+91) 9431432366, (+91) 9007216781
pratik.dutta@stonybrook.edu
My Webpage
Github *Linkedin* *duttaprat16*

Research Experience

Department of Biomedical Informatics, State University of New York, Stony Brook.

April 2021 –Present : **Postdoctoral Research Associate**.
Cancer Genomics

Advisor : **Prof. Ramana V. Davuluri,** Department of Biomedical Informatics, Stony Brook Cancer Center, Stony Brook University.

Strand Life Sciences, Bangalore, India.

Aug 2020 – Mar 2021: **Research Intern**.
Developing a BERT-based model for Phenotype-driven Rare Disease Diagnosis from the PubMed articles.

Advisors : **Dr. Vamsi Veeramachaneni,** *CSO*, Strand Life Sciences, India and **Dr. Rajesh Sundaresan**, *Professor*, Department of Electrical Communication Engineering, IISc, Bangalore.

Education

Jan 2016– Dec 2020 : **PhD(Visvesvaraya Research Fellow), Computer Science & Engineering**, *Indian Institute of Technology*, Patna.
Thesis Title: *Computational Approaches Leveraging Protein Interaction Information for Biomedical Tasks.*

Advisor : **Dr. Sriparna Saha**, *Associate Professor*, Computer Science & Engineering, IIT Patna.

2013–2015 : **Master of Engineering, Information Technology**, *Indian Institute of Engineering Science & Technology*, Shibpur(*Formerly* Bengal Engineering and Science University, Shibpur).

Advisor : **Dr. Hafizur Rahaman**, *Professor*, Department of Information Technology and School of VLSI, IIEST Shibpur.

2009–2013 : **Bachelor of Engineering, Computer Science & Technology**, *Indian Institute of Engineering Science & Technology*, Shibpur(*Formerly* Bengal Engineering and Science University, Shibpur).

Research Experience

Broad Institute of MIT and Harvard

August 2020 – Jan 2021 ***Label-free Prediction of 3D Fluorescent Images***.
Developing a deep Learning model to do label-free prediction of Cell Painting stains. More specifically, given the "brightfield" channel that is acquired together with the 5 Cell Painting channels, the task is to predict each of the 5 channels.

Advisors: **Dr. Shantanu Singh**(*Senior group leader, Imaging Platform*) and **Dr. Anne Carpenter** (*Institute Scientist, Senior Director of the Imaging Platform*), Broad Institute of MIT and Harvard

The University of North Carolina at Chapel Hill

March, 2020 – July 2020 ***Developing a AI Platform for Understanding Various Emerging Viral Diseases and their Virulence in Humans***.
Developing an efficient learning model using machine learning and deep learning techniques, which can be applied in a real life scenario that can extrapolate information from the genomic data of the individual. Different underlying diseases and conditions which aggravate the disease process will be studied.

Collaborators: **Dr. Sanjay Sarkar***(Research Associate*, Virologist, Department of Genetics, School of Medicine, UNC- Chapel Hill)

SystemOnSilicon Corporation(USA)

July,2019 – present ***Developing a Cloud-based Interactive AI Platform for Digital Health Analytics***.
Designing & developing a disruptive cloud-based interactive AI platform for solving various social impactful problems. I am currently leading Artificial Intelligence and Data Science/Analytics R&D for creating novel ML/DL powered solutions for precision and smart agriculture, personalized food recommendation, developing of a conversational AI platform for wellness guidance .

Collaborators: **Raj Ray***(Founder and CEO*, SystemOnSilicon Corporation, USA) and **Dr. Dave Ray***(Co-founder and CSO*, SystemOnSilicon Corporation; Assistance Professor, Alcorn State University)

Indian Institute of Technology, Patna

June,2019 – present ***Identifying Protein-protein Interaction from Biomedical text***.
Developing a deep multi-modal architecture for accurately predicting protein interaction information from biomedical text. Here, along with the textual modality, we utilize structural and underlying sequence information of the interacting proteins for predicting protein interaction in any biomedical corpus.

Advisor : **Dr. Sriparna Saha**, *Associate Professor, Department of Computer Science & Engineering*, IIT Patna (*Personal Web-page*)

July,2018 – June 2020 ***Developing Deep Multi-modal Architecture for Biomedical Problems***.
Analyzing different modalities of genes like gene expression profiles, protein 3D structure, underlying amino acid sequence using popular deep learning models to obtain deeper insight into the underlying biological system. Utilizing this insightful information, we developed few efficient deep models for different interesting problems of computational biology like protein function prediction, improving gene clustering, disease gene prognosis, etc.

Advisor : **Dr. Sriparna Saha**, *Associate Professor, Department of Computer Science & Engineering*, IIT Patna (Personal Web-page)

January,2016 – June, 2018 ***On Protein-protein Interaction identification : Application to Gene Clustering***.
Developed different machine learning-based architecture for improving gene clustering by prudently utilizing protein interaction information with different machine learning methods. In this regard, we have utilized a genetic algorithm(GA) based multi-objective optimization(MOO) technique, graph mining algorithm, ensemble method and generative models.

Advisor : **Dr. Sriparna Saha**, *Associate Professor, Department of Computer Science & Engineering*, IIT Patna (*Personal Web-page*)

Indian Institute of Engineering Science Technology, Shibpur

January,2015 – Dec,2015 ***Design and Synthesis of Reversible Multi-dimentional Nearest-Neighbour(NN) Quantum Circuit***.
Proposed an approach for designing and physically implementing of the multi-dimensional quantum circuits maintaining nearest-neighbor complacency that use minimal number of SWAP gates.

Advisor : **Dr. Hafizur Rahaman**, *Professor, Department of Information Technology; Director, School of VLSI Technology*, IIEST Shibpur(Personal Web-page)

October,2013 – July,2015 ***Implementation and Synthesis of Reversible Logic using Mach-Zehnder Interferometer swithches***, (Available Here).
In this research project, CLA adder and various sequential memory elements (Flip-Flops, Counters, shift register) in all-optical domain using MZI switches are efficiently designed. All the designs are implemented using minimum number of MZI switches and all design can be generalized for n-bit also.

Advisor : **Dr. Hafizur Rahaman**, *Professor, Department of Information Technology; Director, School of VLSI Technology*, IIEST Shibpur(*Personal Web-page*)

2012 – 2013 ***Text Document Clustering with Semantic Similarity through Wordnet***.
Improvement of the text document clustering task over conventional methods by introducing WORDNET and some better clustering algorithms.

Advisor : **Dr. Asit Kumar Das**, *Professor, Department of Computer Science and Technology*, IIEST Shibpur

Publications

Journal Articles

2020 **Pratik Dutta**, Sriparna Saha, Sanket Pai, and Aviral Kumar. A protein interaction information-based generative model for enhancing gene clustering. ***Scientific Reports***, volume 10, pages 1–12. Nature Publishing Group, 2020, **(Impact Factor: 4.122)**, **(h5-index: 178)**.

2020 **Pratik Dutta**, Sriparna Saha, and Sukanya Naskar. A multi-objective based pso approach for inferring pathway activity utilizing protein interactions. *Multimedia Tools and Applications*, pages 1–21. Springer, 2020, **(Impact Factor: 2.60)**.

2020 **Pratik Dutta**, Piyush Mishra, and Sriparna Saha. Incomplete multi-view gene clustering with data regeneration using shape boltzmann machine. *Computers in Biology and Medicine*, page 103965. Elsevier, 2020, **(Impact Factor: 3.434)**.

2020 Swagarika Jaharlal Giri, **Pratik Dutta**, Parth Halani, and Sriparna Saha. Multipredgo: Deep multi-modal protein function prediction by amalgamating protein structure, sequence, and interaction. ***IEEE Journal of Biomedical and Health Informatics(IEEE Transactions on Information Technology in Biomedicine)***. IEEE, 2020, **(Impact Factor: 5.180)**.

2019 **Pratik Dutta**, Sriparna Saha, and Saurabh Gulati. Graph-based hub gene selection technique using protein interaction information: Application to sample classification. ***IEEE Journal of Biomedical and Health Informatics(IEEE Transactions on Information Technology in Biomedicine)***, volume 23, pages 2670–2676. IEEE, 2019, **(Impact Factor: 4.217)**.

2019 **Pratik Dutta**, Sriparna Saha, Saraansh Chopra, and Varnika Miglani. Ensembling of gene clusters utilizing deep learning and protein-protein interaction information. ***IEEE/ACM transactions on computational biology and bioinformatics***. IEEE, 2019, **(Impact Factor: 2.896)**.

2018 Chandan Bandyopadhyay, **Pratik Dutta**, Rakesh Das, and Hafizur Rahaman. Improved designs for all-optical adder circuit using mach–zehnder interferometers (mzi) based optical components. *Journal of The Institution of Engineers (India)*, pages 1–15. Springer, 2018.

2017 **Pratik Dutta** and Sriparna Saha. Fusion of expression values and protein interaction information using multi-objective optimization for improving gene clustering. ***Computers in Biology and Medicine***, volume 89, pages 31–43. Elsevier, 2017, **(Impact Factor: 3.434)**.

Communicated Journal Article

2020 **Pratik Dutta, Aditya Prakash Patra, and Sriparna Saha**, DeePROG: An Attention based Deep Multi-modal Architecture for Disease Gene Prognosis, In *IEEE/ACM Transactions on Computational Biology and Bioinformatics*(Under major revision).

2020 **Pratik Dutta, Nupur Shah, and Sriparna Saha**, A Multi-Objective Optimization-based Clustering Approach for CORD-19 Scholarly Articles, In *Elsevier Journal of Biomedical Informatics*.

In Conference Proceedings

2020 **Pratik Dutta** and Sriparna Saha. Amalgamation of protein sequence, structure and textual information for improving protein-protein interaction identification. In ***Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics(ACL)***, pages 6396–6407. Association for Computational Linguistics, 2020, **(h5-index: 135)**, **(*Core ranking: A**)**.

2019 **Pratik Dutta** and Sriparna Saha. A weak supervision technique with a generative model for improved gene clustering. In ***IEEE Congress on Evolutionary Computation (CEC)***, pages 2521–2528. IEEE, 2019, **(h5-index: 70)**, **(*Core ranking: B*)**.

2018 **Pratik Dutta**, Sriparna Saha, and Agni Besh Chauhan. Predicting degree of relevance of pathway markers from gene expression data: A pso based approach. In ***International Conference on Neural Information Processing***, pages 3–14. Springer, 2018, **(*Core ranking: A*)**.

2016 Ripla Roy Chowdhury, Chandan Bandyopadhyay, **Pratik Dutta**, and Hafizur Rahaman. A boolean expression based template matching technique for optical circuit generation. In ***Proceedings of the International Conference on Advances in Information Communication Technology & Computing***, page 36. ACM, 2016.

2015 **Pratik Dutta**, Chandan Bandyopadhyay, and Hafizur Rahaman. All optical implementation of mach-zehnder interferometer based reversible sequential counters. In *VLSI Design (VLSID)*, pages 232–237. IEEE, 2015.

2014 **Pratik Dutta**, Chandan Bandyopadhyay, and Hafizur Rahaman. All optical implementation of mach-zehnder interferometer based reversible sequential circuit. In ***18th International Symposium on VLSI Design and Test***, pages 1–2. IEEE, 2014.

2014 **Pratik Dutta**, Chandan Bandyopadhyay, Chandan Giri, and Hafizur Rahaman. Mach-zehnder interferometer based all optical reversible carry-lookahead adder. In ***IEEE Computer Society Annual Symposium on VLSI (ISVLSI)***, pages 412–417. IEEE, 2014.

Deployed Web-server

2020 **Deep Multi-modal Architecture for Protein Function Prediction using Protein Structure & Sequence**.
Website: `http://multipred.co.in/`

Tutorials and Academic Recognitions

2020 Presenter of the tutorial **Multi-modality for Biomedical Problems: Theory and Applications** in ***IEEE World Congress on Computational Intelligence*** **(WCCI 2020)**, Glasgow, UK.

2018 **Session Chair of the session "Prediction"** in 25^{th} ***International Conference of Neural Information Processing*** **(ICONIP 2018)**, Siem Reap, Cambodia.

2018 Invited to conduct lab sessions in ***"Training Program on Machine Learning For Ocean Acoustics and Climate Data Analysis"***, during 22-36 October 2018 at **Defence R&D Organization-Naval Physical & Oceanographic Laboratory (DRDO-NPOL), Kochi, Kerala**.

2018 Invited to conduct lab sessions in ***"Faculty Development Program on Machine Learning and Applications"***, during 26-30 March 2018 at **AVB-Indian Institute of Information Technology and Management, Gwalior**.

2017 Invited as a keynote speaker in ***"Symposium on Research Trends in Machine Learning"***, held on 21^{st} December,2017 at **University Of Petroleum & Energy Studies (UPES), Dehradun**.

Fellowships & Awards

2020 Receipt of the ***ACM India-IARCS travel grant*** to attend ***58th Annual Meeting of the Association for Computational Linguistics(ACL), 2020*** at Seattle, Washington, USA.

Jan, 2016 – Dec 2020 ***Visvesvaraya Fellowship*** of Ministry of Electronics and Information Technology (MeitY), Government of India, as a PhD research scholar in Indian Institute of Technology Patna.

2019 Receipt of ***Visvesvaraya Travel Grant*** to attend a international conference ***IEEE Congress on Evolutionary Computation, 2019*** at Wellington, New Zealand.

2018 Recipient of ***SciGenome Research Foundation (SGRF) GYAN Scholarship*** to participate ***Nextgen Genomics, Biology, Bioinformatics and Technologies-2018*** meeting at Jaipur India from 30^{th} September to 2^{nd} October 2018.

2015 Awarded under ***Students Reward Programme*** at the Annual General Meeting of **Global Alumni Association of Bengal Engineering and Science University(GAABESU).**

2015 Recipient of ***Student Fellowship*** for attending and presenting a paper in ***28th International Conference on VLSI Design***(VLSID-2015), held from 2nd January – 7th January 2015 at The Leela Palace, Bangalore, India.

2014 Recipient of ***Student Fellowship*** for attending and presenting a paper in ***18th International Symposium on VLSI Design and Test***(VDAT-2014), held from 16th July-18th July 2014 at PSG College of Technology, Coimbatore, India.

2013 – 2015 Recipient of ***GATE Scholarship*** from AICTE as a postgraduate student in IIEST Shibpur.

2012 Award of Achievement, ORACLE'S WORKFORCE DEVELOPMENT PROGRAM.

2009 Secured a place among top **0.9%** over 1,10,000 students in highly competitive ***West Bengal Joint Entrance Examination***(WBJEE).

2006 Secured ***58th rank*** in the merit list under the ***National Merit Scholarship Scheme*** of the Government of India in MADHYAMIK PARIKHSHA.

2005 Secured ***3rd position*** in quiz competition organized by ***Indian Physical Society*** collaboration with ***S.N. Bose National Centre for Basic Science***.

Computer skills

Programming Languages: Python, PyTorch, keras, R, C, C++, Advanced JAVA

Web Technologies: HTML 5, PHP, JSP, Javascript

Database: SQL, MySQL, Apache, Neo4j

Position of Responsibility

Regular reviewer of the esteemed journal like *IEEE JBHI* and the top-tier conferences like *ACL 2020, JCDL 2020* and many more.

2016-2020 **Executive member of IEEE Student Branch**, IIT Patna.

April 1-5, 2019 **Organizer, GIAN Workshop on Introduction to Unsupervised Data Mining: From Batch to Stream Mining Algorithms**, IIT Patna.

Oct 14-15, 2017 **Volunteer, #OpenGovDataHack:24Hrs Hackathon**, *Organized by National Informatics Centre(NIC), Government of India and IAMAI*, IIT Patna.

Dec 15-23, 2016 **Organizer, GIAN Workshop on Multi-objective Optimization**, IIT Patna.

May 2-7, 2016 **Organizer, GIAN Workshop on Introduction to Natural Language Processing**, IIT Patna.

Teaching Assistantship

Fall, 2019 : **CS564: Foundations of Machine Learning**, IIT Patna.

Spring, 2019 : **CS342: Operating System Lab**, IIT Patna.

Fall, 2018 : **CS564: Foundations of Machine Learning**, IIT Patna.

Spring, 2018 : **CS342: Operating System Lab**, IIT Patna.

Fall, 2017 : **CS345: Advance Database Lab**, IIT Patna.

Fall, 2016 : **CS345: Advance Database Lab**, IIT Patna.

Spring, 2016 : **CS101: Introduction to Computing**, IIT Patna.

Referees

Dr. Sriparna Saha
Associate Professor, Department of Computer Science & Engineering,
Indian Institute of Technology Patna, India.
✉ sriparna.saha@gmail.com

Prof. Hafizur Rahaman
Professor, Department of Information Technology and School of VLSI,
Indian Institute of Engineering Science & Technology, Shibpur.
✉ rahaman_h@yahoo.co.in

Dr. Sanjay Sarkar
Research Associate, Department of Genetics, School of Medicine
The University of North Carolina, Chapel Hill.
✉ sanjaysarkarvet@gmail.com

Dr. Raj Ray
High Tech R&D Leader and Industry Veteran; Former Senior R&D leader in Synopsis, Intel & Sun Microsystems; *Founder & Chief Innovator of* ***SystemOnSilicon Corporation***, USA.
☎ +(650) 352-3968
✉ raj.ray@systemonsilicon.com

Dr. Dave Ray
Assistant Professor,
Department of Biological Sciences,
Alcorn State University, USA.
☎ +(601) 877-6236
✉ droy@alcorn.edu

Dr. Asif Ekbal
Associate Professor, Department of Computer Science & Engineering,
Indian Institute of Technology Patna, India.
✉ asif.ekbal@gmail.com