



## Dr. Sumanta Ray

PhD (Engg.)

- ▶ Male
- ▶ Indian
- ▶ Married
- ▶ DOB: 09/08/1986

## Contact

- 📍 Genome Data Science  
Universität Bielefeld, Hauptgebäude (main building), Universitätsstraße 25, 33615 Bielefeld
- ☎ +91 9231879956
- ✉ sray@cebitec.uni-bielefeld.de
- ✉ sumantababai86@gmail.com
- ✉ sumanta.ray@aliah.ac.in
- 🌐 [github.com/sumantaray](https://github.com/sumantaray)  
[website/sumantaray](https://website/sumantaray)

## Biography

I am currently working as Associate Professor at Department of Computer Science and Engineering, Ghani Khan Choudhury Institute of Engineering & Technology (A CFTI, under MoE, Govt. of India). Before that I worked as Junior Professor at Universität Bielefeld, Bielefeld, Germany. I also worked as an Assistant Professor at Department of Computer Science and Engineering, Aliah University, Kolkata for 10 years. I worked as a stipend scholar (remotely from India) at Genome Data Science, Universität Bielefeld, Bielefeld, Germany for six months. Before that, I spent one year and two months at Life Science and Health group, Centrum Wiskunde and Informatica, The Netherlands as an ERCIM (the European Research Consortium for Informatics and Mathematics) postdoctoral fellow. I received Ph.D from Jadavpur University while working at Machine Intelligence Unit, Indian Statistical Institute, Kolkata, India, where I spent one year and two months as a Junior Research Fellow. My research interests includes deep learning model in single cell genomics, explainable AI, and big data systems in life sciences and health informatics.

## Work experience

### Associate Professor

November/2023 - Today

Department of Computer Science and Engineering,  
Ghani Khan Choudhury Institute of Engineering &  
Technology,  
(A Centrally Funded Technical Institute(CFTI) under  
Ministry of Education Govt of India)

Here, I work on teaching various undergraduate and postgraduate courses and conduct research in various domains of machine and deep learning applications in genomics.

### Junior Professor

August/2023 - Nov./2023

Genome Data Science  
Universität Bielefeld, Bielefeld, Germany

I worked on developing several computational models (Deep/ML) to solve different problems in bioinformatics

### Assistant Professor

Nov./2013 - July/2023

Department of Computer Science and Engineering  
Aliah University, kolkata, India

Here, I taught different course (in UG and PG) like Data Science, Algorithms, Theory of Computation, AI and Soft Computing etc. I have four PhD research scholars (one completed his pre-submission) working in different fields of Bioinformatics and Data Mining.

### Research Associate (Adjunct)

Feb./2022 - Today

Health Analytics Network, Pittsburgh, PA, USA.

### Stipend Scholar

Nov./2020 - May/2021

Genome Data Science, University of Bielefeld, Germany.

Worked on analysis of single cell RNA-seq data using advanced deep learning technique.

## Education

2010 - 2012

### Master of Technology (Computer Science and Engineering)

University of Kalyani

**CGPA-9.18 out of 10, 1st Rank in the University**

Master's thesis: Incorporating Gene Ontology Information in Protein Complex Detection: A Multi-objective Evolutionary Approach.

2005 - 2009

### Bachelors of Technology (Computer Science and Engineering)

Kalyani Government Engineering College, West Bengal University of Technology

CGPA-8.23 out of 10

2002 - 2004

### Higher Secondary (Science)

Boinchi BL. Mukherjee Free Institution, Hooghly

Percentage of Marks: 81.2

2000 - 2002

### Madhyamik (Secondary)

Kalna Maharaja High School, Kalna, Nadia

Percentage of Marks: 82.3

### Postdoctoral Fellow (ERCIM grant)

July/2019 - August/2020

Life Science and Health Group  
Centrum Wiskunde and Informatica, The Netherlands

Here we have developed explainable capsule network model to learn large single cell expression data and predict cell types of unknown cell samples.

### PhD (in Engg)

March/2013 - Oct./2017

Department of Computer Science and Engineering  
Jadavpur University  
(Thesis Guide: Prof. Sanghamitra Bandyopadhyay, Indian Statistical Institute, Kolkata, Prof. Ujjwal Maulik, Jadavpur University, Kolkata, Prof. Anirban Mukhopadhyay, University of Kalyani, Kalyani, Nadia)  
(Thesis Title: Development of New Computational Methods for Predicting Modules in Large Networks)

I devised a series of new algorithms for analysing and predicting modular structures of biological networks of both intra- and inter-species cases. Particularly, our main aim is to develop new techniques for detecting topological changes of co-expression network during progression of viral disease in human host cell. Besides this we also developed multi-objective algorithm for solving some emerging problem in computational biology: such as predicting protein complexes from protein-protein interaction network and predicting differentially co-expressed modules from microarray gene expression data.

### Junior Research Fellow

Sep/2012 - Nov/2013

Machine Intelligence Unit (Prof. Sanghamitra Bandyopadhyay)  
Indian Statistical institute, Kolkata

During this time I am working in the field of evolutionary multi-objective algorithm for solving some emerging problem in computational biology

### Visiting Scientist

March/2015 - April/2015

Machine Intelligence Unit  
Indian Statistical institute, Kolkata

### Visiting Scientist

June/2017 - July/2017

Machine Intelligence Unit  
Indian Statistical institute, Kolkata

## Awards and Fellowship

- Received ERCIM Alain Bensoussan Fellowship (Postdoctoral) for the year 2019-2020.
- Received DST-Inspire Fellowship, by MHRD, Govt. of India in 2013.
- Received Junior Research Fellowship in Swarnajayanti Project at Indian Statistical Institute, Kolkata in 2012.
- Received CSIR-Senior Research Fellowship from Council of Scientific Research (CSIR), India in 2013.
- University merit certificate for ranking 1st class 1st in Master of Computer Science and Engineering from University of Kalyani (2012)
- National Fellowship during Master of Engineering on the basis of GATE-2011.

## Reviewing Activities

## Area of Interests

**Bioinformatics** 7+ yrs.

**Machine Learning** 7+ yrs.

**Deep Learning/ Model Development** 4+ yrs.

**Singe Cell Genomics** 2+ yrs.

**Big data/ Data science** 2+ yrs.

**Feature Engineering** 2+ yrs.

**Health Analytics** 1+ yrs.

## Skills

- ▶ Python Programming
- ▶ R Language
- ▶ Matlab

- Acting as a reviewer of the following journals:
  - Briefings in Bioinformatics
  - Bioinformatics
  - IEEE/ACM Transactions on Computational Biology and Bioinformatics
  - BMC Bioinformatics
  - Nature Communications Biology
  - PLoS ONE
  - Algorithms for Molecular Biology
  - IEEE Transactions on Nanobioscience
  - Scientific Reports
  - Information Science
  - IEEE Access
  - Acted as sub reviewer at ISMB/ECCB 2021
  - Acted as sub reviewer at RECOMB-2021
- Member of Program Committee of the following National and International Conferences:
  - International Conference on Computational Intelligence: Modeling, Techniques and Applications (CIMTA2013), Kalyani, India, September 2013.
  - International Conference on Information Systems Design and Intelligent Applications (INDIA-2015), Kalyani, India, January 2015.
  - International Conference on Recent Trends in Information Systems (RETIS-2015), July 2015, Kolkata, India.
  - International Conference on Emerging Technology in Modelling and Graphics (IEMGraph 2018)

## Invited Lectures

- "Preparing presentation using LATEX", at University of Kalyani, Kalyani, India 13th January, 2021.
- "Supervised framework for cell-types detection using CapsNet", at Genome Data Science Research Seminar, Universität Bielefeld - Bielefeld University, Germany, 25 May, 2020.
- "CODC: A copula based model to identify differential coexpression", at Institute of Informatics, University of Warsaw, Poland, 11 March, 2020.
- "Capsule Network: A new way to analyze single cell RNA-seq data". at Institute of Informatics, University of Warsaw, Poland, 9 March, 2020.
- "Copula based model to identify differential coexpression in TCGA data", at Life Science and Health Seminar, CWI, The Netherlands, February 11, 2020.
- Matlab and its Application in Data Science: Two days Workshop in Kalyani Mahavidyalaya, University of Kalyani, Kalyani, India, August, 2018.
- Application of Data Mining:one day workshop in Department of Computer Science, New Alipore College, West Bengal, India, 2017.
- Association rule mining technique: An application to protein-protein interaction prediction. Three Days Seminar Program on "Machine Learning and Data Mining, MCET, Murshidabad, 2016.

## Publication summary

- Total publications - 38, *h* index-11, *i10* index-13
- Google scholar citation :394
- link: <https://scholar.google.co.in/citations?user=P40MbskAAAAJhl=en>

## Publications

\* – Corresponding author/Main supervisor

- **S. Ray\***, M. Desai, and S. Pyne, Systematic mining of patterns of polysubstance use in a nationwide population survey, **Computers in Biology and Medicine**, Volume 151, Part A, December 2022, 106175
- **S. Ray\***, S. Lall, A. Mukhopadhyay, S. Bandyopadhyay, and A. Schönhuth, Deep variational graph autoencoders for novel host-directed therapy options against COVID-19, **Artificial Intelligence in Medicine**, Volume 134, December 2022, 102418.
- **S. Ray\***, S. Lall, and S. Bandyopadhyay, A deep integrated framework for predicting SARS-CoV2–Human protein-protein interaction, **IEEE Transactions on Emerging Topics in Computational Intelligence**, doi: 10.1109/TETCI.2022.3182354., 2022
- P. Upadhyay and **S. Ray\***, A regularized multi-task learning approach for cell type detection in single RNA sequencing data, **Frontiers in Genetics**, 13 April 2022, <https://doi.org/10.3389/fgene.2022.788832>
- S. Lall, **S. Ray\***, and S. Bandyopadhyay, A topology preserving graph convolution network for clustering of single-cell RNA seq data, **PLoS Computational Biology**, 18(3): e1009600. <https://doi.org/10.1371/journal.pcbi.1009600>, 2022.
- S. Lall, **S. Ray\***, S. Bandyopadhyay, LSH-GAN enables in-silico generation of cells for small sample high dimensional scRNA-seq data., **Nat. Communication Biology**, 577 (2022). <https://doi.org/10.1038/s42003-022-03473-y>.
- S. Lall, A. Ghosh, **S. Ray\***, S. Bandyopadhyay, sc-REnF: An entropy guided robust feature selection for clustering of single-cell rna-seq data, **Briefings in Bioinformatics**, bbab517, <https://doi.org/10.1093/bib/bbab517>, 2022.
- SMM Hossain, L. Khatun, **S. Ray\***, and A. Mukhopadhyay, Pan-cancer classification by regularized multi-task learning., **Scientific Reports**, 11, 24252 (2021).
- S. Lall, **S. Ray\***, S. Bandyopadhyay, RgCop-A regularized copula based method for gene selection in single cell rna-seq data, **PLoS Computational Biology**, 17(10): e1009464. <https://doi.org/10.1371/journal.pcbi.1009464>, 2021.
- SM. Hossain, L. Khatun, **S. Ray\***, A. Mukhopadhyay Identification of key immune regulatory genes in HIV-1 progression, **Gene**, Volume 792, 5 August 2021, 145735
- P. Gharani, **S. Ray**, M. Aruru and S. Pyne Differential Patterns of Social Media Use Associated with Loneliness and Health Outcomes in Selected Socioeconomic Groups, **J. technol. behav. sci.** 2021 (<https://doi.org/10.1007/s41347-021-00208-4>)
- M. Hossain, AA. Halsana, L. Khatun, and **S. Ray\***, and A. Mukhopadhyay Discovering Key Transcriptomic Regulators in Pancreatic Ductal Adenocarcinoma using Dirichlet Process Gaussian Mixture Mode, **Scientific Reports** (7853 (2021), <https://doi.org/10.1038/s41598-021-87234-7>)
- M. Islam, M. Hossain and **S. Ray\***, "DTI-SNNFRA: Drug-target interaction prediction by shared nearest neighbors and fuzzy-rough approximation" **PLoS ONE** <https://doi.org/10.1371>, February 19, 2021
- F. Nielsen, G. Marti, **S. Ray** and S. Pyne, "Clustering patterns connecting COVID-19 dynamics and Human mobility using optimal transport" **Sankhya B**, February, 2021.
- **S. Ray**, S.Lall and S.Bandyopadhyay "CODC: A copula based model to identify differential coexpression", **npj System Biology and Applications**, 6, 20
- S. Pyne, **S. Ray**, R. Gurewitsch , M. Aruru, "Transition from Social Vulnerability to Resiliency vis-à-vis COVID-19", **Statistics and Applications**, Vol.18, pp-197-208, 2020
- S. Biswas, **S. Ray\*** and S. Bandyopadhyay, "Colored Network Motif Analysis by Dynamic Programming Approach: An Application in Host-Pathogen Interaction Network", **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, DOI: 10.1109/TCBB.2019.2923173, 2019.

- **S. Ray** A. Alberuni, and U.Maulik "Computational prediction of HCVhuman protein-protein interaction via topological analysis of HCV infected PPI module", **IEEE Transaction on Nanobioscience**, 17:55-61, 2018,
- **S. Ray**, A.Mukhopdhyay and U. Maulik, "A Review of Computational Approaches for Analysis of Hepatitis C Virus (HCV)-mediated Liver Diseases", **Briefings in Functional Genomics**, doi: 10.1093/bfpg/elx040, 2017.
- SM. Hossain, **S. Ray\*** and A. Mukhopadhyay, "A Comprehensive Analysis on Preservation Patterns of Gene Co-Expression Networks during Alzheimer's Disease Progression", **BMC Bioinformatics**, doi: 10.1186/s12859-017-1946-8, 2017
- U. Maulik, S. Basu and **S. Ray\*** "Identifying protein complexes in PPI network using non-cooperative sequential game", **Scientific Reports**, 7:8410, 2017.
- M. Hossain and **S. Ray\*** and A. Mukhopadhyay "Preservation affinity in consensus modules among stages of HIV-1 Progression", **BMC Bioinformatics**, DOI: 10.1186/s12859-017-1590-3., 2017
- **S.Ray** and U. Maulik, "Identifying differentially coexpressed module during HIV disease progression: A multiobjective approach", **Scientific Reports**, doi: 10.1038/s41598017-00090-2, 2017
- **S. Ray** and U. Maulik "Discovering perturbation of modular structure in HIV progression by integrating multiple data sources through non-negative matrix factorization", **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, 13:6:1086-1099, 2016
- **S. Ray** and S. Bandyopadhyay "A NMF based approach for integrating multiple data sources to predict HIV-1-human PPIs", **BMC Bioinformatics**, doi: 10.1186/s12859-016-0952-6, 2016.
- **S. Ray** and S. Bandyopadhyay, "Discovering condition specific topological pattern changes in coexpression network: an application to HIV-1 progression", **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, 13:1086 - 1099, 2016
- S. Bandyopadhyay, **S. Ray**, A. Mukhopadhyay and U. Maulik "A Multiobjective Approach for Identifying Protein Complexes and Studying their Association in Multiple Disorders", **Algorithms for Molecular Biology**doi: 10.1186/s13015-015-0056-2, 2015.
- S. Bandyopadhyay, **S. Ray**, A. Mukhopadhyay and U. Maulik, A Review of In Silico Approaches for Analysis and Prediction of HIV-1-Human Protein-Protein Interactions, **Briefings in Bioinformatics**, 16:5: 830-851, 2015.
- A. Mukhopadhyay, **S. Ray** and U. Maulik Incorporating the type and direction information in predicting novel regulatory interactions between HIV-1 and human proteins using a biclustering approach **BMC Bioinformatics**, 15:26, 2014.
- A. Mukhopadhyay, **S. Ray** and M. De Detecting Protein Complexes in PPI Network: A Gene Ontology-based Multiobjective Evolutionary Approach, **Molecular BioSystems**, 8:3036-3048, 2013.

## bioRxiv/arXiv preprints/Accepted /revised

- **S. Ray\*** and A. Schönhuth, MarkerCapsule: Explainable Single Cell Typing using Capsule Networks, BioRxiv preprint., doi: <https://doi.org/10.1101/2020.09.22.307512>, 2020

## Conference proceedings

- M. Aruru, S. Pyne, and **S. Ray** A computational framework for identifying patterns of association and transition in addictive substances use over five decades in the US, **APHA's (American Public Health Association) Annual Meeting and Expo, 2021**

- **S. Ray**, S Lall, A Mukhopadhyay, S Bandyopadhyay and A Schönhuth, Deep variational graph autoencoders for novel host-directed therapy options against COVID-19, **29th Conference on Intelligent Systems in Molecular Biology ISMB/ECCB 2021, COVID-19 special track.**
- SM. Hossain, **S. Ray** and A. Mukhopadhyay Detecting Overlapping Gene Communities during Stomach Adenocarcinoma: A Discrete NMF-based Integrative Approach, **2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI), p:1-6, 2021**
- S. Lall, **S. Ray** and S. Bandyopadhyay, Identifying novel SARS-CoV2–human protein interactions using graph embedding, **28th Conference on Intelligent Systems in Molecular Biology (ISMB-2020, COVID-19 special track), 2020**
- SM. Hossain, A. Mukhopadhyay and **S. Ray** Detecting hub genes and key modules in stomach adenocarcinoma using nsNMF based data integration technique **IEEE 18th International Conference on Information Technology, 2019**
- D. Bhattacharjee, SM. Hossain, R. Sultana and **S. Ray** Topological Inquisition into the PPI Networks Associated with Human Diseases Through Graphlet Frequency Distribution, **Pattern Recognition and Machine Intelligence. PReMI 2017, volume 10597:431-437, 2017**
- SM Hossain, **S. Ray**, TS Tannee and A. Mukhopadhyay, Analyzing Prognosis Characteristics of Hepatitis C using a Biclustering Based Approach Number, **7th International Conference on Advances in Computing and Communications (ICACC-20170), Vol:215, pp-282-289, 2017**
- SA. Khan and **S.Ray** Integrating multi-view data: A hypergraph based approach, **First International Conference on Computational Intelligence, Communications, and Business Analytics (CICBA-2017), Vol:776, 347-357, 2017**
- **S. Ray**, SM. Hossain and L. Khatun, Discovering preservation pattern from co-expression modules in progression of HIV-1 disease: An eigengene based approach, **International Conference on Advances in Computing, Communications and Informatics (ICACCI-2016), DOI: 10.1109/ICACCI.2016.7732146, 2016.**
- SM. Hossain, Z. Mahboob, R. Chowdhury, A. Sohel and **S. Ray**, Protein complex detection in PPI network by identifying mutually exclusive protein-protein interactions, **In 6h International Conference on Advances in Computing and Communications (ICACC-2016), Volume 93, pp-1054-1060, 2016.**
- **S. Ray**, A. Hossain and U.Maulik Disease associated protein complex detection: a multi-objective evolutionary approach **International Conference on Microelectronics, Computing and Communications (MicroCom), DOI: 10.1109/Micro-Com.2016.7522420, 2016.**
- **S. Ray**, S. Chakrabarty and A. Mukhopadhyay DCoSpect: a novel differentially co-expressed gene Module detection algorithm using spectral clustering, **4th International Conference on Frontiers in Intelligent Computing: Theory and Applications(FICTA), Vol.404,pp:59-77, 2016.**
- **S. Ray**, S. Biswas, A. Mukhopadhyay and S. Bandyopadhyay Detecting Perturbation in Co-expression Modules Associated with Different Stages of HIV-1 Progression: A Multi-objective Evolutionary Approach **Fourth International Conference of Emerging Applications of Information Technology, DOI:10.1109/EAIT.2014.34, 2015.**
- **S. Ray**, S. Bandyopadhyay, A. Mukhopadhyay, U. Maulik Incorporating Fuzzy Semantic Similarity Measure in Detecting Human Protein Complexes in PPI Network: A Multiobjective Approach, **IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), DOI: 10.1109/FUZZ-IEEE.2013.6622483, 2013.**
- **S. Ray**, A. Mukhopadhyay and U. Maulik Predicting Annotated HIV-1 Human PPIs using a Biclustering Approach to Association Rule Mining, **In Proc. Third International Conference on Emerging Applications of Information Technology, DOI: 10.1109/EAIT.2012.6407854, pp-28-31, 2012**

## Reference:

---

- Prof. Dr. Alexander Schönhuth,  
Professor, Genome Data Science, Universität Bielefeld, Germany.  
Contact: +49 521 106 3362  
email:-aschoen@cebitec.uni-bielefeld.de
- Prof. Saumyadipta Pyne,  
Adjunct (Full) Professor, Department of Statistics and Applied Probability  
University of California Santa Barbara, CA 93106.  
Contact: +1-412-652-2396  
email:-spyne@ucsb.edu
- Prof. Sanghamitra Bandyopadhyay,  
Director, Indian Statistical Institute, Kolkata  
Contact: +91 33 2575 3114  
email:-sanghami@gmail.com

Date, 7th December 2023

---

Dr. Sumanta Ray