Resume

Personal Information

Name: Sourangshu Bhattacharya
Address: Assistant Professor, Dept. of Computer Science and Engineering, IIT Kharagpur, Kharagpur, West Bengal – 721302.
Phone: 03222-282364
E-mail: sourangshu@gmail.com
Date of Birth: 26th June 1980.

Current Position: *Assistant Professor,* Dept. of Computer Science and Engg., IIT Kharagpur.

Education

August 2003 – Jan 2009: *PhD Computer Science*, Indian Institute of Science, Bangalore.
August 2001 – July 2003: *M. Tech. (Distinction), Computer Science*, Indian Statistical Institute, Kolkata.
August 1997 – June 2001: *B. Tech. (First Class), Civil Engg.,* Indian Institute of Technology, Roorkee.

Employment

- **June 2013 onwards:** Assistant Professor, Dept. of Computer Science and Engg., IIT Kharagpur.
- **March 2012 – May 2013**: *Scientist, Yahoo! Labs Bangalore.*
- **December 2008 – February 2012**: *Senior Associate Scientist, Yahoo! Labs, Bangalore.*
- **February – May 2008**, *Visiting Researcher at Laboratory of Computer and Information Science, Helsinki University of Technology, Helsinki, Finland.*

Theses

PhD thesis title: Computational Protein Structure Analysis: Kernel and Spectral Methods
MTech thesis title: Application of Evolutionary Computation to Protein-Ligand Docking Problem.

Publications (Peer Reviewed)

- **Discriminative Link Prediction using Local, Community, and Global Signals.** Abir De, Sourangshu Bhattacharya, Sourav Sarkar, Niloy Ganguly, and Soumen Chakrabarti. *IEEE Transactions On Knowledge And Data Engineering (TKDE), Vol. 28, No. 8, August 2016.*
- **Learning a Linear Influence Model from Transient Opinion Dynamics.** Abir De, Sourangshu Bhattacharya, Parantapa Bhattacharya, Niloy Ganguly, Soumen Chakrabarti. *ACM International Conference on Information and Knowledge Management (CIKM), 2014.*
- **Segmenting Web-Domains and Hashtags using Length Specific Models.** Sriram Srinivasan, Sourangshu Bhattacharya, Rudrasis Chakraborty. *ACM International Conference on Information and Knowledge Management (CIKM), 2012.*
- **Mechanism Design for Cost Optimal PAC Learning in the Presence of Strategic Noisy Annotators**. Dinesh Garg, Sourangshu Bhattacharya, Sundararajan Sellamanickam, Shirish Shevade. *Conference on Uncertainty in Artificial Intelligence (UAI), 2012*.
- **Learning to tokenize web domains.** Sriram Srinivasan, Sourangshu Bhattacharya. *International Conference on World Wide Web, WWW 2011 (poster)*.
- **Robust Formulations for Handling Uncertainty in Kernel Matrices**. Sahely Bhadra, Sourangshu Bhattacharya, Chiranjib Bhattacharyya, Aharon Ben-Tal. *International Conference on Machine Learning (ICML), 2010.*
- **Budgeted PAC Learning with Two Noisy Annotators.** Dinesh Garg, Sundararajan Sellamanickam, Sourangshu Bhattacharya, Shirish Shevade. *Budgeted Learning Workshop, ICML*

2010. http://www.dmargineantu.net/workshop/BudgetedLearning_ICML-2010/.

- **Kernels on Attributed Pointsets with Applications.** Mehul Parsana, Sourangshu Bhattacharya, Chiranjib Bhattacharya, K. R. Ramakrishnan. In Proceedings of *21st Annual Conference on Neural Information Processing Systems (NIPS)*, 2007.
- **Structural Alignment based Kernels for Protein Structure Classification.** Sourangshu Bhattacharya, Chiranjib Bhattacharyya and Nagasuma R Chandra. In Proceedings of *24th International Conference on Machine Learning (ICML)*, 2007.
- **Comparison of protein structures by growing neighborhood alignments.** Sourangshu Bhattacharya, Chiranjib Bhattacharyya and Nagasuma R Chandra. *BMC Bioinformatics*. 2007 Mar 6;8:77.
- **Projections for fast protein structure retrieval.** Sourangshu Bhattacharya, Chiranjib Bhattacharyya and Nagasuma R. Chandra. *BMC Bioinformatics*. 2006 Dec 18;7 Suppl 5:S5.

Book Chapters

- **Kernels on Protein Structures**. Sourangshu Bhattacharya, Chiranjib Bhattacharyya, and Nagasuma R. Chandra. In Computational Intelligence and Pattern Analysis in Biology Informatics, John Wiley & Sons, Inc., New York, USA.

Patents (Filed)

- **Dynamic pricing model for online advertising**. Chaitanya Bandi, Sourangshu Bhattacharya, Krishna Prasad Chitrapura, Dinesh Garg, Sachin Garg, Krishna Pal Singh Rathore.
- **Retrieving engaging user generated comments.** Vidit Jain, Sourangshu Bhattacharya.
- **Learning to Tokenize Web Domains.** Sourangshu Bhattacharya, Sriram Srinivasan. (Defensive Publication).

Awards and Prizes

- Group "You Rock!" award for delivering the decision tree based solution for click through rate prediction for non-guaranteed display exchange.
- *Best student entry* in 2 out of the 3 tasks in KDD Cup 2006.
  http://www.cs.unm.edu/kdd_cup_2006

Projects at Yahoo! Labs

- **Big Data Learning**, implemented **Logistic Regression** on spark that could train logistic regression model in a distributed manner using 100 cores. This is part of Yahoo! Code base now.
- **Decision Tree** based algorithm for click through rate prediction for NGD exchange. The data was extremely sparse (1 in 1000 impressions are clicked), and the features include high cardinality categorical variables. This was implemented using **Hadoop, a Big data platform**.
- A system (called throttling) for allocating impressions for new advertisements (cold starts) in the NGD system, such that they don't incur heavy losses. This is a modified version of epsilon greedy strategy for **multi-armed bandit problems**.

Teaching Experience

- "Machine Learning" elective course in Autumn 2013 and 2014 with 50 and 160 student respectively.
- "Advanced Machine Learning" elective course in spring 2014.
- Teaching Assistant for course **Linear and Nonlinear Optimization** (2005).
- Taken a series of lectures on Graphical Models and Inference for Topics in Machine Learning course at IISc (2007).
- Taken a series of 3 lectures on "Web Intelligence", as a part of the "Business Inteliigence" course at IIM, Calcutta (2012).

Professional Services

Reviewer for:
- Neural Information Processing Systems (NIPS) 2011.
- Pacific Asia Conference on Knowledge Discovery and Datamining (PaKDD) 2010.
- IEEE/ACM Transactions on Computational Biology and Bioinformatics.
- BMC Structural Biology.
- IEEE International Conference on Data Mining (ICDM) 2007.
- ACM International Conference on Web Search and Data Mining (WSDM) 2008.

Skills

- **Languages**: C/C++, Java, Matlab, Perl.
- **Packages / Libraries**: LAPACK, GNU GSL, LIBSVM, SVMSTRUCT.
- **Distributed Computation:** Hadoop, Pig, Spark, MPI.

Technical Reports, Invited/Nominated Papers, Posters, Talks

- Sourangshu Bhattacharya. **Throttling in NGD Predict.** Yahoo Technical Report 2011.
- Kohinoor Dasgupta, Sourangshu Bhattacharya. **Explore Exploit policy for Gamma Reward Distribution** Yahoo Technical Report 2009.
- Sourangshu Bhattacharya, Chiranjib Bhattacharyya, Nagasuma Chandra. **Matchprot: Projections for Fast Protein Structure Retrieval**. Poster in *TechVista*. Jan 2007, Microsoft Research India, Bangalore, India.
- Sourangshu Bhattacharya, Chiranjib Bhattacharyya, Nagasuma Chandra. **Projections for Fast Protein Structure Retrieval**. *6th Annual Inter Research Institute Student Seminar in Computer Science (IRISS)*. 2007. International Institute of Information Technology, Hyderabad, India.
- Sourangshu Bhattacharya, Chiranjib Bhattacharyya, Nagasuma Chandra. **Matchprot: Projections for Fast Protein Structure Retrieval**. Poster in *5th International Conference on Bioinformatics (InCoB)*. Dec 2006. New Delhi, India.
- Sourangshu Bhattacharya. **A talk on Protein Structure Comparison**. *Showcase 2005*. Dept. of Computer Science and Automation, Indian Institute of Science, Bangalore, India.
- Sourangshu Bhattacharya. **Protein Structure Comparison and Neighborhood Preserving Representation.** *Doctoral Symposium*, Dec 2004, Amrita Vishwa Vidyapeetham, Coimbatore, India. *(The paper also got the best presentation award)*
- Sourangshu Bhattacharya, Chiranjib Bhattacharyya. **Using Neighborhood Preserving Projections for Comparing Protein Structures.** *Technical report IISc-CSA-TR-2004-14,* Nov 2004, http://archive.csa.iisc.ernet.in/TR/2004/14/.