Prithwish Chakraborty

Data Scientist
IBM Watson Health, Cambridge, USA

February 10, 2017
prithwish.chakraborty@ibm.com
https://prithwi.github.io/

Education

Fall 2010 - Spring 2016	**PhD Student** in Computer Science, Virginia Tech, USA. **GPA:** 3.97/4.00 **Advisor** Dr. Naren Ramakrishnan
2006 - 2010	**B.E.** in Electronics and Telecomm. Engg., Jadavpur University, India. **GPA:** 9.27/10

Professional Experience

- **IBM Watson Health** Cambridge, MA
  Data Scientist. Manager: Dr. Faisal Farooq *Aug 2016 - Current*

  – Working in the RWE team as a data scientist. Designing and developing data analytics workbooks to demonstrate usage of EMR data towards health prediction. Also involved in day-to-day development activities towards consumer driven product.
  – Invovled in general research acroos teams towards public health forecasting

- **Virginia Tech** Arlington, VA
  GRA, Discovery Analytics Center. Advisor: Dr. Naren Ramakrishnan *Fall 2011 - Spring 2016*

  – Worked on Disease Forecasting. Create streaming data ingestion module for Disease surveillance, Weather and News. Implemented big-date disease forecasting pipeline for EMBERS in *python* over EMBERS *AWS* cluster framework. Forecasts sent in real-time without human supervision and continuously evaluated
  – Interfaced with several agencies such as IARPA and CDC for disease forecasting problems. Collaborated with several institutes such as *NDSSL*, Virginia Tech, *Healthmap*, Harvard Medical School and *YeLab*, University of Michigan, Ann Arbor
  – Helped organize and run flu forecasting market in collaboration with Scicast, George Mason University
  – Collaborated on several projects such as pattern detection in Ionospheric Radiation, with ECE VT and solar Photo-voltaic energy prediction with HP Labs

- **Microsoft Research India** Bangalore, IN
  Visiting Researcher, invited by Dr. Srivatsan Laxman *June 2012*

  – Worked on formalization of histogram estimation using random resampling with possible applications towards distribution prediction under out-of-core data estimation problems

- **Amazon Web Services** Seattle, WA
  SDE Intern, mentored by Paul Sharpe *Summer 2011*

  – Worked on updating an Internal Console of AWS Beanstalk towards more real-time status monitoring of the same. Implemented in *java* using *Spring-Hibernate* framework using a combination of *JSP, js, css, and html5*

- **Virginia Tech** Blacksburg, VA
  GTA in Dept. of Computer Science. Instructor: Mr. N.D. Barnette *Fall 2010 - Spring 2011*

  – Grading and tutoring duties for C++ programming class of ≈ 200 students.

- **Indian Institute of Technology, Delhi** Delhi, IN
  Research Intern, advised by Prof. B.K. Panigrahi *Summer 2009*

– Worked on application of heuristic optimization and machine learning methods to power system characterization and early Epilepsy detection using wavelets. Published several papers to several journals and conferences.

Research Statement

Research Focus: Data-driven spatio-temporal modeling under weakly correlated signals
- *Application area:* Infectious Disease modeling
- *Broad Focus:* Data Science, Machine Learning and Pattern Recognition.

Activities

Tutorial	P. Chakraborty, M. Marathe, N. Ramakrishnan, and A. Vulikanti. Computational Epidemiology and Public Health Policy Planning, Feb 2016. Tutorial, AAAI 2016
Invited Talks	P. Chakraborty. Data Driven Model for Disease Forecasting, 10 2014. Invited Talk, First IEEE international Workshop on Big Data in Computational Biology (BCDE 2014)
Patent	M. Marwah, M. Arlitt, P. Chakraborty, and N. Ramakrishnan. Predicting near-future photovoltaic generation, Sept. 28 2012. US Patent App. 13/631,480
Opinion Piece	P. Chakraborty. US Flu Forecast: Exploring links between national and regional level seasonal characteristics. `http://bit.ly/1CSHTk7`, 2014. Accessed: 2015-03-21

Publications

Current Publications

2016	P. Chakraborty, S. Muthiah, R. Tandon, and N. Ramakrishnan. Hierarchical Quickest Change Detection via Surrogates. *arXiv preprint arXiv:1603.09739*, 2016
	S. Ghosh, P. Chakraborty, E. Cohn, J. S. Brownstein, and N. Ramakrishnan. Designing Domain Specific Word Embeddings: Applications to Disease Surveillance. *arXiv preprint arXiv:1603.00106*, 2016
	H. Wu, Y. Ning, P. Chakraborty, J. Vreeken, N. Tatti, and N. Ramakrishnan. Generating Realistic Synthetic Population Datasets. *arXiv preprint arXiv:1602.06844*, 2016
2015	P. Khadivi, P. Chakraborty, R. Tandon, and N. Ramakrishnan. Time Series Forecasting via Noisy Channel Reversal. In *Machine Learning for Signal Processing (MLSP), 2015 IEEE 25th International Workshop on*, pages 1–6. IEEE, 2015
	Z. Wang, P. Chakraborty, S. R. Mekaru, J. S. Brownstein, J. Ye, and N. Ramakrishnan. Dynamic Poisson Autoregression for Influenza-Like-Illness Case Count Prediction. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1285–1294. ACM, 08 2015
	H. Wu, P. Chakraborty, S. Ghosh, and N. Ramakrishnan. Forecasting Influenza in Senegal with Call Detail Records, 04 2015. NETMOB 2015

2014	P. Chakraborty, P. Khadivi, B. Lewis, A. Mahendiran, J. Chen, P. Butler, E. O. Nsoesie, S. R. Mekaru, J. S. Brownstein, M. V. Marathe, and N. Ramakrishnan. Forecasting a Moving Target: Ensemble Models for ILI Case Count Predictions. In *Proceedings of the 2014 SIAM International Conference on Data Mining, Philadelphia, Pennsylvania, USA, April 24-26, 2014*, pages 262–270, 2014
2012	P. Butler, P. Chakraborty, and N. Ramakrishan. The Deshredder: A visual analytic approach to reconstructing shredded documents. In *Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on*, pages 113–122. IEEE, 2012
	P. Chakraborty, M. Marwah, M. Arlitt, and N. Ramakrishnan. Fine-Grained Photovoltaic Output Prediction Using a Bayesian Ensemble. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*, page online, 2012

Other Publications

2015	S. Ghosh, P. Chakraborty, E. O. Nsoesie, S. R. Mekaru, J. S. Brownstein, and N. Ramakrishnan. A Temporal Topic Modeling Analysis to Assess the Association between Trends in News Reports and Infectious Disease Outbreaks. to be submitted, 2016
	F. S. Tabataba, P. Chakraborty, N. Ramakrishnan, M. Marathe, J. Chen, and B. Lewis. Standard Measures and Quality Metrics for Evaluating the Performance of Forecasting Methods: Special Study on Influenza in US. to be submitted, 2016
	P. Chakraborty, B. Lewis, S. Eubank, M. Marathe, J. Brownstein, and N. Ramakrishnan. How not to forecast the flu. to be submitted, 2016
2012	P. Chakraborty, G. G. Roy, B. Panigrahi, R. Bansal, and A. Mohapatra. Dynamic economic dispatch using harmony search algorithm with modified differential mutation operator. *Electrical Engineering*, 94(4):197–205, 2012
	T. K. Gandhi, P. Chakraborty, G. G. Roy, and B. K. Panigrahi. Discrete harmony search based expert model for epileptic seizure detection in electroencephalography. *Expert Systems with Applications*, 39(4):4055–4062, 2012
2011	G. G. Roy, S. Das, P. Chakraborty, and P. N. Suganthan. Design of non-uniform circular antenna arrays using a modified invasive weed optimization algorithm. *Antennas and Propagation, IEEE Transactions on*, 59(1):110–118, 2011
	P. Chakraborty, S. Das, G. G. Roy, and A. Abraham. On convergence of the multi-objective particle swarm optimizers. *Information Sciences*, 181(8):1411–1425, 2011
2010	G. G. Roy, P. Chakraborty, and S. Das. Designing fractional-order $PI^\lambda D^\mu$ controller using differential harmony search algorithm. *International Journal of Bio-Inspired Computation*, 2(5):303–309, 2010
	G. G. Roy, P. Chakraborty, S.-Z. Zhao, S. Das, and P. N. Suganthan. Artificial foraging weeds for global numerical optimization over continuous spaces. In *IEEE Congress on Evolutionary Computation*, pages 1–8, 2010
2009	P. Chakraborty, G. G. Roy, S. Das, and B. Panigrahi. On population variance and explorative power of invasive weed optimization algorithm. In *Nature & Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on*, pages 227–232. IEEE, 2009
	G. G. Roy, B. Panigrahi, P. Chakraborty, and M. K. Mallick. On optimal feature selection using modified harmony search for power quality disturbance classification. In *Nature & Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on*, pages 1355–1360. IEEE, 2009
	P. Chakraborty, G. G. Roy, S. Sinha, S. Bose, A. Mondal, and S. Das. Automatic Shape Independent Clustering Inspired By Ant Dynamics. In *The Proceedings of International Workshop on Machine Intelligence Research organized by Machine Intelligence Research Labs*, pages 64–74, 2009

P. Chakraborty, G. G. Roy, S. Das, D. Jain, and A. Abraham. An improved harmony search algorithm with differential mutation operator. *Fundamenta Informaticae*, 95(4):401–426, 2009

D. Jain, G. G. Roy, P. Chakraborty, and S. Das. Fuzzy Entropy-based Object Segmentation with an Inertia-Adaptive PSO. In *Advanced Computing and Communications, 2008. ADCOM 2008. 16th International Conference on*, pages 13–18. IEEE, 2008

Participation in Conferences

- **KDD 2015**, 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, Australia, August 2015.
- **SDM 2014**, 2014 SIAM International Conference on Data Mining, Philadelphia, USA, April 2014.
- **DDD 2013**, 2nd International Conference on Digital Disease Detection, San Francisco, USA, September 2013.
- **AAAI 2012**, 26th AAAI Conference on Artificial Intelligence, Toronto, Canada, July 2012.

Awards and Honors

Award	Who's Who in American College's and Universities, Virginia Tech	2014
	Student Travel Award, AAAI	2012
Honor	Vice President, Alpha Epsilon Lambda	2013
	Upsilon Pi Epsilon, Inducted	2013
	Phi Kappa Phi, Invited	2014
	Golden Key,Invited	2014
	Tau Beta Pi,Invited	2013
Other	President, Indian Student Association at Virginia Tech	2012-2013

Technical Skills

Programming Python, **C/C++**, R, Java, Matlab

Javascript, Perl, HTML/CSS

Frameworks NOSQL: MongoDB, MapReduce, Django, NodeJS, Spring