Sudipta Pathak

80 Cisar Road, Apartment No. 20
Willington, CT 06279

+1 765-702-1832
sudipto.pathak@gmail.com
www.linkedin.com/in/sudipta-pathak-a6910117
https://github.com/sudiptap

Education

- **University of Connecticut** — Storrs, CT
  MS and PhD, Computer Engineering (Grades: 3.72) — *Aug 2011 - Aug 2017*
  - Thesis : Applied Machine Learning Algorithms to Improve Top-k Recommendations
  - My research focuses on the following
    * Algorithms to Improve prediction quality of recommendation systems.
    * Propose and Implement algorithm to addess motif search problem in bioinformatics.
    * Application of Neural Network for BigData compression.

- **West Bengal University of Technology** — India, West Bengal
  BS Computer Engineering (Grades: 4.00) — *Aug 2004 - Aug 2008*

Skills

Research: Data mining, machine learning, recommendation systems, algorithm and complexity, collaborative filtering, big data analytics, high performance computing, matrix factorization, bioinformatics.

Languages: C/C++, Java and worked on several other languages and technologies(PHP, Python, C-sharp, General Purpose GPU, CUDA, MPI, OpenMP, R, HTML, CSS, XML, JavaScript, .NET, Linux, Oracle).

Operating Systems: Linux.

Applications: MatLab.

Work Experience

- **Bentley Systems Inc.** — Watertown, CT
  Deep Learning Research Intern — *Feb 2015 - May 2015*
  - Applied deep learning to smart water networks to better predict water usage and identify abnormal events.
  - Improved prediction accuracy by 2.2 percent on previous ANN based framework there by reducing water leakage and False Alarm on Real dataset.
  - Scaled up application using GPU technology and improved runtime by several times.

- **University of Connecticut** — Storrs, CT
  Graduate and Research Assistant — *Aug 2011 - Dec 2016*
  - Developed machine learning algorithms to improve item recommendation.
    * Proposed and implemented algorithm to improve prediction accuracy of collaborative filtering based recommendation system.
    * Proposed new model to apply locally weighted regression in presense of auxiliary information to predict missing entries in sparsely populated datasets using matrix factorization framework.
  - Combine multiple predictive models to improve item recommendation.
    * Proposed algorithm to combine multiple kernels using Ada-Boost in to predict missing entries in sparsely populated datasets.
    * Proposed model offers better prediction accuracy than linear multiple kernel learning on many real datasets.
  - Developed and implemented algorithms for mining motifs in DNA and protein sequences.

* Proposed and implemented algorithm to locate approximately repeated patterns in DNA and protein sequences.
  * Introduced edit distance based motif search to locate all d-neighborhood motifs.
  - Developed and implemented algorithm for lossless compression of FASTQ files for big biological data.
    * Proposed lossless compression algorithm for FASTQ files.
    * Achieved higher compression ratio than all well known algorithms by applying neural network based prediction.

- **University Information Technology Services**   University of Connecticut, Storrs, CT
  Software Engineer   *Aug 2015 -*
  - Software Developer in the UConns Facilities Asset Management Information System (FAMIS) support team.
  - implementing Oracle technology in FAMIS, developing university-wide Visual Map implementation, developing maintenance management and capital project management utilities.

- **Cognizant Technology Solutions**   India, Chennai
  Software Engineer   *Jan 2009 - Jun 2011*
  - Coded extensively using .Net Framework 3.5.
  - Developed and maintained webpages.
  - Wrote SQL queries, stored procedures, cursors and triggers as a part of backend development.
  - Worked on ASP.Net based reporting system.
  - Migrated code bases from C++ to C-sharp.
  - Extensively worked on databases.
  - Worked on ANT scripts and Install Shield for build delivery.
  - Implemented and practiced scrum in all the projects.

- **Indian Statistical Institute**   India, West Bengal
  Undergraduate Research Intership   *Jan 2008 - Aug 2008*
  - Partial image encryption for secured multimedia communications.

Awards and Achievements

- Competitive Pre-doctoral fellowship

- ICCABS 2013 Travel Award

- Reviewer of IEEE Transactions on Big Data Conference

- Associate of the Month Award, Cognizant Technology Solution

Publications

- Soumitra Pal, Sudipta Pathak, Sanguthevar Rajasekaran. "On Speeding-up Parallel Jacobi Iterations for SVDs", 18th IEEE International Conference on High Performance Computing and Communications (HPCC 2016)

- Zheng Yi Wu, Mahmaud El-Maghraby, Sudipta Pathak. Applications of Deep Learning for Smart Water Networks. Computing and Control for the Water Industry (CCWI2015) Sharing the best practice in water management, Volume 119, 2015, Pages 479485

- Sudipta Pathak [1]; Vamsi Kundeti [2]; Martin Schiller [3] and Sanguthevar Rajasekaran [1]. A Structure Based Algorithm for Improving Motifs Prediction. Pattern Recognition in Bioinformatics. Volume 7986 of the series Lecture Notes in Computer Science pp 242-252.

- Sudipta Pathak [1]; Sanguthevar Rajasekaran [1] and Marius Nicolae [1], EMS1 An Elegant Algorithm for Edit Distance Based Motif Search. International Journal of Foundation of Computer Science, Volume 24, Issue 04, June 2013.

- Subrata Saha[1]; SanguthevarRajasekaran[2]; Jinbo Bi[2], and Sudipta Pathak[3], Efficient Techniques for Genotype-Phenotype Correlational Analysis. BMC Medical Informatics and Decision Making 2013, 13:41 doi:10.1186/1472-6947-13-41

- Sanguthevar Rajasekaran [1], and Sudipta Pathak [2], Efficient Algorithms for the Closest Pair Problem and Applications Unpublished Manuscript.

- Nicolae M[1], Pathak S[1], Rajasekaran S[1], LFQC: a lossless compression algorithm for FASTQ files. Bioinformatics. 2015 Oct 15;31(20):3276-81. doi: 10.1093/bioinformatics/btv384.

- A. Mitra, S. Palit, B. B. Chaudhury, S. Kundu, S. Pathak and R. Datta, A New Partial Image Encryption Method for Secured Multimedia Communication, in Proc. Workshop on Mobile Systems (WoMS), WBUT, Kolkata, July 2008.

Conferences

- 18th IEEE International Conference on High Performance Computing and Communications (HPCC 2016)

- 13th Annual Computing and Control in the Water Industry (CCWI) 2015 Conference

- 8th IAPR International Conference on Pattern Recognition in Bioinformatics, 2013

- 3rd IEEE International Conference on Computational Advances in Bio and Medical Sciences (ICCABS), 2013

References

Dr. Sanguthevar Rajasekaran (Major Advisor)
UTC Chair Professor of CSE and Director of Booth Engineering Center for Advanced Technologies (BECAT)
Web: http://www.engr.uconn.edu/ rajasek/