Avradeep Bhowmik

linkedin.com/in/avradeep1/ • avradeep.1@utexas.edu • avradeep1.github.io

Education

MS–PhD 2013–Present
The University of Texas at Austin, *USA*, CGPA: 3.90/4.00.
Advisor: Prof. Joydeep Ghosh,
Department of Electrical and Computer Engineering.
- PhD, ECE 2016–Present (Expected graduation: Summer 2018),
- MS, ECE 2013–2016.

B. Tech 2009–2013
Indian Institute of Technology, Bombay, India, CGPA: 9.43/10.00.
Department of Electrical Engineering.

PhD Thesis

Dissertation
Learning from Aggregated Data
Data aggregation is a ubiquitous feature in modern data-driven applications due to concerns like privacy, robustness and scalability. However, most existing machine learning algorithms are only designed to work with data at the individual granular level. The focus of my dissertation is to bridge this gap and design algorithms and modelling frameworks that have strong predictive performance at the individual level but can be trained with data only available as aggregates. A variety of both linear and non-linear models are studied and extended to the aggregated data setting using techniques from diverse areas like compressed sensing and Fourier analysis. The methods thus developed in the thesis are applied to domains like healthcare, climate science, ecological studies, etc. where data is frequently available only in a naturally aggregated aggregated form.

Research Interests
I am broadly interested in machine learning, data mining, statistical inference and related fields, with a specific focus on data reconstruction algorithms, ranking and rank aggregation, and learning with obfuscated data. In the past I have also worked on applications involving sparse modelling, hierarchical learning and geometry-aware sparse maps for recommender systems, and submodular optimisation techniques applied to problems in operations research.

Selected Coursework
Special Topics in Machine Learning, Large Scale Optimisation, Statistical Relational Learning, Information Theory, Graphical Models, Approximation Algorithms, Time Series and Dynamic Models, Monte Carlo Methods, Probability and Stochastic Processes, Genomic Signal Processing, Program Derivation, Topics in Automata and Logic

Publications

submitted **A Bhowmik**, Z Xing, S Rajan, *"A General Framework for Learning with Taxonomy"*, (currently under review)

submitted **A Bhowmik**, M Chen, Z Xing, S Rajan, *"Predictive Modelling with Aggregated Cost-per-Click Data for Online Advertising"*, (currently under review)

2017 **A Bhowmik**, J Ghosh, O Koyejo, *"Frequency Domain Predictive Modelling with Aggregated Data"*, Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017, Fort Lauderdale, Florida, USA, April 20-22, 2017

2017 **A Bhowmik**, J Ghosh, *"LETOR Methods for Unsupervised Rank Aggregation"*, Proceedings of the 26th International World Wide Web Conference (WWW) 2017, Perth, Australia, April 3-7, 2017

2016 **A Bhowmik**, J Ghosh, O Koyejo, *"Sparse Parameter Recovery from Aggregated Data"*, Proceedings of the 33rd International Conference on Machine Learning (ICML) 2016, New York City, NY, USA, June 19-24, 2016

914 E 52nd Street – Austin, TX 78751 – USA
+1 (512) 300 4487 • avradeep.1@utexas.edu

2016 **A Bhowmik**, N Liu, E Zhong, B N Bhaskar, S Rajan, *"Geometry Aware Mappings for High Dimensional Sparse Factors"*, Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS) 2016, Cadiz, Spain, May 9-11, 2016

2015 **A Bhowmik**, J Ghosh, O Koyejo, *"Generalized Linear Models for Aggregated Data"*, Proceedings of the 18th International Conference on Artificial Intelligence and Statistics, (AISTATS) 2015, San Diego, California, USA, May 9-12, 2015, *(Oral presentation)*

2014 **A Bhowmik**, V Borkar, D Garg, M Pallan, *"Submodularity in the Team Formation Problem"*, Proceedings of the 2014 SIAM International Conference on Data Mining (SDM), Philadelphia, Pennsylvania, USA, April 24-26, 2014

Research Experience

2013–Present **The University of Texas Austin**, *USA*, Graduate Research Assistant.
Worked on multiple NSF and industry funded projects.
Supervisor: Dr. Joydeep Ghosh
- Fourier analysis for linear modelling with non-uniform spatio-temporally aggregated data
- Sparse linear estimation and parameter recovery with group-wise aggregates
- Generalised linear modelling with histogram aggregation
- Monotone Retargeting and LETOR methods for rank aggregation with object features

2012–2013 **Indian Institute of Technology (IIT) Bombay**, *India*, Undergraduate Research.
Supervisor: Dr. Vivek Borkar
- Sub-modular formulation and analysis of ad-hoc team formation problem in social networks *(joint work with Dr. Dinesh Garg, IBM Research)*
- Adaptive clustering with the weighted majority algorithm for cold-start recommendation systems *(joint work with Dr. Onkar Dabeer, Qualcomm Research)*

Summer '12 **Institute of Science and Technology (IST)**, *Austria*, Visiting Researcher.
Supervisor: Dr. Christoph Lampert
- Randomised hashing techniques for clustering images in a large database
- Designing new similarity metrics for efficient neighbourhood graph construction

Industry Experience

Summer '17 **Criteo Research, Palo Alto**, *USA*, Research Intern.
Manager: Dr. Suju Rajan
Designed a novel modelling and algorithmic framework that can learn supervised generalised linear models with taxonomy. The framework was designed to be compatible with missing and noisy taxonomic information, both for training and in deployment, and was able to learn accurate models even with more than half the taxonomic labels corrupted or replaced with random errors.
In a parallel project, developed a novel predictive modelling framework for estimating individual-level cost-per-click (CPC) that can be trained using CPC data only available as daily or hourly aggregates over clickstreams. The framework was generalised to incorporate GLMs and demonstrated applicability to domains beyond online advertising

Summer '16 **Verizon Labs, Palo Alto**, *USA*, Research Intern.
Manager: Dr. Santanu Das
Designed feature selection algorithms for scalable time series recommendation and extracting relevant information from customer purchase history to learn top ranked items for future purchase interest. Technique thus developed reduced the complexity of training by 92% with almost no loss in accuracy and provided actionable marketing insights on relevant customer information for future promotional campaigns

914 E 52nd Street – Austin, TX 78751 – USA
+1 (512) 300 4487 • *avradeep.1@utexas.edu*

Summer '15 **Yahoo Labs, Sunnyvale**, *USA*, Research Intern.

Manager: Dr. Suju Rajan

Designed algorithms for high dimensional sparse factor embeddings for real time search and recommendation. Developed a novel framework that used a mapping schema between a high dimensional permutation space to tiles on a tessellated unit sphere to obtain sparse parameters for factorisation models. Empirical tests on real datasets found the technique to outperform standard tree-based or hashing techniques in both speed and accuracy

Academic Honors

Recipient of the George J. Heuer, Jr. Ph.D. Endowed Graduate Fellowship for year 2013

Qualified IIT JEE with an All India Rank of 34 (top 0.008%) and rank 1 in the IIT Guwahati Zone

All India Rank of 212 (top 0.0002%) and State rank 1 in AIEEE year 2009

6th rank in the state (top 0.006%) in WBJEE examination 2009

Programming Skills

Python, R, Matlab, C/C++.

Academic Service

Reviewer: IEEE Transactions on Network Science and Engineering, Journal of Ecological Modelling, Transactions on Knowledge and Data Engineering, IEEE Access

PC Member: ACM International Conference on the World Wide Web 2017

Miscellaneous

Volunteer for National Service Scheme (NSS) for the year 2009–2010

Member of the organising committee for TechFest 2010 at IIT Bombay

References

Available upon request

914 E 52nd Street – Austin, TX 78751 – USA

+1 (512) 300 4487 • *avradeep.1@utexas.edu*