

Debolina Paul

Baburam Ghosh Road
Government Housing Estate
Block-I, Flat no.-8
Kolkata-700040, India
☎ +91 9123838542
✉ debolinap8@gmail.com

"All models are wrong, but some are useful" - George Box

About Me

I am an M.Stat second-year student of Indian Statistical Institute (ISI), Kolkata, India. I will be joining the PhD program in Statistics at Stanford University this fall. I am broadly interested in Machine Learning, Statistical Learning Theory, Bayesian Statistics and Optimization Techniques. I am particularly interested in Cluster Analysis, Classification, Pattern Recognition and Convex Optimization. I mainly use R and Python3 for programming, but am fluent in Matlab and C/C++.

Education

- 2021-Present **PhD. in Statistics**, Stanford University, CA, USA.
- 2019-2021 **Master of Statistics (M.Stat)**, Indian Statistical Institute, Kolkata, West Bengal, India.
Specialization: Applied Statistics and Computational Statistics.
Thesis: Median of Means and its Applications to Unsupervised Learning; Advisor: Dr. Swagatam Das.
- 2016-2019 **Bachelor of Statistics (B.Stat (Hons.))**, *First Division with Distinction, Gold Medalist*, Indian Statistical Institute, Kolkata, West Bengal, India.
Major: Statistics.
- 2013-2015 **Higher Secondary School**, Bethune Collegiate School, Kolkata, West Bengal, India.
- 2001-2013 **Primary and High School**, Holy Child Girls' Higher Secondary School, Kolkata, West Bengal, India.

Peer Reviewed Journal Publications (Accepted/Published)

- 2021 **Paul, D.**, Chakraborty, S. and Das, S., 2021. On the Uniform Concentration Bounds and Large Sample Properties of Clustering with Bregman Divergences. *Stat*. DOI: <https://doi.org/10.1002/sta4.360>.
- 2020 **Paul, D.** and Das, S., 2020. A Bayesian Non-parametric Approach for Automatic Clustering with Feature Weighting. *Stat*. DOI: <https://doi.org/10.1002/sta4.306>.
- 2020 Chakraborty, S., **Paul, D.** and Das, S., 2020. Hierarchical clustering with optimal transport. *Statistics & Probability Letters*, p.108781. DOI: <https://doi.org/10.1016/j.spl.2020.108781>.

Peer Reviewed Conference Publications (Accepted/Published)

- 2021 Chakraborty, S.*, **Paul, D.*** and Das, S., 2021, July. t -Entropy: A New Measure of Uncertainty with Some Applications. In *2021 IEEE International Symposium on Information Theory* (to appear). Preprint: <https://arxiv.org/abs/2105.00316> (*denotes equal contribution).
- 2021 Chakraborty, S.*, **Paul, D.*** and Das, S., 2021, February. Automated Clustering of High-dimensional Data with a Feature Weighted Mean Shift Algorithm. In *The 35th International Conference of Association for the Advancement of Artificial Intelligence (AAAI), 2021* (to appear). Preprint: <https://arxiv.org/abs/2012.10929> (*denotes equal contribution).
- 2020 Chakraborty, S.*, **Paul, D.***, Das, S. and Xu, J., 2020, June. Entropy Weighted Power k-Means Clustering. In *The 23rd International Conference on Artificial Intelligence and Statistics (AISTATS), 2020* (pp. 691-701). DOI: <http://proceedings.mlr.press/v108/chakraborty20a.html>(*denotes equal contribution).

- 2019 **Paul, D.**, Chakraborty, S., Das, S. and Zelinka, I., 2019, July. On the non-convergence of differential evolution: some generalized adversarial conditions and a remedy. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion (GECCO), 2019* (pp. 265-266). ACM. DOI: <https://doi.org/10.1145/3319619.3322007>.
- 2019 **Paul, D.** and Chakraborty, S., 2019, October. A New Visual Cryptography Scheme with Perfect Contrast using Galois Fields. In *2019 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE), 2019* (pp. 7-11). IEEE. DOI: <https://ieeexplore.ieee.org/document/8979050>.

Preprints

- 2021 **D. Paul**, S. Chakraborty, S. Das. Robust Principal Component Analysis: A Median of Means Approach. arXiv preprint arXiv:2011.06461. <https://arxiv.org/abs/2011.06461>.
- 2020 **Paul, D.***, Chakraborty, S.*, Das, S. and Xu, J., 2020. Kernel k -Means, By All Means: Algorithms and Strong Consistency. arXiv preprint arXiv:2011.06461. <https://arxiv.org/abs/2011.06461>. (*denotes equal contribution).
- 2020 **Paul, D.**, Chakraborty, S., Li, D. and Dunson, D., 2020. Principal Ellipsoid Analysis (PEA): Efficient non-linear dimension reduction & clustering. arXiv preprint arXiv:2008.07110. <https://arxiv.org/abs/2008.07110>.

Summer Schools

- Sep 7-Sep 10, 2021 **Research with Impact: Developing Skills as a Community-Engaged Scholar**, organized by *Stanford Graduate Summer Institute (SGSI)*, Stanford University, Stanford, California, USA.
- Jul 27-Aug 4, 2021 **Deep Learning Theory Summer School at Princeton**, *Princeton University*, Princeton, New Jersey, USA.

Research Positions

- May-August, 2020 **Duke Statistical Science Summer Research Internship**, (*online due to COVID*) organized by *Statistical Science*, Duke University, Durham, North Carolina, USA.
Supervisor: Dr. David Dunson, Arts and Sciences Distinguished Professor of Statistical Science & Mathematics, Duke University.
Title: Identifying Correlation between Human Brain Structural and Functional Connectomes.
- May-July, 2019 **Duke Statistical Science Summer Research Internship**, organized by *Statistical Science*, Duke University, Durham, North Carolina, USA.
Supervisor: Dr. David Dunson, Arts and Sciences Distinguished Professor of Statistical Science & Mathematics, Duke University.
Title: Principal Ellipsoid Analysis (PEA): Efficient non-linear dimension reduction & clustering.
- May-July, 2018 **RC Bose Summer Internship in Cryptology**, organized by the *R. C. Bose Centre for Cryptology and Security*, Indian Statistical Institute, Kolkata, supported and funded by Microsoft Research Labs, India.
Supervisor: Dr. Bimal Kumar Roy, ISI Kolkata.
Title: A New Visual Cryptography Scheme with Perfect Contrast using Galois Fields.

Academic Awards and Achievements

- 2021 **EDGE: Enhancing Diversity in Graduate Education Doctoral Fellowship**, Stanford University.
- 2020 Recipient of the *Indian Association for Research in Computing Science (IARCS) travel grant* by *ACM India* for attending the conference AISTATS, 2020.
- 2020 *Usri Gangopadhyay Memorial Gold Medal* for outstanding academic performance in B.Stat 2016-19.

- 2019 **Ranked 1st in Madhava Mathematics Competition 2019 (All India Rank), organized by the Homi Bhaba Centre for Science Education, Tata Institute of Fundamental Research (TIFR), Mumbai, India.** According to the organizers, this is the first and only time a female candidate to have received this honor.
- 2019 Mukul Chaudhuri Memorial Award for academic excellence in B.Stat 2nd year.
- 2018 Ranked within top quartile at Simon Marais Mathematics Competition for Asia-Pacific Region and ranked 7th in India.
- 2018 Participant at the Madhava Mathematic Camp 2018, organized by the Homi Bhaba Centre for Science Education, Tata Institute of Fundamental Research (TIFR), Mumbai, India.
- 2018 Mukul Chaudhuri Memorial Award for academic excellence in B.Stat 1st year.
- 2017 Participant at the Madhava Mathematics Camp 2017, organized by the Homi Bhaba Centre for Science Education, Tata Institute of Fundamental Research (TIFR), Mumbai, India.

Selected Workshops

- June 6, 2019 Presenter of the Session of *Hands-on Demonstration of Machine Learning*.
Supervisor: Dr. David Dunson, Duke University.
Workshop: Leveraging Artificial Intelligence and Machine Learning to Advance Environmental Health Research and Decisions, a workshop convened by the Standing Committee on the Use of Emerging Science for Environmental Health Decisions, organized at Washington, DC, USA, by *The National Academies of Sciences, Engineering and Medicine*.
- June 13, 2019 Presenter of *Convex Mixture Regression for Quantitative Risk Assessment*, published at *Biometrics*, Volume 74, Issue 4.
Supervisor: Dr. David Dunson, Duke University.
Workshop: Mixtures Meeting organized by the *National Institute of Environmental Health Sciences (NIEHS)* at the Research Triangle Park, NC, USA.

Programming Skills

- R
- Python
- Matlab
- C/C++

Interests

- Academia Machine Learning, Data Science, Cluster Analysis, Metric Learning, Image Processing, Classification, Optimization, Cryptography.
- Personal Singing, Reading Novels, Solving Puzzles.

Professional Activities

- 2021-Present Reviewer at *BMC bioinformatics*
- 2021-Present Reviewer at *Neurocomputing*
- 2021-Present Reviewer at *Engineering Applications of Artificial Intelligence*
- 2021-Present Reviewer at *Indonesian Journal of Electrical Engineering and Computer Science*
- 2020-Present Reviewer at *Pattern Recognition*

Relevant Links

- Google Scholar <https://scholar.google.co.in/citations?user=TDakv1QAAAAJ&hl=en>
- ORCID ID <https://orcid.org/0000-0002-3981-9596>
- Website <https://sites.google.com/view/debolinap8/home>