Shounak Datta

Duke University
Durham, North Carolina
27705 USA

email: shounak.jaduniv@gmail.com
shounak.datta@duke.edu
Phone: (+1)-919-396-1988
URL: https://shounak-d.github.io/
GitHub: https://github.com/Shounak-D

Areas of specialization

Machine Learning * Deep Learning * Imbalanced Classification * Missing Features * Clustering * Constrained Optimization * Multi-objective Optimization * Evolutionary Computation

Research interests & experience

* Working on machine learning, data mining and stochastic optimization for 8+ years now.
* Proficient in data analysis mechanisms like classification, clustering, feature learning, dimensionality reduction, etc.
* Expertise on machine learning with irregularities characterizing real-world data, such as class imbalance, presence of small disjuncts within classes, missing feature values, etc.
* Knowledgeable about convex as well as non-convex optimization techniques (including evolutionary computation).
* Expertise on multi-objective optimization methods.
* Current research interests include development of deep learning algorithms for computer vision, natural language processing, and medical applications.
* Supervised 20+ undergraduate and postgraduate research interns (from various notable academic institutes in India) on projects leading to papers, technical reports, and dissertations since 2013.

Selected research projects

2019-now **Deep learning models for Glaucoma Diagnosis**, Dept. of Electrical and Computer Engineering, Duke University, Durham, NC, USA.

2019-now **Deep natural language processing models for generating syntactically proper sentences**, Dept. of Electrical and Computer Engineering, Duke University, Durham, NC, USA.

2018-2019 **Deep generative network based resampling of data to tackle class imbalance in classification**, Electronics and Communication Sciences Unit, Indian Statistical Institute, Kolkata, India.

2013-2019 **Developing learning algorithms resilient to data irregularities such as class imbalance,**

small disjuncts, missing features, Electronics and Communication Sciences Unit, Indian Statistical Institute, Kolkata, India.

2015-2019 **Developing fuzzy partitional clustering methods which can automatically determine the required level of fuzziness**, Electronics and Communication Sciences Unit, Indian Statistical Institute, Kolkata, India.

2011-2013 **Analysis and predicting of stock market indices**, Department of Electronics and Telecommunication, Jadavpur University, Kolkata, India.

2011-2013 **Developing electrooculogram based control mechanisms for wheelchair to facilitate rehabilitation**, Department of Electronics and Telecommunication jointly with the School of Bioscience and Engineering, Jadavpur University, Kolkata, India.

Positions held

2019-now Postdoctoral Associate, Duke University, Durham NC, USA.

2015-2019 Senior Research Fellow, Indian Statistical Institute, Kolkata, India.

2013-2015 Junior Research Fellow, Indian Statistical Institute, Kolkata, India.

Education

2019 PH.D. in Computer Science, *On the Design of Learning Systems with Resilience to Data Irregularities*, Indian Statistical Institute, Kolkata, India, under the supervision of Prof. (Dr.) Swagatam Das

2013 M.E. in Electronics and Telecommunication Engineering (Specialization: Control Engineering), Jadavpur University, Kolkata, India. CGPA: **9.78**

2011 B.TECH. in Electronics and Communication Engineering, Maulana Abul Kalam Azad University of Technology, Kolkata, India. CGPA: **9.01**

Programming Languages & Tools

Tensorflow, Keras, PyTorch, Python, MATLAB.
Extensive experience of preparing articles and presentations using LaTeX.

Publications

14 journal- and 5 peer-reviewed conference-papers; for more details please visit Google Scholar page at https://scholar.google.co.in/citations?user=qtW4ugoAAAAJ

JOURNAL ARTICLES

2020 "Appropriateness of Performance Indices for Imbalanced Data Classification: An Analysis", Sankha Subhra Mullick, Shounak Datta, Sourish Gunesh Dhekane, Swagtam Das, *Pattern Recognition (Elsevier)*.

2020 "Artificial Intelligence Mapping of Structure to Function in Glaucoma", Eduardo Mariottoni, Shounak Datta, David Dov, Alessandro Jammal, Samuel Berchuck, Ivan Tavares, Lawrence Carin, Felipe Medeiros, *Translational Vision Science and Technology (ARVO)*.

2019 "Boosting with Lexicographic Programming: Addressing Class Imbalance without Cost Tuning", Shounak Datta, Sayak Nag, Swagatam Das, *IEEE Transactions on Knowledge and Data Engineering*.

2019 "Fuzzy Clustering to Identify Clusters at Different Levels of Fuzziness: An Evolutionary Multi-Objective Optimization Approach", Avisek Gupta, Shounak Datta, Swagatam Das, *IEEE Transactions on Cybernetics*.

2018 "Clustering with Missing Features: A Penalized Dissimilarity Measure based approach", Shounak Datta, Supritam Bhattacharjee, Swagatam Das, *Machine Learning*.

2018 "Multi-Objective Support Vector Machines: Handling Class Imbalance with Pareto Optimality", Shounak Datta, Swagatam Das, *IEEE Transactions on Neural Networks and Learning Systems*.

2018 "Fast Automatic Estimation of the Number of Clusters from the Minimum Inter-Center Distance for Center-Based Clustering", Avisek Gupta; Shounak Datta; Swagatam Das, *Pattern Recognition Letters (Elsevier)*.

2018 "Handling data irregularities in classification: Foundations, trends, and future challenges", Swagatam Das, Shounak Datta, Bidyut B. Chaudhuri, *Pattern Recognition* 81, 674-693.

2018 "Adaptive Learning-Based k-Nearest Neighbor Classifiers With Resilience to Class Imbalance", Sankha Subhra Mullick, Shounak Datta, Swagatam Das, *IEEE Transactions on Neural Networks and Learning Systems*.

2017 "Generalized mean based back-propagation of errors for ambiguity resolution", Shounak Datta, Sankha Subhra Mullick, Swagatam Das, *Pattern Recognition Letters* 94, 22-29.

2017 "A Radial Boundary Intersection aided interior point method for multi-objective optimization", Shounak Datta, Abhiroop Ghosh, Krishnendu Sanyal, Swagatam Das, *Information Sciences* 377, 1-16.

2016 "A feature weighted penalty based dissimilarity measure for k-nearest neighbor classification with missing features", Shounak Datta, Debaleena Misra, Swagatam Das, *Pattern Recognition Letters* 80, 231-237.

2015 "Near-Bayesian Support Vector Machines for imbalanced data classification with equal or unequal misclassification costs", Shounak Datta, Swagatam Das, *Neural Networks* 70, 39-52.

2012 "Development strategy of eye movement controlled rehabilitation aid using Electrooculogram", Anwesha Banerjee, Shounak Datta, Amit Konar, D. N. Tibarewala", *International Journal of Scientific and Engineering Research* 3 (6), 1-6.

Conference articles

2019 "Generative Adversarial Minority Oversampling", Sankha Subhra Mullick, Shounak Datta, Swagatam Das, *ICCV 2019*, Accepted.

2015 "Rough-Fuzzy Collaborative Multi-level Image Thresholding: A Differential Evolution Approach", Sujoy Paul, Shounak Datta, Swagatam Das, *MENDEL 2015, Proceedings of*, 329-341.

2013 "Real time electrooculogram driven rehabilitation aid", Anwesha Banerjee, Pratyusha Das, Shounak Datta, Amit Konar, Ramadoss Janarthanan, D. N. Tibarewala", *International Conference on Advances in Computing, Proceedings of the*, 435-440.

2012 "Single channel electrooculogram (EOG) based interface for mobility aid", Anwesha Banerjee, Sumantra Chakraborty, Pratyusha Das, Shounak Datta, Amit Konar, D. N. Tibarewala", *Intelligent Human Computer Interaction (IHCI), Proceedings of the 4th International Connference on*, 1-6.

2012 "Electrooculogram based online control signal generation for wheelchair", Anwesha Banerjee, Shounak Datta, Pratyusha Das, Amit Konar, D. N. Tibarewala, Ramadoss Janarthanan, *Electronic*

System Design (ISED), Proceedings of the International Symposium on, 251-255.

Theses & Dissertations

2018 "On the Design of Learning Systems with Resilience to Data Irregularities", Shounak Datta, under the guidance of Prof. (Dr.) Swagatam Das, *Ph.D. Thesis*, Indian Statistical Institute, Kolkata, India.

2013 "Analysis and prediction of time series indices obtained from stock market indices", Shounak Datta, under the guidance of Prof. (Dr.) Amit Konar, *M.E. Dissertation*, Jadavpur University, Kolkata, India.

Preprints

2017 "Diversifying Support Vector Machines for Boosting using Kernel Perturbation: Applications to Class Imbalance and Small Disjuncts", *communicated to Pattern Recognition Letters (Elsevier)*, Shounak Datta, Sayak Nag, Sankha Subhra Mullick, Swagatam Das, *arXiv* 1712.08493.

Professional activities

Reviewer Duties

2018-now IEEE Transactions on Knowledge and Data Engineering
2018-now IEEE Transactions on Neural Networks and Learning Systems
2017-now Transactions on Knowledge Discovery from Data (ACM)
2017-now Pattern Recognition (Elsevier)
2015-now Information Sciences (Elsevier)
2014-now Engineering Applications of Artificial Intelligence (Elsevier)
2014-2015 IEEE Transactions on Systems, Man, and Cybernetics: Systems
2014-2015 Neurocomputing (Elsevier)
2017 Ninth International Conference on Advances in Pattern Recognition
2015 Eight International Conference on Advances in Pattern Recognition

Grants Obtained

I have participated in the drafting of a successful grant application for the award of an NVIDIA Titan Xp GPU to my Ph.D. supervisor Prof. Swagatam Das.

Talks & lectures

2018 "Data Irregularities in Pattern Classification", *SSCVGIP 2018*, Electronics and Communication Sciences Unit, Indian Statistical Institute, Kolkata.

2014 "Evolutionary Algorithms", *Lectures on Bio-Inspired Computing*, Electronics and Communication Sciences Unit, Indian Statistical Institute, Kolkata.

Personal information & contact details

Born	September 03, 1989, Kolkata, India.
Nationality	Indian
Languages known	English (fluent) * Bengali (mother language) * Hindi (fluent)
Extracurricular interests	Painting, Calligraphy, Micrography, Graphic narratives
Residence	1315 Morreene Road, Apt. 14K, Durham, NC, USA, 27705.