Subhabrata Majumdar

Address:
1112 8th Street SE Apt 7
Minneapolis, MN 55414

Phone: (612) 806-9996
e-mail: zoom.subha@gmail.com
Web: http://stat.umn.edu/~majum010

Research interests

Theory: Multivariate quantiles and statistical depth functions, depth-based robust estimators;

Applications: Statistical computing, machine learning algorithms, large-scale optimizations, high-dimensional models in statistical chemistry, public health and human genomics.

Education

PhD Statistics, University of Minnesota - Twin Cities, Fall 2012 - present (*expected Spring 2017*);
Advisor: Snigdhansu Chatterjee

M.S. Statistics, Indian Statistical Institute, 2012.
(Specialization: Actuarial science and Genetics)
Masters thesis: Adjusting for Treatment Effects in Studies of Quantitative Traits

B.S. Statistics, Indian Statistical Institute, 2010.

Publications

- Basak, S.C. and **Majumdar, S.** Prediction of Mutagenicity of Chemicals from Their Calculated Molecular Descriptors: A Case Study with Structurally Homogeneous versus Diverse Datasets. *Current Computer-Aided Drug Design*, **2015**, *upcoming*;
- With 9 authors. Predictive Modeling for Public Health: Preventing Childhood Lead Poisoning. *Proceedings of the 21st ACM SIGKDD conference on Knowledge Discovery and Data Mining, Sydney, Australia*, **2015**, *upcoming*;
- (Invited) Basak, S.C. and **Majumdar, S.** Editorial: The Importance of Rigorous Statistical Practice in the Current Landscape of QSAR Modelling. *Current Computer-Aided Drug Design*, **2015**, *11*, 2–4;
- Mukherjee, U., **Majumdar, S.** and Chatterjee, S. Fast and Robust Supervised Learning in High Dimensions Using the Geometry of the Data. In: *Advances in Data Mining: Applications and Theoretical Aspects*, ser. *Lecture Notes in Computer Science*, **2015**, *9165*, 109–123;
- **Majumdar, S.** and Basak, S.C. Adapting Interrelated Two-Way Clustering Method for Quantitative Structure-Activity Relationship (QSAR) Modeling of Mutagenicity/ Non-Mutagenicity of a Diverse Set of Chemicals. *Current Computer-Aided Drug Design*, **2013**, *9*, 463–471;
- (Submitted) **Majumdar, S.** and Chatterjee, S. Robust estimation of principal components from depth-based multivariate rank covariance matrix;
- (Submitted) **Majumdar, S**. Copula-based tail dependence networks for multivariate data.
- (Submitted) **Majumdar, S.** and Mallik, A. A fast and accurate denoising technique for high-resolution image streams using data depth-based robust kernel PCA.

Conference presentations

- (Upcoming) ASA Joint Statistical Meetings, Seattle, WA: Aug 2015;
- 20th ACM SIGKDD conference on Knowledge Discovery and Data Mining, New York City, NY: Aug 2014;
- ASA Joint Statistical Meetings, Boston, MA: Aug 2014.

Relevant coursework

Theory of statistics and probability, linear algebra, real and complex analysis, measure theory, stochastic processes;

Regression methods, bayesian methods, design of experiments, machine learning, statistical computing;

Statistics in climate statistics, human genetics, bioinformatics methods, mathematical biology, survival analysis.

Skills

R, Matlab, Stata, Mathematica, Python; C/C++, SQL; HTML, CSS, PHP; LaTeX, Microsoft office.

Major awards

- University of Minnesota School of Statistics travel award, 2014 – 2015;
- KVPY national fellowship, Department of Science and Technology, Govt. of India, 2008 – 2012;
- National scholar, National Council of Educational Research and Training, Govt. of India, 2005 – 2008;
- Best Project award in state-level conference of the Association of Surgeons in India, 2011.

Academic and professional experience

Santander Consumer USA, Dallas, TX.
Data Science intern, May 2015 – Aug 2015.

Data Science for Social Good fellowship, University of Chicago. June 2014 – Aug 2014.
Collaborated with the Chicago Department of Public Health in building a predictive model for lead poisoning prevention in the city of Chicago;

National Marrow Donor Program, Minneapolis, MN.
Statistician Intern, Bioinformatics division, June 2013 – Aug 2013.
Designed a spatial algorithm for data-driven marrow donor recruitment for Leukemia patients with rare alleles;

Educational Initiatives, Bangalore, India.
Summer Intern, June 2011 – July 2011,
Analyzed educational survey data on teacher aptitude collected by four state governments in India as well as govts. of Nepal and Bhutan;

Saha Institute of Nuclear Physics, Kolkata, India.
Undergraduate Research Associate in biophysical sciences, Jan 2008 – June 2010.

Teaching experience

Teaching Assistant at School of Statistics, Univ. of Minnesota, Fall 2012 – Present;

STAT 8051 - Advanced Regression Techniques; Fall 2014;
STAT 3022 - Data Analysis; Spring 2014;
STAT 5021 - Statistical Analysis, STAT 5031 - Statistical Methods for Quality Improvement; Fall 2013;
STAT 5303 - Designing Experiments, STAT 5401 - Applied Multivariate Methods; Spring 2013;
STAT 3011 - Introduction to Statistical Analysis; Fall 2012.

Professional activities

Journal Referring: Current Computer-Aided Drug Design, Australasian Medical Journal;

Statistical Consulting;

Member of the Institute of Mathematical Statistics (IMS).