SARBOJIT ROY

Postdoctoral Research Fellow, Biostatistics Group
Department of Computer, Electrical, and Mathematical Sciences & Engineering(CEMSE)
King Abdullah University of Science and Technology(KAUST), Saudi Arabia
sarbojit.roy@kaust.edu.sa, rsarbojit@gmail.com

RESEARCH INTERESTS

Pattern Recognition, Nonparametric Methods, Time Series Analysis, High-Dimensional Asymptotics, Robust Statistical Methods, Variable Selection, Neuroimaging.

EDUCATION

Ph.D., Statistics, IIT Kanpur, India..	CPI: 8.7/10	2016–2022
– Title: *On Some Modifications of Distance Based Classifiers for High-Dimensional Data.*		
M.Sc., Statistics, IIT Kanpur, India.	CPI: 8.3/10	2012–2014
B.Sc., Statistics, University of Calcutta, India.	CGPA: 7.2/10	2009–2012
Harvard Higher Education Teaching Certificate course.		Apr–Jun, 2023

PUBLICATIONS

- **Sarbojit Roy**, Malik S. Sultan, Tania R. Vallejo, Leena A. Ibrahim, and Hernando Ombao. Classification of high-dimensional time series in spectral domain using explainable features with applications to neuroimaging data. In *Proceedings of The 28th International Conference on Artificial Intelligence and Statistics*, 2025.
- Anass B. El-Yaagoubi, Sipan Aslan, Farah Gomawi, Paolo V. Redondo, **Sarbojit Roy**, Malik S. Sultan, Mara S. Talento, Francine T. Tarrazona, Haibo Wu, Keiland W. Cooper, Norbert J. Fortin, and Hernando Ombao. Methods for Brain Connectivity Analysis with Applications to Rat Local Field Potential Recordings. *Entropy*, 27(4), 2025.
- **Sarbojit Roy**, Soham Sarkar, Subhajit Dutta, and Anil K. Ghosh. On exact feature screening in ultrahigh-dimensional binary classification. *Journal of Computational and Graphical Statistics*, 33, 2024.
- Jyotishka Ray Choudhury, Aytijhya Saha, **Sarbojit Roy**, and Subhajit Dutta. Robust Classification of High-Dimensional Data Using Data-Adaptive Energy Distance. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 86–101. Springer, 2023.
- **Sarbojit Roy**, Jyotishka Ray Choudhury, and Subhajit Dutta. On some fast and robust classifiers for high dimension, low sample size data. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, pages 9943–9968., 2022.
- **Sarbojit Roy**, Soham Sarkar, Subhajit Dutta, and Anil K Ghosh. On generalizations of some distance based classifiers for HDLSS data. *Journal of Machine Learning Research*, 23(14):1–41., 2022.

PREPRINTS

- Mara S. Talento, **Sarbojit Roy**, and Hernando Ombao. KenCoh: A Rank-based Canonical Coherence, 2024 [arXiv:2412.10521].

MANUSCRIPTS UNDER PREPARATION

- **Sarbojit Roy**, Suhasini Subba Rao, and Hernando Ombao. On classification of high-dimensional time series based on Whittle approximation of likelihood.
- **Sarbojit Roy**, Mara S. Talento, and Hernando Ombao. Robust canonical analysis of time series under elliptic symmetry.
- **Sarbojit Roy** and Debraj Das. A two-sample test for equality of distributions in ultrahigh dimensions.
- Mahesh Halder, **Sarbojit Roy**, and Avijit Saha. A statistical approach to discriminate the crust versus mantle sources and their contribution to the origin of silicic rocks in igneous provinces.

SEMINAR PRESENTATIONS

- *'Interpretable Classification of High-dimensional Time Series in Spectral Domain'*
  – International Indian Statistical Association Conference, Kochin, India. Dec, 2024
- *'Classification of High-dimensional Time Series in Spectral Domain Using Explainable Features'*
  – Research Presentation Series, KAUST ASA Student Chapter. Oct, 2024
- *'Robust Classification of High-Dimensional Data using Data-Adaptive Energy Distance'*
  – European Conference on Machine Learning and Data Mining, Torino, Italy. Sep, 2023
- *'On the use of inverse spectral density matrices for classifying high dimensional time series'*
  – Joint Statistical Meetings, Toronto, Canada. Aug, 2023
- *'On some fast and robust classifiers for high dimension, low sample size data'*
  – 25th Artificial Intelligence and Statistics Conference, Valencia, Spain. Mar, 2022
- *'Adjustments of the nearest neighbor classifier in HDLSS settings'*
  – Innovations in Data and Statistical Sciences, Mumbai, India. Dec, 2019
  – 62nd World Statistics Congress, Kuala Lumpur, Malaysia. Aug, 2019
- *'A revisit to Tukey's ninther method of finding a low effort robust estimator of location parameter in large sample'*
  – 100th Indian Science Congress, Kolkata, India. Jan, 2013

SCHOOLS/WORKSHOPS ATTENDED

- Statistics for Learning from Complex Data: Foundations and Applications, KAUST. Nov 17-21, 2024
- Frontier Statistics and Data Science for a Sustainable World, KAUST. Nov 06-09, 2023
- Indo-French Centre for Applied Mathematics (IFCAM) Winter School 2018 held at Indian Institute of Scientific Education and Research Kolkata. Dec 3-14, 2018
- 'Advanced Techniques and Methodologies of Big Data in Science and Engineering' offered by the Global Initiative for Academic Networks (GIAN) held at IIT Kanpur. Feb 3-7, 2018

AWARDS/FELLOWSHIPS

- Travel support given by the Indian Association for Statistics in Clinical Trials to attend the conference Innovations in Data and Statistical Sciences in Mumbai, India. Dec, 2019
- The Asian Development Bank (ADB) funding for participants from ADB developing member economies to attend 62nd World Science Congress in Kuala Lumpur, Malaysia. Aug, 2019
- Institute fellowship during Ph.D. by IIT Kanpur, India. July, 2016-May, 2022
- Inspire scholarship by Department of Science & Technology (DST), Government of India during B.Sc. and M.Sc. in Statistics. Jun, 2009-May, 2014

SUPERVISION

I have supervised or am currently supervising, along with Professor Hernando Ombao, the following MS/PhD students for their research projects, at KAUST, Saudi Arabia.

- 2023-2024 Malik Shahid Sultan (MS)
- 2024-present Mara Sherlin Talento (PhD)

REVIEWER

I have reviewed articles for the journals– Annals of the Institute of Statistical Mathematics (AISM), Data Science in Science, IEEE Transactions on Medical Imaging, Journal of the American Statistical Association (JASA), Journal of Time Series Analysis (JTSA), and Scientific Reports.

RESEARCH VISITS

Department of Mathematics, University of Hamburg, Germany.	Feb 04-06, 2025
Institute for Mathematical Stochastics, Otto von Guricke University Magdeburg	Jan 28- Feb 03, 2025
Department of Bio-Science and Bio-Engineering, IIT Kanpur	Dec 18- 20, 2024

TEACHING ASSISTANCE

Time Series Analysis with Prof. Hernando Ombao.	Spring 2023
Introduction to Bayesian Analysis with Prof. Debasis Kundu.	2nd semester 2020-2021
Applied Stochastic Process with Prof. Debasis Kundu.	1st semester 2020-2021
Probability and Statistics with Prof. Neeraj Misra.	2nd semester 2019-2020
Introduction to Complex Analysis with Prof. Kaushik Bal.	1st semester 2019-2020
Partial Differential Equations with Prof. T Muthukumar.	1st semester 2019-2020
Non-Linear Regression with Prof. Debasis Kundu.	2nd semester 2018-2019
Applied Stochastic Process with Prof. Debasis Kundu.	1st semester 2018-2019
Probability Theory with Prof. Subhajit Dutta.	2nd semester 2017-2018
Introduction to Complex Analysis with Prof. Rama Rawat.	1st semester 2017-2018

WORK EXPERIENCES

Post-doctoral Researcher — Jan, 2023 - present

Biostatistics Group, CEMSE, KAUST.

Visiting Scientist — Nov - Dec, 2022

Worked at Theoretical Statistics and Mathematics Unit, Indian Statistical Institute (ISI) Delhi.

Project Research Assistant — Jun - Nov, 2022

Worked on a project titled 'Bootstrapping Regression Models' (funded by DST) at the Department of Mathematics at IIT Bombay under supervision of Prof. Debraj Das.

Project Linked Person — Oct, 2015 - Mar, 2016

Worked on a project titled 'Estimation of demand for banknotes and coins' (funded by the Reserve Bank of India) at the Sampling and Official Statistics Unit at ISI Kolkata.

Econometrician at Bravo Lucy AS — Jul, 2014 - May, 2015

– Developed a probabilistic model to provide the odds for tailor-made bets in football matches.
– Worked on video analytics-driven business solutions.

Summer Intern — May - July, 2013

Worked on a project titled 'Analysis of integrated complaint management system data' at the Institute for Development and Research in Banking Technology (IDRBT), Hyderabad.

TECHNICAL EXPERTISE

Statistical software: `R` (experienced), `python` (beginner)

Documentation: LaTeX