



An introduction to robust data analysis and its applications

Panos M. Pardalos¹ · Hossein Moosaei² · Milan Hladík³ · M. Tanveer⁴

Published online: 26 July 2024
© The Author(s) 2024

In the era of burgeoning data science, the application of robust methodologies for data analysis has become paramount across various crucial domains, including healthcare systems, financial predictions, and beyond. Real-world data challenges, characterized by uncertainty, voluminous databases, and high-dimensional samples, necessitate robust techniques that not only handle such complexities efficiently but also yield reliable models. The convergence of robust optimization and data analysis paves the way for the development of sophisticated and innovative solutions to tackle contemporary problems.

In the dynamic landscape of data analysis, robust methodologies have emerged as indispensable tools for navigating the uncertainties inherent in real-world datasets. This special issue of the *Annals of Operations Research* serves as a testament to the growing significance of robust data analysis and its multifaceted applications across diverse domains.

The compilation of papers within this issue showcases the breadth and depth of research in the field, covering a spectrum of topics ranging from robust optimization to the application of machine learning techniques in handling uncertainty. These contributions not only

✉ Hossein Moosaei
hossein.moosaei@ujep.cz

Panos M. Pardalos
pardalos@ise.ufl.edu

Milan Hladík
hladik@kam.mff.cuni.cz

M. Tanveer
mtanveer@iiti.ac.in

¹ Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA

² Department of Informatics, Faculty of Science, Jan Evangelista Purkyně University, Ústí nad Labem, Czech Republic

³ Department of Applied Mathematics, School of Computer Science, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

⁴ Department of Mathematics, Indian Institute of Technology Indore, Simrol, Indore 453552, Madhya Pradesh, India

advance the theoretical foundations of robust data analysis but also offer practical insights that can inform decision-making processes in various industries.

The journey through this special issue begins with critical examination of uncertainty in medical data, as articulated in the paper “Handling of uncertainty in medical data using machine learning and probability theory techniques: a review of 30 years (1991–2020)” by Alizadehsani, Roshanzamir, Hussain, Khosravi, Koohestani, Zangoeei, Abdar, Beykikhoshk, Shoeibi, Zare, Panahiazar, Nahavandi, Srinivasan, Atiya, and Acharya. This comprehensive review delves into the intersection of machine learning and probability theory, spanning three decades of research, and elucidates methodologies for effectively managing uncertainty inherent in medical data—an essential endeavour for advancing healthcare analytics and decision-making.

The voyage continues with a focus on optimization techniques, with papers such as “The stochastic multi-gradient algorithm for multi-objective optimization and its application to supervised machine learning” by Liu and Vicente, providing innovative approaches to address complex optimization problems.

Moving forward, the exploration extends to performance assessment methodologies, as demonstrated in “A novel robust network data envelopment analysis approach for performance assessment of mutual funds under uncertainty” by Peykani, Emrouznejad, Mohammadi, and Gheidar-Kheljani. This paper exemplifies how robust data envelopment analysis can enhance decision-making processes in financial settings amid uncertain market conditions.

The intersection of robust data analysis and healthcare is prominently featured, with studies such as “Infection vulnerability stratification risk modelling of COVID-19 data: a deterministic SEIR epidemic model analysis” by Kumar, Choi, Wamba, Gupta, and Tan, shedding light on the application of deterministic epidemic models to mitigate the impact of infectious diseases.

Additionally, in “Robust optimization approaches for portfolio selection: a comparative analysis” Georgantas, Doumpos, and Zopounidis conduct a comprehensive evaluation of robust optimization models for portfolio selection using US market data from 2005 to 2020. They compare different robust optimization models, including robust versions of mean-variance models, conditional value-at-risk, and the Omega ratio, to offer insights into their efficacy through various portfolio performance metrics.

In 2007, Jayadeva et al. introduced a novel machine learning approach known as the Twin Support Vector Machine (TWSVM), designed for binary classification and inspired by GEPSVM (Mangasarian & Wild, 2005). Owing to its outstanding performance, TWSVM has found applications in diverse domains such as speaker recognition, medical detection, and more, attracting attention from both theoretical and practical perspectives (Wang et al., 2016; Moosaei et al., 2021; Moosaei & Hladik, 2023; Moosaei et al., 2024). Consequently, an overview of these advancements becomes essential. Therefore, the paper “Comprehensive review on twin support vector machines” authored by Tanveer, Rajani, Rastogi, Shao, and Ganaie offers valuable insights into these methods and their extensions, providing a comprehensive understanding of TWSVM and its diverse applications.

As the reader delves deeper into the issue, they encounter a diverse array of topics, including pattern classification, efficiency evaluation, supply chain management, and beyond. Each paper not only contributes to the expanding body of knowledge in its respec-

tive area but also underscores the versatility and relevance of robust data analysis across various domains.

We extend our heartfelt gratitude to Editor-in-Chief Professor Endre Boros, as well as Publications Manager Ann Pulido, for their invaluable support in curating this special issue. We also express our sincere appreciation to all the authors for their invaluable contributions, and to the reviewers for their meticulous evaluation and insightful feedback.

In conclusion, this special issue of the *Annals of Operations Research* showcases a collection of 25 meticulously selected papers, each addressing significant challenges and presenting innovative solutions. It stands as a testament to the enduring significance of robust data analysis in addressing the complexities and uncertainties of the modern world. It is our hope that the insights gleaned from these contributions will inspire further advancements in the field and foster collaborations that transcend disciplinary boundaries.

Funding Open access publishing supported by the National Technical Library in Prague.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Jayadeva, Khemchandani, R., & Chandra, S. (2007). Twin support vector machines for pattern classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(5), 905–910.
- Mangasarian, O. L., & Wild, E. W. (2005). Multisurface proximal support vector machine classification via generalized eigenvalues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1), 69–74.
- Moosaei, H., & Hladík, M. (2023). Sparse solution of least-squares twin multi-class support vector machine using ℓ_0 and ℓ_p -norm for classification and feature selection. *Neural Networks*, 166, 471–486.
- Moosaei, H., Ketabchi, S., Razzaghi, S., & Tanveer, M. (2021). Generalized twin support vector machines. *Neural Processing Letters*, 53(2), 1545–1564.
- Moosaei, H., Bazikar, F., & Hladík, M. (2024). Multi-task twin support vector machine with Universum data. *Engineering Applications of Artificial Intelligence*, 132, 107951.
- Wang, S., Lu, S., Dong, Z., Yang, J., Yang, M., & Zhang, Y. (2016). Dual-tree complex wavelet transform and twin support vector machine for pathological brain detection. *Applied Sciences*, 6, 6: 169.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.