

**WISDOM presents YoungWomen4OR in:
“Data Science and Optimization”
Join us for a coffee and a chat!**

What: EURO WISDOM Forum YoungWomen4OR Talks¹

Where: Zoom - Register in this [Google Form](#) to receive the Zoom link - The webinar is going to be recorded and made available afterwards

When: 24th March 2022, 15:00 – 16:30 (Central European Time)

Webinar Format

- Introductions/Webinar etiquette –
Prof. Dr. Dolores Romero Morales, *Copenhagen Business School, Denmark.*
- Data Science and Optimization– YoungWomen4OR Talks:
 - *On fair and explainable optimal regression trees*, Cristina Molero-Río. - 10 minutes
 - *Dealing with complex data: a tree-based linear regression model for hierarchical categorical variables*, M. Remedios Sillero-Denamiel. - 10 minutes
 - *New models and methods for data science in a nutshell*, Sandra Benítez-Peña. - 10 minutes
- Meeting the challenges - Overview/Current Challenges, synergies with existing work
 - Prof. Dr. Ilker Birbil, *University of Amsterdam, Netherlands.* - 10/15 minutes
- Moderated open discussion with Coffee and Networking – 15 minutes

YoungWomen4OR Speakers



Cristina Molero-Río, *PhD student at the University of Seville, Spain.*

Title: On fair and explainable optimal regression trees.

Abstract: In this talk, we model an optimal regression tree through a continuous optimization problem, where a compromise between prediction accuracy and sparsity is sought. Our approach can easily accommodate other desirable properties for the regression task, such as fairness, by avoiding the discrimination of groups that share sensitive features; and local explainability, by providing in a natural way the impact that continuous predictor variables have on each individual prediction; as well as handle complex data including functional data. We illustrate the performance of our approach on real-world datasets.

See <https://scholar.google.com/citations?user=rUU3M38AAAAJ&hl=en>.

¹ WISDOM is a forum to support, empower, and encourage the participation of all genders in Operational Research and Management Science. It is an initiative supported by EURO, the Association of European Operational Research Societies. Please visit: <https://www.euro-online.org/web/pages/1654/wisdom>



Dr. M. Remedios Sillero-Denamiel, *Postdoctoral Fellow at School of Computer Science and Statistics, Trinity College Dublin, Ireland.*

Title: Dealing with complex data: a tree-based linear regression model for hierarchical categorical variables.

Abstract: Complex data stands for datasets with many variables, dependence structure among the variables, non-identical population class sizes or different misclassification costs, among others. Specifically, many real-life applications consider nominal categorical predictor variables that have a hierarchical structure, e.g. economic activity data in Official Statistics. This talk focuses on linear regression models built in the presence of this type of nominal categorical predictor variables, and studies the consolidation of their categories to have a better tradeoff

between interpretability and fit of the model to the data. The so-called Tree based Linear Regression (TLR) model is presented, which optimizes both the accuracy of the reduced linear regression model and its complexity, measured as a cost function of the level of granularity of the representation of the hierarchical categorical variables.

See: <https://www.scss.tcd.ie/personnel/sillero>



Dr. Sandra Benítez-Peña, *Postdoctoral Fellow at University Carlos III of Madrid, Spain.*

Title: New models and methods for data science in a nutshell.

Abstract: In this talk we present a brief review of new models and methods we have developed to face different data science challenges, focusing particularly in the tasks of Classification, Regression and Benchmarking. With such a purpose, we make use of Optimization and Statistical tools. In order to address the different problems and be able to obtain knowledge from data, we combine instruments from both disciplines. Regarding the task of Classification, we have modelled novel Support Vector Machines (SVM) classifiers capable of managing with different misclassification costs, in which a Feature Selection (FS) can be

embedded and that provides probabilistic outputs. From the Regression perspective, we handle specifically the problem of data linkage, since information can be obtained from many sources that contain the same individuals but there is a lack of information about how to merge them. Finally, we move to Benchmarking, in which the main objective is to compare different entities through an efficiency score. Here, our aim is to perform FS, improving interpretability and comprehension of the obtained model and efficiencies. The reported numerical experience on real-world datasets demonstrates that our models overcome those existing ones.

See: <https://www.researchgate.net/profile/Sandra-Benitez-Pena>.

Subject matter expert.



Prof. Dr. Ilker Birbil, University of Amsterdam, Netherlands.

Ilker Birbil is a professor in AI & Optimization Techniques for Business & Society at University of Amsterdam, where he is currently the head of Business Analytics Section. He received his PhD in 2002 from North Carolina State University with a major in Industrial Engineering, and minors in Operations Research and Mathematics. He was a postdoctoral research fellow at the Erasmus Research Institute of Management for two years. From 2004 to 2018, he was a faculty member at Sabanci University in Istanbul, where he was one of the founders of the Data Analytics professional degree program. He held the Chair in Data Science and Optimization at Erasmus University Rotterdam from 2018 to 2021. With his colleagues and students, he authored research articles on

mathematical optimization, operations management, and data science. These articles appeared in leading journals such as, Mathematical Programming, SIAM Journal on Optimization, Mathematical Methods of Operations Research, Management Science, European Journal on Operational Research, and Transportation Science. He completed various research projects as the principal investigator and participated in several initiatives as a researcher or as a committee member. He also collaborated with industry to work on business problems ranging from cloud computing to airline crew planning. His research interests center around optimization methods in data science and decision making. Lately, he is interested in interpretable machine learning and data privacy in operations research.

See: <https://www.uva.nl/en/profile/b/i/s.i.birbil/s.i.birbil.html?cb>