

Andrea Cappozzo

Curriculum Vitae
February 2021

📍 Department of Statistics & Quantitative Methods,
University of Milano-Bicocca, Italy
🏠 andreacappozzo.rbind.io
✉ andrea.cappozzo@unimib.it
🐦 @AndreaCappozzo
🌐 AndreaCappozzo

Research Interests

Mixture models, Robust statistics, Model-based clustering and classification, Variable selection, Statistical learning

Work Experience

2020/2 – Pres. **Postdoctoral researcher**, Department of Statistics & Quantitative Methods, University of Milano-Bicocca
2017/9 – Pres. **Volunteer data scientist**, Heartindata, Milan
2020/1 – 2020/4 **Freelance Data Scientist**, DCG, Milan
2015/9 – 2016/9 **Business analyst and planner**, HP Inc, Barcelona

Teaching Experience

2019/3 – Pres. **Teaching Assistant**, BSc courses in Statistics and Statistical Methods, University of Milano-Bicocca
2019/3 – 2019/9 **Academic Tutor**, BSc courses in Statistical Models and Computer Science, University of Milano-Bicocca
2017/9 – 2018/2 **Teaching Assistant**, BSc courses in Statistics, Bocconi University

Education

2020/2 **Ph.D. in Statistics and Mathematical Finance** (Doctor Europaeus) University of Milano-Bicocca
2015/4 **M.Sc. in Statistical Sciences** (with honors) University of Padua
2012/7 **B.Sc. in Statistics and Management** (with honors) University of Padua

Visiting Periods

2018/3 – 2019/2 **Visiting PhD student**, Insight Centre for Data Analytics, University College Dublin, Ireland
2014/1 – 2014/6 **Exchange Semester**, School of Economics and Management, Tilburg University, The Netherlands

Publications

Refereed journals

1. Cappozzo, A, F Greselin, and TB Murphy (2021). Robust variable selection for model-based learning in presence of adulteration. *Computational Statistics & Data Analysis* **158**, 107186.
2. Cappozzo, A, L Duponchel, F Greselin, and TB Murphy (2021). Robust variable selection in the framework of classification with label noise and outliers: applications to spectroscopic data in agri-food. *Analytica Chimica Acta*, 338245.
3. Cappozzo, A, F Greselin, and TB Murphy (2020). Anomaly and Novelty detection for robust semi-supervised learning. *Statistics and Computing* **30**(5), 1545–1571.
4. Cappozzo, A, F Greselin, and TB Murphy (2020). A robust approach to model-based classification based on trimming and constraints. *Advances in Data Analysis and Classification* **14**(2), 327–354.

Submitted and working papers

1. Cappozzo, A, F Greselin, and TB Murphy (2020). Robust model-based learning to discover new wheat varieties and discriminate adulterated kernels in X-ray images.
(Just accepted In: *Statistical Learning and Modeling in Data Analysis - Methods and Applications*)
2. Denti, F, A Cappozzo, and F Greselin (2020+). A Two-Stage Bayesian Nonparametric Model for Novelty Detection with Robust Prior Information (Revision submitted) <https://arxiv.org/abs/2006.09012>

Monographs and refereed conference proceedings

1. Cappozzo, A, F Greselin, and TB Murphy (2020). "Variable selection for robust model-based learning from contaminated data". In: *Book of Short Papers SIS 2020*. Pearson, pp. 1117–1122.
2. Cappozzo, A and F Greselin (2019). "Detecting wine adulterations employing robust mixture of factor analyzers". In: *Statistical Learning of Complex Data*. Springer Berlin Heidelberg, pp. 13–21.
3. Cappozzo, A, F Greselin, and G Manzi (2019). "Predicting and improving smart mobility: a robust model-based approach to the BikeMi BSS". In: *Smart Statistics for Smart Applications 2019 - Book of Short papers*. Milano: Pearson, pp. 737–742.
4. Cappozzo, A, F Greselin, and TB Murphy (2019). "Supervised learning in presence of outliers, label noise and unobserved classes". In: *Book of short papers | Cladag2019*. Centro Editoriale di Ateneo Università di Cassino e del Lazio Meridionale, pp. 104–107.

5. Cappozzo, A, F Ferraccioli, M Stefanucci, and P Secchi (2018). "An object oriented approach to multimodal imaging data in neuroscience". In: *Studies in Neural Data Science : Startup Research 2017, Siena, Italy, June 25-27*. Vol. 257. Springer New York LLC, pp. 57–73.
6. Cappozzo, A, F Greselin, and TB Murphy (2018). "The role of trimming and variable selection in robust model-based classification for food authenticity studies". In: *COMPSTAT 2018 Book of Abstracts*. COMPSTAT and CRoNoS, pp. 35–35.
7. Greselin, F and A Cappozzo (2017). "Wine authenticity assessed via trimming". In: *Book of short papers | Cladag2017*. Mantova: Universitas Studiorum Srl.

Presentations

Invited

- 2019/9 **Scientific meeting CLADAG 2019**, University of Cassino and Southern Lazio, Italy
 2018/8 **COMPSTAT 2018**, Unirea Hotel, Iasi, Romania

Seminars

- 2020/9 **IR-group meeting**, virtual seminar
 2019/12 **Mock talk II**, University of Milano-Bicocca, Italy
 2019/7 **Mock talk I**, University of Milano-Bicocca, Italy
 2018/10 **Working Group on Statistical Learning**, University College Dublin, Ireland

Contributed

- 2021/2 **e-CHIMIOMETRIE 2021**, virtual conference
 2020/9 **MBC² Workshop on Models and Learning for Clustering and Classification**, virtual conference
 2020/6 **e-Rum 2020**, virtual conference
 2019/12 **CMStatistics 2019**, Senate House University of London, United Kingdom
 2019/6 **Smart Statistics for Smart Applications - SIS conference**, Catholic University of the Sacred Heart, Italy
 2017/9 **Scientific meeting CLADAG 2017**, University of Milano-Bicocca, Italy

Poster sessions

- 2021/1 **DS³ Data science summer school**, virtual conference
 2020/7 **Young-ISA Twitter Poster Conference**, virtual conference
 2019/7 **Working Group on Model-Based Clustering**, Wirtschaftsuniversität Wien, Austria
 2018/9 **MBC² Workshop on Model-Based Clustering and Classification**, University of Catania, Italy

Summer Schools and Workshops

- 2018/4 **Robust Statistics: Foundations and Recent Developments**, University of Milano-Bicocca, Italy
 2017/11 **Introduction to Functional Data Analysis**, Università degli Studi di Bergamo, Italy
 2017/6 **Startup Research**, Certosa di Pontignano, Italy
 2017/5 **International Summer School on Classification and Data Analysis**, University of Bologna, Italy

Awards

- 2018/9 **Best poster presentation**, MBC² Workshop on Model-Based Clustering and Classification, Catania
 2017/9 **Member of the third winning team**, Young CLADAG - Data science competition, Politecnico di Milano
 2017/6 **Member of one of the four winning teams**, Stats Under the Stars 3 - Data science competition, University of Florence

Service to profession

Referee service

Advances in Data Analysis and Classification, Computational Statistics and Data Analysis, Statistics and Computing, Statistical Methods & Applications.

Membership

Institute of Mathematical Statistics, Italian Statistical Society.

Computer skills

R (advanced), latex (advanced), markdown (advanced), bash (intermediate), C++ (intermediate), git (intermediate), python (intermediate), mathematica (basic), matlab (basic), SQL (basic).

Languages

Italian (mother tongue), English (proficient), Spanish (good knowledge).