

Andrea Cappozzo

Curriculum Vitae
December 2021

📍 Department of Mathematics,
Politecnico di Milano, Italy
🏠 andreacappozzo.rbind.io
✉ andrea.cappozzo@polimi.it
🐦 @AndreaCappozzo
🌐 AndreaCappozzo

Research Interests

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

Work Experience

2021/4 – Pres. **Assistant Professor** (Ricercatore SECS-S/01, tipo A), Department of Mathematics, Politecnico di Milano
2017/9 – Pres. **Volunteer data scientist**, Heartindata, Milan
2020/2 – 2021/3 **Postdoctoral researcher**, Department of Statistics & Quantitative Methods, University of Milano-Bicocca
2020/1 – 2020/4 **Freelance Data Scientist**, DCG, Milan
2015/9 – 2016/9 **Business analyst and planner**, HP Inc, Barcelona

Teaching Experience

2022/2 – Pres. **Lecturer**, MSc course in Biostatistics, Politecnico di Milano
2021/9 – Pres. **Teaching assistant**, MSc course in Nonparametric statistics, Politecnico di Milano
2021/9 – 2021/9 **Lab with R**, Mixture models module, PhD in Economics, Statistics and Data Science, University of Milano-Bicocca
2021/5 – 2021/5 **Lab with R**, Modelos de Mixtura con aplicaciones en R, Universidad Nacional de San Agustín de Arequipa
2019/3 – 2021/9 **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
2018/8	Deep Learning A-Z: Hands-on Artificial Neural Networks	udemy.com
2018/4	Python for Data Science and Machine Learning Bootcamp	udemy.com
2017/2	Statistical Learning online course (completed with distinction)	Stanford University
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, García Escudero, LAG, Greselin, F, & Mayo-Isar, A. (2021). Parameter Choice, Stability and Validity for Robust Cluster Weighted Modeling. *Stats*, 4(3), 602–615. <https://doi.org/10.3390/stats4030036>
2. Denti, F, Cappozzo, A, & Greselin, F. (2021). A two-stage Bayesian semiparametric model for novelty detection with robust prior information. *Statistics and Computing*, 31(4), 42. <https://doi.org/10.1007/s11222-021-10017-7>
3. Cappozzo, A, Greselin, F, & Murphy, TB. (2021). Robust variable selection for model-based learning in presence of adulteration. *Computational Statistics & Data Analysis*, 158, 107186. <https://doi.org/10.1016/j.csda.2021.107186>
4. Cappozzo, A, Duponchel, L, Greselin, F, & Murphy, TB. (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. <https://doi.org/10.1016/j.aca.2021.338245>
5. Cappozzo, A, Greselin, F, & Murphy, TB. (2020). Anomaly and Novelty detection for robust semi-supervised learning. *Statistics and Computing*, 30(5), 1545–1571. <https://doi.org/10.1007/s11222-020-09959-1>
6. Cappozzo, A, Greselin, F, & Murphy, TB. (2020). A robust approach to model-based classification based on trimming and constraints. *Advances in Data Analysis and Classification*, 14(2), 327–354. <https://doi.org/10.1007/s11634-019-00371-w>

Submitted and working papers

1. Cappozzo, A, Ieva, F, & Fiorito, G. (2021+). A general framework for penalized mixed-effects multitask learning with application on DNAm biomarkers creation
(Submitted) <https://arxiv.org/abs/2112.12719>
2. Cappozzo, A, Casa, A, & Fop, M. (2021+). Group-wise shrinkage for multiclass Gaussian Graphical Models
(Submitted) <https://arxiv.org/abs/2105.07935>

Monographs and refereed conference proceedings

1. Cappozzo, A, Casa, A, & Fop, M. (2021). Penalized model-based clustering for three-way data structures. In *Book of short papers SIS 2021* (pp. 758–763). Pearson.
2. Cappozzo, A, & Greselin, F. (2021). Monitoring tools for robust estimation of cluster weighted models. In *Book of short papers SIS 2021* (pp. 1245–1250). Pearson.
3. Cappozzo, A, Greselin, F, & Murphy, TB. (2021). Robust model-based learning to discover new wheat varieties and discriminate adulterated kernels in x-ray images. In S Balzano, GC Porzio, R Salvatore, D Vistocco, & M Vichi (Eds.), *Statistical learning and modeling in data analysis* (pp. 29–36). Springer International Publishing. https://doi.org/10.1007/978-3-030-69944-4_4
4. Cappozzo, A, Greselin, F, & Murphy, TB. (2020). Variable selection for robust model-based learning from contaminated data. In *Book of short papers SIS 2020* (pp. 1117–1122). Pearson.
5. Denti, F, Cappozzo, A, & Greselin, F. (2020). Bayesian nonparametric adaptive classification with robust prior information. In *Book of short papers SIS 2020* (pp. 655–660). Pearson.
6. Cappozzo, A, & Greselin, F. (2019). Detecting wine adulterations employing robust mixture of factor analyzers. In *Statistical learning of complex data* (pp. 13–21). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-030-21140-0_2
7. Cappozzo, A, Greselin, F, & Manzi, G. (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* (pp. 737–742). Pearson.
8. Cappozzo, A, Greselin, F, & Murphy, TB. (2019). Supervised learning in presence of outliers, label noise and unobserved classes. In *Book of short papers | Cladag2019* (pp. 104–107). Centro Editoriale di Ateneo Università di Cassino e del Lazio Meridionale.
9. Cappozzo, A, Ferraccioli, F, Stefanucci, M, & Secchi, P. (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, pp. 57–73). Springer New York LLC. https://doi.org/10.1007/978-3-030-00039-4_4
10. Cappozzo, A, Greselin, F, & Murphy, TB. (2018). The role of trimming and variable selection in robust model-based classification for food authenticity studies. In *COMPSTAT 2018 book of abstracts* (pp. 35–35). COMPSTAT; CRoNoS.
11. Greselin, Francesca, & Cappozzo, A. (2017). Wine authenticity assessed via trimming. In *Book of short papers | Cladag2017*. Universitas Studiorum Srl.

Presentations

Invited

- 2021/9 **22nd European Young Statisticians Meeting**, virtual conference
 2019/9 **Scientific meeting CLADAG 2019**, University of Cassino and Southern Lazio, Italy
 2018/8 **COMPSTAT 2018**, Unirea Hotel, Iasi, Romania

Seminars

- 2021/5 **Mox seminar series**, Politecnico di Milano, Italy
 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

- 2021/9 **Scientific meeting CLADAG 2021**, virtual conference
 2021/6 **SIS 2021- 50th Meeting of the Italian statistical society**, virtual conference
 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/software sessions

- 2021/10 **Working Group on Model-Based Clustering**, Athens, Greece
- 2021/4 **International workshop on Spectroscopy and Chemometrics**, virtual conference
- 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, CLAssification and Data Analysis Group (CLADAG)

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).