Roberta Siciliano's curriculum vitae (short description)

Full Professor of Statistics
Laboratory and Research Group **Stad**
***S*tatistics, *t*echnology, *a*nalysis of *d*ata**
Department of Industrial Engineering
University of Naples Federico II
P.le Tecchio, 80 - Napoli
80126 Napoli
Tel. +39-081-675120
E-mail: roberta@unina.it
Mob +39-3209225895

http://wpage.unina.it/roberta http://www.stad.unina.it http://www.docenti.unina.it

Roberta Siciliano was born in Naples on May 12, 1964. She received the Laurea degree in Economics and the PhD in Computational Statistics and Data Analysis from the University of Naples where she is currently full professor of Statistics at the Department of Industrial Engineering, University of Naples Federico II. She is member of *ICAROS (Interdepartmental Center for Advances in Robotic Surgery)* of the University of Naples Federico II, of the Scientific Board of *L.U.P.T. (Centro Interdipartimentale di Ricerca "Laboratorio di Urbanistica e di Pianificazione del Territorio)*, member of the *Board for the PhD in Mathematics and Computer Sciences* and member of the Scientific Board of the *Quality Center* of University of Naples Federico II. She is in the Faculty Board of IPE (Istituto per le Attività Educative e la Ricerca) of Naples. She was member of the Scientific Committee of ISTAT for Big Data (2012-2015). She is *mentore* of the Laboratory and Research Group *STAD (Statistics, Technology, Analysis of Data*).

Her research activity is in the field of *Statistical Learning and Multivariate Data Analysis*. She has published more than sixty papers on international journals, *encyclopaedia*, conference proceedings, and revised monographs. Main topics include *Partitioning Algorithms and Tree-based decision methods for Classification and Regression, Non Parametric Regression, Data Editing, Categorical Data Modeling, Web Mining*.

She is in the *Editorial Board of Journal of Classification*. She acts as referee of many international journals (*Computational Statistics and Data Analysis, Journal of Royal Statistical Society, Journal of Probability and Statistics, Pattern Recognition, Journal of Biostatistics,* etc.).

She is ordinary member of ISI, IASC, IFCS, SIS, CLADAG. She has served as chair and scientific/organizing member committee for numerous international conferences, including Co-Chair of the Scientific and Organizing Committe at *the 68th Meeting of the Psychometric Society* (Chia, July 7-10, 2003), member of the Scientific Committee at *the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Edmonton, Alberta in Canada, July 23-26, 2000), Area Chair of Statistics for *the Fourth International Sympoium on Intelligent Data Analysis* (Lisbona, September 13-15, 2001*),* member of the Scientific Committee of *EURISBIS 2009 - European Regional Meeting of the International Society for Business and Industrial Statistics"* (University of Cagliari, May 30 - June 3, 2009), member of the Programme Committe of *the Annual International Conferences on Intelligent Data Analysis.* Actually, she is in the Programme Committee of CLADAG2017.

She was keynote lecturer on "The Speed-Interpretability-Accuracy Trade-Off in Decision Tree Learning", *European Conference on Data Analysis 2014* (Jacobs University Bremen, July 2- 4, 2014, keynote lecturer on "Exploratory versus Decision Trees" at *COMPSTAT 1998* (Bristol), invited lecturer on "Multivariate Data Analysis through Classification and Regression Trees" at *the Second World Conference of the International Association for Statistical Computing: "Computational Statistics and Data Analysis on the Eve of the 21st Century"* (Pasadena, February 19-22, 1997).

She was lecturer in "Probability models and statistical inference" and in "Multivariate Statistics" for the Course "HDMCuRF Highway Design and Management: Curricular Reform for Russian Federation Design and Implementation of Higher Education Master Courses in Russia", TEMPUS IV - Project n. 516888 from 11-06-2012 to 22-06-2012.

She has been awarded numerous grants for her statisticians group.

She is Scientific Responsabile of the department unit for HORIZON2020 - MAGIC Project (*Moving Towards Adaptive Governance in Complexity: Informing Nexus Security* – https://magic-nexus.eu –, GA 689669, 2016-2020), was Scientific Responsible of the department unit for iWebCare Project (2006-2008): "Integrated Web Services Platform for the facilitation of fraud detection in health care e-government services" (FP6-2004-IST-4-28055), for INSPECTOR project (2000-2003): "Quality in the Statistical Information Life-Cycle: A Distributed System for Data Validation" (IST-2000-26347). She contributed in other Projects, such as in the Esprit programme TESS (1997-2000) - Automatic system for time-series analysis and forecasting; in the IST Programme VITAMIN-S (Visual Data Mining System – 2001-2003).

She was national coordinator of MIUR (Ministry of University Scientific Research) research projects: "Classification and Regression Trees: Methods, Open Source Technologies and Case Studies" (2005-2007), "Knowledge Discovery Methods and Informational Statistics for decision-making" (2003-2005), and "Statistical models for classification and segmentation of complex data structures" (2001-2003).

She was Scientific Responsible of research projects funded by Local Institutions (Region, Province, City Hall) in the fields of Tourism, Social Sciences and Economics: "Campagna D'Ascolto dei cittadini", Piano Strategico della Città di Napoli (2006-07); "SOCIALNET: Osservare la realtà per combattere l'esclusione: Focus sulla terza età" Assessorato alle politiche sociali della Provincia di Napoli (2007-08); "MONITUR: Monitoraggio e analisi statistica dei dati sul turismo e sulla fruizione dei beni culturali", Assessorato al Turismo e dei Beni Culturali della Regione Campania (2008-09).

Actually, she is Principal investigator of the project "*Moving Towards Adaptive Governance in Complexity: Informing Nexus Security*" (MAGIC), nr. 689669 funded by: H2020-EU.3.5.4. - Enabling the transition towards a green economy and society through eco-innovation.

She is member of the scientific staff for the following projects in the field of Robotics: RODYMAN (Robotic Dynamic MANipulation), Grant agreement no: 320992 since 01-07-2013; SAPHARI (Safe and Autonomous Physical Human-Aware Robot Interaction), IP 287513, call FP7-ICT-2011-7 from 01-01-2014 to 30-09-2015; EuRoC (European Robotics Challenges), n.608849 from 01-01-2014 to 31-07-2014; ARCAS (Aerial Robotics Cooperative Assembly Sistem), Contract number: 287617 from 10-01-2014 to 30-12-2014; SHERPA (Smart collaboration between Humans and ground-aErial Robots for imProving rescuing activities in Alpine environments), Contract number: IP - 600958 since 01-01-2015.

Publications:

Submitted papers

Iorio, C., Aria, M., D'Ambrosio, A., Siciliano, R. (2018). Informative Trees by Visual Pruning. Submitted, Conditionally Accepted.

Amodio, S., D'Ambrosio, A., Iorio, C. and Siciliano, R. (2018). Adjusted Concordance Index, an extension of the Adjusted Rand index to fuzzy partitions. Submitted, Conditionally Accepted. Technical report available as arXiv preprint arXiv:1509.00803 (Cornell University Library, 16/03/2016)

Pandolfo, G., Iorio, C., Siciliano, R. (2018) Robust mean-variance portfolio through the weighted L^p depth function. Submitted.

Iorio, C., Frasso, G., D'Ambrosio, A., Siciliano, R. (2019). Boosted-Oriented Probabilistic Smoothing-Spline Clustering of Series. Submitted. Technical report available as arXiv preprint arXiv: 1507.04905v2 (Cornell University Library, 07/10/2015).

Journal papers

D'Ambrosio, A., Iorio, C., Staiano, M., and Siciliano, R. (2019). Median constrained bucket order rank aggregation. *Computational Statitstics*, pp.1-16. DOI: 10.1007/s00180-018-0858-z

Aria, M., Capaldo, G., Iorio, C., Orefice, C.I., Riccardi, M., Siciliano, R. (2018). PLS Path Modeling for causal detection of project management skills: A research field in National Research Council in Italy. *Electronic Journal of Applied Statistical Analysis*, vol.11(2), pp.516-545, ISSN: 2070-5948. DOI: 10.1285/i20705948v11n2p516

Aria, M., D'Ambrosio, A., Iorio, C., Siciliano, R., and Cozza, V. (2018). Dynamic recursive tree-based partitioning for malignant melanoma identification in skin lesion dermoscopic images. *Statistical papers*, pp.1-17, ISSN:0932-5026. DOI: 10.1007/s00362-018-0997-x

Iorio, C., Frasso, G., D'Ambrosio, A. and Siciliano, R. (2018). A P-spline based clustering approach for portfolio selection. *Expert systems with applications*, vol. 95, pp. 88-103, ISSN: 0957-4174. DOI: 10.1016/j.eswa.2017.11.031. Online first: November 14, 2017.

Siciliano, R., D'Ambrosio, A., Aria, M. and Amodio, S. (2017). Analysis of web visit histories, part II: Predicting navigation by Nested STUMP Regression Trees. *Journal of Classification*, https://doi.org/10.1007/s00357-017-9239-5

D'Ambrosio, A., Mazzeo, G., Iorio, C., and Siciliano, R. (2017). A differential evolution algorithm for finding the median ranking under the Kemeny axiomatic approach. *Computers and Operations Research*, vol. 82, pp. 126-138. DOI: 10.1016/j.cor.2017.01.017.

D'Ambrosio, A., Aria, M., Iorio, C and Siciliano, R. (2017). Regression trees for multivalued numerical response variables, *Expert systems with applications*, vol. 69, pp. 21-28, DOI: 10.1016/j.eswa.2016.10.021

Siciliano, R., D'Ambrosio, A., Aria, M. and Amodio, S. (2016) Analysis of web visit histories, part I: Distance-based visualization of sequence rules. *Journal of Classification*, vol. 33(2), pp. 298-324 DOI: 10.1007/s00357-016-9204-8.

Iorio, C., Frasso, G., D'Ambrosio, A., and Siciliano R. (2016). Parsimonious Time Series Clustering using P-Splines, *Expert Systems with Applications*, vol. 52, pp. 26-38. DOI: 10.1016/j.eswa.2016.01.004.

Amodio, S., D'Ambrosio, A. and Siciliano, R. (2016) Accurate algorithms for identifying the median ranking when dealing with weak and partial rankings under the Kemeny axiomatic approach. *European Journal of Operational Research,* 292(2), pp. 667-676. DOI: 10.1016/j.ejor.2015.08.048.

Faggiano, A., et al. (2016). Somatostatin Analogues according to Ki67 index in neuroendocrine tumours: an observational retrospective-prospective analysis from real life. *Oncotarget*, 7.5: 5538.

D'Ambrosio A., Aria M. and Siciliano R. (2012). Accurate Tree-based Missing Data Imputation and Data Fusion within the Statistical Learning Paradigm, *Journal of Classification*, vol. 29(2), pp. 227-258. DOI: 10.1007/s00357-012-9108-1.

Conversano, C., and Siciliano, R. (2009). Incremental tree-based missing data imputation with lexicographic ordering. *Journal of classification*, 26(3), 361-379.

Aria, M., Gallo, S., Murolo, F., and Siciliano, R. (2008). Le leve discriminanti della soddisfazione: il caso SEPSA. *Rivista di economia e statistica del territorio*.

Petrakos, G., Conversano, C., Farmakis, G., Mola, F., Siciliano, R., and Stavropoulos, P. (2004). New ways of specifying data edits. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 167(2), 249-274.

Conversano, C., Siciliano, R., and Mola, F. (2002). Generalized additive multi-mixture model for data mining. *Computational statistics & data analysis*, 38(4), 487-500.

Cappelli, C., Mola, F., and Siciliano, R. (2002). A statistical approach to growing a reliable honest tree. *Computational statistics & data analysis*, 38(3), 285-299.

Conversano, C., Mola, F., and Siciliano, R. (2001). Partitioning algorithms and combined model integration for data mining. *Computational Statistics*, 16(3), 323-339.

Siciliano, R., and Mola, F. (2000). Multivariate data analysis and modeling through classification and regression trees. *Computational Statistics & Data Analysis*, 32(3), 285-301.

Palumbo, F., and Siciliano, R. (1999). Factorial discriminant analysis and probabilistic models. *Metron*, 56, 186-198.

Siciliano, R., and Mooijaart, A. (1997). Three-factor association models for three-way contingency tables. *Computational statistics & data analysis*, 24(3), 337-356.

Mola, F., and Siciliano, R. (1997). A fast splitting procedure for classification trees. *Statistics and Computing*, 7(3), 209-216.

Siciliano, R., and Van der Heijden, P. G. (1994). Simultaneous latent budget analysis of a set of two-way tables with constant row sum data. *Metron*, 53, 155-179.

Siciliano, R., Mooijart, A., and van der Heijden, P. G. (1993). A probabilistic model for nonsymmetric correspondence analysis and prediction in contingency tables. *Statistical Methods & Applications*, 2(1), 85-106.

Siciliano, R. (1992). Reduced-rank models for dependence analysis of contingency tables. *Statistica Applicata*, 4, 481-501.

Siciliano, R. (1990). Asymptotic distribution of eigenvalues and statistical tests in non symmetric correspondence analysis. *Statistica applicata*, 2(3), 259-276.

Lauro, N. C., and Siciliano, R. (1989). Exploratory methods and modelling for contingency tables analysis: an integrated approach. *Statistica applicata*, 1(1), 5-32.

Book Chapters

Iorio, C., Frasso, G., D'Ambrosio, A. and Siciliano, R. P-splines based clustering as a general framework: some applications using different clustering algorithms. To appear in Mola, F., Conversano, C., and Vichi, M. Eds., *Classification, (Big) Data Analysis and Statistical Learning*, . Springer series: Studies in Classification, Data Analysis, and Knowledge Organization. Springer-Verlag.

Siciliano R. and D'Ambrosio A. (2012). Statistical monitoring of tourism in the knowledge era. *In Morvillo A. (Ed.). Advances in Tourism Studies.* McGrow-Hill, pp. 231-258.

Siciliano R., Aria M., D'Ambrosio A. and Tutore V.A. (2011). Indagine statistica sulle aspettative e priorità per soddisfare il turista a Napoli, in Becheri E., Maggiore G. (a cura di), *XVII rapporto sul turismo italiano*, Franco Angeli, pp. 449-470.

Siciliano, R., and Aria, M. (2011). Two-Class trees for non-parametric regression analysis. In *Classification and Multivariate Analysis for Complex Data Structures* (pp. 63-71). Springer Berlin Heidelberg.

Pecoraro, M., and Siciliano, R. (2008). Statistical Methods for Profiling Users in Web Usage Mining. In Song, M., and Wu, Y.B. (Eds.), *Handbook of Research on Text and Web Mining Technologies,* pp- 359-368, ICI Global.

Siciliano, R., and Conversano, C. (2008). Decision Tree Induction. In Wang, J. (Ed.) *Encyclopedia of data warehousing and mining*, pp. 624-630. IGI Global.

Conversano, C. and Siciliano, R. (2008). Statistical Data Editing. In Wang, J. (Ed.) *Encyclopedia of data warehousing and mining*, pp. 1835-1840. IGI Global.

Siciliano, R., Aria, M. and D'Ambrosio, A. (2008). Posterior Prediction Modelling of Optimal Trees, in *Proceedings in Computational Statistics* (COMPSTAT 2008), 18th Symposium Held in Porto, Portugal, Brito, Paula (Ed.), Springer-Verlag, pp. 323-334

Tutore, V. A., Siciliano, R., & Aria, M. (2007, September). Conditional classification trees using instrumental variables. In *International Symposium on Intelligent Data Analysis* (pp. 163-173). Springer Berlin Heidelberg.

D'Ambrosio A., Aria M. and Siciliano R. (2007), Robust Tree-based Incremental Imputation Method for Data Fusion. *Lecture notes in computer science 4723 (Advances in Intelligent Data Analysis), Springer-Verlag, pp 174-183.*

Siciliano R., Aria, M. and D'Ambrosio A. (2006), Boosted incremental tree-based imputation of missing data. *Data Analysis, Classification and the Forward Search*. Springer series: Studies in Classification, Data Analysis, and Knowledge Organization. Springer-Verlag, pp. 271-278.

Aria, M., Mooijaart, A., and Siciliano, R. (2003). Neural Budget Networks of Sensorial Data. *Between Data Science and Applied Data Analysis*, 369-377.

Aria, M., Mola, F., and Siciliano, R. (2002). Growing and Visualizing Prediction Paths Trees in Market Basket Analysis. In *Compstat 2002* (pp. 377-382). Physica-Verlag HD.

Siciliano, R., and Mooijaart, A. (2001). Unconditional latent budget analysis: a neural network approach. In *Advances in Classification and Data Analysis* (pp. 127-134). Springer Berlin Heidelberg.

Cappelli, C., Mola, F., and Siciliano, R. (2001). Selecting Regression Tree Models: a Statistical Testing Procedure1. In *Advances in Classification and Data Analysis* (pp. 249-256). Springer Berlin Heidelberg.

Conversano, C., Siciliano, R., and Mola, F. (2000, June). Supervised classifier combination through generalized additive multi-model. In *International Workshop on Multiple Classifier Systems* (pp. 167-176). Springer Berlin Heidelberg.

Cappelli, C., Mola, F., and Siciliano, R. (2000). A third Stage in regression tree growing: searching for statistical reliability. In *Data Analysis, Classification, and Related Methods* (pp. 193-198). Springer Berlin Heidelberg.

Klaschka, J., Siciliano, R., and Antoch, J. (1998). Computational enhancements in tree-growing methods. In *Advances in Data Science and Classification* (pp. 295-302). Springer Berlin Heidelberg.

Siciliano R. (1998) Exploratory Versus Decision Trees. In: Payne R., Green P. (eds) *COMPSTAT 1998*. pp. 113-124, Physica, Heidelberg

Mola, F., Klaschka, J., and Siciliano, R. (1996). Logistic classification trees. In *COMPSTAT 1996* (pp. 373-378). Physica-Verlag HD.

Mola, F., and Siciliano, R. (1994). Alternative strategies and CATANOVA testing in two-stage binary segmentation. In *New Approaches in Classification and Data Analysis* (pp. 316-323). Springer Berlin Heidelberg.

Siciliano, R., and Mola, F. (1994). Modelling for recursive partitioning and variable selection. In *COMPSTAT 1994* (pp. 172-177). Physica-Verlag HD.

Mola, F., and Siciliano, R. (1992). A two-stage predictive splitting algorithm in binary segmentation. In Dodge, Y. et al. (eds.) *Computational statistics* (pp. 179-184). Physica-Verlag HD.

Siciliano, R. (1989). Non Symmetrical Logarithmic Analysis for Contingency Tables. In Decarli A., Francis B.J., Gilchrist R., Seeber G.U.H. (eds) *Statistical Modelling. Lecture Notes in Statistics*, vol 57, (pp. 278-285). Springer, New York.

Proceedings

Aria, M., Iorio, C., Siciliano, R. and Staiano, R. (2017). Visual tools for interactive clustering of EU state members via metabolic patterns. In *Greselin, F., Mola, F. and Zenga, M. (Eds.), 11th scientific meeting of the Classification and Data Analysis Group, Book of Abstract,* p. 1-5, Universitas Studiorum S.r.l. Casa Editrice, Mantova – Italy.

D'Ambrosio, A., Iorio, C., and Siciliano, R. (2017). Constrained consensus bucket order. In *Greselin, F., Mola, F. and Zenga, M. (Eds.), 11th scientific meeting of the Classification and Data Analysis Group, Book of Abstract*, p. 1-5, Universitas Studiorum S.r.l. Casa Editrice, Mantova – Italy.

D'Ambrosio, A., Frasso, G., Iorio, C. and Siciliano R. (2015). Probabilistic boosted-oriented clustering of time series. In *Mola, F. and Conversano, C. (Eds.), 10th scientific meeting of the Classification and Data Analysis Group, Book of Abstract,* p. 61-64, CUEC Editrice.

Mazzeo, G., D'Ambrosio, A. and Siciliano, R. (2015). Accurate Algorithms for Consensus Ranking Detection. In *Mola, F. and Conversano, C. (Eds.), 10th scientific meeting of the Classification and Data Analysis Group, Book of Abstract,* p. 255-258, CUEC Editrice

Iorio, C., D'Ambrosio, A., Frasso, G. and Siciliano R. (2015). Parsimonious Clustering of Time Series. In *Mola, F. and Conversano, C. (Eds.), 10th scientific meeting of the Classification and Data Analysis Group, Book of Abstract,* p. 226-229, CUEC Editrice.

D'Ambrosio A. (2012). Missing Data Imputation within the Statistical learning Paradigm. *Proceedings of the 46th Scientific Meeting Of The Italian Statistical Society.*

Siciliano R., Tutore V.A., Aria M., D'Ambrosio A. (2010). Trees with leaves and without leaves. In *45th scientific meeting of the Italian Statistical Society*.

Siciliano, R. (2008). Two-class trees: a tree-based methodological framework. In *First joint meeting of the société francophone de classification and the classification and data analysis group of the Italian statistical society*, pp. 27-30, ESI.

Aria M., D'Ambrosio A. and Siciliano R. (2007), Robust Incremental Trees for Missing Data Imputation and Data Fusion.*Classification and Data Analisys 2007, Book of short papers* (Macerata, September 12-14, 2007), EUM macerata, 287-290.

Siciliano R., Aria. and D'Ambrosio A. (2005), Boosted stump algorithm for missing data incremental imputation. *Invited talk: CLADAG 2005, Book of Short Papers* (Parma, June 6-8, 2005), MUP, Parma, 161-164.

Siciliano, R., Aria, M., and Conversano, C. (2004). Harvesting trees: methods, software and applications. In *Proceedings in Computational Statistics: 16th Symposium of IASC*. COMPSTAT2004, held Prague.

Aria, M., and Siciliano, R. (2003). Learning from trees: two-stage enhancements. *Proceedings of Classification and Data Analysis Group* (CLADAG 2003), 22-24.

Cappelli, C., Mola, F., and Siciliano, R. (1998). An alternative pruning method based on the impurity-complexity measure. In *COMPSTAT 1998* (pp. 221-226). Physica-Verlag HD.