

STEFANO NEMBRINI

01/12/1982

Brescia

+39 339 74 65 240

e-mail: stefanonembrini@gmail.com

Education

2009-2013 UNIVERSITY OF MILAN – BICOCCA

PhD in Methodological and Applied Statistics - scholarship.

PhD Thesis: *Machine Learning Methods for Feature Selection and Rule Extraction in Genome-wide Association Studies (GWASs)*.

Supervisor: Prof. Paola Zuccolotto, University of Brescia,

Co-supervisor: Cristian Pattaro, PhD, European Accademy of Bolzano (Eurac)

Co-supervisor: Marco Sandri, MSc, IULM University - Milan.

The work includes a theoretical and applied approach to machine learning methods based on CARTs (Classification and Regression Trees) with the application to regression problems in GWASs (GenomeWide Association Studies) with the use of a large dataset (SNPs > 300,000). Particular interest is given to feature selection and rule extraction algorithms for genetics.

2009 UNIVERSITY OF BRESCIA - FACULTY OF ECONOMICS

MSc in Economics

Thesis: *Interpreting Multivariate Data: Visual Representations and Data Reduction Techniques* (Prof. Paola Zuccolotto, PhD).

2008 UNIVERSITY OF BRESCIA - FACULTY OF ECONOMICS

BSc in Economics

Thesis: *Application of Multivariate Techniques for the Analysis of Students' Abilities and Perspectives* (Prof. Paola Zuccolotto, PhD).

2005 CATHOLIC UNIVERSITY (BRESCIA)

FACULTY OF FOREIGN LANGUAGES AND LITERATURES

BA in English and German Studies

Thesis: *Globalization: A Semantic Approach* (Prof. Annalisa Zanola, PhD).

2004 UNIVERSITY OF LEEDS (U.K.)

Exchange programme (Socrates/Erasmus project) with exams taken in the following departments: *Leeds University Business School (LUBS)*, *School of Modern Languages and Cultures* (German Linguistics and Phonetics), *School of Classics*.

2001 LICEO SCIENTIFICO A. CALINI HIGH SCHOOL (Brescia)

Traineeships

- 2012 EUROPEAN ACADEMY OF BOLZANO (EURAC)
Training (April-June) for the application of genome-wide studies (GWASs) - Centre of Biomedicine.
- 2008 STUDIO SYNTHESIS S.R.L.
Consulting, Business Intelligence, CRM solutions, software selection, quantitative analysis.
- 2004 OMB SALERI S.P.A.
Valve manufacturing for automotive industry.
Public Relations, market research, technical translation.

Work Experience

- 2017 **UNIVERSITY OF BRESCIA** - DEPARTMENT OF CIVIL, ENVIRONMENTAL, ARCHITECTURAL ENGINEERING, AND MATHEMATICS
Postdoctoral Researcher - *Assessment of the state of water, air, soil and health impacts in the town of Brescia.*
- 2016 **UNIVERSITY OF BRESCIA** - DEPARTMENT OF MEDICAL AND SURGICAL SPECIALTIES, RADIOLOGICAL SCIENCES AND PUBLIC HEALTH
Postdoctoral Researcher (MAPEC_LIFE - Monitoring air pollution effects on children for supporting Public Health Policy), reference: Prof. Rosella Levaggi rosella.levaggi@unibs.it.
- 2015-2017 **UNIVERSITY OF BRESCIA** - DEPARTMENT OF ECONOMICS AND MANAGEMENT
University Lecturer for the course of Econometrics (Prof. Raffaele Miniaci).
- 2015 **UNIVERSITY OF BRESCIA** - DEPARTMENT OF ECONOMICS AND MANAGEMENT
Postdoctoral Researcher (*Quantile Regression Models for the study of Tax Determinants of Subsidiaries' Capital Structure*), reference: Prof. Raffaele Miniaci raffaele.miniaci@unibs.it.
- 2013 **BROWN UNIVERSITY** (Providence, RI, USA) DEPARTMENT OF EPIDEMIOLOGY, CENTER FOR POPULATION HEALTH & CLINICAL EPIDEMIOLOGY - 07/2013 to 12/2013.
Postdoctoral Research Associate: *Test Statistics For The Analysis Of Genetic Pathways*. Reference: Dr. Yen-Tsung Huang yen-tsung_huang@brown.edu.
- 2012-2013 **UNIVERSITY OF BRESCIA** - DEPARTMENT OF ECONOMICS AND MANAGEMENT
University Lecturer for the course of Statistics For Marketing Research (Prof. Paola Zuccolotto).
- 2012 **EUROPEAN ACADEMY OF BOLZANO (EURAC)** - 06/2012 - 12/2012
Machine Learning Methods for the selection of *single-nucleotide polymorphisms* for the prediction of *estimated glomerular filtration rate* using the data from the MICROS Study (Pattaro et al. BMC Med Genet 2007).

2010-2012 UNIVERSITY OF BRESCIA - FACULTY OF ECONOMICS

University Lecturer for the course of Market Analysis (Prof. Paola Zuccolotto).

Computer Skills

- Windows (98, XP, Vista, 7) and Linux (Ubuntu);
- MS Office (Word, Access, Excel, PowerPoint) (ECDL);
- Openoffice/Libreoffice (Writer, Calc, Base, Impress);
- Statistical packages: R (e.g.: Boruta, homals, boot, BayesTree, caret, lars, parcor, randomForest, randomForestSRC, RRF,rqpd, randomForestSRC, party, bartMachines, igraph), Stata, SPSS, Systat, Random Jungle, Matlab, WinBUGS, SAS.
- Java, JavaScript, Php, Microsoft SQL Server, Joomla, HTML, CSS.
- Use of large databases/datasets (e.g. AMADEUS)
- LaTeX (T_EXnicCenter, L_AT_EX).

Language Skills

English: Fluent,

German: Good,

Spanish: Moderate,

Portuguese: Moderate,

French: Basic.

Research Interests and Statistical Skills

- Linear And Non Linear Regression, ANOVA, Penalised Regression (Ridge Regression, Lasso, Elastic Net),

- Vector autoregression (VAR), seemingly unrelated regressions (SUR), and other techniques applied in econometrics (e.g. fixed effects estimator),
- Unconditional and Conditional Quantile Regression (for panel data), Bayesian Quantile Regression For binary outcomes,
- Unsupervised (Clustering Methods) and Supervised Learning (Regression And Classification Problems),
- Machine Learning (CART, Conditional Inference Trees, Bagging, Random Forest (Quantile, Oblique, Regularized, Weighted Subspace, Uniform RF), Randomized Trees, Conditional Forest, Random Ferns, Neural Networks, Boosted Trees, Support Vector Machines), Multivariate Adaptive Regression Splines (MARS), Bayesian Additive Regression Trees,
- Testing procedures based on Variable Importance Measures for GWASs (GenomeWide Association Studies) and other extremely large datasets,
- Variable Selection (e.g. exploiting Variable importance measures and Recursive feature elimination),
- Resampling Techniques (Bootstrap, Jackknife, Cross Validation),
- Multiple Testing Techniques,
- Social Network Analysis,
- ANOVA, ANCOVA, MANOVA, MANCOVA,
- Stochastic Processes, Markov Switching Models,
- Multivariate Analysis (MultiDimensional Scaling, Clustering Techniques (K-means, Hierarchical Clustering, Biclustering, RForest Clustering, Partitioning Around Medoids), (Non Linear) Principal Components Analysis, Independent Component Analysis, Factor Analysis).
- Others: Logistic/Logic Regression, Regression for count data (zero-inflated models).

Publications

1. Menoncin, F. & Nembrini, S. **Stochastic continuous time growth models that allow for closed form solutions** J Econ (2017)
2. `uqr`: **Unconditional Quantile Regression: Estimation and Inference for Unconditional Quantile Regression for cross-sectional and panel data.** R package for Econometrics available at <https://CRAN.R-project.org/package=uqr>
3. `IPMRF`: **Intervention in Prediction Measure (IPM) for Random Forests.** <https://CRAN.R-project.org/package=>

4. Pedrazzani, R., Menoni, L., Nembrini, S., Manili, L., & Bertanza, G. (2016). **Suitability of Sludge Biotic Index (SBI), Sludge Index (SI) and filamentous bacteria analysis for assessing activated sludge process performance: the case of piggery slaughterhouse wastewater.** *Journal of Industrial Microbiology & Biotechnology*, 1-12.
5. Rinaldi, F., Nembrini, S., Concoreggi, C., Magoni, M., & Padovani, A. (2016). **Neurological diseases and health care utilization among first-generation immigrants.** *Journal of neurology*, 263(4), 714-721. Chicago
6. Papa, M., Pedrazzani, P., Nembrini, S.: **Should rheological properties of activated sludge be measured?** *Appl. Rheol.* 25:2 (2015)

Conferences/Talks

- Miniaci, R., Nembrini, S., Panteghini, P.: **Dynamic quantile regression models and capital structure: heterogeneous reaction to tax incentives and profitability** - Seventh Italian Congress of Econometrics and Empirical Economics (ICEEE 2017) - 25-27 January 2017, Messina (Italy).
- **A novel testing approach for random forests suitable for GWASs based on the Gini importance.** June 7 2017. Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck. (invited)

Working Papers

1. Bua, G., Nembrini, S. **On ECB Trust: An Unconditional Quantile Regression approach - Evidence from Panel Data implementing Shi's Double Bootstrap.** 2016
2. Vaccari, Foladori, Vitali, Nembrini: **Benchmarking Of Energy Consumption In Municipal Wastewater Treatment Plants.** Working Paper. 2015

Referee/Reviewer

Journal of Public Health Research