41003 – Statistical Machine Learning

INSTRUCTOR **Daniele Durante**

CONTACT INFORMATION
Department of Decision Sciences Sciences
Via Roentgen, 1, 20136 Milano, 3-D1-05
✉: daniele.durante@unibocconi.it
web: https://danieledurante.github.io/web/

OBJECTIVES

The course will combine methods, applications, theory and computational aspects in Statistical Machine Learning. More specifically, methods will be motivated and evaluated with a focus on applied problems, but the overarching goal will be on the general methodological framework, including theoretical results and statistical properties. The topics will be presented via a careful discussion of the original papers.

SYLLABUS

1. An Introduction to Statistical and Machine Learning [L1 – L2]

- Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science.* 16(3): 199–231.
- Jordan, M.I. (2019). Artificial intelligence—the revolution hasn't happened yet. *Harvard Data Science Review.* 1(1) + Candes, E., Duchi, J., Sabatti, C. (2019). Statistics and the oncoming AI revolution. *Harvard Data Science Review.* 1(1) + Donoho, D. (2019) What's missing from today's machine intelligence juggernaut? *Harvard Data Science Review.* 1(1)
- Efron, B. (2020). Prediction, estimation, and attribution. *Journal of the American Statistical Association.* 115(530): 636–655.

2. Regression Trees, Bagging, Boosting and (Causal) Random Forests [L3 – L4 – L5]

- Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and Regression Trees.* Chapman & Hall.
- Breiman, L. (1996). Bagging predictors. *Machine Learning.* 24(2): 123–140.
- Breiman, L. (2001). Random forests. *Machine Learning.* 45(1): 5–32.
- Friedman, J., Hastie, T. and Tibshirani, R. (2000). Additive logistic regression: A statistical view of boosting. *The Annals of Statistics*, 28(2): 337–407.
- Wager, S., and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523): 1228–1242.

3. Bayesian Trees and Bayesian Additive Regression Trees [L6 – L7]

- Chipman, H. A., George E. I. and McCulloch R. E. (1998). Bayesian CART model search. *Journal of the American Statistical Association.* 93(443): 935–948.
- Chipman, H. A., George E. I. and McCulloch R. E. (2010). BART: Bayesian additive regression trees. *The Annals of Applied Statistics.* 4(1): 266–298.

4. Conformal Inference/Prediction [L8 – L9]

- Angelopoulos, A.N., and Bates, S. (2023). Conformal prediction: A gentle introduction. *Foundations and Trends® in Machine Learning*, 16(4): 494–591.

5. Topics in Nonparametric Density Regression [L10]

- Quintana, F. A., Müller, P., Jara, A., and MacEachern, S. N. (2022). The dependent Dirichlet process and related models. *Statistical Science*, 37(1), 24-41.

6. Topics in Bayesian Unsupervised Learning [L11 – L12]

- Hoff, P. (2021). Additive and multiplicative effects network models. *Statistical Science.* 36: 34–50.
- Legramanti, S., Rigon, T., Durante, D. and Dunson, D.B. (2022). Extended stochastic block models with application to criminal networks. *Annals of Applied Statistics.* 16(4): 2369–2395.
- Gopalan, P., Hofman, J.M., and Blei, D.M. (2015). Scalable recommendation with hierarchical Poisson factorization. *In UAI 2015 Proceedings*, pp. 326-335. + Gopalan, P., Ruiz, F.J., Ranganath, R., and Blei, D. (2014). Bayesian nonparametric Poisson factorization for recommendation systems. *In AISTATS 2014 Proceedings*, pp. 275-283.

Additional references to specific articles will be suggested during the course.

GRADING

The evaluation is based on an individual project in which the student is asked to provide a critical and thoughtful discussion of an interesting topic considered during the course.