Machine Learning with Macroeconomics and Microeconomics Applications

Center for Monetary and Financial Economics (CMFE) Workshop

A typical day consists of the lecture in the morning part followed by applications and the practical exercises in the second part. Applications in economics and finance will be considered to illustrate the different methodologies. Exchanges and discussions are encouraged.

Outline

Day 1: Machine Learning and Big Data with Micro Applications

Methods:
Linear regression and statistical learning

- Principals of machine learning
    - Prediction vs causal analysis
    - The metrics of prediction
    - Training set, test sets and validation sets
- Linear regression and statistical learning
- Bootstrap and cross-validation
- Regularized regression
    - Ridge
    - Lasso
    - Elastic-net

- Classification and Regression trees
    - Trees
    - Bagging
    - Random Forest
    - Boosting
    - Ranking variables as predictors
- Deep learning (neural networks)

Applications in class:

- Machine learning with R
- Predicting housing prices (continuous variable), defaults on loans (discrete variable)

Day 2: Machine Learning and Big Data with Macro Applications

Methods

- Theory of forecasting
- Factor models
- Principal component analysis
- Regularized regressions in time series
- Cross-validation in time series
- Support vector machine
- Complete subset regressions

Applications in class:

- High-dimensional (data-rich) macroeconomic panels
- Forecasting industrial production growth (real activity) and CPI inflation with Matlab

References

- James, Witten, Hastie, Tibshirani (JWHT), *An Introduction to Statistical Learning*, Springer texts, http://www-bcf.usc.edu/~gareth/ISL/ISLR%20First%20Printing.pdf

- Hastie, Tibshirani, Friedman, (HTF), *The Elements of Statistical Learning*, Springer Series in Statistics, https://web.stanford.edu/~hastie/Papers/ESLII.pdf

- Aiello, Eckstrand, Fu, Landry, Lanford, Aboyoun, *Machine Learning with R and H20*¸ H20 booklet, http://h2o-release.s3.amazonaws.com/h2o/master/3233/docs-website/h2o-docs/booklets/R_Vignette.pdf

- Nykodym, Rao, Wang, Krajevic, Lanford, Hussami, *Generalized Linear Modeling with H20*¸ H20 Booklet, https://h2o-release.s3.amazonaws.com/h2o/rel-slater/9/docs-website/h2o-docs/booklets/GLM_Vignette.pdf

- Malohlava, Candel, *Gradient Boosting Machine with H20*, H20 Booklet, http://docs.h2o.ai/h2o/latest-stable/h2o-docs/booklets/GBMBooklet.pdf

- Candel, Landford, LeDell, Parmar, Arora, *Deep Learning with H20*, H20 Booklet, https://h2o-release.s3.amazonaws.com/h2o/rel-slater/9/docs-website/h2o-docs/booklets/DeepLearning_Vignette.pdf

- Bengio, Yoshua (2012), "Practical recommendations for gradient-based training of deep architectures", https://arxiv.org/abs/1206.5533

- Additional references for Support Vector Machine Regression:
    - http://kernelsvm.tripod.com
    - Smola, A. et B. Scholkopf (2004), « A tutorial on support vector regression », *Statistics and Computing* 14, 199-222.

- Additional references for Dynamic factor model:
    - Stock, J.H. et M.H. Watson (2016), "Factor Models and Structural Vector

Autoregressions in Macroeconomics" in *Handbook of Macroeconomics,* Vol2A, John B. Taylor and Harald Uhlig (eds), 2016, Chapter 8, pp 415-526

- Bai, J. et Ng, S. (2008), "Forecasting economic time series using targeted predictors", *Journal of Econometrics* 146, 304–317
- Bai, J. et Ng, S. (2008), "Boosting Diffusion Indices", *Journal of Applied Econometrics*, 24:4, 607-629

- Additional references for forecasting:
  - Leroux, M., Kotchoni, R. et D. Stevanovic (2017), « Macroeconomic forecasts accuracy in data-rich environment », UQAM
  - Giannone, D., Lenza, M. and G. Primiceri (2018), "Economic Predictions with Big Data: The Illusion of Sparsity", Staff Report No. 847. NYFed
  - Diebold, F.X., *Forecasting*, http://www.ssc.upenn.edu/~fdiebold/Textbooks.html
  - Stock, J.H. et M.H. Watson (2017) "Twenty Years of Time Series Econometrics in Ten Pictures" *Journal of Economic Perspectives*, Vol. 31 No. 2.