Readings in Statistics and Econometrics 2016: Bayesian statistical learning

Organizers: Hedibert Freitas Lopes & Paulo C. Marques F.

Objective: In this Second Readings in Statistics and Econometrics we will study and discuss, through a series of well established papers, the broad topic of Statistical Learning with an emphasis on its natural Bayesian solutions. The 5 lectures and 8 seminars will take place on Fridays between 10am and 12pm from January 29th to April 8th 2016. Paulo and I will give lectures discussing traditional Statistical Learning techniques, alternated with seminars given by the participants on papers presenting Bayesian counterparts to the techniques discussed in the lectures.

Outline of the meetings (5 lectures and 8 seminars)

- Lecture 1 on k-nearest neighors (k-NN)
- Seminar 1: Holmes and Adams (2002,2003)
- Seminar 2: Cucala, Marin, Robert and Titterington (2009)
- Lecture 2 on LASSO regularization
- Seminar 3: Griffin and Brown (2010,2012,2014)
- Seminar 4: Polson, Scott and Windle (2014)
- Lecture 3 on Random Forests
- Seminar 5: Chipman, George and McCulloch, (2010)
- Lecture 4 on Supporting Vector Machines
- Seminar 6: Tipping (2001)
- Seminar 7: Polson and Scott (2011)
- Lecture 5 on k-means Clustering
- Seminar 8: Paulo's brief review of Dirichlet Processes
- Seminar 9: Kulis and Jordan (2012)

Books

- An Introduction to Statistical Learning (James, Witten, Hastie and Tibshrani)
- Applied Predictive Modeling (Kuhn and Johnson)
- Bayesian Reasoning and Machine Learning (Barber)
- The Elements of Statistical Learning (Hastie, Tibshrani and Friedman)
- Machine Learning: A Probabilistic Perspective (Murphy)
- Pattern Recognition and Machine Learning (Bishop)
- Pattern Classification (Duda, Hart and Stork)
- Probability and Measure (Billingsley)
- Probability and Measure Theory (Ash and Doleans-Dade)
- Optimal Statistical Decisions (DeGroot)
- Theory of Statistics (Schervish)
- Principles of Uncertainty (Kadane)

Papers

- Chipman, George and McCulloch (2010) BART: Bayesian Additive and Regression Trees. AOAS, 4, 266-298.
- Cucala et al. (2009) A Bayesian Reassessment of Nearest-Neighbor Classification. JASA, 104, 263-273.
- Griffin and Brown (2010) Inference with normal-gamma prior distributions in regression problems. BA, 5, 171-188.
- Griffin and Brown (2012) Structuring shrinkage: some correlated priors for regression. Biometrika, 99, 481-487.
- Griffin and Brown (2013) Some priors for sparse regression modelling. BA, 8, 691-702.
- Holmes and Adams (2002) A Probabilistic NN Method for Statistical Pattern Recognition. JRSS-B, 64, 295-306.
- Holmes and Adams (2003) Likelihood Inference in NN Classification Models. Biometrika, 90, 99-112.
- Kulis and Jordan (2012) Revisiting k-means: New Algorithms via Bayesian Nonparametrics. Proc. XXIX ICML.
- Polson and Scott (2011) Data Augmentation for Support Vector Machines. BA, 6, 1-24.
- Polson, Scott and Windle (2014) The Bayesian Bridge. JRSS-B, 76, 713-733.
- Tipping (2001) Sparse Bayesian learning and the Relevance Vector Machine. JMLR, 1, 211-244.
- Marques and Pereira (2013) Predictive Analysis of Microarray Data.