Project in Structured Data : learning, prediction, dependency, testing

M2 Data Science - 2016-17
Florence d'Alché-Buc – Zoltán Szabó

Important dates and deadline :

- Feb 27 : list of projects
- Apr 2 : projects to be sent no later than Midnight, use the following subject in the email [M2 Structured data project]

No more than 4 students per project

Subjects :

- Using the following papers and topics listed below, choose among the listed papers as well as those given at the end of each set of slides (each lecture, see the moodle), a learning task studied in this course, you can EITHER consider a single task and implement two approaches OR consider several tasks and study the same method.
- The project has to be written as a scientific article, (introduction, problem , method, description of the code and numerical experiments) – Appendix : A script to launch a demo. Also describe the library you used and the code you wrote yourself.

Part I : Structured output prediction (non exhaustive list)

P1 Tsochantaridis, I. and Joachims, T. and Hofmann, T. and Altun (2005), Y., Large margin methods for structured and interdependent output variables, JMLR, 6, 2005 (et ICML)

P2 Collins, Michael, Discriminative Training Methods for Hidden Markov Models: Theory and Experiments with Perceptron Algorithms, Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10, 2002. (structured perceptron)

P3 Blaschko, Lampert (2008) Learning to localize objects with structured output regression, ECCV 2008.

P4 Geurts, Wehenkel, d'Alché-Buc (2006), Kernelizing outputs of tree-based methods, ICML 2006.

P5 Argyriou, E vgeniou, Pontil (2007) Convex Multi-task feature Learning, NIPS 2006, published in 2007.

P6 Dinuzzo, Ong, Gehler, Pillonetto (2011): Learning Output Kernels with Block Coordinate Descent. ICML 2011: 49-56.

P7 Su, Heinonen, Rousu (2010), Structured Output Prediction of Anti-cancer Drug Activity, PRIB 2010.

P8 C. H. Lampert, H. Nickisch, and S. Harmeling. "Learning To Detect Unseen Object Classes by Between-Class Attribute Transfer". In CVPR, 2009

P9 Deep Structured Output Learning for Unconstrained Text Recognition, Max Jaderberg,

Karen Simonyan, Andrea Vedaldi, Andrew Zisserman, arxiv,2015.

P10 Liang-Chieh Chen, Alexander G. Schwing, Alan L. Yuille, Raquel Urtasun:
Learning Deep Structured Models. ICML 2015: 1785-1794 .

P11 Structured prediction via output space search.
JR Doppa, A Fern, P Tadepalli - Journal of Machine Learning Research, 2014 – jmlr.org

Part II : dependency, testing (non-exhaustive)

P12 Independent Subspace Analysis: blind signal separation. Example:
demixing of ECG signals (mum & her baby), mixed music bands/songs.
Refs:
-Jean-Francois Cardoso. Multidimensional independent component analysis. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 1941–1944, 1998.
(separation principle: ISA = ICA up to permutation)
-Fabian. J. Theis, Blind signal separation into groups of dependent signals using
joint block diagonalization, ISCA, pp. 5878-5881, 2005.

P13 Outlier-robust image registration: information theoretical
registration of photos when standard L^p based measures perform poorly.
Refs:
-Huzefa Neemuchwala, Alfred Hero, Sakina Zabuawala, and Paul Carson. Image registration methods in high dimensional space. International Journal of Imaging Systems and Technology, 16:130–145, 2007.
-Jan Kybic. High-dimensional mutual information estimation for image registration. In IEEE International Conference on Image Processing (ICIP), pages 1779–1782, 2004.

P14 Feature selection. Example: which attributes (criminal rate, local
taxes, ...) affect most the price of a house.
Refs:
-Hanchuan Peng, Fuhui Long, and C. Ding. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(8):1226-1238, 2005.
-Le Song, Alex Smola, Arthur Gretton, Justin Bedo, Karsten Borgwardt. Feature Selection via Dependence Maximization. 13(May):1393−1434, 2012.
-Barnabás Póczos, Zoubin Ghahramani, Jeff Schneider. Copula-based Kernel Dependency Measures. ICML, pages 1635-1642, 2012.

P15 Distribution regression: examples include wheat/corn yield or aerosol estimation from satellite images, inference of non-analytically available statistics of a random variable (say entropy, or number of mixture components), or prediction of the expected life-time of a hard drive.
Refs:
-Zhuang Wang, Liang Lan, and Slobodan Vucetic. Mixture model for multiple instance regression and applications in remote sensing. IEEE Transactions on Geoscience and Remote Sensing, 50:2226–2237, 2012. (dataset: http://www.dabi.temple.edu/~vucetic/data/MIR_datasets.zip)
-Joseph F. Murray, Gordon F. Hughes, Kenneth Kreutz-Delgado. Machine Learning Methods for Predicting Failures in Hard Drives: A Multiple-Instance Application. Journal of Machine Learning Research 6 (2005) 783-816.
-Junier Oliva, Willie Neiswanger, Barnabás Póczos, Jeff Schneider, Eric Xing. Fast Distribution To Real Regression. JMLR W&CP 33: 706-714, 2014.
-Barnabás Póczos, Aarti Singh, Alessandro Rinaldo, Larry Wasserman. Distribution-Free

Distribution Regression. JMLR W&CP 31: 507-515, 2013.
-Zoltán Szabó, Bharath K. Sriperumbudur, Barnabás Póczos, Arthur Gretton. Learning Theory for Distribution Regression. Journal of Machine Learning Research, 17(152):1-40, 2016.
-Ali Rahimi and Ben Recht. Random features for large-scale kernel machines. In NIPS, pages 1177–1184, 2007. (random Fourier feature approximation/acceleration)

P16 High-power hypothesis testing with interpretable features. Examples:
song-title or video-caption dependency testing, discrimination of face
emotions or topics of documents, translation testing [texts in 2
languages; say how 'good' Google Translate is;)].
Refs:
-Wittawat Jitkrittum, Zoltán Szabó, Arthur Gretton. An Adaptive Test of Independence with Analytic Kernel Embeddings. TR, 2016 (https://arxiv.org/abs/1610.04782)
--Wittawat Jitkrittum, Zoltán Szabó, Kacper Chwialkowski, Arthur Gretton. Interpretable Distribution Features with Maximum Testing Power. In NIPS-2016, pages 181-189, 2016.
-Kacper P. Chwialkowski, Aaditya Ramdas, Dino Sejdinovic, Arthur Gretton. Fast Two-Sample Testing with Analytic Representations of Probability Measures. NIPS, pages 1981-1989, 2015.