





Bibliography	<p>A syllabus will be written based on the following sources (not exhaustive):</p> <p>Amari, S. and Wu, S. (1999). Improving support vector machine classifiers by modifying kernel functions. <i>Neural Networks</i>, 12(6):783-789.</p> <p>Chitta, R., Jin, R., Havens, T.C. and Jain, A.K. (2011). Approximate kernel k-means: Solution to large scale kernel clustering. In <i>Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining</i>, pages 895-903. ACM.</p> <p>Devroye, L., Györfi, L. and Lugosi, G. (2013). <i>A probabilistic theory of pattern recognition</i>, volume 31. Springer Science &amp; Business Media.</p> <p>Fan, J., Liao, Y. and Mincheva, M. (2011). High dimensional covariance matrix estimation in approximate factor models, <i>The Annals of Statistics</i>, 147, 186–197. Fan, J., Liao, Y. and Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements, <i>Journal of the Royal Statistical Society: Series B</i>, 75, 603-680.</p> <p>Gonen, M. and Alpaydyin, E. (2011). Multiple kernel learning algorithms. <i>Journal of machine learning research</i>, 12:2211-2268.</p> <p>Grandvalet, Y. and Canu, S. (2003). Adaptive scaling for feature selection in SVMs. In: <i>Advances in neural information processing systems</i>, pages 569-576.</p> <p>Guyon, I. and Elisseev, A. (2006). An introduction to feature extraction. <i>Feature extraction</i>, pages 1-25. Fine, S. and Scheinberg, K. (2001). Efficient SVM training using low-rank kernel representations. <i>Journal of Machine Learning Research</i>, 2(Dec):243-264.</p> <p>Hagen, L. and Kahng, A.B. (1992). New spectral methods for ratio cut partitioning and clustering. <i>IEEE transactions on computer-aided design of integrated circuits and systems</i>, 11(9):1074-1085.</p> <p>Hardle, W., Dwi Prastyo, D. and Hafner, C.M. (2014). Support Vector Machines with Evolutionary Feature Selection for Default Prediction, <i>Handbook of Applied Nonparametric and Semiparametric Econometrics and Statistics</i>, Oxford UP, edited by A. Ullah, J. Racine and L. Su.</p> <p>Hardle, W. and Simar, L. (2015). <i>Applied Multivariate Statistical Analysis</i>, Springer Verlag.</p> <p>Jain, A.K. (2010). Data clustering: 50 years beyond k-means. <i>Pattern recognition letters</i>, 31(8): 651-666.</p> <p>Johnson, W.B. and Lindenstrauss, J. (1984). Extensions of lipschitz mappings into a hilbert space. <i>Contemporary mathematics</i>, 26(189-206):1.</p> <p>Keerthi, S.S. and Lin, Ch-J. (2003). Asymptotic behaviors of support vector machines with gaussian kernel. <i>Neural computation</i>, 15(7):1667-1689.</p> <p>Kloft, M., Brefeld, U., Laskov, P., Müller, K.-R., Zien, A., and Sonnenburg, S. (2009). Efficient and accurate lp-norm multiple kernel learning. In <i>Advances in neural information processing systems</i>, pages 997-1005.</p> <p>Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices, <i>Journal of Multivariate Analysis</i>, 88, 365-411.</p> <p>Ledoit, O. and Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices, <i>Annals of Statistics</i>, 40, 1024-1060.</p> <p>Ledoit, O. and Wolf, M. (2015). Spectrum estimation: a unified framework for covariance matrix estimation and PCA in large dimensions, <i>Journal of Multivariate Analysis</i>, 139, 360-384.</p> <p>Ledoit, O. and Wolf, M. (2020). Direct nonlinear shrinkage estimation of large dimensional covariance matrices, <i>Annals of Statistics</i>.</p> <p>Lee, Y.-J. and Huang, S.-Y. (2007). Reduced support vector machines: A statistical theory. <i>IEEE Transactions on Neural Networks</i>, 18(1):1-13.</p> <p>Lee, S.-W. and Bien, Z. (2010). Representation of a Fisher criterion function in a kernel feature space. <i>IEEE transactions on neural networks</i>, 21(2):333-339.</p> <p>Mohar, B., Alavi, Y., Chartrand, G. and Oellermann, OR (1991). The laplacian spectrum of graphs. <i>Graph theory, combinatorics, and applications</i>, 2(871-898):12.</p> <p>Neumann, J., Schnorr, C. and Steidl, G. (2005). Combined svm-based feature selection and classification. <i>Machine learning</i>, 61(1):129-150.</p> <p>Ng, A.-Y., Jordan, M.I. and Weiss, Y. (2002). On spectral clustering: Analysis and an algorithm. In <i>Advances in neural information processing systems</i>, pages 849-856.</p> <p>Peters, G. W., <i>Statistical Machine Learning and Data Analytic Methods for Risk and Insurance (Version 8, 2017)</i>. Available at SSRN: <a href="https://ssrn.com/abstract=3050592">https://ssrn.com/abstract=3050592</a>.</p> <p>Scholkopf, B. and Smola, A.J. (2001). <i>Learning with kernels: support vector machines, regularization, optimization, and beyond</i>. MIT press, 2001.</p> <p>Yao, J., Zheng, S. and Bai, Z. (2015). <i>Large sample covariance matrices and highdimensional data analysis</i>, Cambridge UP</p>
Faculty or entity in charge	LSBA

Programmes containing this learning unit (UE)				
Program title	Acronym	Credits	Prerequisite	Learning outcomes
Master [120] in Data Science : Statistic	DATS2M	3		
Master [120] in Statistics: General	STAT2M	3		
Certificat d'université : Statistique et science des données (15/30 crédits)	STAT2FC	3		