

Course 11	SVMs and Regularized Learning
Program	<p>1. Introduction:</p> <ul style="list-style-type: none"> • Classification and Margins. • Classical Unconstrained Optimization. • Hard Margin Linear SVMs: primal problem. • Lagrangian formulation and dual problem. • KKT conditions and optimal solution. • Linear SVMs for nonlinear problems • Support Vector Regression <p>2. SVM models:</p> <ul style="list-style-type: none"> • Kernelization and non-linear SVMs: <ul style="list-style-type: none"> ◦ Non-linear classification in feature spaces. ◦ Mercer's Theorem. Kernel trick. ◦ Common kernel choices. ◦ Parameter tuning. • Non-linear Support Vector Regression: <ul style="list-style-type: none"> ◦ Primal and dual formulations, kernelization. • One-class Support Vector Machine: <ul style="list-style-type: none"> ◦ Density estimation problems. ◦ Primal and dual formulations, kernelization. ◦ Kernels for non-vector data. • Other SVM-related models. <p>3. SVM learning algorithms:</p> <ul style="list-style-type: none"> • Brief introduction to convex optimization. • Non-linear SVM learning algorithms: <ul style="list-style-type: none"> ◦ Chunking and decomposition methods. SVMlight. ◦ Sequential Minimal Optimization. LIBSVM. • Linear SVM learning algorithms: <ul style="list-style-type: none"> ◦ Primal solver: Pegasos. ◦ Dual solver: LIBLINEAR. <p>4. Regularized Models and Convex optimization:</p> <ul style="list-style-type: none"> • Convex optimization problems. • Subgradients and subdifferential calculus. • Proximal optimization. • Proximal methods: Forward-backward Splitting, Douglas-Rachford, Dykstra, etc. • The ISTA and FISTA algorithms. • Application to regularized learning: lasso, elastic nets, group variants, fused lasso. • Application to image processing. • Application to projection problems. <p>5. Practical Sessions:</p> <ul style="list-style-type: none"> • Python with sklearn.
Bibliography	<ul style="list-style-type: none"> • <i>Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond.</i> B. Schölkopf and A. J. Smola. MIT Press, 2002. • <i>A Tutorial on Support Vector Machines for Pattern Recognition.</i> C. J. C. Burges. <i>Data Mining and Knowledge Discovery, Volume 2, 2002.</i> • <i>A Tutorial on Support Vector Regression.</i> A. J. Smola and B. Schölkopf. <i>Statistics and Computing, Volume 48, 2003.</i> • <i>Proximal Splitting Methods in Signal Processing.</i> P. L. Combettes and J. C. Pesquet. <i>Recherche, Volume 49, 2009.</i> • <i>A Fast Iterative Shrinkage–Thresholding Algorithm for Linear Inverse Problems.</i> A. Beck and M. Teboulle. <i>SIAM Journal on Imaging Sciences, Volume 2, 2009.</i> • <i>Convex Optimization.</i> S. Boyd and L. Vandenberghe. Cambridge University Press, 2004.