New Mexico State University
Klipsch School of Electrical Engineering

EE565 Pattern Recognition and Machine Learning
Fall 2016 – Project #6
Due: 5:00pm Thu. Nov. 10

Name: ______________________

Grade: ______________________

Project

The goal of this project is to gain experience with

- kernel methods (Chapter 6)
- support vector machine (SVM) (Chapter 7)
- relevance vector machine (RVM) (Chapter 7)

for regression and classification problems. Several toolboxes can be used for this project.

For MATLAB, please see the Statistics toolbox and the Statistical Pattern Recognition (STPR) toolbox

`http://cmp.felk.cvut.cz/cmp/software/stprtool/`

for SVM functions and

`http://www.vectoranomaly.com/downloads/downloads.htm`

for RVM functions.

For Python, please see scikit-learn for SVM functions and

`https://github.com/JamesRitchie/scikit-rvm`

(or other available library) for RVM functions.

Companion files for this project may be found at

`http://www.ece.nmsu.edu/~pdeleon/Teaching/EE565/Projects/CompanionFiles6.zip`

Report

Please submit a printed report with your results including commentary and plots. Please `mailto:pdeleon@nmsu.edu` a zip file containing your report and all MATLAB code to recreate your results.

Notes

Students are encouraged to discuss detailed, technical aspects with each other and Prof. De Leon. Students are encouraged to utilize existing MATLAB functions in this project. However, students must write all other required codes on an *individual* basis.

1 Kernel Method for Regression

Recreate Fig. 1.4 (single plot) in the textbook, using the author's curvefitting.txt file containing the ten training points, but with a kernel-based regressor where $k(x_n, x_m) = \boldsymbol{\phi}^T(x_n)\boldsymbol{\phi}(x_m)$ and

$\boldsymbol{\phi}(x)$ is the feature space mapping; the specific kernels are listed below. You must implement (6.8) as part of the training code (set $\lambda = 10^{-4}$) and (6.9) as part of the test code.

a Linear kernel, $k(x_n, x_m) = x_n x_m$

b Kernel is constructed from a 3rd order polynomial feature mapping, $\phi_j(x) = x^j$ where $0 \leq j \leq 3$

c Kernel is constructed from a 9th order polynomial feature mapping, $\phi_j(x) = x^j$ where $0 \leq j \leq 9$

d Gaussian kernel, $k(x_n, x_m) = \exp(-|x_n - x_m|^2/2\sigma^2)$ where $\sigma^2 = 1$

e Radial basis function (RBF) or distance kernel, $k(x_n, x_m) = |x_n - x_m|$

f Provide a table indicating the RMS training and test errors for each of the above kernels.

2 Kernel Method for Classification

a Use the author's classification.txt file (2-D input) to train a Gaussian kernel-based binary classifier with $\sigma^2 = 1$. You must implement (6.8) as part of the training code. The classification data is described in Appendix A p. 682. Plot training data (denote $\mathcal{C}_1$ data with a blue 'o' and $\mathcal{C}_2$ data with a red 'x') and decision boundary.

b Evaluate classifier accuracy using 1000 test points generated with the supplied code. You must implement (6.9) as part of the test code. Plot test data (denote $\mathcal{C}_1$ decisions with a blue 'o' and $\mathcal{C}_2$ decisions with a red 'x') and decision boundary.

3 Support Vector Machine for Classification

a Use the author's classification.txt file (2-D input) to train a SVM classifier (Gaussian kernel) and (almost perfectly) recreate Figure 7.4. Report the number of support vectors and compare to the number of training data points.

b Evaluate the SVM classifier using the test set from Problem 2(b) and compare the results. Plot test data and denote $\mathcal{C}_1$ decisions with a blue 'x' and $\mathcal{C}_2$ decisions with a red 'x'.

4 Relevance Vector Machine for Classification

a Use the author's classification.txt file (2-D input) to train a RVM classifier (Gaussian kernel) and (almost perfectly) recreate Figure 7.12(a). Report the number of relevance vectors and compare to the result in Problem 3(a).

b Evaluate the RVM classifier using the test set from Problem 2(b) and compare the results. Plot test data and denote $\mathcal{C}_1$ decisions with a blue 'o' and $\mathcal{C}_2$ decisions with a red 'x'.

BONUS (up to +5) Plot the posterior probabilities given by the RVM output and (almost perfectly) recreate Figure 7.12(b).