CS6140 Machine Learning

HW6 - SVM and Kernels

Make sure you check the syllabus for the due date.

PROBLEM 1 SVM library [40 points]

A) Run an SVM from a package or library of your choice on the Spambase dataset. Try several kernels, including the polynomial and the RBF ones. Report the results. Use one of these packages: SVMlight, SGDSVM, osu SVM, LIBSVM, Matlab SVMtrain, or other software (here, here).

B) Run an SVM from a package or library of your choice on the Digits Dataset (Training data, labels. Testing data, labels). Use your extracted HAAR features from HW5. If you choose an SVM packge that does not provide multi-class implementations, you should write a wrapper that will make the code run for each class versus the others, separately.

PROBLEM 2 Implement your own SVM with SMO solver for Spambase [50 points]

Instead of using a SVM package, implement your own using SMO optimization and run on Spambase dataset with k-folds cross validation. Compare with Problem 1 results. Here is an SMO article other than the ones mentioned in class.

PROBLEM 3 Implement your own SVM with SMO solver for Digit Dataset [70 points]

Instead of using a SVM package, implement your own using SMO optimization and run it on the Digits Dataset (Training data, labels. Testing data, labels) with HAAR features extracted on HW5. Compare with Problem 1 results. The Digits dataset is very learnable, so to speed up the computation you can sample the training set at 10% or 20% per class (but make sure to use the entire testing set for measuring performance).

Since the data has a range of 10 labels (multiclass) while your SVM is a binary classifier, you will have to implement a wrapper on top of the SVM. You can choose one of the following:

- One-vs-the rest approach and train 10 SVM classifiers (one per class)
- Run ECOC on top of SVMs (similar with HW5 setup, only with SVM instead of boosting)
- We suggest a voting schema: train all possible one-to-one SVM classifiers, for a total of (10 choose 2) = 45 models. Each one of these will train/test only on labeled data for the two particular classes is made

for : for example 7vs9 SVM will only train/test on datapoints labeled 7 or 9. To obtain a multiclass classifier: first run (for a given test-datapoint) all 45 models and get their scores; then you would need a voting strategy in order to decide a prediction or a ranking among all 10 classes. Such voting strategy can be to predict the class with most wins, and if there is tie for the most wins to use the direct "match" one-to-one to break the tie.

PROBLEM 4 [20 points]

Explain why $0 \leq \alpha \leq C/m$ is a constraint in the dual optimization with slack variables. (HINT: read Chris Burges tutorial first) Distinguish three cases, and explain them in terms of the classification and constraints: a) $0 = \alpha$; b) $0 < \alpha < C/m$; c) $\alpha = C/m$.

This has been discussed in class and in SVM notes; a detailed rigurous explanation is expected.

PROBLEM 5 [20 points]

Consider the following 6 points in 2D, for two classes:

class 0: (1,1) (2,2) (2,0)

class 1: (0,0) (1,0) (0,1)

a) Plot these 6 points, construct the optimal hyperplane by inspection and intuition (give the W,b) and calculate the margin.

b) Which points are support vectors ?

c) [Extra Credit] Construct the hyperplane by solving the dual optimization problem using the Lagrangian. Compare with part (a).

PROBLEM 6 Implement better SMO [extra credit]

Extra points will be given for an SMO implementation for both PB2 and PB3 that is reasonable fast.
Extra points will be given for an implementation for both PB2 and PB3 that works with kernels (for example Gaussian Kernel)

PROBLEM 7 [Extra Credit]

Run your SMO-SVM on other datasets.

PROBLEM 8 [Extra Credit]

Same problem as 2, but dont use SMO. Instead use a built in (or existing library) quadratic solver from Matlab, Python, Java, C etc, inn order to solve the dual problem.

PROBLEM 9 [Extra Credit]

What is the VC dimension of the SVM with linear kernel ?