MLCV: Programming Assignement 2

Mohamed N'AITN'BARK

November 18, 2014

1 Support Vector Machine classifiers:
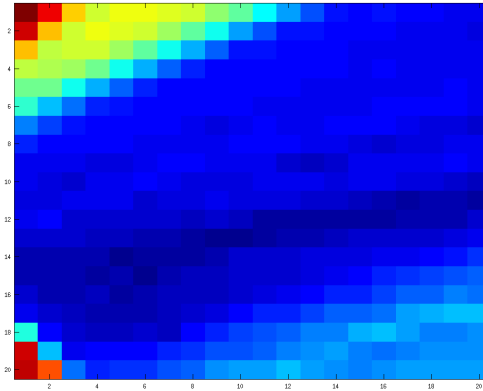

Figure 0.1: Misclassification error as a function of γ(Kernel bandwith) and Cthe cost.
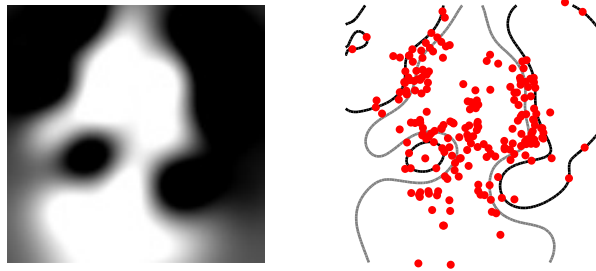

Figure 0.2: Right: The posterior dustribution of our data, Left: SVM support vectors and boundary decision.
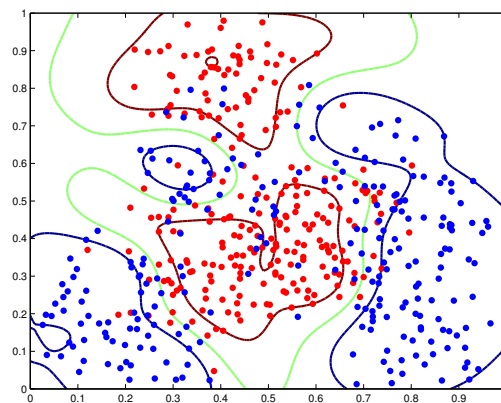

Figure 0.3: Labeled data and SVM separator.

Using a SVM model with soft margin and using a gaussian kernel, we find using the 10-fold cross-validation that the best estimation is made with $\gamma = 33{,}60$ and C = 4,83. The misclassification error on the training set is then 14,8%.

On the test set, the cross validated model commits 90 errors with error rate of 18%.

2 Adaboost

strong learner – round 10 strong learner – round 50 strong learner – round 100strong learner – round 400
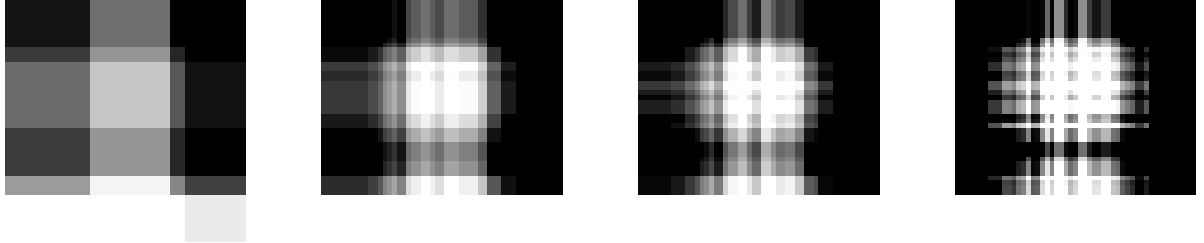

Figure 0.4: Strong learner scores displayed after the 10th, 50th, 100th and 400th iteration.
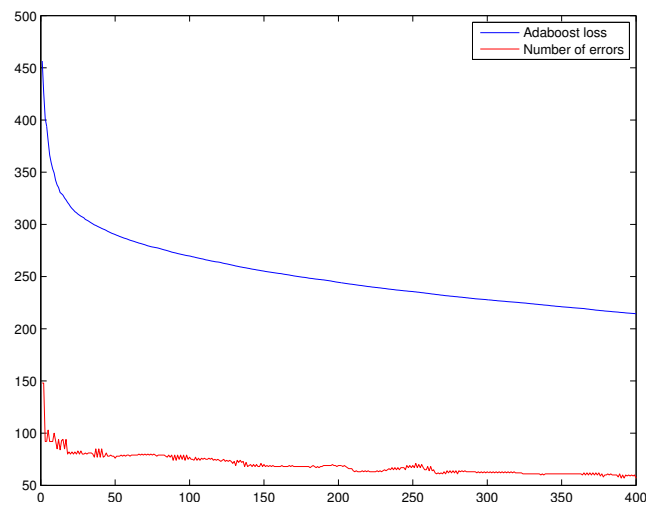

Figure 0.5: Loss function of Adaboost (exponential lost) at each iteration. As expected its value is decreasing round and upper bounds the number of errors at each round.
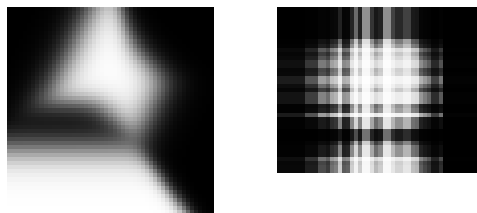

Figure 0.6: Right: The posterior dustribution of our data, Left: The distribution given by Adaboost.

4

3 SVM vs Adaboost

	Adaboost	SVM
Training Set	11,4%	14,8%
Test Set	29,8%	18%

Table 0.1: Misclassification error by model (SVM/Adaboost) and set (training/test).

The Adaboost model performs better than SVM on the training test and less on the test set, this suggests that Adaboost overfits the data.

4 Fast week learner selection

For weak learners of the form: $h_{\theta,s}(x) = s(2[x \geq \theta] - 1)$ where $s \in \{-1, 1\}$. We can implement a faster algorithm by sorting our features and using the proposed recursion. In matlab the naive algorithm requires 26 seconds while the fast implementation takes only 390 ms, the ratio is 54.

Algorithm 0.1

For each $s \in \{-1, 1\}$:

For each dimension d of our features space:

Compute π the sorting permutaion of the dimension d

We form the vector $V = (v_i)_i$ s.t. $v_0 = \sum_i w_{\pi(i)}[y_{\pi(i)} = -s]$ and $\forall i \in [1, n-1] : v_i = sw_{\pi(i)}y_{\pi(i)}$

Then if $W = (w_i)_i$ is the cumulative sum of the vector V the error commited for $h_{(s, \theta_{\pi(i)})}$ is w_i

---