ECE523 Engineering Applications of Machine Learning and Data Analytics
Homework 4

1 Support Vector Machines (Revisited)

$$\text{Ł}(W_T, b, \xi, \alpha, r) = \frac{1}{2}\|W_T\|^2 + C\sum_{i=1}^{n}\xi_i^2 - BW_T^T W_S - \sum_{i=1}^{n} a_i[y_i(W_T^T X_i + b) + \xi_i - 1] - \sum_{i=1}^{n} r_i\xi_i$$

$$\frac{\partial \text{Ł}}{\partial W_T} = 0 \quad \Rightarrow \quad W_T = BW_S + \sum_{i=1}^{n}\alpha_i y_i X_i \tag{1}$$

$$\frac{\partial \text{Ł}}{\partial b} = 0 \quad \Rightarrow \quad \sum_{i=1}^{n}\alpha_i y_i = 0 \tag{2}$$

$$\frac{\partial \text{Ł}}{\partial \xi_i} = 0 \quad \Rightarrow \quad C - \alpha_i - r_i = 0 \tag{3}$$

Substitute (1), (2), (3) into Ł, We will get:

$$\underset{\alpha}{\arg\max} \ \sum_{i=1}^{n}\alpha_i(1 - By_i X_i^T W_S) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j X_i^T X_j$$

$$s.t. \quad 0 \le \alpha_i \le C \quad \forall i \in [n] \quad and \quad \sum_{i=1}^{n}\alpha_i y_i = 0$$

2 AdaBoost: How much do you understand the algorithm?

Schapire first proved that strongly learnable is equivalent to weakly learnable.

In the AdaBoost algorithm, after each round, we are going to increase the weight of the points which we labelled a wrong class. Such that in the next round, the classifier is more likely to fix this mistake.

Further more, the error of AdaBoost has an upper bound:

$$2^T \prod_{t=1}^{T} \sqrt{\epsilon_t(1 - \epsilon_t)}$$

Because $\epsilon_t < \frac{1}{2}$, when T increases, this upper bound will decrease. Therefore, boosting works.

3 Support Vector Machines (Revisited)

The code for this problem is in directory ./code/SVM.m
The source hyperplane parameters:

$$W_S = [1.9279, 1.8208] \qquad b = -0.6746$$

The target hyperplane parameters:

$$\boldsymbol{W}_T = [2.3150, 2.1426] \qquad b = -0.7971$$

The classification error for target_train is 0.060.
The classification error for target_test is 0.035.

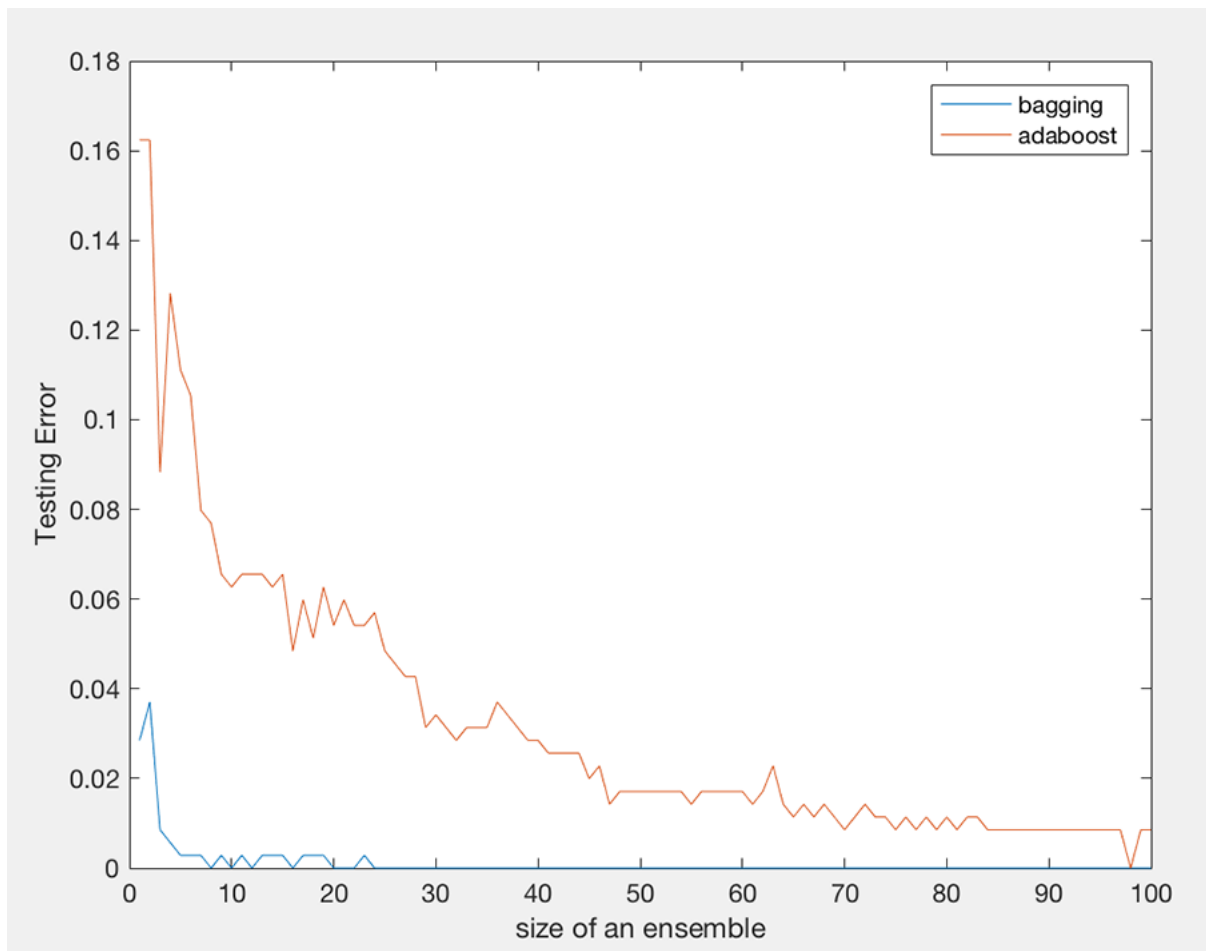# 4   An Experiment with Ensembles

The code for this problem is in directory ./code/ensemble.m

Fig.1 Testing error with ensemble size