

## Running Adaboost on Naive Bayes

$D_1$ : Assign uniform probabilities to training data  $S$ .

For  $t = 1$  to  $NumBoostingIterations$  {

    Sample 300 examples from training data with replacement,  
    according to distribution  $D_t$  to get  $S_t$

    Train Naive Bayes on that  $S_t$  to get (binary) classifier  $h_t$

    Calculate error rate  $\epsilon_t$  (fraction of misclassifications) of  $h_t$  on entire  
    training set  $S$ .

    Calculate  $D_{t+1}$  for each example in the entire training set  $S$ .

}

Output final classifier.  
Test on test set.

## Adaboost Formulas

$\epsilon_t$  = Fraction of misclassifications of  $h_t$  over  $S$

$$\alpha_t = \frac{1}{2} \ln \left( \frac{1 - \epsilon_t}{\epsilon_t} \right)$$

$$D_{t+1}(i) = \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

$$Z_t = \sum_{i=1}^{|S|} D_t(i) \exp(-\alpha_t y_i h_t(x_i))$$