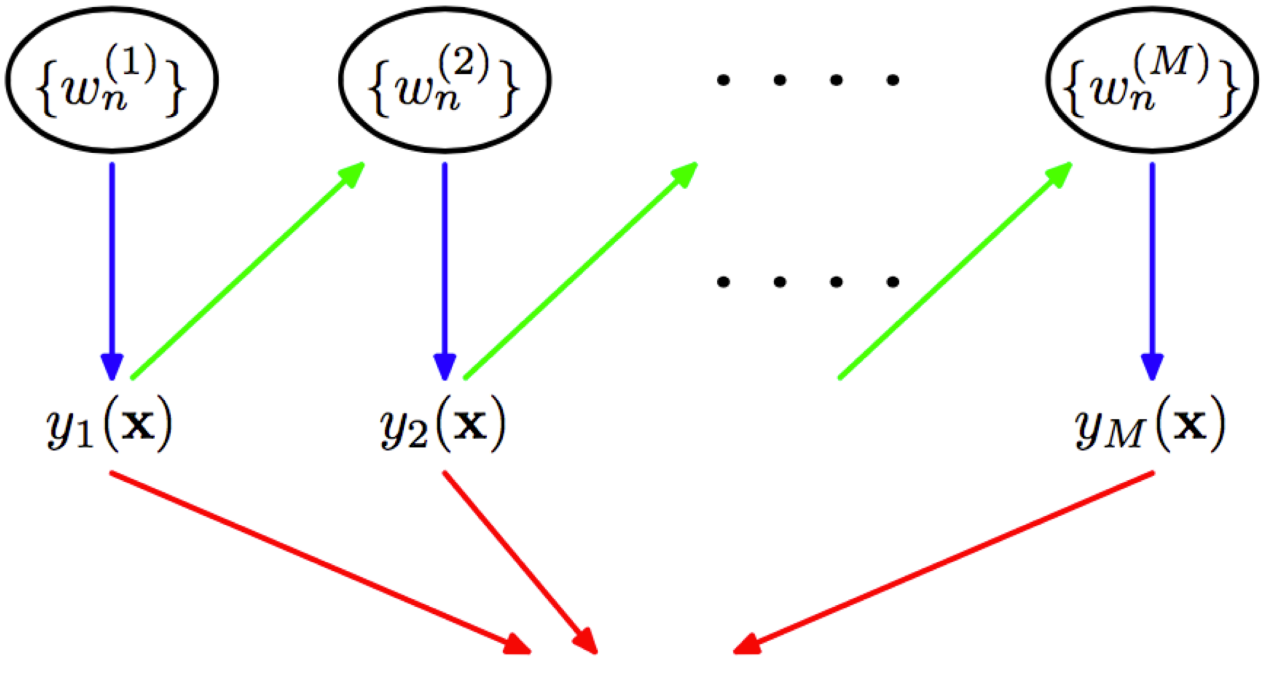
AdaBoost

AdaBoost

Each base classifier $y_m(\mathbf{x})$ is trained on a weighted form of the training set (blue arrows) in which the weights $w_n^{(m)}$ depend on the performance of the previous classifier $y_{m-1}(\mathbf{x})$ (green arrows). Once all base classifiers are trained, they are combined to give the final classifiers $Y_M(\mathbf{x})$ (red arrows).



$$Y_M(\mathbf{x}) = \operatorname{sign}\left(\sum_{m}^{M} \alpha_m y_m(\mathbf{x})\right)$$

Figure from: Christopher Bishop, “Pattern Recognition and Machine Learning”, Springer, 2006

AdaBoost

1. Initialize $w_n^{(1)} = \frac{1}{N}$ for $n = 1, 2, \ldots, N$
2. For $m = 1, 2, \ldots, M$
   a. Fit a classifier to the training data by minimizing the weighted error function

$$J_m = \sum_{n=1}^{N} w_n^m I(y_m(\mathbf{x}_n) \neq t_n)$$

where

$$I(y_m(\mathbf{x}_n) \neq t_n) = \begin{cases} 1 & \text{if} \quad y_m(\mathbf{x}_n) \neq t_n \\ 0 & \text{otherwise} \end{cases}$$

AdaBoost

b. Evaluate

$$\epsilon_m = \frac{\sum_{n=1}^{N} w_n^m I(y_m(\mathbf{x}_n) \neq t_n)}{\sum_{n=1}^{N} w_n^m}$$

and

$$\alpha_m = \log(\frac{1-\epsilon_m}{\epsilon_m})$$

c. Update the weighting

$$w_n^{(m+1)} = w_n^{(m)} \exp\{\alpha_m I(y_m(\mathbf{x}_n) \neq t_n)\}$$

3. Make predictions using the final model given by:

$$Y_M(\mathbf{x}) = \mathrm{sign}(\sum_{m=1}^{M} \alpha_m y_m(\mathbf{x}))$$

AdaBoost

Observations:

- We start by giving equal weights to all data points but then give greater emphasis to points that keep being misclassified
- ϵ_m represents the weighted measures of the error rates of each of the base classifier on the data set
- α_m give greater weight to the more accurate classifiers.