Boosting

Boosting

- Base classifiers are trained in sequence
- Misclassified data points get added weight in the training of subsequent classifiers.
- Once all classifiers are trained, their predictions are combined by a weighted majority scheme.

Example of boosting

- AdaBoost – (Adaptive Boosting) : based on a set of linear models
- XGboost : based on a set of decision trees

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).

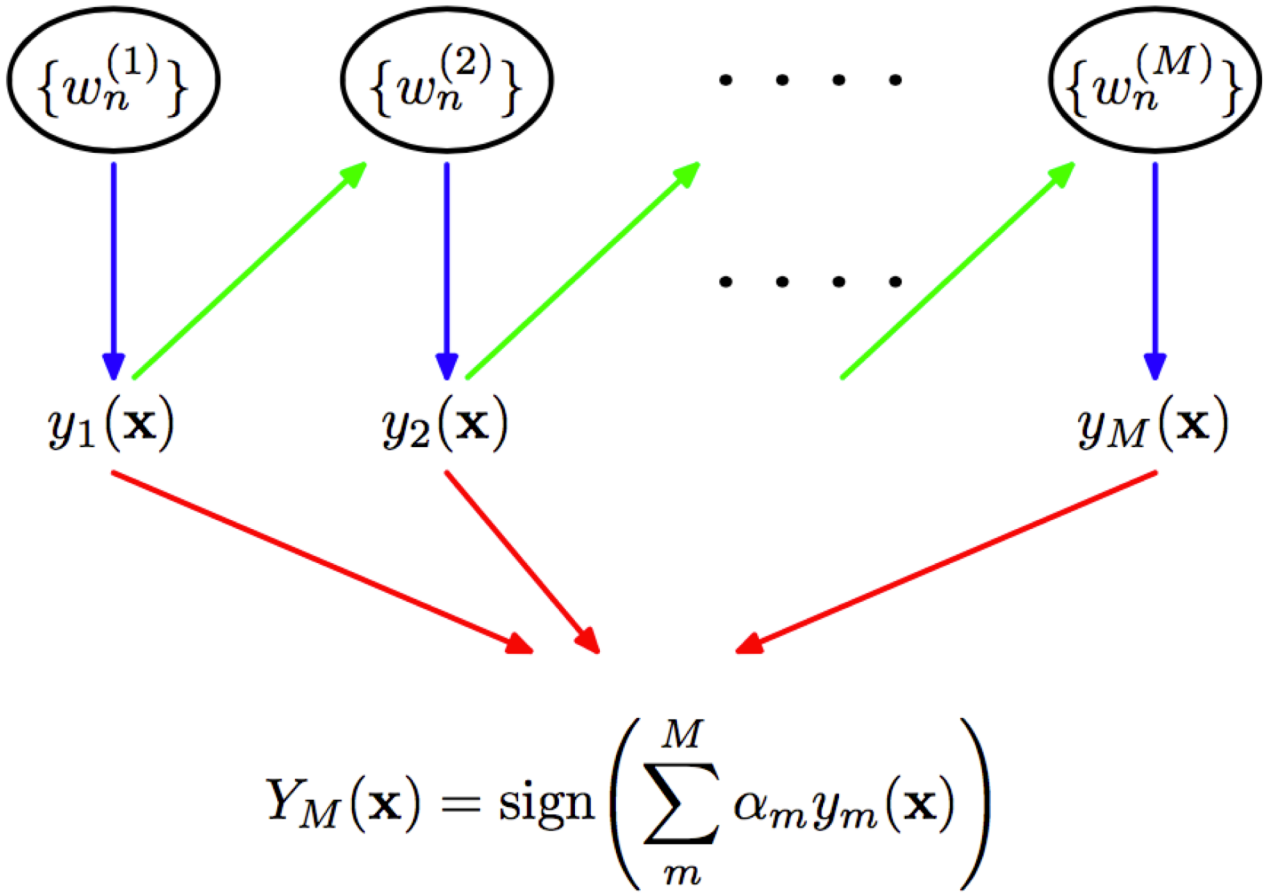

Figure from: Christopher Bishop, "Pattern Recognition and Machine Learning", Springer, 2006