

A note on ensemble learning

Yapi Donatien Achou

December 14, 2018

1 Introduction

Ensemble learning models are obtained by combining simple base models in order to obtain a much more performant model.

2 Bagging methods

3 Boosting methods

3.1 AdaBoost

The idea of AdaBoost is to train a weak classifier on weighted versions of the original training dataset and make the final prediction based on weighted majority vote. Let $G_m, m = 1, 2, \dots, M$ be a weak classifier defines such that its performance is slightly worse or better than random guess. Let D_1 be the training sample and D_2, D_3, \dots, D_M be weighted samples obtained from the original training data set. By sequentially applying G_1, G_2, \dots, G_M to the data set, D_2, D_3, \dots, D_M , a much better model G is obtained by combining G_1, G_2, \dots, G_M through a weighted majority vote such that

$$G(x) = \text{sign} \left(\sum_{m=1}^M \alpha_m G_m(x) \right) \quad (1)$$

where the weights $\alpha_m, m = 1, 2, \dots, M$ are obtained from the boosting algorithm and weight the contribution of each G_m [3]. At iteration m , misclassified training examples by classifier G_{m-1} have their weight increased, while correctly classified training examples have their weight decrease. This insures that training examples that are difficult to classify receive more attentions, by making sure that the next classifier concentrates on the training examples that are missed by previous one [3]. This procedure improves the performance of the weak classifier at each iteration.

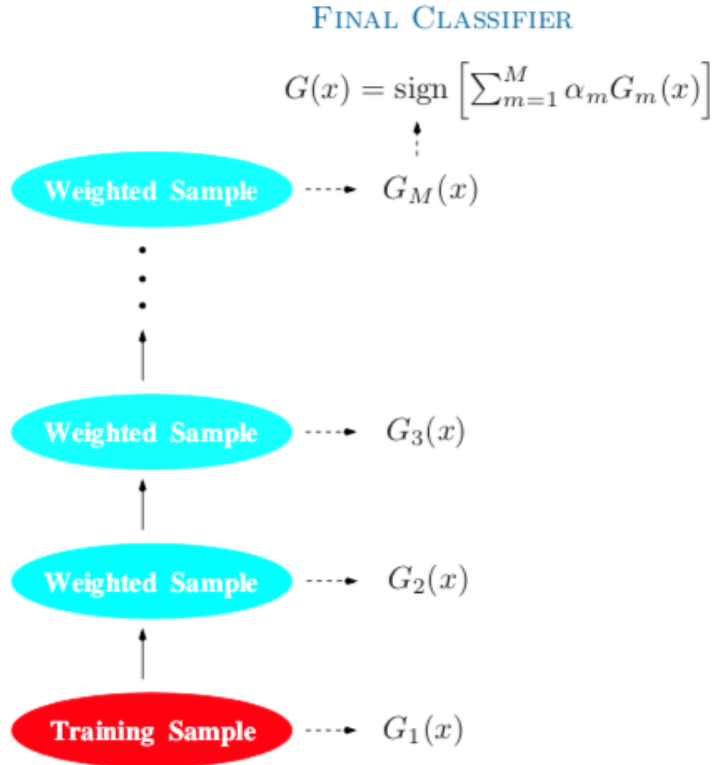


Figure 1: Schematic description of AdaBoost algorithm. Picture taken from [3]

Figure 1 shows the description of AdaBoost

References

- [1] Chollet, François and others *Keras*, <https://keras.io>. 2015
- [2] Ian Goodfellow and Yoshua Bengio and Aaron Courville *Deep Learning*. <http://www.deeplearningbook.org>, 2016
- [3] Trevor Hastie, Robert Tibshirani, Jerome Friedman *The Elements of Statistical Learning*. Springer Series in Statistics. 2008