



## CHAPTER 1

---

# Boosting

Boosting is a machine learning ensemble meta-algorithm primarily used to reduce bias, and to a lesser extent variance, in supervised learning. It works by iteratively learning weak classifiers and adding them to a final strong classifier in a way that the subsequent weak learners try to correct the mistakes of the previous ones.

### 1.1 AdaBoost

AdaBoost, short for Adaptive Boosting, is one of the first and simplest boosting algorithms. Given a set of  $n$  training examples  $(x_1, y_1), \dots, (x_n, y_n)$  where  $y_i$  are binary outputs, the algorithm works as follows:

1. Initialize weights  $w_i = 1/n$  for  $i = 1, \dots, n$ .
2. For  $t = 1$  to  $T$ :
  - Train a weak learner  $h_t$  using the weighted examples.
  - Compute the weighted error  $\epsilon_t = \sum_{i: h_t(x_i) \neq y_i} w_i$ .

- Set  $\alpha_t = \frac{1}{2} \log \left( \frac{1-\epsilon_t}{\epsilon_t} \right)$ .
  - Update the weights:  $w_i = w_i \exp(-\alpha_t y_i h_t(x_i))$  for  $i = 1, \dots, n$ , and normalize them so that they sum to one.
3. The final model is  $H(x) = \text{sign} \left( \sum_{t=1}^T \alpha_t h_t(x) \right)$ .

## 1.2 Gradient Boosted Trees

Gradient Boosted Trees is a generalization of boosting to arbitrary differentiable loss functions. It works by sequentially adding predictors to an ensemble, each one correcting its predecessor by fitting the new predictor to the residual errors.



## APPENDIX A

---

# Answers to Exercises





---

# INDEX

AdaBoost, [1](#)

Boosting, [1](#)

Gradient Boosted Trees, [2](#)