Ensemble Methods

Christos Dimitrakakis

November 28, 2023

Bagging

Algorithm

- Input: Data D, bags K, base learner λ
- For $k = 1, \ldots, K$

– Sample with replacement $D_k \sim \mathrm{Unif}(D)$ – Obtain predictor $\pi_k = \lambda(D_k)$.

- Return $\{\pi_k\}$

The bagged predictor

$$\pi = f\left(\sum_k \pi_k\right)$$

Bagging classifiers

Classification setting

- Weak learner $\lambda : D \to \Pi$
- Base hypotheses $\pi_k : X \to \{-1, 1\}$

with

$$\pi_k = \lambda(D_k), \qquad D_k \sim D$$

- Aggregate hypothesis

$$\pi(x) = \text{sgn}\left(\sum_{k=1}^{K} \pi_k(x)\right)$$

PAC property

For any $\delta \in (0, 1)$, and any $\pi^* : X \to \{-1, 1\}$ and a hypothesis class Π with with VC dimension d, for T data points, and $K \in [0.02T, T]$ bootstrap samples, then

$$\mathbb{L} \in O\left(\frac{1}{T}[d + \ln(1/\delta)]\right), \qquad \text{w.p.} 1 - \delta.$$

Sub-sample-and-aggregate

Algorithm

- Input: Data D, number of experts K, base learner λ
- For $k = 1, \ldots, K$

– Sample without replacement $D_k \sim \mathrm{Unif}(D)$ – Obtain predictor $\pi_k = \lambda(D_k)$.

- Return $\{\pi_k\}$

The aggregated predictor

$$\pi = f\left(\pi_1, \ldots, \pi_k\right)$$