Ensemble Learning

03/10/2023

Error of an Algorithm: A

$$E[(h_D(x) - y)^2] = E[(h_D(x) - \bar{h}(x))^2] + E[(\bar{h}(x) - \bar{y}(x))^2] + E[(\bar{y}(x) - y(x))^2]$$

Variance — Bias^2 — Noise

⇑ minimize variance

$h_D(x) \rightarrow \bar{h}$

Weak law of large number (WLLN)

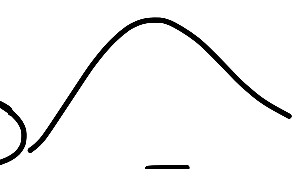
$\frac{1}{m}\sum_{i=1}^{m} x_i \rightarrow \bar{x}$

$m \rightarrow \infty$

m

x_i ← i.i.d

$i = 1 .. m$



$\bar{x}$

$D_i \sim P^n$

$i \rightarrow \infty$

$D \sim P^n$ $P(x, y)$

$\downarrow$

h_D

$D_i \sim P^n$ $D_1, D_2, \dots D_m$

$\mathcal{H}$ = hypothesis space

$\hat{h} = \frac{1}{m}\sum_{i=1}^{m} h_{D_i} \rightarrow \bar{h}$ $m \rightarrow \infty$ $\bar{h}$ $|D_i| = n$

Simulated Sampling

$|D| = n$ P^n $P(x, y)$

Given D $Q(x, y | D) = \frac{1}{n}$

pick a training sample from D uniformly at random following $Q \rightarrow$ $D_i \sim D$

$D_i \sim Q^n$

Sample with replacement

$|D_i| = n$

$D_1, D_2, \dots, D_m \sim Q^n$

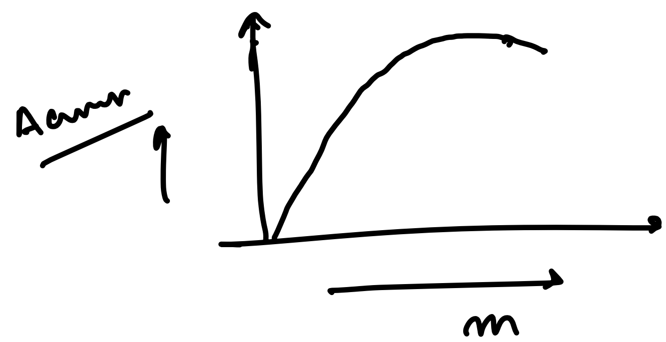
$$\hat{h}_D = \frac{1}{m}\sum_{i=1}^{m} h_{D_i} \not\to \bar{h} \quad m \to \infty$$

in practice, $(\hat{h}_D - \bar{h})^2 \to 0$

D = large (n)
$m \to$ resample large.

BAGGING
Bootstrap Aggregating.

$m \to \infty$
diminishing return



Adv

1) Easy to implement
2) Reduces variance
3) mean, variance → uncertainty of prediction

$1\sigma \quad \pm 2\sigma$

4) unbiased estimate of the test error.

training instance $(x_i, y_i) \in D$

$$S_i = \{ D_k \mid (x_i, y_i) \notin D_k \} \quad k = 1, \ldots, m$$

$$\bar{h}_i(x) = \left(\sum_{k \in S_i} h_k(x) \right) \frac{1}{|S_i|}$$

$h_i(x_i) \to$ Prediction

$$\epsilon_{OOB} = \frac{1}{n} \sum_{(x_i, y_i) \in D} l\left(\bar{h}_i(x_i), y_i \right)$$

out of Bag error

Random Forest

d = no. of features

1) $D_i \sim D$
2) $k < d$
3) $DT(k, D_i)$
