Pranava Raman BMS

Machine Learning

2019103555

29/11/2021

Week-13-Obs

Aim

To implement ensemble classifier

Adaboost algorithm

- Init all weights → $\left[\frac{1}{N}\right]$, N = No. of datapoints

→ while $0 < \varepsilon_t < \frac{1}{2}$ [and $t < T$]:-

- train classifier on $\{S, w^{(t)}\}$, getting hypotheses $h_t(x_n)$

- compute error

$$\varepsilon_t = \sum_{n=1}^{N} w_n^{(t)} I(y_n \neq h_t(x_n))$$

- set $\alpha_t = \log\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)$

- Update weights using:-

$$w_n^{(t+1)} = w_n^{(t)} \exp(\alpha_t I(y_n \neq h_t(x_n)))/Z_t$$

Z_t normalization constant

- Output

$$f(x) = \operatorname{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$$

Bagging classifier

classifier generation:-

-> Let N be the size of training set

-> for each of t iterans:-

* sample N instances with replacement from the original set
* apply Learning algorithm on sample
* Stor resulting classifie.

classification:-

for each of the classifier:-

predict class of instance using classifier

return class that was predicted most often

Random forest Training algorithm

-> For each of N trees:-

* create new bootstrap sample of the Training set
* use this bootstrap to train decision tree
* at each node of decision tree
* randomly select m features
* compute information gain only on the set of features,
* select the optimal feature
* Repeat until tree is completed.