South China University of Technology

The Experiment Report of Machine Learning

SCHOOL: SCHOOL OF SOFTWARE ENGINEERING

SUBJECT: SOFTWARE ENGINEERING

Author:
钟恩俊 王泽众 诸俊浩

Supervisor:
Mingkui Tan

Student Name：
201530613818
201530612958
201530613962

Grade:
Undergraduate

December 22, 2017

Face Classification Based on AdaBoost Algorithm

Abstract—In this experiment, we will study about the Adaboost algorithm and use Adaboost to solve the face classification problem, and combine the theory with the actual project.

I. Introduction

To understand AdaBoost algorithm further, we use AdaBoost algorithm to classify face picture and non-face picture. We choose DecisionTreeClassifier as weak classifier ,implementing an adaboost algorithm which can improve the accuracy of base learners by combining them with different weight. During this experiment,we try to conduct the adaboost algorithm,and

II. Methods and Theory

Example algorithm

With

Samples:

$$x_1 , x_2 \ldots x_{\mathrm{n}}$$

Desired outputs:

$$y_1, y_2 \ldots y_n, y \in \{-1,1\}$$

Initial weights :

$$\omega_1, \omega_2 \ldots . \omega_n \quad set\ to\ \frac{1}{n}$$

Weak learners:

$$h_m(X): X \to \{-1,1\}$$

Error function:

$$\mathrm{E(f(x), y, i)} = e^{-y_i f(x_i)}$$

For t in 1….T:

Find weak learner $\mathrm{h_t}(x)$ that minimizes ϵ_{t}, that

$$\epsilon_{\mathrm{t}} = \sum_{i=1}^{n} \omega_{i,t} \quad (h_t(x_i) \neq y_i)$$

Choose α_{t}:

$$\alpha_{\mathrm{t}} = \frac{1}{2} \ln(\frac{1-\varepsilon_t}{\varepsilon_t})$$

Update weight ω_m of sample in the m iteration

$$\omega_{m+1}(i) = \frac{\omega_m(i)}{z_m} e^{-\alpha_t y_i h_t(x_i)}, where\ i = 1,2,3 \ldots, n\ and$$

$$z_m = \sum_{i=1}^{n} \omega_m(i) e^{-\alpha_t y_i h_t(x_i)}$$

Final stronge learner

$$H(X) = sign(\sum_{t=1}^{T} \alpha_t h_t(X))$$

Choosing α_{t}

α_{t} is chosen as it can be analytically shown to be the minimizer of the exponential error function for Discrete AdaBoost.

Minimize:

$$\sum_i \omega_i e^{-y_i h_i a_t}$$

Using the convexity of the exponential function, and assuming that $\forall \mathrm{i}, \mathrm{h_i} \in [-1,1]$ we have :

$$\sum_i \omega_i e^{-y_i h_i a_t} \sum_i \left(\frac{1-y_i h_i}{2}\right) \omega_i e^{a_t} + \sum_i \left(\frac{1+y_i h_i}{2}\right) \omega_i e^{-a_t}$$

$$= \left(\frac{1+\epsilon_{\mathrm{t}}}{2}\right) e^{a_t} + \left(\frac{1-\epsilon_{\mathrm{t}}}{2}\right) e^{-a_t}$$

We then differentiate that expression with respect to α_{t} and set it to zero to find the minimum of the upper bound:

$$\left(\frac{1+\epsilon_{\mathrm{t}}}{2}\right) e^{a_t} + \left(\frac{1-\epsilon_t}{2}\right) e^{-a_t} = 0$$

Got :

$$\alpha_{\mathrm{t}} = \frac{1}{2} \ln(\frac{1-\varepsilon_t}{\varepsilon_t})$$

III. Experiment

Dataset

This experiment provides 1000 pictures, of which 500 are human face RGB images, stored in datasets/original/face; the other 500 is a non-face RGB images, stored in datasets/original/nonface.

Implementation

Algorithm : AdaBoost

Input: D={$(x_1, y_1), \ldots, (x_n, y_n)$}where $x_i \in X, y_i \in \{-1,1\}$

Initialize: Sample distribution ω_m

Base learner: $\mathcal{L}$

$\omega_1(i) = \frac{1}{n}$

$for\ m = 1,2,\ldots,M\ do$

$h_m(x) = \mathcal{L}(\mathrm{D}, \omega_m)$

$\epsilon_m = p(h_m(X_i) \neq y_i) = \sum_{i=1}^{n} \omega_m(i) I(h_m(X_i) \neq y_i)$

If $\epsilon_m > 0.5\ then$

Break

end

$\alpha_m = \frac{1}{2} \log \frac{1-\epsilon_m}{\epsilon_m}$

$\omega_{m+1}(i) = \frac{\omega_m(i)}{z_m} e^{-\alpha_m y_i h_m(x_i)}, where\ i =$
$1,2,3 \ldots, n\ and\ z_m = \sum_{i=1}^{n} \omega_m(i) e^{-\alpha_m y_i h_m(x_i)}$

end

Output: $H(X) = sign(\sum_{m-1}^{m} \alpha_m h_m(X))$

Result：

	precision	recall	f1-score	support
-1.0	0.89	0.95	0.92	168
1.0	0.94	0.88	0.91	162
avg / total	0.92	0.92	0.92	330

IV. CONCLUSION

Boosting can be seen as minimization of a convex loss function over a convex set of functions.Each iterations of AdaBoost make a simply classify and make the data which is classified wrong more important in the next iteration. Make the base learner with high accuracy is more important.