Ensemble

> Pattern Recognition Homeworks
>
> Student: the-zxs

Solutions

Problem 1 - Boosting

1.1

$$E = e^{-\alpha_m/2} \sum_{n\in\Gamma_m} w_n^{(m)} + e^{\alpha_m/2} \sum_{n\in\mathcal{M}_m} w_n^{(m)}$$

let:

$$\frac{\partial E}{\partial \alpha_m} = -\frac{1}{2} e^{-\alpha_m/2} \sum_{n\in\Gamma_m} w_n^{(m)} + \frac{1}{2} e^{\alpha_m/2} \sum_{n\in\mathcal{M}_m} w_n^{(m)} = 0$$

resulting in:

$$\begin{aligned}
\alpha_m &= \ln \frac{\sum_{n\in\Gamma_m} w_n^{(m)}}{\sum_{n\in\mathcal{M}_m} w_n^{(m)}} \\
&= \ln \frac{\sum_{n=1}^{N} w_n^{(m)} I(y_m(\mathbf{x}_n) = t_n) / \sum_{n=1}^{N} w_n^{(m)}}{\sum_{n=1}^{N} w_n^{(m)} I(y_m(\mathbf{x}_n) \neq t_n) / \sum_{n=1}^{N} w_n^{(m)}} \\
&= \ln \frac{1-\epsilon}{\epsilon}
\end{aligned}$$

1.2

Given:

$$w_n^{(m+1)} = w_n^{(m)} \exp\{-\frac{1}{2} t_n \alpha_m y_m(\mathbf{x}_n)\}$$

which is:

$$w_n^{(m+1)} = w_n^{(m)} \exp\{\alpha_m I(y_m(\mathbf{x}_n) \neq t_n)\}$$

Where $I(true) = 1$ and $I(false) = -1$

Programming

2.1

```
minErr = inf;
[n, p] = size(X);
epsilon = 1e-5;
for dim = 1:p
    [X_cur, index] = sort(X(:, dim));
    y_cur = y(index);
    w_cur = w(index);

    for d_cur = [1, -1]
        match = (-d_cur*ones(n, 1)) ~= y_cur;

        for j = 0:n
            if j ~= 0
                match(j) = ~match(j);
            end
            if w_cur.'*match < minErr
                minErr = w_cur.'*match;
                k = dim;
                if j == 0
                    a = X_cur(1) - epsilon;
                else
                    a = X_cur(j);
                end
                d = d_cur;
            end
        end
    end
end
```

the overall time complexity is $O(p\log(n)n)$

2.2

`update_weights.m`

```
err = 2*d*((X(:, k) <= a)-0.5) ~= y;
index = 2*(err-0.5);

w_update = w.*exp(alpha*index);
```

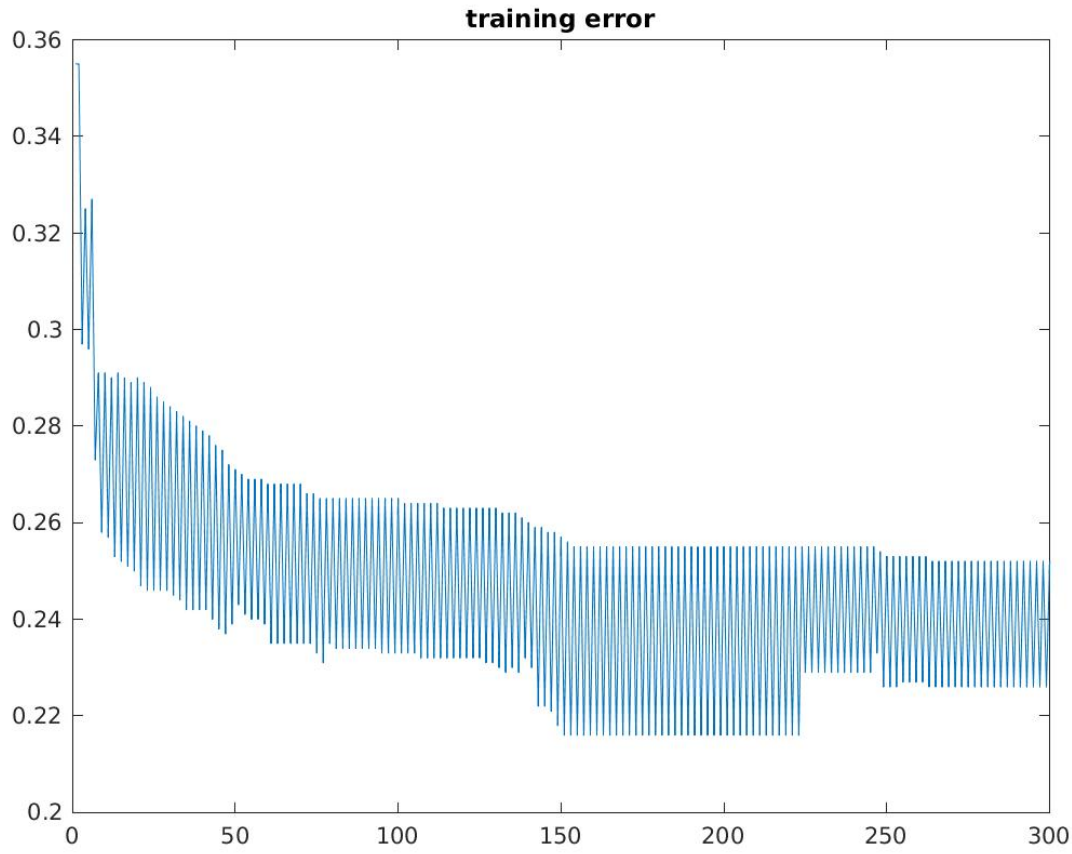
`adaboost_error.m`

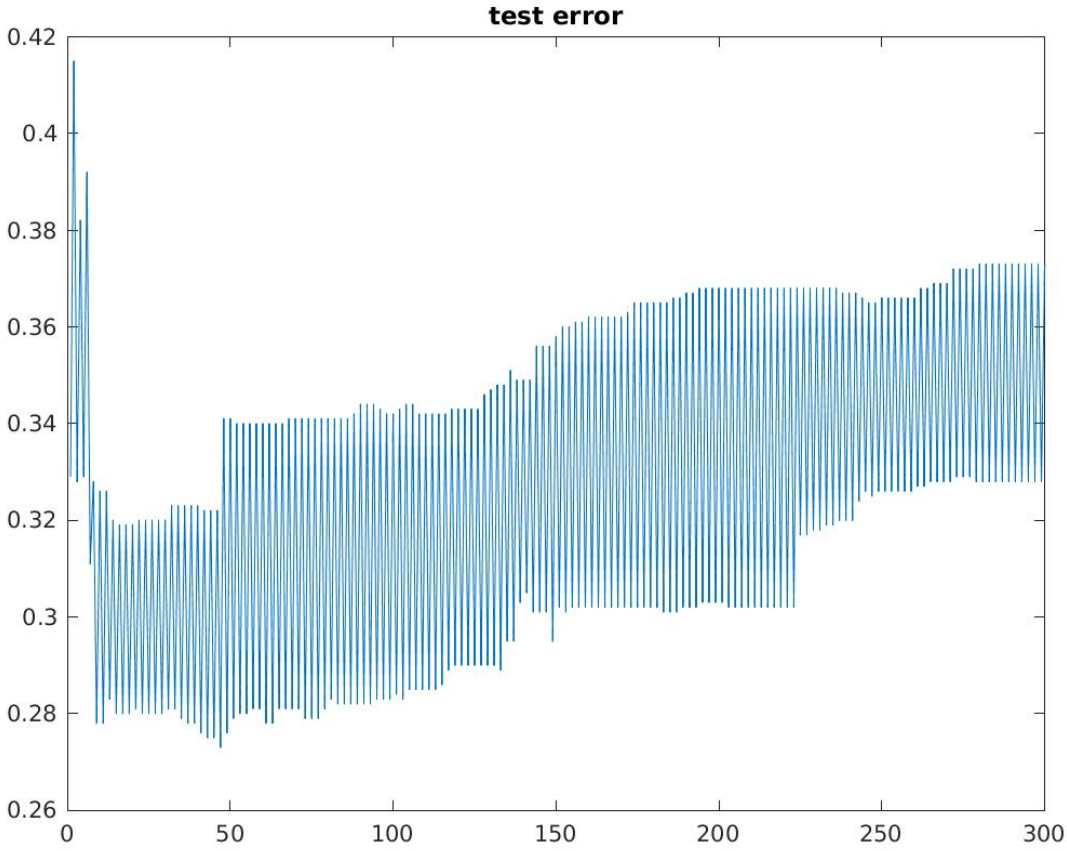
```
for i = 1:length(d)
    if d(i) == 0
        n = i-1;
        break;
    end
```

```
end

p = zeros(length(y),1);
if d(end) ~= 0
    n = length(d);
end

for i = 1:n
    p = p + 2*alpha(i)*d(i)*((X(:, k(i)) <= a(i))-0.5);
end

p = 2*((p>0)-0.5) ~= y;
e = sum(p)/length(y);
```

2.3

```
>> [e_train, e_test] = adaboost(X_train, y_train, X_test, y_test, 300);
>> plot(1:300, e_train');
>> title("training error");
>> saveas(gcf,'./train.jpg');
>> plot(1:300, e_test');
>> title("test error");
>> saveas(gcf,'./test.jpg');
```




test error
0.42
0.4
0.38
0.36
0.34
0.32
0.3
0.28
0.26
0
50
100
150
200
250
300