10-701/15-781, Machine Learning: Homework 5

Eric Xing, Tom Mitchell, Aarti Singh
Carnegie Mellon University
Updated on March 24, 2010

- The assignment is due at 10:30am (beginning of class) on **Mon, April 26, 2010**.
- Separate you answers into three parts, one for each TA, and put them into 3 piles at the table in front of the class. Don't forget to put both your name and a TA's name on each part.

1 AdaBoost [Ni, 30 pt]

1.1 Sequential Optimization [5 pts]

Solutions:

$$
\begin{aligned}
E &= \sum_{i=1}^{N}(y_i - f_t(x_i))^2 && (1)\\
&= \sum_{i=1}^{N}\{y_i - f_{t-1}(x_i) - \alpha_t h_t(x_i)\}^2 && (2)\\
&= \sum_{i=1}^{N}\left\{r_i^2 + \alpha_t^2 h_t^2(x_i) - 2r_i\alpha_t h_t(x_i)\right\}, && (3)
\end{aligned}
$$

where $r_i = y_i - f_{t-1}(x_i)$.

$$
\frac{\partial E}{\partial \alpha_t} = 2\sum_{i=1}^{N}\left\{\alpha_t h_t^2(x_i) - r_i h_t(x_i)\right\} \quad (4)
$$

Therefore, $\alpha_t = \frac{\sum_{i=1}^{N} r_i h_t(x_i)}{\sum_{i=1}^{N} h_t^2(x_i)} = \frac{1}{N}\sum_{i=1}^{N} r_i h_t(x_i)$

1.2 [5 pts]

Yes it will always get zero training error, and the minimum number of weak classifier is 2.

1.3 About Margin[5 pts]

see lecture slides

1.4 [5 pts]

"Margin" means non-zero penalty even when a sample is correctly classified (inside the margin). All three learners have it. Logistic regression and Adaboost have infinitely wide margin.

1.5 Overfitting [5 pts]

The test error curve should first go down then go up.

1.6 [5 pts]

- use log loss instead of exp loss
- restrict the number of weak classifiers.
- cross validation
- etc