Reference Page

Decision Trees

Let D be the data, C be the class attribute, and A be an attribute (which could be C). The accessor $A.values$ denotes the values of attribute A.

$$\begin{aligned}
Info(D) &= \sum_{i \in C.values} -\frac{|D_i|}{|D|} * \log_2\left(\frac{|D_i|}{|D|}\right) && (1)\\
Info(A, D) &= \sum_{j \in A.values} \frac{|D_j|}{|D|} * Info(D_j) && (2)\\
Gain(A, D) &= Info(D) - Info(A, D) && (3)\\
GainRatio(A, D) &= \frac{Gain(A, D)}{SplitInfo(A, D)} && (4)\\
SplitInfo(A, D) &= Info(D) \text{ considering } A \text{ as the class attribute } C. \\
&= \sum_{i \in A.values} -\frac{|D_i|}{|D|} * \log_2\left(\frac{|D_i|}{|D|}\right) && (5)
\end{aligned}$$

Support Vector Machines

$$\text{Primal:} \quad \min_{w,b} \frac{1}{2}\|w\|^2 \quad \text{s.t.} \quad \forall i \; y_i(w \cdot x - b) \geq 1 \tag{6}$$

$$\text{Dual:} \quad \min_{\alpha} \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^{N}\alpha_i \quad \text{s.t.} \quad \sum_{i=1}^{N} y_i\alpha_i = 0 \text{ and } \forall i \; \alpha_i \geq 0 \tag{7}$$

Probability Theory

$$P(A_1, \ldots, A_m \mid B_1, \ldots, B_n) = \frac{P(B_1, \ldots, B_n \mid A_1, \ldots, A_m)P(A_1, \ldots, A_m)}{P(B_1, \ldots, B_n)} \tag{8}$$

$$P(A_1, \ldots, A_m \mid B_1, \ldots, B_n) = \frac{P(A_1, \ldots, A_m, B_1, \ldots, B_n)}{P(B_1, \ldots, B_n)} \tag{9}$$

Logistic Regression

$$P(y = 1 \mid x) = \frac{\exp(x\beta)}{1 + \exp(x\beta)} \tag{10}$$

$$P(y = 0 \mid x) = 1 - P(y = 1|x) \tag{11}$$

$$l(\beta) = \sum_{i=1}^{n} [y_i \log P(y_i = 1|x_i) + (1 - y_i)\log P(y_i = 0|x_i)] \tag{12}$$

AdaBoost

$$h_t = \arg_{h_t} \min \epsilon_t \quad \text{where} \quad \epsilon_t = \sum_{i=1}^{N} w_t(x_i)\mathbb{1}[y_i \neq h_t(x_i)] \tag{13}$$

$$\alpha_t = \frac{1}{2}\ln\frac{1 - \epsilon_t}{\epsilon_t} \tag{14}$$

$$w_{t+1}(x_i) = \frac{w_t(x_i)\exp(-\alpha_t y_i h_t(x_i))}{Z_t} \quad \text{where} \quad Z_t = \sum_{i=1}^{N} w_t(x_i)\exp(-\alpha_t y_i h_t(x_i)) \tag{15}$$

$$H(x) = \text{sign}\left(\sum_{t=1}^{T}\alpha_t h_t(x)\right) \tag{16}$$