فاصله اقلیدسی

$$Dist(x_1,x_2)=\sqrt{\sum_{i=1}^{d}(x_{1i}-x_{2i})^2}$$

$x_1, x_2, x_3, \ldots, x_N$

متوسط :

$$\mu=\frac{1}{N}*\sum_1^N x_i$$

واریانس :

$$\sigma=\frac{1}{N-1}*\sum_1^N (x_i-\mu)^2$$

کوواریانس

$$Cov(x,y)=\sigma_{xy}^2=\frac{1}{N-1}*\sum_1^N (x_i-\mu_x)^T(y_i-\mu_y)$$

D: ماتریس شامل M داده nبعدی،

ماتریس کوواریانس

$$\sigma_{i,j}^2=\frac{1}{M-1}*\sum_{k=1}^{M}(D_{k,i}-\mu_i)^T(D_{k,j}-\mu_j)$$

$$\mu_r=\frac{1}{M}*\sum_{k=1}^{M}D_{k,r}$$

نرمال سازی گوسی

$D_{k,i}^{new}=D_{k,i}^{old}-\mu_i$; i=1,2,…,n ; k=1,2,…,M

$$\mu_i=\frac{1}{M}*\sum_{k=1}^{M}D_{k,i}$$

$$\sigma_i^2=\frac{1}{M-1}*\sum_{k=1}^{M}(D_{k,i}-\mu_i)^2$$

خطای طبقه‌بند

K : تعداد برگهای درخت،

N : تعداد کل داده‌های آموزشی،

m : تعداد داده‌های آموزشی که توسط طبقه بند غلط دسته‌بندی شده‌اند :

$E(T)=\frac{m}{N}$ خطای آموزشی

$E'(T)=\frac{m+k*0.5}{N}$ خطای تست

Cost(Model, Data)=Cost(Model)+Cost(Data|Model)

Bayesian

Class of x = $arg\ \max_{1\le i\le k}\{p(c_i)\prod_{j=1}^{d}p(A_j|C_i)\}$

$p(c_i)=\frac{N^i}{N}$; $p(A_j|c_i)=\frac{N_{A_j}^{(i)}+1}{N^{(i)}+K}$ Aj :گسسته

$p(A_j|c_i)=\frac{1}{\sqrt{2\pi\delta_{ij}^2}}\exp(-\frac{\|A_j-\mu_{ij}\|^2}{2\delta_{ij}^2})$ Aj :پیوسته

$$m-estimate=\frac{N_{A_j}^{(i)}+mp}{N^{(i)}+m}$$

N : تعداد کل رکوردها

$N^{(i)}$: تعداد رکوردهایی که برچسب کلاس آنها C_i است

$N_{A_j}^{(i)}$: تعداد رکوردهایی که برچسب کلاس آنها C_i و مقدار صفت آنها A_jاست.

μ_{ij} : میانگین صفت jام تمام رکوردهایی است که برچسب کلاس آنها C_i است.

δ_{ij}^2 : واریانس صفت jام تمام رکوردهایی است که برچسب کلاس آنها C_i است.

مجموع خطای طبقه‌بندها $\varepsilon_{ensemble}=\sum_{i=Aksariat}^{N}\binom{N}{i}\varepsilon^i(1-\varepsilon)^{N-i}$

Class label={1,2,3,…c}

t:گره مورد بررسی

D_t:رکوردهای آموزشی در گره

D_t^j: تعداد رکوردهایی در گره t که برچسب کلاس آنها j است

$P(j|t)=\frac{|D_t^j|}{|D_t|}$; j=1,2,…,c

$Gini(t)=1-\sum_{j=1}^{c}[p(j|t)]^2$

--

اگر گره t، کاملا خالص باشد

$p(j|t)=\begin{cases}1, & j=i\\ 0, & O.W\end{cases}$; Gini(t)=0

--

اگر گره t، کاملا ناخالص باشد

$P(j|t)=\frac{1}{c}$ j=1,2,3,…,c ; $Gini(t)=1-\frac{1}{c}$

--

$Gini_{split}=-\sum_{i=1}^{k}\frac{n_i}{n}Gini(t_i)$

--

$Entropy(t)=-\sum_{j=1}^{c}p(j|t)\ log_2^{(p(j|t))}$

اگر گره t، کاملا خالص باشد

$p(j|t)=\begin{cases}1, & j=i\\ 0, & O.W\end{cases}$; Entropy(t)=0

--

اگر گره t، کاملا ناخالص باشد

$P(j|t)=\frac{1}{c}$ j=1,2,3,…,c ; $Entropy(t)=log_2^c$

--

$Gain_{split}=Entropy(t)-\sum_{i=1}^{k}\frac{n_i}{n}Entropy(t_i)$

--

$Split_{Info}=-\sum_{i=1}^{k}\frac{n_i}{n}log_2^{(\frac{n_i}{n})}$

$Gain_{ratio}=\frac{Gain_{Split}}{Split_{Info}}$

--

$Error(t)=1-\max_{1\le j\le c}\{p(j|t)\}$

--

اگر گره t، کاملا خالص باشد

$p(j|t)=\begin{cases}1, & j=i\\ 0, & O.W\end{cases}$; Error(t)=0

--

اگر گره t، کاملا ناخالص باشد

$P(j|t)=\frac{1}{c}$ j=1,2,3,…,c ; $Error(t)=1-\frac{1}{c}$

Svm خطی با داده‌های جداپذیر خطی:

$$\begin{cases}\min\left\{\frac{1}{2}\|w\|^2\right\}\\ S.t{:}\, \underset{i=1,2,..,N}{y_i(w^T.x_i+b)\geq 1}\end{cases}$$

$$L=\frac{1}{2}\|w\|^2+\sum_{i=1}^{N}\lambda_i\left(1-y_i(w^T.x_i+b)\right)$$

$$\begin{cases}\frac{\partial L}{\partial w}=w-\sum_{i=1}^{N}\lambda_i y_i x_i=0;\, w=\sum_{i=1}^{N}\lambda_i y_i x_i;\\ \frac{\partial L}{\partial b}=-\sum_{i=1}^{N}\lambda_i y_i=0;\, \sum_{i=1}^{N}\lambda_i y_i=0;\end{cases}$$

$y_i(w^T.x_i+b)\geq +1;\, i=1,2,\dots,N$

$\lambda_i\geq 0;\, \mathrm{i}=1,2,\dots,N$

$\lambda_i\left(1-y_i(w^T.x_i+b)\right)=0$

$$\begin{cases}L=\sum_{i=1}^{N}\lambda_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j y_i y_j\overrightarrow{x_i x_j};\\ \lambda_i\geq 0\end{cases}$$

مسائل بهینه‌سازی مقید با قید تساوی

$$\begin{cases}\min\{f(x_1,\dots,x_n)\}\\ S.t{:}\, \underset{i=1,2,..,p}{g_i(x_1,\dots,x_n)=0}\end{cases}$$

$$L(x,\lambda)=f(x)+\sum_{i=1}^{p}\lambda_i\, g_i(x)$$

$$\frac{\partial L}{\partial x_i}=0;\, i=1,2,\dots,n$$

$$\frac{\partial L}{\partial \lambda_i}=0;\, i=1,2,\dots,p$$

مسائل بهینه‌سازی مقید با قید نامساوی

$$\begin{cases}\min\{f(x_1,\dots,x_n)\}\\ S.t{:}\, \underset{i=1,2,..,q}{h_i(x_1,\dots,x_n)\leq 0}\end{cases}$$

$$L(x,\lambda)=f(x)+\sum_{i=1}^{q}\lambda_i\, h_i(x)$$

$\frac{\partial L}{\partial x_i}=0; \quad i=1,2,\dots,n$

$h_i(x)\leq 0; \quad i=1,2,\dots,q$

$\lambda_i(x)\geq 0; \quad i=1,2,\dots,q$

$\lambda_i h_i(x)=0;\, i=1,2,\dots,q$

Svm خطی با داده‌های جداناپذیر خطی:

$$\begin{cases}\min\left\{\frac{1}{2}\|w\|^2+c\sum_{i=1}^{N}\varepsilon_i{}^k\right\}\\ S.t{:}\, \underset{i=1,2,..,N\,;\,\varepsilon_i\geq 0}{y_i(w^T.x_i+b)\geq 1-\varepsilon_i}\end{cases}$$

$$L=\frac{1}{2}\|w\|^2+c\sum_{i=1}^{N}\varepsilon_i\sum_{i=1}^{N}\lambda_i\left(1-\varepsilon_i-y_i(w^T.x_i+b)\right)$$

$$+c\sum_{i=1}^{N}\mu_i(-\varepsilon_i)$$

$$\begin{cases}\frac{\partial L}{\partial w}=w-\sum_{i=1}^{N}\lambda_i y_i x_i=0;\, w=\sum_{i=1}^{N}\lambda_i y_i x_i;\\ \frac{\partial L}{\partial b}=-\sum_{i=1}^{N}\lambda_i y_i=0;\, \sum_{i=1}^{N}\lambda_i y_i=0;\\ \frac{\partial L}{\partial \varepsilon_i}=c-\lambda_i-\mu_i=0;\; \lambda_i+\mu_i=c \Rightarrow \begin{cases}0\leq\lambda_i\leq c\\ 0\leq\mu_i\leq c\end{cases}\end{cases}$$

$\lambda_i\geq 0, \mu_i\geq 0$

$$\begin{cases}\varepsilon_i\geq 0\\ \underset{i=1,2,..,N}{y_i(w^T.x_i+b)\geq 1-\varepsilon_i}\end{cases}$$

$$\begin{cases}\mu_i\varepsilon_i=0\\ \lambda_i\left(1-\varepsilon_i-y_i(w^T.x_i+b)\right)=0\end{cases}$$

$$\begin{cases}L=\sum_{i=1}^{N}\lambda_i-\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j y_i y_j\overrightarrow{x_i x_j};\\ 0\leq\lambda_i\leq c\end{cases}$$

Svm غیرخطی:

$$\begin{cases}\min\left\{\frac{1}{2}\|w\|^2+c\sum_{i=1}^{N}\varepsilon_i{}^k\right\}\\ S.t{:}\, \underset{i=1,2,..,N\,;\,\varepsilon_i\geq 0}{y_i(w^T.\emptyset(x)+b)\geq 1-\varepsilon_i}\end{cases}$$

$$\begin{cases}L_D=\sum_{i=1}^{N}\lambda_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j y_i y_j\overrightarrow{\emptyset(x_i)\emptyset(x_j)};\\ 0\leq\lambda_i\leq c\end{cases}$$

$K(u,v)=\overrightarrow{\emptyset(u).\emptyset(v)}=<\emptyset(v).\emptyset(u)>$

$k(x,y)=(\vec{x}.\vec{y}+1)^p, K(x,y)=\exp(-\frac{\|x-y\|^2}{2\delta^2}$

$k(x,y)=\tanh(k.\vec{x}.\vec{y}-\delta)$

$$max\begin{cases}L_D=\sum_{i=1}^{N}\lambda_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j y_i y_j k(x_i,x_j);\\ 0\leq\lambda_i\leq c\end{cases}$$

$\mu_i+\lambda_i=0;\, i=1,2,\dots,N$

$$\sum_{i=1}^{N}\lambda_i y_i k(x_i,x)+b=0$$

$$\begin{cases}\mu_i\varepsilon_i=0\\ \lambda_i\left(y_i(w.\emptyset(x_i)+b-1+\varepsilon_i)\right)=0\end{cases}$$

$$\begin{cases}\mu_i\varepsilon_i=0\\ \lambda_i\left(y_i\left(\sum_{j=1}^{N}\lambda_j\, y_j.\, k(x_j,x_i)+b-1+\varepsilon_i\right)\right)=0\end{cases}$$

الگوریتم bagging :

$$C(x) = \arg\max_{y} \left\{ \sum_{i=1}^{k} \delta(c_i(x) = y) \right\}$$

$$\delta(p) = \begin{cases} 1, & if\ p\ is\ True \\ 0, & O.W \end{cases}$$

قوانین انجمنی:

$$S(item\ Set) = \frac{\sigma(item\ Set)}{|T|}$$

$$S(x \to y) = \frac{\sigma(x \cup y)}{|T|}$$

$$C(x \to y) = \frac{\sigma(x \cup y)}{\sigma(x)}$$

$$\sum_{k=1}^{d-1} \left(\binom{d}{k} \sum_{j=1}^{d-k} \binom{d-k}{j} \right) = 3^d - 2^{d+1} - 1$$

$$Support(x \to y) = \frac{\sigma(x \cup y - x)}{|T|} = \frac{\sigma(y)}{|T|} = s(y)$$

$$Confidence(x \to y - x) = \frac{\sigma(x \cup y - x)}{\sigma(x)} = \frac{\sigma(y)}{\sigma(x)}$$

$$Conf(x' \to y - x') < Min_{Conf}$$

الگوریتم Adaboost :

$$D = \{(x_i, x_j) | i = 1,2, \dots, N\}$$

$$W = \{w_j | \ w_j = \frac{1}{n} \ \ j = 1,2, \dots, N\}$$

$$\varepsilon_i = \frac{1}{N} \left\{ \sum_{j=1}^{N} w_j \delta(c_i(x_j) \neq y_j \right\}$$

$$\propto_i = \frac{1}{2} \ln \frac{1 - \delta_i}{\varepsilon_i}$$

$$w_j^{(i+1)} = \frac{w_j^{(i)}}{z_i} \begin{cases} \exp(-\propto_i), & c_i(x_j) = y_j \\ \exp(\propto_j), & c_i(x_j) \neq y_j \end{cases}$$

$$C^*(x) = \arg\max_{y} \left\{ \sum_{i=1}^{k} \propto_i \delta(c_i(x) = y) \right\}$$

خوشه‌بندی، الگوریتم k-means :

K : تعداد خوشه‌ها

$D = \{x_i \epsilon \mathbb{R}^d \ |i = 1,2, \dots, N\}$

$$C_i = \frac{1}{|cluster\ i|} \sum_{x_j \epsilon\ cluster\ i} x_j \ \ ; i = 1,2, \dots, k$$

$$\begin{cases} f(c_1, c_2, \dots, c_k) = \sum_{i=1}^{k} \sum_{x_j \epsilon\ cluster\ i} d(x_j, c_i) \\ d(x_j, c_i) : \text{فاصله } x_j \text{ از } c_i \end{cases}$$