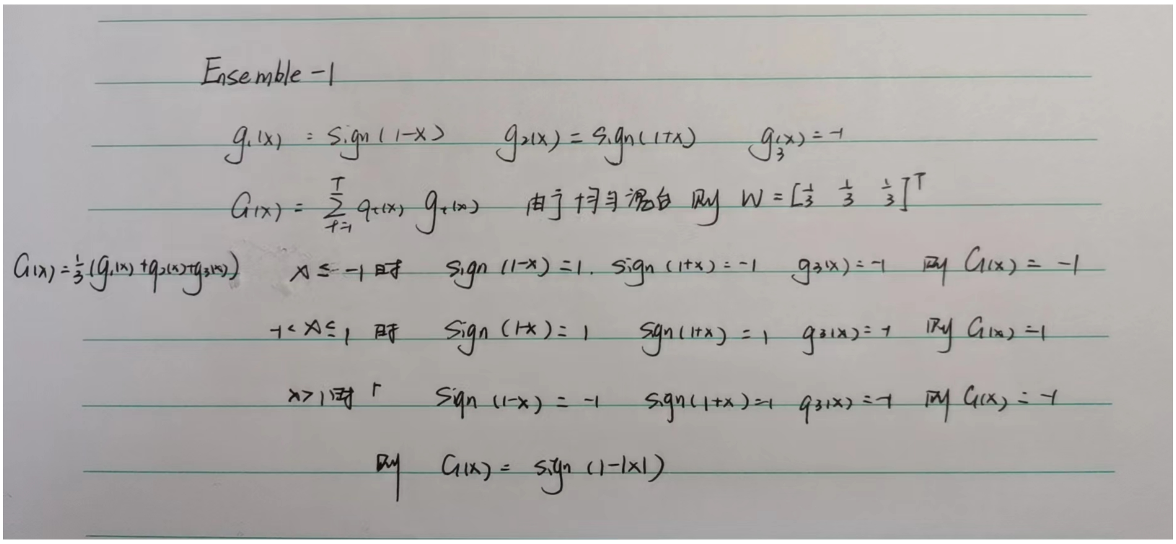
机器学习 后续 作业

姓名： 张泽群 学号： 19049100002 **班级：2班**

Ensemble-1

答:



Ensemble-1

$g_1(x) = sign(1-x)$ $g_2(x) = sign(1+x)$ $g_3(x) = -1$

$G(x) = \sum_{t=1}^{T} q_t(x) g_t(x)$ 由于均匀混合 则 $W = [\frac{1}{3}\ \frac{1}{3}\ \frac{1}{3}]^T$

$G(x) = \frac{1}{3}(g_1(x) + g_2(x) + g_3(x))$ $x \le -1$ 时 $sign(1-x) = 1$, $sign(1+x) = -1$ $g_3(x) = -1$ 则 $G(x) = -1$

$-1 < x \le 1$ 时 $sign(1-x) = 1$ $sign(1+x) = 1$ $g_3(x) = -1$ 则 $G(x) = 1$

$x > 1$ 时 $sign(1-x) = -1$ $sign(1+x) = 1$ $g_3(x) = -1$ 则 $G(x) = -1$

则 $G(x) = sign(1-|x|)$

Regress-1

答:

交叉验证**留一验证 Leave-One-Out Cross Validation**

正如名称所建议， 留一验证(LOOCV)意指只使用原本样本中的一项来当做验证资料， 而剩余的则留下来当做训练资料。 这个步骤一直持续到每个样本都被当做一次验证资料。 事实上，这等同于和K-fold 交叉验证，其中K为原本样本个数。 在某些情况下是存在有效率的演算法，如使用kernel regression 和 Tikhonov regularization。 **优点**：每一个分类器或模型都是用几乎所有的样本来训练模型，最接近样本，这样评估所得的结果比较可靠。实验没有随机因素，整个过程是可重复的。 **缺点**：计算成本高，当N非常大时，计算耗时,因为需要建立的模型数量与原始数据样本数量相同,当原始数据样本数量相当多时,LOO-CV在实作上便有困难几乎就是不显示,除非每次训练分类器得到模型的速度很快,或是可以用并行化计算减少计算所需的时间.

Regress -1

a、线性回归得出的最优解为：

$$y = 1.5 + 0 \cdot x = 1.5$$

则 $MSE = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 = \frac{1}{3}[0.5^2 \times 3] = 0.25$

$MAE = \frac{1}{m}\sum_{i=1}^{m}|(y_i - \hat{y}_i)| = \frac{1}{3}(0.5 \times 3) = 0.5$

b. MSE of LOOCV :

① (1.1) 作 validation : $y = x - 1$，则 $(y_v - \hat{y}_v)^2 = 1$

② (2.2) 作 validation : $y = 1$ 则 $(y_v - \hat{y}_v)^2 = 1$

③ (3.1) 作 validation : $y = x$ 则 $(y_v - \hat{y}_v)^2 = 2^2 = 4$

则 $MSE = \frac{1}{3}(1 + 1 + 4) = 2$

Regress-2

答:

regress-2

二维输入及其输出对 $\{x_{i1}, x_{i2}, y_i\}$ 设存在m个

回归模型：$\hat{y}_i = w_1^2 x_{i1} + w_2^2 x_{i2}$

$$LSM: \quad E(w_1, w_2) = \sum_{i=1}^{m} (y_i - w_1^2 x_{i1} - w_2^2 x_{i2})^2$$

$$\begin{cases} \dfrac{\partial E(w_1, w_2)}{\partial w_1} = 4(w_1^3 \sum_i^m x_{i1}^2 + w_2^2 w_1 \sum_i^m x_{i1} x_{i2} - w_1 \sum_i^m y_i x_{i1}) \\ \dfrac{\partial E(w_1, w_2)}{\partial w_2} = 4(w_2^3 \sum_i^m x_{i2}^2 + w_1^2 w_2 \sum_i^m x_{i1} x_{i2} - w_2 \sum_i^m y_i x_{i2}) \end{cases}$$

令上二式均 $= 0$

则有
$$\begin{cases} (w_1^2 \sum_i^m x_{i1}^2 + w_2^2 \sum_i^m x_{i1} x_{i2}) = \sum_i^m y_i x_{i1} \\ (w_2^2 \sum_i^m x_{i2}^2 + w_1^2 \sum_i^m x_{i1} x_{i2}) = \sum_i^m y_i x_{i2} \end{cases}$$

$$\begin{bmatrix} \sum_i^m x_{i1}^2 & \sum_i^m x_{i1} x_{i2} \\ \sum_i^m x_{i2}^2 & \sum_i^m x_{i1} x_{i2} \end{bmatrix} \begin{bmatrix} w_1^2 \\ w_2^2 \end{bmatrix} = \begin{bmatrix} \sum_i^m y_i x_{i1} \\ \sum_i^m y_i x_{i2} \end{bmatrix}$$

则
$$\begin{bmatrix} w_1^2 \\ w_2^2 \end{bmatrix} = \begin{bmatrix} \sum_i^m x_{i1}^2 & \sum_i^m x_{i1} x_{i2} \\ \sum_i^m x_{i2}^2 & \sum_i^m x_{i1} x_{i2} \end{bmatrix}^{-1} \begin{bmatrix} \sum_i^m y_i x_{i1} \\ \sum_i^m y_i x_{i2} \end{bmatrix}$$

则
$$w_1 = \sqrt{\frac{\sum_i^m y_i x_{i1} - w_2^2 \sum_i^m x_{i1} x_{i2}}{\sum_i^m x_{i1}^2}}$$

$$w_2 = \sqrt{\frac{\sum_i^m y_i x_{i2} - w_1^2 \sum_i^m x_{i1} x_{i2}}{\sum_i^m x_{i2}^2}}$$

CNN-1

答:

CNN 习题

1. 尺寸 = $\frac{3-2}{1} + 1 = 2.$

图片 = $\begin{bmatrix} 25 & 31 \\ 43 & 49 \end{bmatrix}$

CNN-2

答:

若考虑bias，则在(·)内+1

2.

	输出尺寸	层参数量
INPUT LAYER	224×224×3	0
conv3-64	224×224×64	(3×3×3)×64
conv3-64	224×224×64	(3×3×64)×64
maxpool	112×112×64	0
conv3-128	112×112×128	(3×3×64)×128
conv3-128	112×112×128	(3×3×128)×128
max pool	56×56×128	0
conv3-256	56×56×256	(3×3×128)×256
conv3-256	56×56×256	(3×3×256)×256
conv3-256	56×56×256	(3×3×256)×256
conv3-256	56×56×256	(3×3×256)×256
max pool	28×28×256	0
conv3-512	28×28×512	(3×3×256)×512
conv3-512	28×28×512	(3×3×512)×512
conv3-512	28×28×512	(3×3×512)×512
conv3-512	28×28×512	(3×3×512)×512
maxpool	14×14×512	0
conv3-512	14×14×512	(3×3×512)×512
conv3-512	14×14×512	(3×3×512)×512
conv3-512	14×14×512	(3×3×512)×512
conv3-512	14×14×512	(3×3×512)×512
maxpool	7×7×512	0
FC-4096	1×1×4096	(7×7×512)×4096
FC-4096	1×1×4096	4096×(4096)
FC-1000	1×1×1000	1000×(4096)

总参数 = 143652544 (未考虑bias)

Bayes-1

Bayes-1

正常 $P(w_1)=0.9$　异常 $P(w_2)=0.1$

$p(x|w_1)=0.2$　$p(x|w_2)=0.4$

则 $p(w_1|x)=\dfrac{P(w_1)P(x|w_1)}{P(w_1)p(x|w_1)+p(w_2)p(x|w_2)}=\dfrac{P(x w_1)}{P(x)}=\dfrac{0.18}{0.18+0.04}=\dfrac{9}{11}$

同理 $p(w_2|x)=\dfrac{2}{11}$　则 该细胞为正常细胞

$p(w_1|x)>p(w_2|x)$

Bayes-2

Bayes-2.

x_1取值 $A_1=\{1.2.3\}$.　x_2取值 $A_2=\{S.M.L\}$　$Y=\{-1.1\}$

解：$p(-1)=\frac{6}{15}=0.4$　$P(1)=0.6$　，设两特征属性独立

$P(A_1=2|1)=\frac{1}{3}$　$p(A_2=S|1)=\frac{1}{9}$　则 $P(A_1=2.A_2=S|1)=\frac{1}{27}$

$p(A_1=2|-1)=\frac{1}{3}$　$p(A_2=S|-1)=\frac{1}{2}$　$P(A_1=2.A_2=S|-1)=\frac{1}{6}$

设 $p(A_1=2.A_2=S)=\lambda$ 则 $P(1|A_1=2.A_2=S)=\dfrac{p(1)P(A_1=2.A_2=S|1)}{P(A_1=2.A_2=S)}=\dfrac{0.6\times\frac{1}{27}}{\lambda}=\dfrac{1}{45\lambda}$

同理 $P(-1|A_1=2.A_2=S)=\dfrac{0.4\times\frac{1}{6}}{\lambda}=\dfrac{1}{15\lambda}$

$p(-1|(2,S))>P(1|(2.S))$

则 $x=(2.S)^T$ 的类别为 -1.

Bias&Variance-1

Bias & Variance

分析：

(1). ① 偏差较大，方差较小，数据欠拟合

② 偏差、方差都较小，数据正常拟合

③ 偏差较小、方差较大，数据过拟合