

《人工智能引论》课后练习-3

内容：机器学习 提交时间：2023-04-17 姓名：_____ 学号：_____

一、（20 分）对以下数据一维线性回归 $y=wx+b$

| | | | |
|---|---|---|---|
| X | 0 | 2 | 3 |
| Y | 1 | 1 | 4 |

请列出平方损失函数 L ，并直接通过令 $\frac{\partial L}{\partial w} = 0, \frac{\partial L}{\partial b} = 0$ ，求出最小化 L 时 w, b 的数值解。请画出得到的回归曲线。

二、（20 分）课上学的逻辑回归以 $\{1, -1\}$ 作为正负类标签，本题使用 $\{1, 0\}$ 作为正负类标签。给定数据集 $D = \{(x_1, y_1), \dots, (x_n, y_n)\}$ 。设权重 (weight) 为 $w \in \mathbb{R}^d$ 和 偏置 (bias) 为 $b \in \mathbb{R}$ ， σ 表示 sigmoid 函数

1)（6 分）写出 $p(y = y_i | x = x_i)$ 在 $y_i = 0, 1$ 下分别是多少。

2)（14 分）利用 $p(y = y_i | x = x_i) = p(y = 1 | x = x_i)^{y_i} p(y = 0 | x = x_i)^{1-y_i}$ ，推导逻辑回归在 D 上的对数似然函数 (log-likelihood)。

三、（30 分）利用树模型对以下数据进行二分类。id 表示数据编号，A, B, C 是特征，y 是标签。

| | | | | | | | | | |
|----|----|----|----|----|----|----|---|---|---|
| id | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| A | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| B | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| C | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| y | -1 | -1 | -1 | -1 | -1 | -1 | 1 | 1 | 1 |

1)（15 分）在树的根节点，特征 A 的 gain ratio 是多少？(请使用以 2 为底的对数)

2)（15 分）假设在根节点对 A 分裂。在第二层所有结点对 C 分裂，在第三层对 B 分裂。请画出分类树并预测 $x_* = [1, 1, 1]$ 的标签。

四、（30 分）推导 softmax, log softmax 的反向传播公式。设输入 $z \in \mathbb{R}^d$ ，计算图为线性（计算结点之间顺序连接，没有跨层连接），总损失函数为 L 。

1)（15 分）softmax 的输出为 $a \in \mathbb{R}^d, a_i = \frac{e^{z_i}}{\sum_j e^{z_j}}$ 。用 $\frac{\partial L}{\partial a}$ 来表示 $\frac{\partial L}{\partial z}$

2)（15 分）log softmax 的输出为 $a \in \mathbb{R}^d, a_i = \ln \frac{e^{z_i}}{\sum_j e^{z_j}}$ 。用 $\frac{\partial L}{\partial a}$ 来表示 $\frac{\partial L}{\partial z}$

提示：逐分量表示 $\frac{\partial L}{\partial z_i}$ 。先求 $\frac{\partial a_j}{\partial z_i}$ ，再利用使用链式法则 $\frac{\partial L}{\partial z_i} = \sum_j \frac{\partial L}{\partial a_j} \frac{\partial a_j}{\partial z_i}$ 。你可以使用 a 来表示 $\frac{\partial L}{\partial z}$ ，最终表达式中不要出现 z 。