计算题

1．已知训练数据， 正类： $x_1=(1,-1),\ y_1=+1$；

负类：$x_2=(-1,1),\ y_2=-1$， $x_3=(0,2),\ y_3=-1$.

试用感知机学习算法的原始形式求感知机模型 $f(x)=sign(w\cdot x+b)$。（ 设$\eta=1$，求出 w, b ）

解：构建最优化问题：

$$\min_{w,b} L(w,b)=-\sum_{x_i\in M} y_i(w\cdot x_i+b)$$

按照感知机学习算法的原始形式求解w,b

（1）取初值 $w_0=(0,0)$， $b_0=0$

（2）对 $x_1=(1,-1)$， $y_1(w_0\cdot x_1+b_0)=0$， 未能被正确分类，更新 w_1,b_1

$$w_1=w_0+y_1x_1=(1,-1),\ \ b_1=b_0+y_1=1$$

得到线性模型

$$w_1x+b_1=x^{(1)}-x^{(2)}+1$$

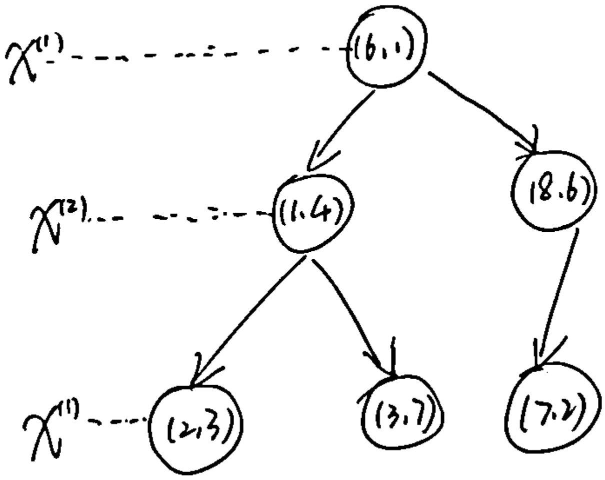
（3）对 x_1,x_2,x_3，显然， $y_i(w_1\cdot x_i+b_1)>0$，所有数据点被正确分类，损失函数达到极小。

分离超平面为： $x^{(1)}-x^{(2)}+1=0$

感知机模型为： $f(x)=\mathrm{sign}(x^{(1)}-x^{(2)}+1)$

2. 给定一个数据集 T={(2,3)，(1,4)，(8,6)，(3,7)，(6,1)，(7,2)}
构造一个平衡 kd 树。

解：根结点对应包含数据集 T 的矩形，选择 $x^{(1)}$ 轴，6 个数据点的 $x^{(1)}$ 坐标的中位数是 6，以平面 $x^{(1)}=6$将空间分为左、右两个子矩形（子结点）；接着，左矩形以 $x^{(2)}=4$ 分为两个子矩形，右矩形以 $x^{(2)}=6$ 分为两个子矩形，如此递归，最后得到如下图所示的 kd 树。



3 已知如下数据

申请人	性别	职业	收入	办理信用卡
1	男	个体	高	会
2	男	个体	高	不会
3	女	事业	中	会
4	男	事业	中	会
5	女	待业	低	不会
6	女	事业	高	会
7	男	待业	中	不会
8	女	个体	低	会
9	女	个体	中	会
10	男	个体	高	不会

请用朴素贝叶斯算法，判断银行是否会给一名收入高的男性个体户办理信用卡，给出理由。

解：根据朴素贝叶斯算，由上述表格，将“会”办理信用卡记为“1”，“不会”办理信用卡记为“-1”，可以计算下列概率：

（1）训练

$P(Y=1)=\frac{6}{10}$，$P(Y=-1)=\frac{4}{10}$

$P(X^{(1)}=男|Y=1)=\frac{2}{6}$，$P(X^{(1)}=女|Y=1)=\frac{4}{6}$

$P(X^{(2)}=个体|Y=1)=\frac{3}{6}$，$P(X^{(2)}=事业|Y=1)=\frac{3}{6}$，$P(X^{(2)}=待业|Y=1)=\frac{0}{6}$

$P(X^{(3)}=高|Y=1)=\frac{2}{6}$，$P(X^{(3)}=中|Y=1)=\frac{3}{6}$，$P(X^{(3)}=低|Y=1)=\frac{1}{6}$

$P(X^{(1)}=男|Y=-1)=\frac{3}{4}$，$P(X^{(1)}=女|Y=-1)=\frac{1}{4}$

$P(X^{(2)}=\text{个体}|Y=-1)=\frac{2}{4}$，$P(X^{(2)}=\text{事业}|Y=-1)=\frac{0}{4}$，$P(X^{(2)}=\text{待业}|Y=-1)=\frac{2}{4}$

$P(X^{(3)}=\text{高}|Y=-1)=\frac{2}{4}$，$P(X^{(3)}=\text{中}|Y=-1)=\frac{1}{4}$，$P(X^{(3)}=\text{低}|Y=-1)=\frac{1}{4}$

（2）测试

对于给定 $X^{(1)}=\text{男}, X^{(2)}=\text{个体}, X^{(3)}=\text{高}$，计算

$$P(Y=1)P(X^{(1)}=\text{男}|Y=1)P(X^{(2)}=\text{个体}|Y=1)P(X^{(3)}=\text{高}|Y=1)=\frac{6}{10}\cdot\frac{2}{6}\cdot\frac{3}{6}\cdot\frac{2}{6}=\frac{1}{30}$$

$$P(Y=-1)P(X^{(1)}=\text{男}|Y=-1)P(X^{(2)}=\text{个体}|Y=-1)P(X^{(3)}=\text{高}|Y=-1)=\frac{4}{10}\cdot\frac{3}{4}\cdot\frac{2}{4}\cdot\frac{2}{4}=\frac{3}{40}$$

（3）判断

因为 $P(Y=-1)P(X^{(1)}=\text{男}|Y=-1)P(X^{(2)}=\text{个体}|Y=-1)P(X^{(3)}=\text{高}|Y=-1)$ 最大，所以 $y=-1$，即判断不会给这名收入高的男性个体户办理信用卡。

4. 已知训练数据，正类：$x_1=(1,1),\ y_1=+1$；

负类：$x_2=(-2,0),\ y_2=-1$，$x_3=(0,-2),\ y_3=-1$.

试用线性可分支持向量机学习算法求出模型 $f(x)=sign(w\cdot x+b)$。

解：根据线性可分支持向量机学习算法和 3 个训练样本

（1）构造并求解约束最优化问题

$$\begin{aligned}&\min_{\alpha}\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j(x_i\cdot x_j)-\sum_{i=1}^{N}\alpha_i\\&=\frac{1}{2}(2\alpha_1^2+4\alpha_2^2+4\alpha_3^2+4\alpha_1\alpha_2+4\alpha_1\alpha_3)-\alpha_1-\alpha_2-\alpha_3\\&\text{s.t. }\alpha_1-\alpha_2-\alpha_3=0\\&\quad\alpha_i\geq 0,\ \ i=1,2,3\end{aligned}$$

解这一优化问题。将 $\alpha_1=\alpha_2+\alpha_3$ 代入目标函数并记为

$$s(\alpha_2,\alpha_3)=5\alpha_2^2+5\alpha_3^2+6\alpha_2\alpha_3-2\alpha_2-2\alpha_3$$

对 α_2,α_3 求偏导数并令为 0，解得 $s(\alpha_2,\alpha_3)$ 在 $\alpha_2=\frac{1}{8},\alpha_3=\frac{1}{8}$ 时取极小值，此时 $\alpha_1=\frac{1}{4}$，同时满足约束条件 $\alpha_1\geq 0,\ \alpha_2\geq 0,\ \alpha_3\geq 0$。

（2）计算

$$\begin{aligned}w^* &= \sum_{i=1}^{N}\alpha_i^* y_i x_i \\ &= \frac{1}{4}\cdot 1\cdot(1,1)+\frac{1}{8}\cdot(-1)(-2,0)+\frac{1}{8}\cdot(-1)(0,-2)=(\frac{1}{2},\frac{1}{2})\end{aligned}$$

即 $w_1^*=\frac{1}{2}, w_2^*=\frac{1}{2}$。

$$\begin{aligned}b^* &= y_1-\sum_{i=1}^{N}\alpha_i^* y_i(x_i\cdot x_1) \\ &= 1-(\frac{1}{4}\cdot 1\cdot 2+\frac{1}{8}\cdot(-1)\cdot(-2)+\frac{1}{8}\cdot(-1)\cdot(-2))=0\end{aligned}$$

（3）求得分离超平面为

$$\frac{1}{2}x^{(1)}+\frac{1}{2}x^{(2)}=0$$

分类决策函数为：

$$f(x)=\mathrm{sign}(\frac{1}{2}x^{(1)}+\frac{1}{2}x^{(2)})$$

5.下列数据包含 5 个二维的样本点，用聚合层次聚类进行聚类(样本之间用欧式距离，类与类之间的距离用最短距离)。五个点分别为：（1，1），（-1，2），（2，3），（4，1），（3，0）。

解：将五个点分别记为 x_1, x_2, x_3, x_4, x_5，则五个点之间的欧式距离由如下矩阵 D

D 表示：

$$D=\left[d_{ij}\right]_{5\times 5}=\begin{bmatrix}0 & \sqrt{5} & \sqrt{5} & 3 & \sqrt{5}\\ \sqrt{5} & 0 & \sqrt{10} & \sqrt{26} & 2\sqrt{5}\\ \sqrt{5} & \sqrt{10} & 0 & 2\sqrt{2} & \sqrt{10}\\ 3 & \sqrt{26} & 2\sqrt{2} & 0 & \sqrt{2}\\ \sqrt{5} & 2\sqrt{5} & \sqrt{10} & \sqrt{2} & 0\end{bmatrix}$$

（1）首先用 5 个样本构建 5 个类，$G_i=\left\{x_i\right\}, i=1,2,\cdots,5$，这样，样本之间的距离也就变成类之间的距离，所以 5 个类之间的距离矩阵亦为 D。

（2）由矩阵 D 可以看出，$D_{45}=D_{54}=\sqrt{2}$ 为最小，所以把 $G_4=G_5$ 合并为一个新类，

记作 $G_6=\left\{x_4,x_5\right\}$。

（3）计算 G_6 与 G_1,G_2,G_3 之间的最短距离，有

$$D_{61}=\sqrt{5},\ D_{62}=2\sqrt{5},\ D_{63}=2\sqrt{2}$$

又注意到其余两类之间的距离是

$$D_{12}=\sqrt{5},\ D_{13}=\sqrt{5},\ D_{23}=\sqrt{10}$$

$D_{61}=D_{12}=D_{13}=\sqrt{5}$ 最小，这里我们取 G_1，将 G_1 与 G_6 合并成一个新类，记作 $G_7=\{x_1,x_4,x_5\}$。

（4）计算 G_7 与 G_2,G_3 之间的最短距离，

$$D_{72}=\sqrt{5},\ D_{73}=\sqrt{5}$$

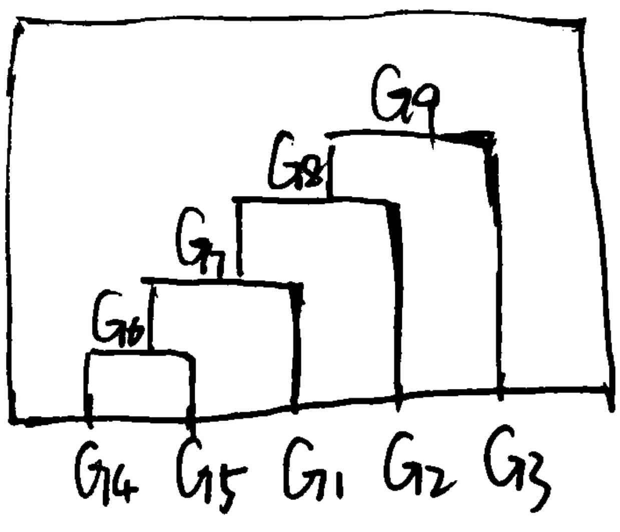
又注意到

$$D_{23}=\sqrt{10}$$

$D_{72}=D_{73}=\sqrt{5}$ 最小，这里我们取 G_2，将 G_2 与 G_7 合并成一个新类，记作 $G_8=\{x_1,x_2,x_4,x_5\}$。

（5）将 G_8 与 G_3 合并成一个新类，记作 $G_9=\{x_1,x_2,x_3,x_4,x_5\}$，即将全部样本聚成 1 类，聚类终止。

上述层次聚类过程可以用下面的层次聚类图表示。



（第（3）和（4）步选择合并的类不同，答案会不同）

6. 已知训练数据包含 5 个二维的样本点 $(1,1),\ (-1,2),\ (2,3),\ (4,1),\ (3,0)$，试用 K

均值聚类算法将样本聚到 2 个类中（设样本到类中心的距离基于欧式距离）。
解： 根据 K 均值聚类算法，

（1）选择两个样本点作为类中心。假设选择 $m_1^{(0)} = x_1 = (1,1), m_2^{(0)} = x_2 = (-1,2)$。

（2）以 $m_1^{(0)}, m_2^{(0)}$ 为类 $G_1^{(0)}, G_2^{(0)}$ 的中心，计算 $x_3 = (2,3), x_4 = (4,1), x_5 = (3,0)$ 与 $m_1^{(0)} = (1,1), m_2^{(0)} = (-1,2)$ 的欧式距离平方。

对 $x_3 = (2,3)$，$d(x_3, m_1^{(0)}) = 5,\ d(x_3, m_2^{(0)}) = 10$，将 x_3 分到类 $G_1^{(0)}$。

对 $x_4 = (4,1)$，$d(x_4, m_1^{(0)}) = 9,\ d(x_4, m_2^{(0)}) = 26$，将 x_4 分到类 $G_1^{(0)}$。

对 $x_5 = (3,0)$，$d(x_5, m_1^{(0)}) = 5,\ d(x_5, m_2^{(0)}) = 20$，将 x_3 分到类 $G_1^{(0)}$。

（3）得到新的类 $G_1^{(1)} = \{x_1, x_3, x_4, x_5\}, G_2^{(1)} = \{x_2\}$，计算类的中心 $m_1^{(1)}, m_2^{(1)}$：

$$m_1^{(1)} = (\frac{5}{2}, \frac{5}{4}), m_2^{(1)} = (-1,2)$$

（3）重复步骤（2）和（3）。

将 x_1 分到类 $G_1^{(1)}$，将 x_2 分到类 $G_2^{(1)}$，x_3 分到类 $G_1^{(1)}$，x_4 分到类 $G_1^{(1)}$，x_5 分到类 $G_1^{(1)}$。

得到新的类 $G_1^{(2)} = \{x_1, x_3, x_4, x_5\}, G_2^{(2)} = \{x_2\}$。

由于得到的新的类没有发生改变，聚类停止。得到聚类结果：

$$G_1^* = \{x_1, x_3, x_4, x_5\}, G_2^* = \{x_2\}$$

（第（1）步初始类中心的选择不同，答案会不同）

7. 求矩阵 $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 0 & 0 \end{pmatrix}$ 的奇异值分解。

解： 根据奇异值分解算法，

（1）求对称矩阵 $A^{\mathrm{T}}A$

$$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

特征值 λ 和特征向量 x 满足特征方程

$$(A^{\mathrm{T}}A - \lambda I)x = 0$$

得到齐次线性方程组

$$\begin{cases}(5-\lambda)x_1+4x_2=0\\4x_1+(5-\lambda)x_2=0\end{cases}$$

该方程组有非零解的充要条件是

$$\begin{vmatrix}5-\lambda & 4\\4 & 5-\lambda\end{vmatrix}=0$$

即

$$\lambda^2-10\lambda+9=0$$

解此方程，得矩阵 $A^{\mathrm{T}}A$ 的特征值 $\lambda_1=9$ 和 $\lambda_2=1$

将特征值 $\lambda_1=9$ 代入线性方程组，得到对应的单位特征向量

$$v_1=\begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}$$

同样得到特征值 $\lambda_2=1$ 对应的单位特征向量

$$v_2=\begin{bmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\end{bmatrix}$$

（2）求正交矩阵 V
构造正交矩阵 V

$$V=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}$$

（3）求对角矩阵 Σ

奇异值为 $\sigma_1=\sqrt{\lambda_1}=3$ 和 $\sigma_2=\sqrt{\lambda_2}=1$。构造对角矩阵

$$\Sigma=\begin{bmatrix}3 & 0\\0 & 1\\0 & 0\end{bmatrix}$$

（4）求正交矩阵 U

基于A的正奇异值计算得到列向量u_1和u_2

$$u_1=\frac{1}{\sigma_1}Av_1=\frac{1}{3}\begin{bmatrix}1&2\\2&1\\0&0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{bmatrix}$$

$$u_2=\frac{1}{\sigma_2}Av_2=1\begin{bmatrix}1&2\\2&1\\0&0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\end{bmatrix}=\begin{bmatrix}-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{bmatrix}$$

列向量u_3是A^{T}的零空间$N(A^{\mathrm{T}})$的一组标注正交基。为此，求解以下线性方程组

$$A^{\mathrm{T}}x=\begin{bmatrix}1&2&0\\2&1&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

即

$$\begin{cases}x_1+2x_2=0\\2x_1+x_2=0\end{cases}$$

取(x_2,x_3)为$(0,1)$，得到$N(A^{\mathrm{T}})$的基$(0,0,1)^{\mathrm{T}}$，则$N(A^{\mathrm{T}})$的一组标注正交基是

$$u_3=\begin{bmatrix}0\\0\\1\end{bmatrix}$$

构造正交矩阵U

$$U=\begin{bmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}&0\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\0&0&1\end{bmatrix}$$

（5）矩阵A的奇异值分解

$$A = U\Sigma V^{\mathrm{T}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$