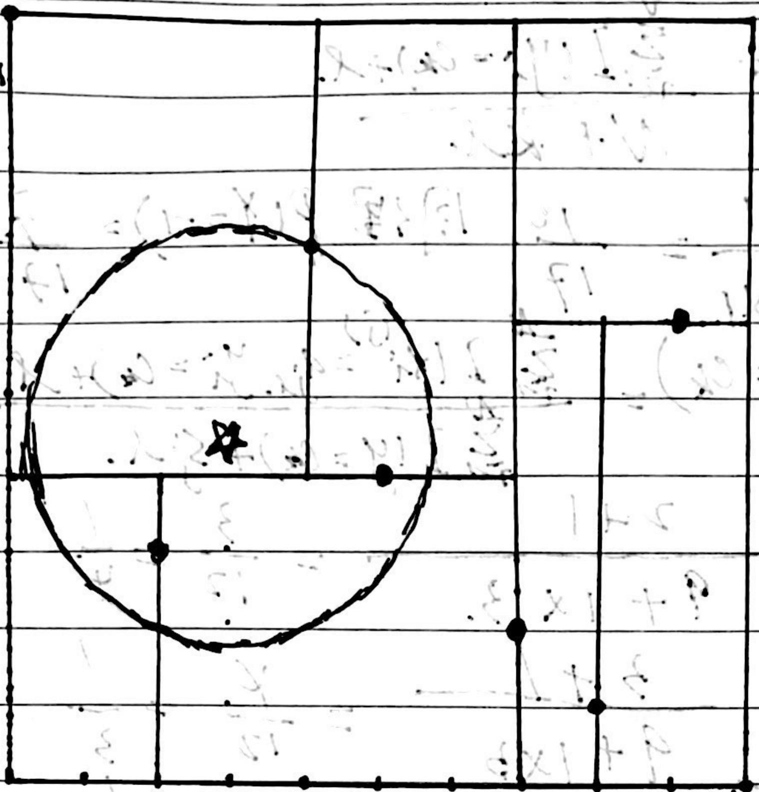
温兆和. 10205501432. 统计方法与机器学习作业5

1.解



目标点(3,4.5)位于(4,7)所在的区域内.故其最近邻的距离小于其到(4,7)的距离.
退到(4,7)的根结点(5,4),在(5,4)的另一子结点(2,3)所在区域搜索.以(3,4.5)到(4,7)的线段为半径,(3,4.5)为圆心的圆与该区域相交,且有(5,4),(2,3)两个点在该区域内.其中(5,4)到(3,4.5)的距离为$\frac{\sqrt{17}}{2}$.
再退到(5,4)的子结点(2,3).(2,3)到(3,4.5)的距离为$\frac{\sqrt{13}}{2}<\frac{\sqrt{17}}{2}$,故目前(2,3)为最近邻点.
再退到(5,4)的父结点(根结点)(7,2).从(7,2)右侧区域寻找最近邻点.但这个区域与圆不相交.不可能有(3,4.5)的最近邻.
综上,要找的最近邻点是(2,3)

2. 解：由于 $P_\lambda(Y=c_k)=\dfrac{\sum_{i=1}^{N} I(y_i=c_k)+\lambda}{N+K\lambda}$

故 $P(Y=1)=\dfrac{9+1}{15+2\times1}=\dfrac{10}{17}$ 同理 $P(Y=-1)=\dfrac{7}{17}$

又 $P_\lambda(X^{(j)}=a_{jl}\mid Y=c_k)=\dfrac{\sum_{i=1}^{N} I(x_i^{(j)}=a_{jl},y_i=c_k)+\lambda}{\sum_{i=1}^{N} I(y_i=c_k)+S_j\lambda}$

故 $P(X^{(1)}=1\mid Y=1)=\dfrac{2+1}{9+1\times3}=\dfrac{3}{12}=\dfrac{1}{4}$

$P(X^{(1)}=2\mid Y=1)=\dfrac{3+1}{9+1\times3}=\dfrac{4}{12}=\dfrac{1}{3}$

同理：$P(X^{(1)}=3\mid Y=1)=\dfrac{5}{12}$，$P(X^{(2)}=S\mid Y=1)=\dfrac{1}{6}$

$P(X^{(2)}=M\mid Y=1)=\dfrac{5}{12}$ $P(X^{(2)}=L\mid Y=1)=\dfrac{5}{12}$

$P(X^{(1)}=1\mid Y=-1)=\dfrac{4}{9}$ $P(X^{(1)}=2\mid Y=-1)=\dfrac{1}{3}$

$P(X^{(1)}=3\mid Y=-1)=\dfrac{2}{9}$ $P(X^{(2)}=S\mid Y=-1)=\dfrac{4}{9}$

$P(X^{(2)}=M\mid Y=-1)=\dfrac{1}{3}$ $P(X^{(2)}=L\mid Y=-1)=\dfrac{2}{9}$

故当 $x=(2,S)$ 欲求

$y=\arg\max_{c_k} P(Y=c_k)\prod_j P(X^{(j)}=x^{(j)}\mid Y=c_k)$

$c_k\in\{-1,1\}$

由于 $P(Y=1)\,P(X^{(1)}=2\mid Y=1)\,P(X^{(2)}=S\mid Y=1)=0.0327$

$P(Y=-1)\,P(X^{(1)}=2\mid Y=-1)\,P(X^{(2)}=S\mid Y=-1)=0.0610$

故 $y=-1$ 最大化 $P(Y=c_k)\prod_j P(X^{(j)}=x^{(j)}\mid Y=c_k)$

故我们认为当 $x=(2,S)$，$Y=-1$