机器学习期末重点

1，（10 分）考虑具有如下两类输入分量的分类问题。这四类输入向量分别是：

第一类：$\xi=\left[\xi_1,\xi_2,\dots,\xi_m\right]^T$ 第二类：$\left\{p_3=\begin{bmatrix}2\\-1\end{bmatrix},p_4=\begin{bmatrix}3\\-2\end{bmatrix}\right\}$

试设计一个 Perceptron 求解该问题。

1） 用图示法确定 Perceptron 的一组权值和偏置。（5 分）

2） 写出网络学习的 Perceptron 规则。（5 分）

2，（10 分）在线性回归模型中，给定训练样本集合$\left\{(x^{(i)},y^{(i)}),i=1,2,\dots,m\right\}$与模型参数$\theta$

1） 求损失函数$J(\theta)=\dfrac{1}{2}\sum_{i=1}^{m}(\theta^r x^{(i)}-y^{(i)})^2$的海森矩阵（5 分）

2） 给出牛顿法求解最小二乘问题的第一次迭代证明过程，证明其一步收敛，即得到$\theta^*=(X^TX)^{-1}X^Ty$（5 分）

3，(10 分)简述在学习理论中，模型的偏差（Bias）与方差（Variance）的定义。分析一个模型的复杂程度与偏差（Bias）,方差的关系。

4，(8 分)给定一组数据集合$S=\left\{x^{(1)},\dots,x^{(d)}\right\}$和一个假设类 H，给出

1） H 把 S 打散的定义？（4 分）

2） S 的 VC 维的定义？（4 分）

5，（10 分）简述 K-means 无监督聚类的原理与过程

6，（10 分）简述机器学习中主要的两类特征选择的方法：wrapper model feature selection 和 Filter feature selection 的原理，说明两者的不同。

7，（10 分）简述模型选择中的“五倍交叉验证”和“留一法”的主要步骤。

8，（10 分）考虑多变量的回归问题，假设训练集中的样本为：

$$\left\{(x^{(i)},y^{(i)}),i=1,2,\dots,m\right\},x^{(i)}\in R^n,y^{(i)}\in R^p$$

用一个线性回归模型来预测输出，即给定模型参数矩阵$\Theta\in\mathbb{R}^{n\times p}$，

$$y=\Theta^Tx$$

1） 该问题的损失函数是：$J(\Theta)=\dfrac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{p}((\Theta^Tx^{(i)})_j-y_j^{(i)})^2$，将 $J(\Theta)$写成矩阵向量形式（2 分）

2） 求解使得$J(\Theta)$最小的Θ的解析解（4 分）

3） 考虑单个变量$y_j^{(i)}$（j=1,...,p）,我们有：$y_j^{(i)}=\Theta_j^T\,x^{(i)},j=1,\dots,p$（这里，每一个$\Theta\in\mathbb{R}^n$）

求 P 个独立方程最小二乘解（4 分）

9，（22 分）采用支撑向量机（SVM）算法分类线性不可分数据时，可以引入松弛变量ξ来放松边界，调整后的算法称为l_2范数软边界 SVM。这种新的算法由下面的优化问题给出：

$$\min_{w,b,\xi}\frac{1}{2}\|w\|^2+\frac{C}{2}\sum_{i=1}^{m}\xi_i^2 \quad , \quad s.t.\, y^{(i)}(w^T x^{(i)}+b)\geq 1-\xi_i, i=1,\ldots m$$

1) 写出l_2范数软边界 SVM 优化问题的拉格朗日函数（4 分）
2) 通过求梯度：$\nabla_w L, \frac{\partial L}{\partial b}, \nabla_\xi L,$ 并令他们等于 0，求解拉格朗日函数关于w,b,ξ的最小值。其中，$\xi=\left[\xi_1,\xi_2,\ldots,\xi_m\right]^T$（6 分）
3) 采用非线性 SVM 算法进行数据分类时，解释其中的结构风险最小化的含义？（4 分）
4) 解释为何要定义核函数？（4 分）
5) 列举常用的两种核函数。（4 分）