**哈工大2019秋机器学习期末考试**

1. 信息增益是什么？举例说明为什么信息增益好。决策树分界面是否是线性的，举例说明。

2. 最小错误概率下的最优分类方法并证明

3. 什么是过拟合，决策树、SVM怎么避免的过拟合。

4. 在第二题的分类方法下，假设类概率分布是伯努利分布，类条件概率分布是协方差相等的两个高斯分布。求分界面方程。求类后验概率。

5. 逻辑回归优化目标函数。在交叉熵的角度理解，交叉熵中的两个概率分布分别是什么。

6. PCA推导。PCA应用举例（2个）

7. GMM模型优化目标函数，（EM算法）主要步骤。在做实验时，协方差矩阵行列式为0原因，如何解决。

8. 在低维空间中线性不可分的样本，在高维空间中线性可分，为什么。举例，你做的实验中有无应用。

9. 机器学习中内积的应用举例（3个）。解释他们的意义。