

1. (Classification) 请问怎么理解机器学习模型的泛化能力 (Generalization Capability) ? 如何提升分类模型的泛化能力? [20分]
2. (Clustering) 在聚类任务当中, 你认为聚类算法重要, 还是样本点之间的距离 (相似度) 定义哪个更重要, 请阐明你的观点。 [20分]
3. (Supervised learning) 请分析比较 SVM 与深度神经网络在分类任务上的优缺点。 [20分]
4. (Semi-Supervised Learning) 请分析阐述: 半监督学习在什么情况下有效。提示: 可结合你所学过的方法进行分析。 [20分]
5. (Regularization) 机器学习的目标函数常常会引入正则项, 请分析比较 L1-范数和 L2-范数所起的作用, 并且你会怎么选择? [20分]