

问答题：六道大题，每道大题有两三道小题，不用带计算器

1. 线性回归

1. 损失函数的构建，推导（求偏导要会）

2. 逻辑回归

1. 逻辑回归与线性回归的区别，对比

3. 线性分类器

1. 逻辑回归、Sigmoid函数、判别边界
2. 交叉熵损失函数、交叉熵损失函数的梯度求解
3. Softmax函数

4. 从概率角度理解回归和分类问题

5. 过拟合和正则化

1. 过拟合问题的概念和解决，欠拟合问题
2. 正则化项

6. 支持向量机（爹）

1. 最大边界分类器：函数的定义、如何推导、最终的数学公式
2. 软的最大边界分类器：松弛变量
3. 支持向量机的推导：核函数

7. 神经网络

1. 激活函数、了解典型神经网络

8. 神经网络的优化和训练技巧

1. Adam

9. Bagging & Boosting

10. PCA（爹）

11. 最小化重构误差的推导、证明

12. 最大化方差的推导、证明

13. （了解）奇异值分解的推导

14. 聚类：K-Means

1. K-means 的目标函数、基本步骤和时间复杂度

15. 决策树

1. 自己学

16. EM 算法（爷爷）

1. 用高斯混合模型的例子讲明白 E 步骤和 M 步骤
2. 了解 EM 和 K-Means 的关系

17. 推荐系统 (爹)

1. 基于内容的推荐系统：三大步骤 (Plan of Action) 、优点和缺点
2. 基于协同滤波的推荐系统：三大步骤、相似性计算方法、基于用户的协同滤波和基于项的协同滤波的区别
3. 冷启动问题、稀疏性问题

18. Ensemble Learning 集成学习

19. PageRank、社区发现、关联规则挖掘、LSH 不考