线性回归，逻辑回归：

损失函数，如何通过梯度下降法求解参数，最优参数表达式（如果存在）

过拟合：

什么是过拟合，降低过拟合风险的方法

训练方法：

训练集-矫正集-测试集，留出法（Hold-out method）, Repeated hold-out method, k-fold cross-validation, k-fold cross-validation with validation and test sets, Bootstrap method

决策树

熵的求解，条件熵的求解，求解对应属性的信息增益，决策树的构建

SVM

SVM的基本思想，SVM的损失函数，SVM中的核函数

PCA

PCA的基本思想，PCA的不同理解角度，PCA的推导过程

聚类

K-Means聚类的思想、步骤、目标函数，K的选取问题，高斯混合聚类算法的思想和步骤

EM算法

EM算法中的E步和M步

推荐系统

基于矩阵分解的推荐系统，基于用户的协同推荐，基于商品的协同推荐，冷启动问题，数据稀疏的问题