**机器学习周末班-课后作业1**

1. 在模型评估过程中，过拟合和欠拟合具体是指什么现象？常用的过拟合解决方法有哪些？
2. 梯度下降法的分类及对应的优缺点？
3. 线性回归优化中，除了梯度下降法，还有什么别的优化方法？
4. 逻辑回归在特征较多时可以进行怎样的优化？
5. 常用的特征选择方法有哪些?
6. 为什么需要对数值型的特征做归一化？
7. 当使用逻辑回归处理多标签的分类问题时，有哪些常见做法？
8. 为什么 L1，L2 正则化可以防止过拟合？
9. 尝试解释：L1正则化使得模型参数具有稀疏性的原理
10. 案例题：利用逻辑回归对酒类数据进行分类，并对模型参数进行优化。

数据说明：Wine葡萄酒数据集是来自UCI上面的公开数据集，这些数据是对意大利同一地区种植的葡萄酒进行化学分析的结果，这些葡萄酒来自三个不同的品种。该分析确定了三种葡萄酒中每种葡萄酒中含有的13种成分的数量。在wine数据集中，这些数据包括了三种酒中13种不同成分的数量。文件中，每行代表一种酒的样本，共有178个样本；一共有14列，其中，第一个属性是类标识符，分别是1/2/3来表示，代表葡萄酒的三个分类。后面的13列为每个样本的对应属性的样本值。剩余的13个属性是，酒精、苹果酸、灰、灰分的碱度、镁、总酚、黄酮类化合物、非黄烷类酚类、原花色素、颜色强度、色调、稀释葡萄酒的OD280/OD315、脯氨酸。其中第1类有59个样本，第2类有71个样本，第3类有48个样本。