**周报**

1. 自学统计学习方法，以及对基于机器学习的人类活动分类数据集进行了训练和测试，测试数据准确率为96.6%。
2. 学习理解机器学习的基本知识，包括：监督学习，统计学习三要素，模型评估与模型选择，正则化与交叉验证，泛化能力，生成模型与判别模型以及三大主要问题模型：分类模型，标注模型和回归模型。
3. 学习理解感知机的实验原理及实现方式。感知机是一个对线性可分空间进行分类的模型，它的学习策略是在既定初始值的基础上，根据每次迭代的损失函数，计算误分类点上的梯度值，根据这一梯度进行权值及偏置的更新，停止条件为所有样本点均被正确分类。感知机有着易于实现，算法简单，而且在样本点是线性可分的情况下，感知机一定是收敛的。缺点就是只能对线性可分的样本进行分类，普适性不强。
4. 学习理解K近邻法处理分类问题。K近邻法更多的是一种“计算”模型，它的实现机制是在给定的所有样本点中，对新来的测试样本点进行逐一的计算距离，这里的距离一般是欧式距离，但也可以是更为普通的Lp距离，不同距离计算方法对最终的分类结果会有影响，在求得的距离中取前K个，即构成了K阶的一个列表，然后对这K个点以多数投票的方式进行最终分类标签的确定，从而实现分类。K近邻方法的原理是非常简单的，但是在大数据量的情况下，逐一的计算距离会非常缓慢，整个“学习”过程会非常漫长，所以就有了各种改进的K近邻算法，它们使用不同的优化搜索算法，包括Kd树等，是的搜索的速度大大加快，但每个算法都不是完美的，以kd树为例，当数据的维度与数据的数量相当时，kd树就会显得毫无优势，所以要针对各自的数据进行针对性的选择优化搜索算法。
5. 学习理解朴素贝叶斯法。朴素贝叶斯法是典型的生成学习方法，生成方法由训练数据学习联合概率分布，然后求得后验概率分布；具体的说，利用训练数据学习P(X|Y)和P(Y)，从而得到联合概率分布P(X, Y)，概率估计方法可以是极大似然估计或者贝叶斯估计。朴素贝叶斯分布的基本假设是条件独立性，这是一个较强的假设，由于这一假设，模型包含的条件概率的数量大为减少，朴素贝叶斯法的学习与预测大为简化，因而朴素贝叶斯法高效，且易于实现，其缺点是分类的性能不一定很高。