模型：类似y=f(x)=w1x+w0(模型参数取值待确定）

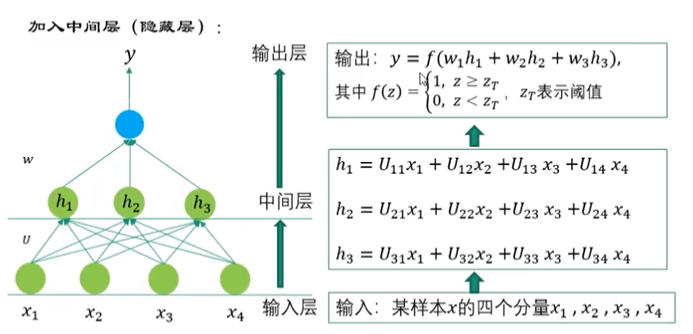
训练过程：实际上就是模型参数求解过程

通过训练模型参数有了最终取值

利用模型对待分类数据进行分类

从简单线性分类器到深度学习：

增加中间层

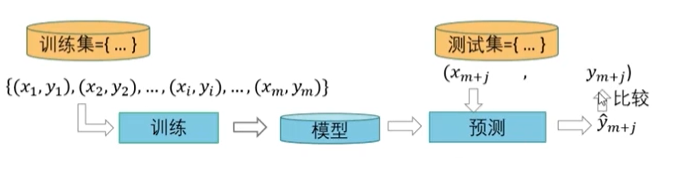


深度学习网络往往包含多个中间层

有监督学习中数据通常被分为训练集与测试集

训练集：用来训练模型，即被用来学习得到系统的参数取值

测试集：用于最终报告模型的评价结果，因此在训练阶段测试集中的样本应该是unseen的



通过训练集训练得到模型参数，通过测试集预测模型得到实验值与真实值比较

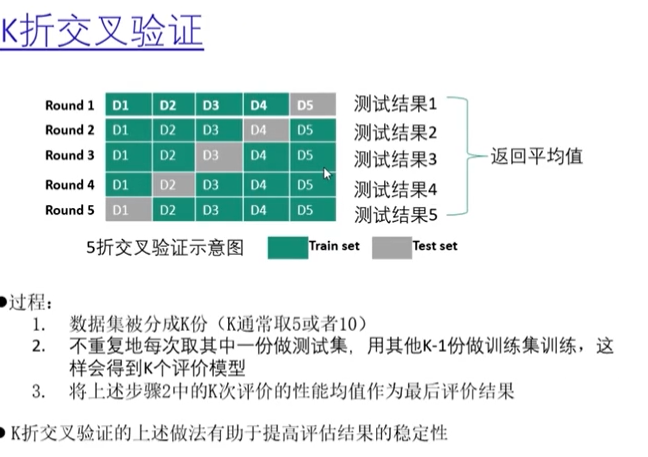
训练集测试集拆分--留出法

留出法数据拆分步骤：

1. 将数据随机分成两组，一组为训练集，一组为测试集
2. 利用训练集训练分类器，然后利用测试集评估模型，记录最后的分类准确率为此分类器的性能标准

优点：处理简单 缺点：基于数据集拆分基础上的性能评价结果不够稳定

k则交叉验证



分层抽样策略

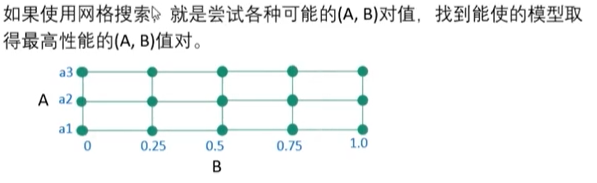
将数据集划分成k份，特点在于，划分的k分中，每份内各个类别数据的比例和原始数据集相同

网格搜索

超参数：学习过程之前需要设置其值的一些变量，而不是通过训练得到的参数数据

两个超参数：A和B

A为离散取值，B在区间[0,1],需进行离散化，如变为{0，.25，,0.50,0.75，1.0}

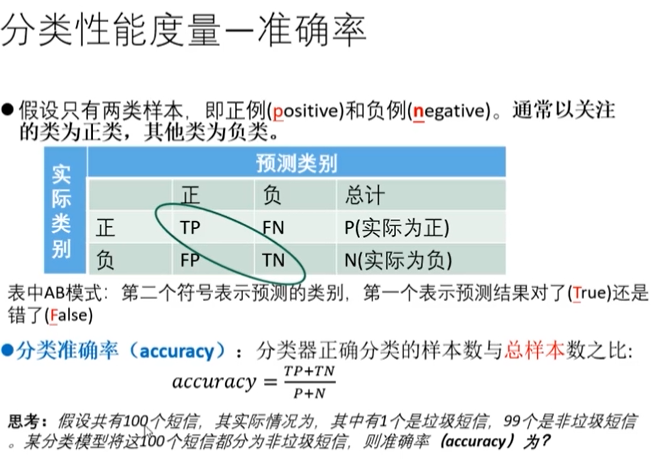


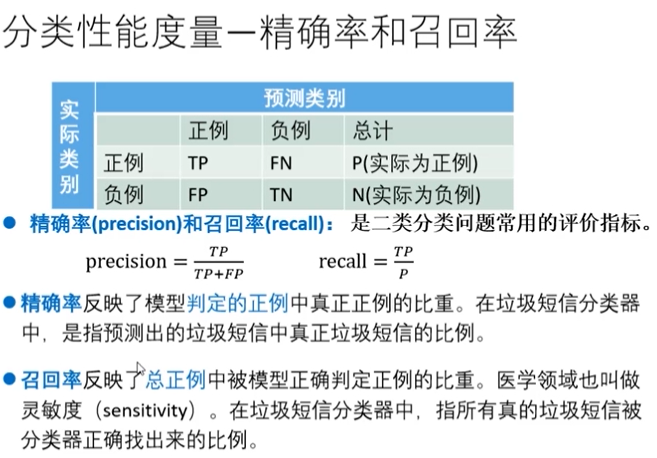
超参数取值的各种组合，在训练集上采用交叉验证的方法，比较哪种性能更好，从而得到最优超参数的取值组合

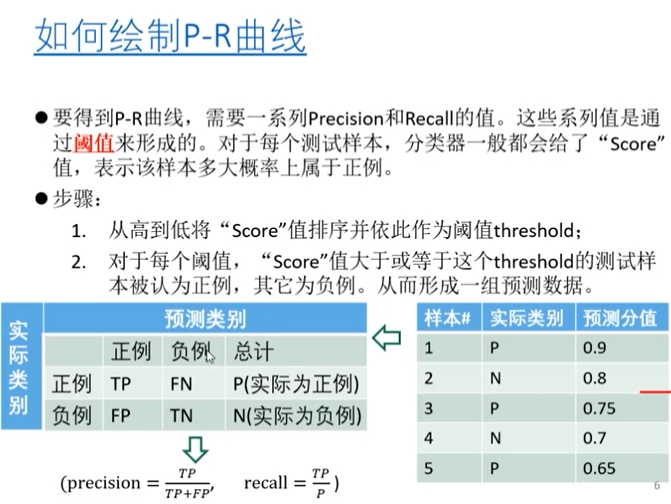
分类问题

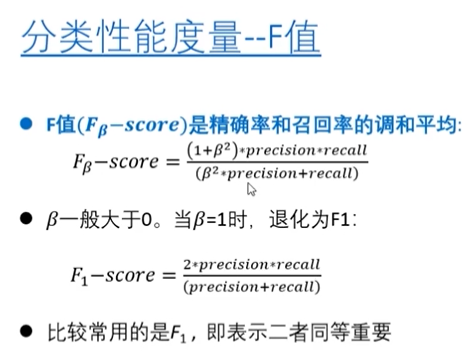
分类解决的是预测样本属于哪个或哪些预定义的类别

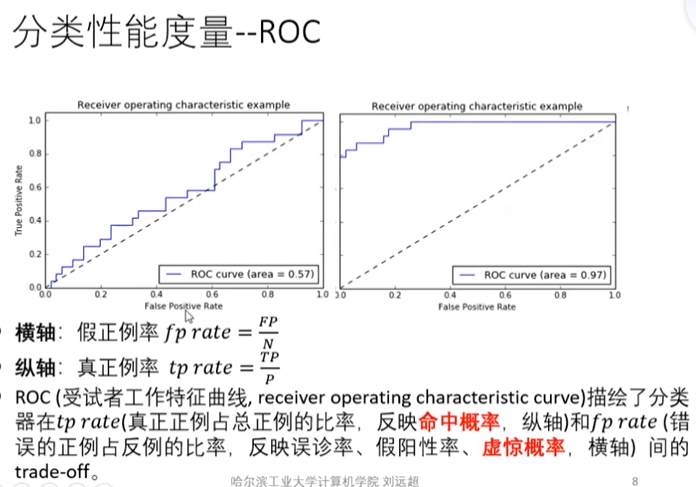
分类的机器学习的两大阶段：1.从训练数据中学习得到一个分类决策函数或分类模型，称分类器2.利用学习得到的分类器对新的样本进行类别预测  
分类性能度量

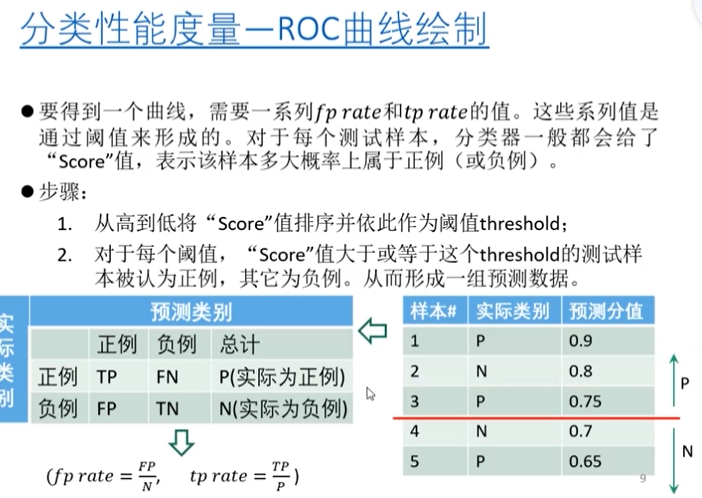


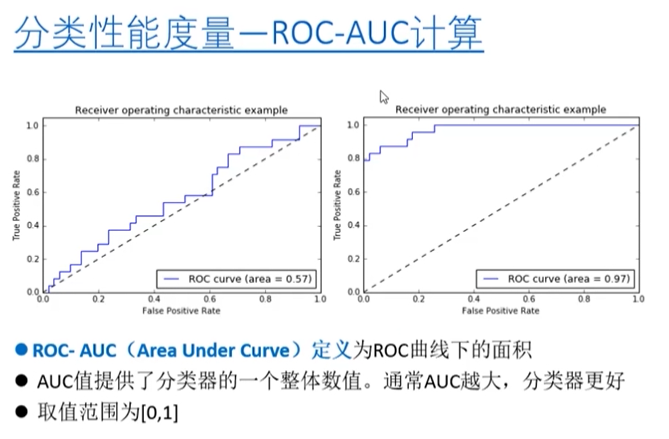












分类报告：显示每个类的分类性能。包括每个类标签的准确率、召回率、F1值等。