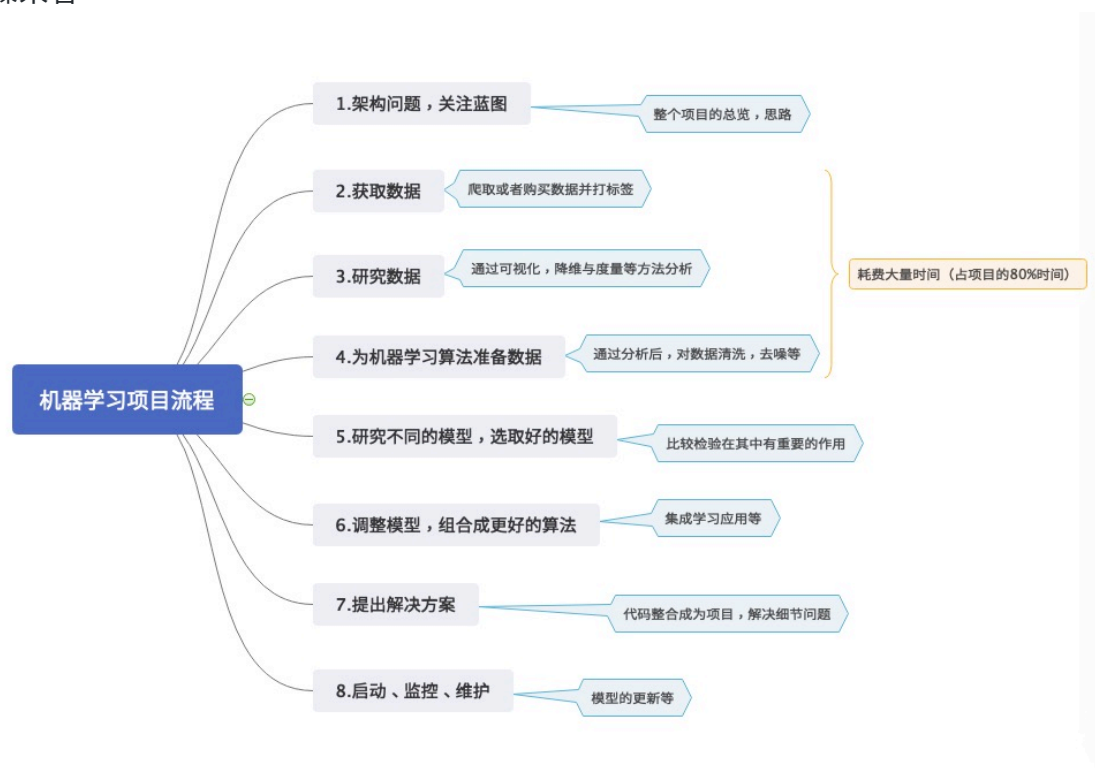


一.前言

比较检验很重要，比较检验很重要，比较检验很重要！重要的事说三遍！

二.为什么重要

我们通过某种实验评估测得学习器的性能度量，然后通过结果比较选择，那么如何比较呢？不是简单的比较性能度量的结果，实际远比这复杂的多,我们先从机器学习的步骤来看：

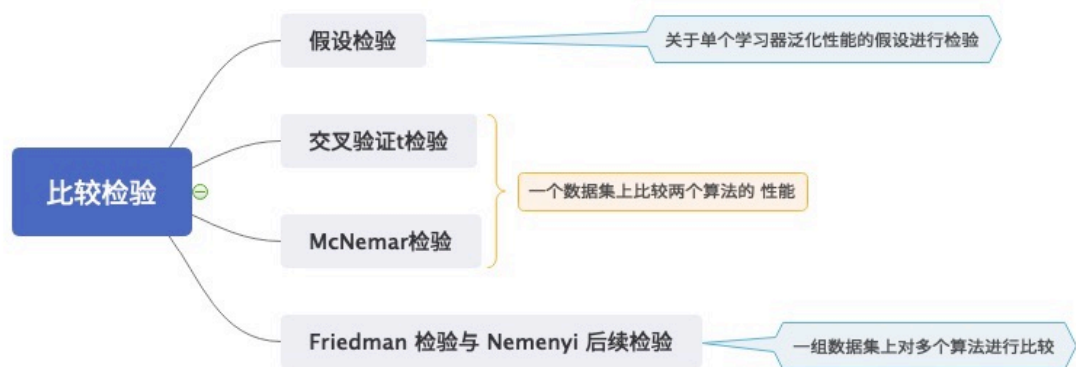


我们再来看看《机器学习》中的描述：

这里面涉及几个重要因素:首先，我们希望比较的是泛化性能，然而通过实验评估方法我们获得的是测试集上的性能，两者的对比结果可能未必相同;第二，测试集上的性能与测试集本身的选择有很大关系，且不论使用不同大小的测试集会得到不同的结果，即使用相同大小的测试集?若包含的测试样例不同，测试结果也会有不同;第二，很多机器学习算法本身有一定的随机性，即使用相同的参数 设置在同一个测试集上多次运行，其结果也会有不同.

那么如何进行比较检验呢？

三.比较检验的方法



假设检验

交叉验证T检验

McNemar验证

Friedman 检验与 Nemenyi 后续检验
