1. 最经典的拆分策略有三种： "一对一" (OvO) 、"一对其余" (OvR) 和 "多对多" (MvM).

2. 训练集和测试集划分常用的方法是 留出法 、 交叉验证法 、 自助法

3. 假设用P 来评估计算机程序在 某任务类T 上的性能，若一个程序通过利用 经验E 在T 中任务丰获得了性能改善，则我们就说关于T 和 P ，该程序对 E 进行了学习。

4. 一般地，令 D = {x1, x2, , , xm} 表示包含m个示例的数据集，每个示例由d个属性描述，则每个示例 xi =( xi1, xi2, , , xid) 是d维样本空间X中的一个向量， ，其中xij是xi在第j个属性上的取值, d称为样本xi的 "维数" .

5. 根据训练数据是否拥有标记信息，学习任务可大致划分为两大类 "监督学习“ 和 "无监督学习" (unsupervised learning) ，分类和回归是前者的代表，而聚类则是后者的代表.

6. 通常我们把分类错误的样本数占样本总数的比例称为"错误率" (error rate) ，即如果在m 个样本中有α 个样本分类错误，则错误率 E= α/m; 相应的，1 一α/m 称为"精度"，即"精度=1-错误率"。

7. 我们把学习器的实际预测输出与样本的真实输出之间的差异称为"误差" (error) ,学习器在训练集上的误差称为"训练误差" (training error)或"经验误差" (empirical error) ，在新样本上的误差称为"泛化误差" (generalization error).

8. 当学习器把训练样本学得"太好"了的时候，很可能巳经把训练样本自身的一些特点当作了所有潜在样本都会具有的一般性质，这样就会导致泛化性能下降这种现象在机器学习中称为"过拟合" (overfitting). 与此相对的是"欠拟合" (underfitting) ，这是指对训练样本的一般性质尚未学好.

9. 根据个体学习器的生成方式，目前的集成学习方法大致可分为两大类，即个体学习器问存在强依赖关系、必须串行生成的序列化方法，以及个体学习间不存在强依赖关系、可同时生成的并行化方法;

10. 个体学习器问存在强依赖关系、必须串行生成的序列化方法，代表算法是Boosting ，以及个体学习间不存在强依赖关系、可同时生成的并行化方法的代表是Bagging 和"随机森林" (Random Forest).

11. 根据预测目标的不同，学习任务可以分为 分类 、 回归 和 聚类 。

12. 决策树的核心技术包含 划分选择 和 剪枝处理 。

13. 卷积神经网络的三个主要计算层 卷积层 、池化层 和\_\_\_\_连接层。

14. 集成学习的一般结构，即先产生一组 基学习器 ，再用某种策略将它们结合起来。

15. 集成学习中的集成主要分为 同质集成 和 异质集成 两种。

16. 高斯混合聚类，采用 概率 模型来表达聚类原型，簇划分则由原型对应 后验概率 确定.。

17. 基于线性变换来进行降维的方法称为 PCA 。

18. 若要获得高维空间对应的低维子空间，最简单是对原始高维空间进行 线性变换 。