# 作业二：模型评估及选择

《机器学习》 周志华著，第二章 模型评估及选择

总分50分。

### 1. 评估方法（留出法）

**题目：**数据集包含1000个样本，其中500个正例、500个反例，将其划分为包含70%样本的训练集和30%样本的测试集用于留出法评估，试估计共有多少种划分方式。**（10分）**

**答案：**

### 2. 评估方法（交叉验证法）

**题目：**数据集包含100个样本，其中正、反例各一半，假定学习算法所产生的模型是将新样本预测为训练样本较多的类别（训练样本数相同时进行随机猜测），试给出用10折交叉验证法和留一法分别对错误率进行评估所得到的结果。**（20分）**

**答案：**

10折交叉验证：交叉验证中每个子集数据分布要尽可能保持一致，那么本题中10次训练中每次正反例各占45，模型训练结果随机猜测，错误率期望为50%。

留一法：若留出样本为正例，训练集中则有50个反例和49个正例，模型预测为反例；反之留出样本为反例，模型预测为正例，错误率为100%。

### 3. 性能指标

**题目：**试述真正例率（$TPR$）、假正例率（$FPR$）与查准率（$P$）、查全率（$R$）之间的联系。**（10分）**

**答案：**

查全率: 真实正例被预测为正例的比例

真正例率: 真实正例被预测为正例的比例

显然查全率与真正例率是相等的。

查准率:预测为正例的实例中真实正例的比例

假正例率: 真实反例被预测为正例的比例

两者并没有直接的数值关系。

### 4. 习题一：性能指标（错误率、精度、查准率、查全率、F1）

**题目：** 在癌症辅助诊断（判断是否患癌症）中，假设有1000名受测者，其中有2位癌症患者（正例）。现有以下两个模型及其对应的预测结果信息：

* 模型I：全部判断为没有癌症；
* 模型II：2位患者预测正确，另有4位反例预测为患病，其余预测正确。

计算模型I、模型II对应的以下性能指标：错误率、精度、查准率、查全率、F1。写出计算过程。**（10分）**

**答案：**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 错误率 | 精度 | 查准率P | 查全率R | F1 |
| 模型I | 2/1000 | 99.8% | 0 | 0 | 0 |
| 模型II | 4/1000 | 99.6% | 1/3 | 1 | 3/2 |