

第八课 机器学习简单介绍

第14节 模型测试及评价

主 讲: Robin

模型测试及评价

评估模型

- 不同的应用有着不同的目标，不同的评价指标
- **准确率 (accuracy)** 是最常见的一种
- 但是准确率越高，模型就越好么？
- 假设，在1000个样本中，有999个正样本，1个负样本（不均衡数据集）
如果全部预测正样本，就可以得到准确率**99.9%**！这样的场景有：信用卡欺诈检测，离职员工检测等。
- 有些任务更关心的是某个类的准确率，而非整体的准确率

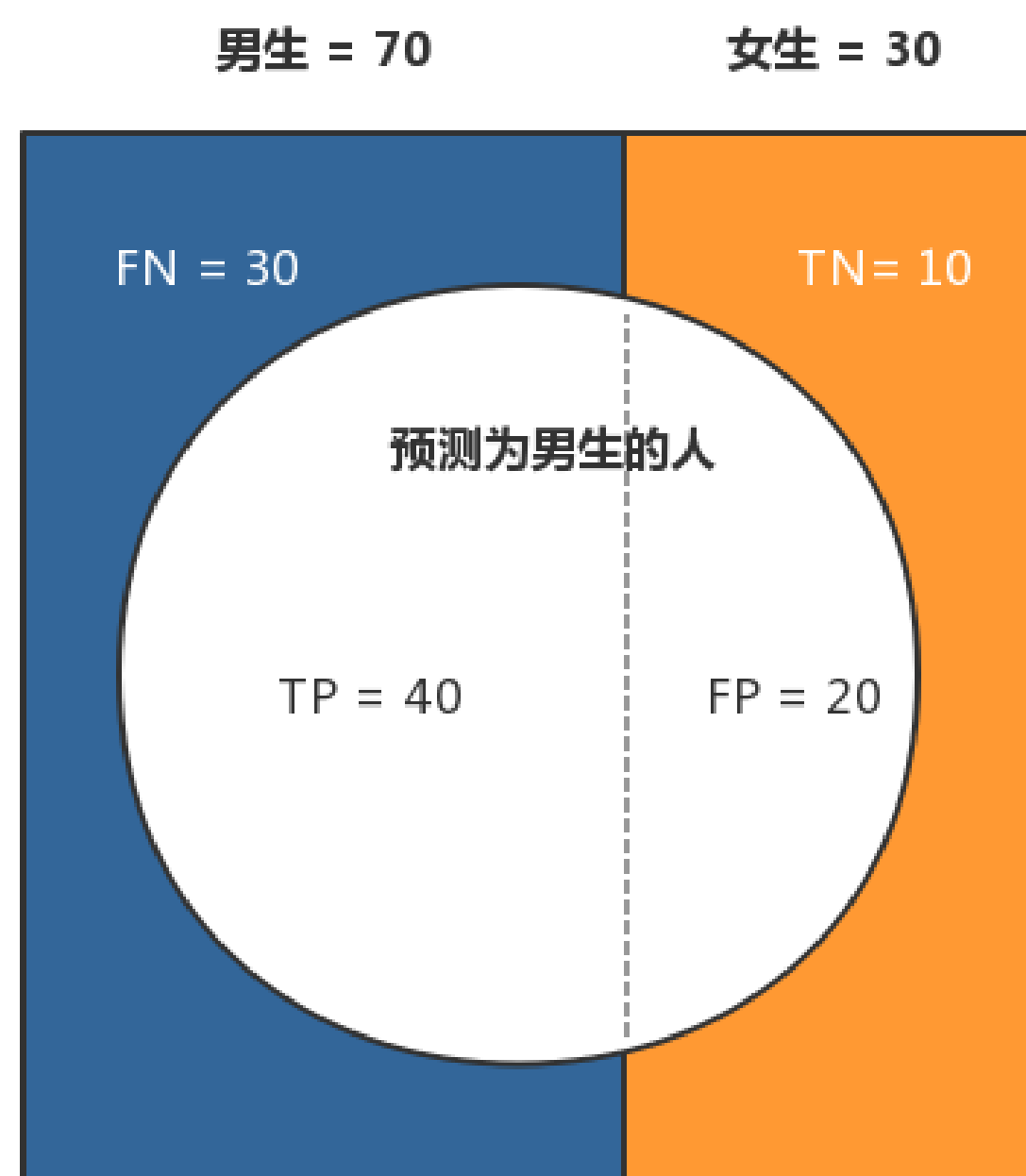
模型测试及评价

- 真正例(TP), 预测值是1, 真实值是1。被正确分类的正例样本。
- 假正例(FP), 预测值是1, 但真实值是0
- 真反例(TN), 预测值是0, 真实值是0
- 假反例(FN), 预测值是0, 但真实值是1
- 召回率(Recall): $TP / (TP + FN)$
 - 表示在所有正样本中, 被预测正确的个数, 即查全率
- 精确率(Precision): $TP / (TP + FP)$
 - 表示在所有预测为正样本中, 被预测正确的个数, 即查准率
- 注意: Recall和precision只适用于二分类问题

		Prediction	
		Positive	Negative
Ground truth	Positive	True positive (TP)	False negative (FN)
	Negative	False positive (FP)	True negative (TN)

模型测试及评价

- 例子：假设一个班级有100个学生，其中
 - 男生70人（蓝色矩形），女生30人（橙色矩形）
 - 一个分类器根据学生的身高和体重预测其性别，结果如下



- TP：实际为男生，预测为男生；
- FP：实际为女生，预测为男生；
- FN：实际为男生，预测为女生；
- TN：实际为女生，预测为女生；



- 准确率 = $(TP + TN) / \text{总样本} = (40 + 10) / 100 = 50\%$
- 精确率 = $TP / (TP + FP) = 40 / 60 = 66.67\%$
- 召回率 = $TP / (TP + FN) = 40 / 70 = 57.14\%$

模型测试及评价

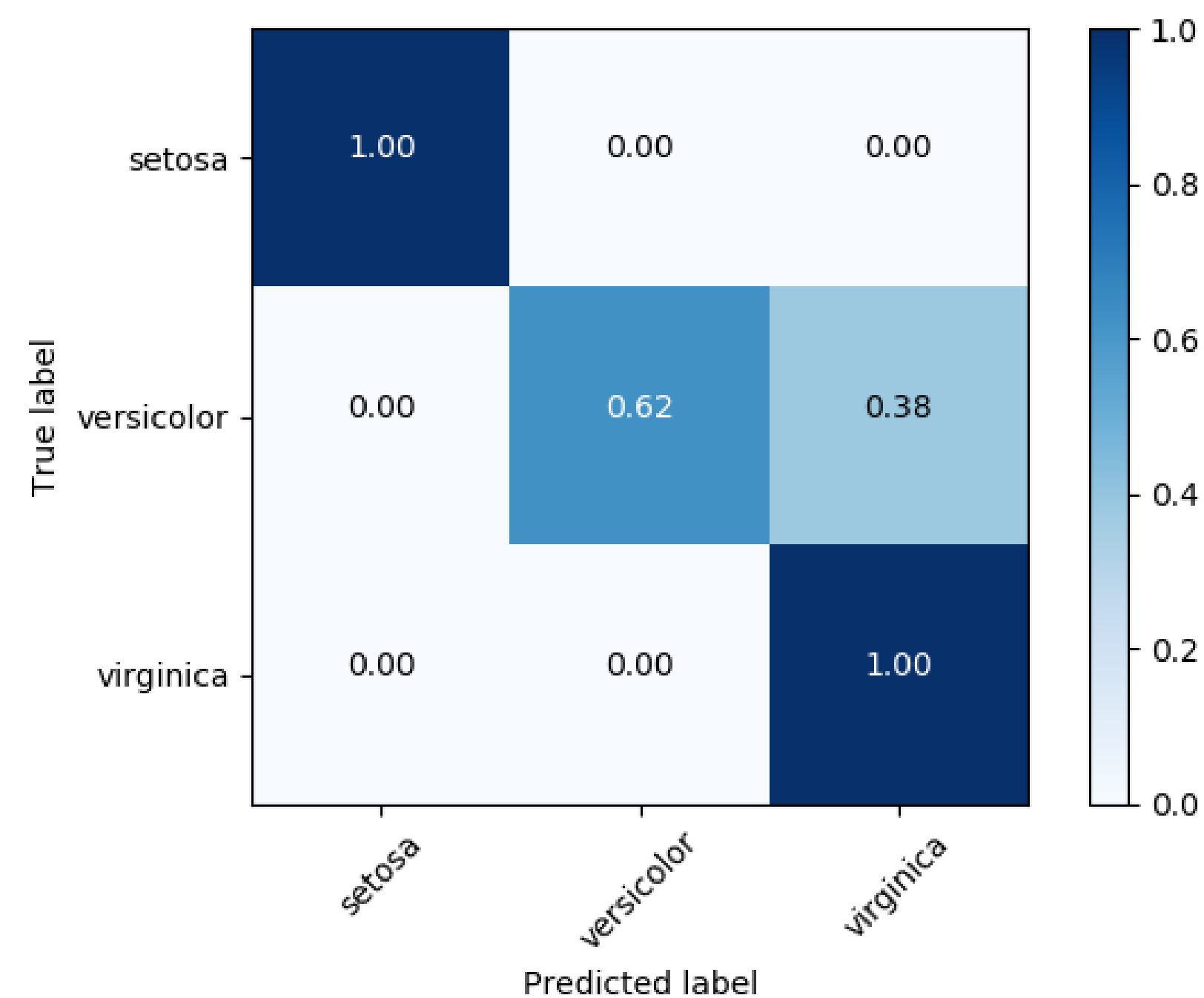
- F1值
- 将召回率和精确率用一个数值表示

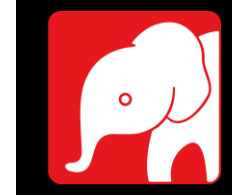
$$F_1 = 2 \cdot \frac{1}{\frac{1}{\text{recall}} + \frac{1}{\text{precision}}} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

- sklearn.metrics中包含常用的评价指标
 - accuracy_score
 - precision_score
 - recall_score
 - f1_score

模型测试及评价

- 混淆矩阵 (confusion matrix)
- 可用于多分类模型的评价
- `sklearn.metrics.confusion_matrix()`





只为遇见明天更优秀的你！