# 评估模型分类能力

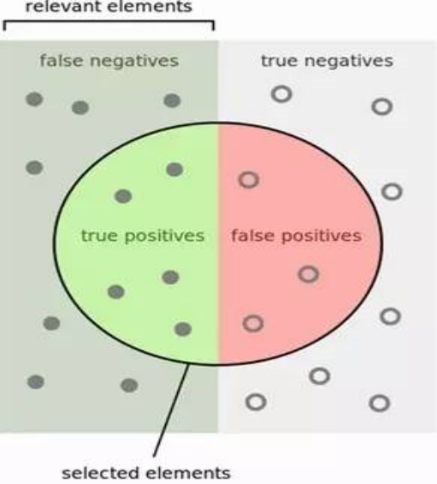
## 1.相关参数

①TP(True Positives)：预测为正样本，实际也为正样本的特征数

②FP(False Positives)：预测为正样本，实际为负样本的特征数

③TN(True Negatives)：预测为负样本，实际也为负样本的特征数

④FN(False Negatives)：预测为负样本，实际为正样本的特征数



绿色的半圆是TP

红色的半圆是FP

左边的灰色长方形（不包括绿色半圆）是FN

右边的浅灰色长方形（不包括红色半圆）是TN

绿色和红色组成的圆内代表我们分类得到模型结果认为是正值的样本。

⑤精确度P: 

⑥召回率R：，灵敏度TPR，正样本中正确识别正样本比例：

⑦特异性S：，FPR为1-S，负样本中错误识别为正样本比例：

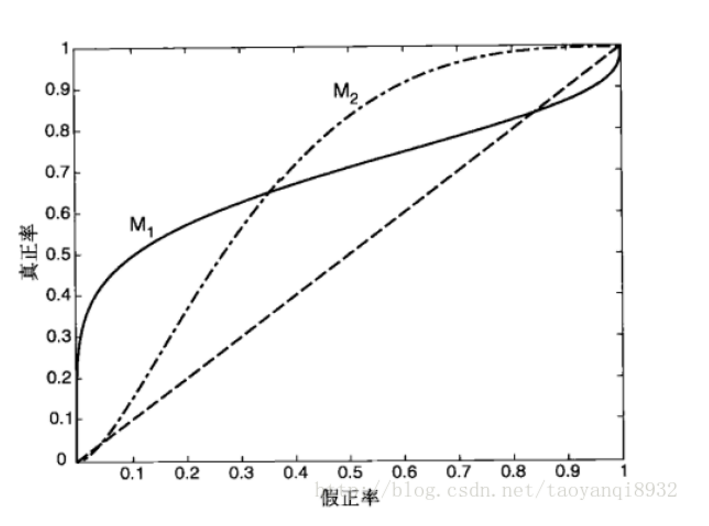
⑧综合评价精确度和召回率 (两者一视同仁)：

⑧综合评价精确度和召回率（两者间有偏好）：

（偏好召回率，偏好精确度）

## 2.ROC曲线和PR曲线

ROC曲线以TPR为y轴，以FPR为x轴。TPR越高，FPR越小，模型和算法就越高效,也就是ROC曲线越靠近左上越好，曲线下方的面积AUC（Area Under Curve）。



PR曲线以精确率P为y轴，以召回率R为x轴。P越高，R越高，模型和算法就越高效,也就是PR曲线越靠近右上越好。

