

BIOF3 FIGCODE · SCI 绘图代码集

二分类评估 (ROC + Confusion)

导出日期: 2026年6月27日



扫码关注微信公众号【生信F3】
获取文章完整内容，分享生物信息学最新知识。

分类: 生存分析 | 依赖包: pROC、ggplot2、patchwork、caret

解决的生物学问题

我的模型在阳性 / 阴性样本上的区分能力如何? 错分主要发生在哪种类型?

应用场景

- 临床诊断模型评估
- 癌症亚型分类性能
- 基于 omics 的风险分层
- 机器学习模型对比

输入数据格式

数据框: truth (0/1), prediction_score (0-1), prediction_class (0/1)

关键参数

- threshold (分类阈值, 默认 0.5)
- positive class (哪类为阳性)
- pROC::ci.auc (AUC 置信区间 bootstrap 次数)

快速开始

1. 直接在线运行

打开 [FigCode 在线绘图](#), 点击本工具卡片的"在线绘图"按钮即可用内置示例数据出图, 零环境配置。

2. 本地复现

下载脚本和示例数据，本地 RStudio 运行：

```
# 下载
curl -O https://<your-site>/figcode/scripts/roc-confusion.R
curl -O https://<your-site>/figcode/data/roc-confusion.csv
```

3. 安装依赖

```
# CRAN 包
install.packages(c("pROC", "ggplot2", "patchwork", "caret"))

# Bioconductor 包 (如需)
# BiocManager::install(c())
```

完整代码

```
library(pROC)
library(ggplot2)
library(patchwork)
library(caret)

# --- Demo data ---
set.seed(42)
n <- 300
truth <- rbinom(n, 1, 0.4)
score <- truth * 0.7 + rnorm(n, 0, 0.4)
score <- (score - min(score)) / (max(score) - min(score))
pred <- as.integer(score > 0.5)

# --- ROC ---
roc_obj <- roc(truth, score, ci = TRUE, plot = FALSE)
auc_str <- sprintf("AUC = %.3f (95% CI: %.3f-%.3f)",
                   auc(roc_obj), ci.auc(roc_obj)[1], ci.auc(roc_obj)[3])

p1 <- ggroc(roc_obj, color = "#dc2626", size = 1.2) +
  geom_abline(intercept = 1, slope = 1, linetype = "dashed", color = "grey60") +
  annotate("text", x = 0.4, y = 0.2, label = auc_str, size = 4) +
  labs(title = "ROC Curve", x = "Specificity", y = "Sensitivity") +
  theme_classic()

# --- Confusion matrix ---
cm <- confusionMatrix(factor(pred, levels = c(0,1)),
```

```
      factor(truth, levels = c(0,1)),
      positive = "1")
cm_df <- as.data.frame(cm$table)

p2 <- ggplot(cm_df, aes(Reference, Prediction, fill = Freq)) +
  geom_tile() +
  geom_text(aes(label = Freq), size = 6, color = "white") +
  scale_fill_gradient(low = "#bfbdbfe", high = "#1d4ed8") +
  labs(title = sprintf("Confusion Matrix (Acc = %.3f)", cm$overall["Accuracy"]),
       x = "Truth", y = "Prediction") +
  theme_classic() +
  theme(legend.position = "none")

p1 + p2
```

替换为自己的数据

脚本中以 `# --- Demo data ---` 标注的段落是示例数据生成代码。替换为自己的数据时，**保持列名一致**即可：

- 输入格式：数据框：truth (0/1), prediction_score (0-1), prediction_class (0/1)
- 使用 `read.csv()` / `readRDS()` 读取本地文件

延伸阅读

- 相关教程：[二分类评估完整流程](#)
- 出现 bug？欢迎在 [FigCode 页面](#) 点击对应卡片，在评论区留言。