

R 语言实战系列之

Practical methods for learning data science in R

古德漠鹰

March 1, 2024

PMLDSR_0004: 临床预测模型中的 二分类逻辑回归最小样本量的评估

- 临床预测模型 (Clinical Prediction Models)
 - 诊断模型 (diagnostic model)
 - 预后模型 (prognostic model)
 - 传统回归方法 (e.g., 逻辑回归) vs. 高级方法 (e.g., 随机森林)
 - 重要问题之一：针对二分类逻辑回归如何评估最小样本量
 - 降低模型过度拟合
 - 确保预测模型关键参数估计准确性
 - 经验法则：10-EPP (events per parameter), 5-EPP, 15-EPP
 - 超越经验法则的方法 (Riley 等人提出的标准)
- 在RStudio中执行实战操作
 - R 包：pmsampsize
 - 举例：二分类逻辑回归模型
 - 检验已建立的模型样本量
 - 建立新模型所需样本量
 - 讨论参数变化对样本量评估的影响
 - 介绍一些相关论文