

广东省公共卫生研究院

粤公卫院函〔2026〕62号

广东省公共卫生研究院关于举办 R 语言 机器学习疾病风险预测与评估技术 培训班的通知

各有关单位：

当前人工智能技术快速发展，以机器学习为核心的分析方法已成为疾病风险预测和评估的重要技术手段。为进一步提升卫生健康领域专业技术人员和科研工作者数据分析及机器学习模型应用能力，我院将于 10 月下旬在广州市举办“R 语言机器学习疾病风险预测与评估技术培训班”。现将相关事项通知如下：

一、培训目标

采用理论授课结合实操演练的教学模式，帮助学员系统掌握 R 语言的数据处理、数据可视化、机器学习算法及模型构建等技能，提升业务与科研工作中的数据技术应用水平。

二、培训内容

- （一）R 语言概述及 AI 辅助编程；
- （二）数据管理及可视化技术；
- （三）机器学习概述与模型评估技术；
- （四）机器学习聚类与降维算法原理与实践；

(五) 机器学习分类预测模型原理和实践;

(六) 机器学习回归预测模型原理和实践。

具体日程安排详见附件。

三、培训对象

各级医疗卫生机构专业技术人员、医学院校及科研院所研究人员、研究生,以及其他对医学健康数据分析感兴趣的相关人员。

四、培训安排

(一) 时间: 10月27—31日, 27日下午报到, 28—30日上课, 31日上午离会。

(二) 培训方式: 线下培训与线上培训同步, 线上培训通过小鹅通平台直播。

(三) 线下培训地点: 美豪丽致酒店(广州番禺长隆南村万博地铁站店), 地址: 广州市番禺区兴南大道1号。

(四) 培训费用

1. 线下培训: 培训费2700元/人, 含学费、场地费、资料费和午餐费。学员的早餐、晚餐、住宿和交通费用需自理。

2. 线上培训: 培训费2200元/人, 包含学费、资料费。

3. 收费方式: 请在开班前通过银行转账缴费, 收款账户如下:

收款单位: 广东省公共卫生研究院

账号: 3602 0887 0920 0021 474 (写全19位)

开户银行: 工商银行广州怡乐支行

请务必备注: R 培训班线上/线下+姓名

(五) 报名方式: 学员请于10月20日前扫描下图二维码填写报名信息。如有疑问, 请联系沈老师或朱老师: 020-31051612;

或在线咨询会务组（咨询 QQ 群号：801116971）。



报名二维码

五、其他事项

（一）学员群组管理。报名成功后，工作人员将组织学员加入培训专用 QQ 群。培训资料、直播方式及相关通知均通过该群发布。培训资料含安装软件、专用讲义，以及实践数据和代码。

（二）软硬件准备。请学员准备 Windows 系统电脑，提前安装好 R 软件和 RStudio 软件。会务组将在 QQ 群上传相关软件和安装指南，请学员按照指引完成预装。

（三）发票开具。培训费发票统一开具增值税电子普通发票，于培训结束后一周内发送至学员预留邮箱。

（四）住宿安排。学员请自行联系酒店预订房间。本次培训酒店为美豪丽致酒店，培训学员可享协议优惠价，预定电话/微信：19875286731（酒店伍经理）。广交会期间房间紧张，建议提前预订。

附件：R 语言机器学习疾病风险预测与评估技术培训班日程表



附件

R 语言机器学习疾病风险预测与评估技术
培训班日程表

时间：2026 年 10 月 27—31 日 地点：广州市 美豪丽致酒店

时间		课程名称	课程内容
10 月 27 日下午		报到	
10 月 28 日	9:00-12:00	R 语言概述及 AI 辅助编程	(1) R 语言基础：R 软件及 R 程序包安装、数据结构、编程基础、数据导入与导出等；(2) AI 辅助编程：AI 助手和 AI 插件辅助编程，如代码解读、生成代码等
	14:15-17:15	数据管理及可视化技术	(1) 数据管理技术及实践：基于 tidyverse 包的健康数据管理和批量处理等；(2) ggplot2 绘图语法和技巧：绘制条图、折线图、散点矩阵图、箱式图、热力图、时序图、森林图等
10 月 29 日	9:00-12:00	机器学习概述与模型评估技术	(1) 机器学习定义、核心特性与分类，优势和价值，及其卫生健康领域的应用场景；(2) 机器学习常用方法和基本分析流程；(3) 机器学习模型评估和选择，核心评估指标，模型选择方法，机器学习全流程工作流等
	14:15-17:15	机器学习聚类与降维	(1) 聚类分析：K-means 聚类 and 层次聚类原理和实践(案例：基于地区社会经济卫生等指标的聚类分析) (2) 主成分分析原理和实践(案例：药物化合物成分质量评价指标的主成分分析) (3) 因子分析原理和实践(案例：健康体检指标体系的因子分析与维度提取)
10 月 30 日	9:00-12:00	机器学习分类	(1) 决策树分类预测(案例：乳腺癌复发风险预测分析) (2) 支持向量机分类预测(案例：儿童百日咳重症的影响因素及预测分析) (3) SHAP 可解释性分析研究预测模型中特征变量的贡献，绘制蜂群图和瀑布图等
	14:15-17:15	机器学习回归	(1) 多项式回归预测(案例：年龄、BMI 等与血压/血糖的非线性关系与预测分析) (2) 最近邻回归(KNN)预测(案例：基于健康体检资料的血压/血糖值预测分析) (3) 随机森林回归预测(案例：登革热传播的环境影响因素重要性排序及其流行趋势预测分析)
10 月 31 日上午		离会	