

# 机器学习分析多组学课程安排

## 第一节课 **AI+多组学 CNS** 文思路解读

- 1.基于 Deepseek 高效阅读多组学的 CNS 生信文章
- 2.AI 整理代谢组、蛋白组、宏基因组等多组学数据分析流程
- 3.Deepseek 从多篇生信文献中快速总结提炼多组学数据分析方法
- 4.Deepseek 评价多组学生信文章思路的创新性和数据分析可行性

## 第二节课 **Deepseek** 辅助多组学生信课题设计

- 1.Deepseek 辅助代谢组、蛋白组、宏基因组联合分析策略
- 2.AI 指导代谢组、蛋白组、宏基因组、转录组交叉验证的思路
- 3.Deepseek 汇总生信论文模板设计的层次和逻辑要点
- 4.Deeseek 辅助生信多组学数据研究热点挖掘
- 5.Deeseek 构建多个公共数据库联合验证有效策略

## 第三节：编程基础学习--R 语言

1.R 和 Rstudio 的安装、环境配置

2.R 语言简单语法及常见命令

3.以 Cell 文章方法描述学习 R 包的安装及使用

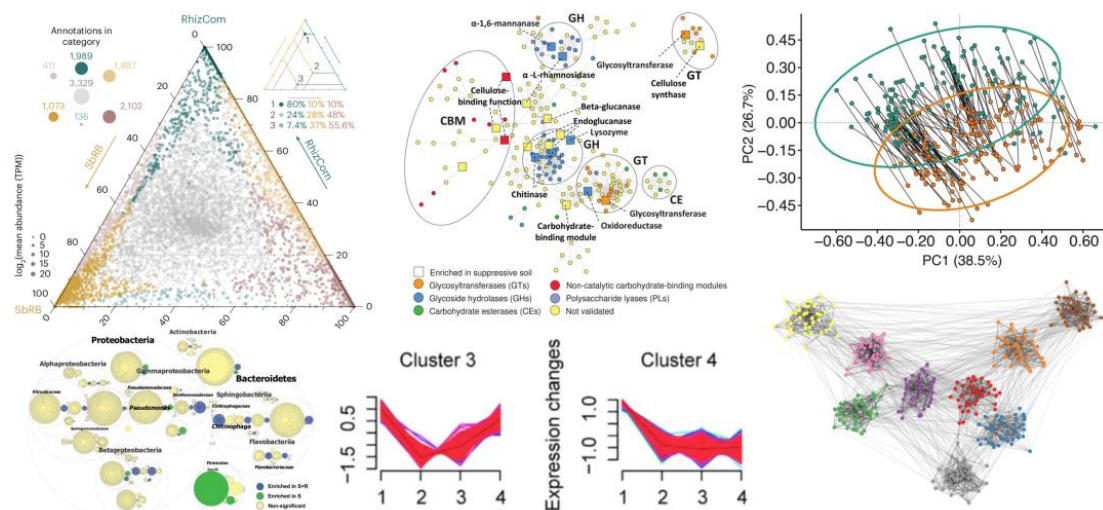
4.以 Nature 文章源代码学习重点函数基础代码

5.向量、矩阵、数据框和列表的创建和索引

6.自定义 Function 函数构建

7.for 循环、字符型数据的处理【Cell】

8.云服务器的使用方法



## 第四节：机器学习代谢组数据分析

1.deepseek 对 CNS 文章代谢组学系统解读

2.三大常见代谢物库：HMDB、METLIN 和 KEGG;

代谢组学原始数据库: **Metabolomics Workbench** 和 **Metabolights**

3.无监督式机器学习在代谢组学数据处理中的应用

4.三种回归用于代谢组学数据分析: 线性回归、**Logistic** 回归与 **Cox** 回归

5.PLS-DA 算法: 决策树, 随机森林和贝叶斯网络模型

6.多尺度嵌入式代谢物共表达网络分析

7.TDEseq: 识别不同分组条件下基因表达趋势的变化

8.Nature 子刊代谢分析的源代码复现

## 第五节: 机器学习蛋白组学数据分析

1.蛋白组无监督聚一致性聚类(**Consensus Clustering**)

2, **PPI** 蛋白互作网络构建

3.构建通路富集网络: 通路基因集之间的相似性, 将富集的通路表示为一个相互连接的簇网络

4.蛋白组加权基因共表达网络分析 **WGCNA**

5.主成分分析不同分组变量差异结果展示



1.meconetcomp: 微生物共生网络比较

2.进化树揭示物种间亲缘关系

3.微生物组的扩增子测序分析

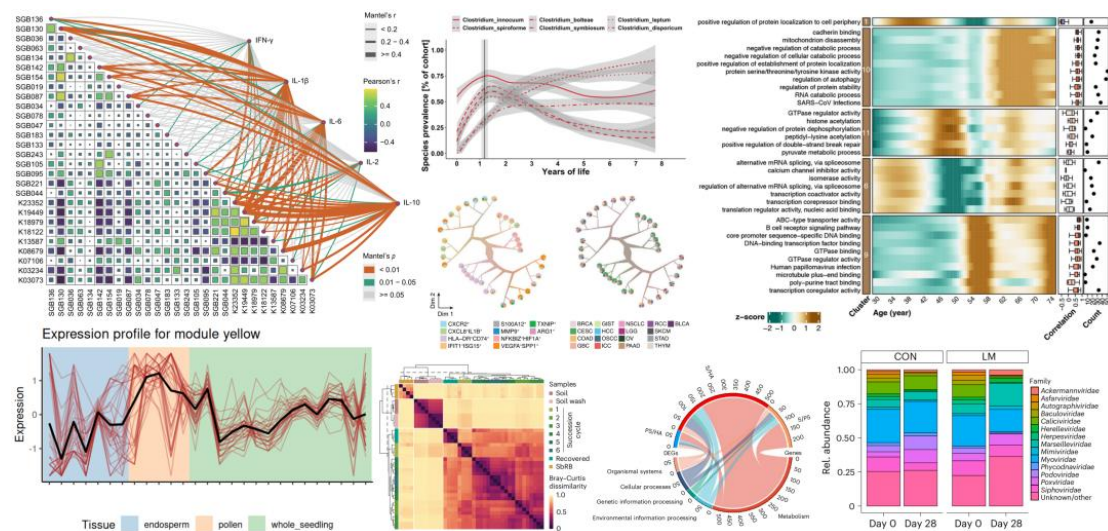
4.不同条件下的微生物群落差异

5.Mantel 检验群落距离矩阵和不同处理组之间的相关性

6.微生物群落和环境数据机器学习分析

7.Anova 进行微生物功能丰度差异检验

8.Nature Microbiology 文章源代码复现



## 第八节：机器学习转录组数据分析

1.随机效应模型 (REM) 方法对差异基因进行 meta 分析

2.差异基因的转录因子富集构建调控网络

3.富集分析: GSEA、GSVA 分析

4.相关性聚类热图识别转录因子共表达模块

5.CIBERSORT 免疫浸润

6.调控网络图展示细胞动静状态--Regulatory network

7.TO-GCN 时序基因共表达网络

8.CELL 主刊文章源代码复现

## 第九节: 机器学习应用于多种组学筛选重要特征

1.图神经网络预后模型构建进行预后基因挖掘

2.xgboost 模型: 基于梯度增强的决策树算法

2.随机森林模型

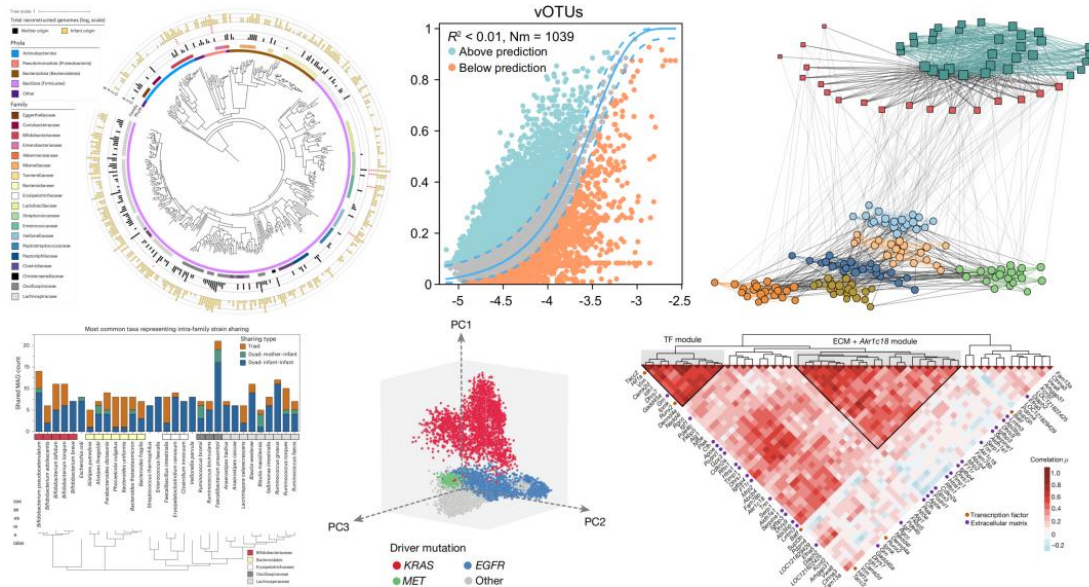
3.lasso 回归等模型: 线性回归的另一种正则化形式

4.弹性网络回归:结合了岭回归和 Lasso

5. 朴素贝叶斯和支持向量机

7, 多种机器学习算法用于各组学筛选重要的特征: 代谢物、肠道微生物、mRNA、蛋白质

8.Nature Medicine 文章源代码复现



## 第十节：网络药理学分析

1. deepseek 解读网络药理学研究思路

2. 药物靶点的查询，复方、单体网络靶标定位

3. 药物成分与疾病的关联分析，构建疾病与中药成分的网络

4. 网络分析（如 PPI 互作、模块挖掘）

5. 拓扑特征（度中心性、介数中心性）预测关键靶点

6. 基于 AI 的组学数据网络关系挖掘

## 第十一节：分子对接

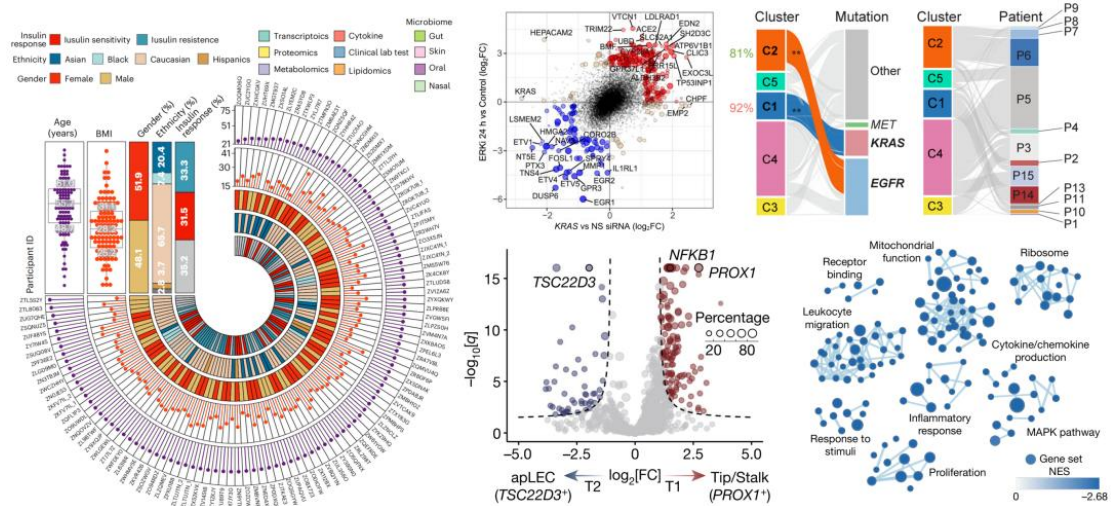
1. AIDD 人工智能药物发现与设计的基本介绍

2. 配体小分子和受体白蛋白大分子的前处理

- 3.建立和运行 Grid 格点文件，建立对接格点空间文件
- 4.使用 Pymol 软件查看详细的结合位点，相互作用的氢键
- 5.结合能的计算，结合位点的确定

## 第十二节：代谢组学与转录组学、蛋白组学联合分析

- 1.主成分分析代谢组学和蛋白组学变量差异
- 2.基于共表达网络（GCN）和调控网络（GRN）探索多组学数据的拓扑特性
- 3.多组学 3D 数据可视化：整合分析
- 4.Metscape 的使用
- 5.crosslink 转录-蛋白-代谢 多组学相关性可视化
- 6.多组学联合分析三维多层网络：一个级别基因互作关系，第二层是蛋白/代谢物，层间边缘（连线）是基因和代谢物的生物学关系
- 7.使用 Deepseek 设计分析基于蛋白组+代谢组的多器官视角揭示脓毒症早期的能量代谢适应性，并生成初步代码框架
- 8.Nature Aging 文章源代码复现



### 第十三节课: Deepseek 解读肠道微生物组学与代谢组学、蛋白组、转录组联合分析课题思路设计

1. Deepseek 设计网络毒理学和多组学的综合分析揭示肝毒性的潜在机制，并生成初步课题思路框架
2. Deepseek 设计微生物+代谢组的小鼠下丘脑回路对肠道微生物群组成的快速调节，并生成课题思路框架
3. Deepseek 设计微生物+转录，肠道细菌促结直肠癌新机制，并生成课题思路框架
4. Deepseek 设计微生物+蛋白，灵芝多糖对阿尔茨海默病预防作用的机制，并生成课题思路框架

### 第十四节课: AI 辅助生信论文投稿与发表

1. Deepseek 辅助科研论文选刊 (实操+指令)

2.GPT 辅助文本的润色、改写与扩写

3.AI 自动摘要、删减与提炼关键词句

4.Deepseek 辅助拒稿处理、推荐更适合期刊, 提供投稿建议 (实操+指令)