- ☑课程助教
  - ☑李钰心、伍智广、曲希然
- ☑项目过程
  - ☑每人报名一个助教的题目、每位助教带人上限36人
  - 図6人一组选一个任务,组内每人选取一种模型(不重复)
  - **三具体细节还可以协商解决**
- ☑ 评分细则
  - ☑问题分析与数据集(5′)、数据预处理(10′)、方法调节 (10′)、结果分析与结论(10′)、创新思维(5′)
  - ☑优秀设计可以作为论文发表、毕业设计、项目经历

☑任务-李钰心: 大脑理解

问题

数据集

方 法

1.预测受试者性别 (男/女)

2.预测受试者诊断类别(患病/健康)

3.预测受试者年龄群体 (儿童/成年) 1.ADHD数据集

(研究注意力缺陷/多动障碍)

2.ABIDE数据集

(研究自闭症)

1. 决策树

2. 朴素贝叶斯

3. K近邻

4. 线性回归

5. 支持向量机

6. 随机森林

7. 多层神经网络

8.卷积神经网络

9.图神经网络

10.循环神经网络

11.....



☑任务: 伍智广: 语言理解

数据集 问题 THUCNews数据集 游戏体育新闻分类 娱乐教育新闻分类 ChnSentiCorp\_htl\_all数据集 酒店评论分类 waimai\_10k数据集 外卖评价分类 weibo\_senti\_100k数据集 微博评论分类 HC3数据集 问题答案分类

### 方法

#### 传统机器学习:

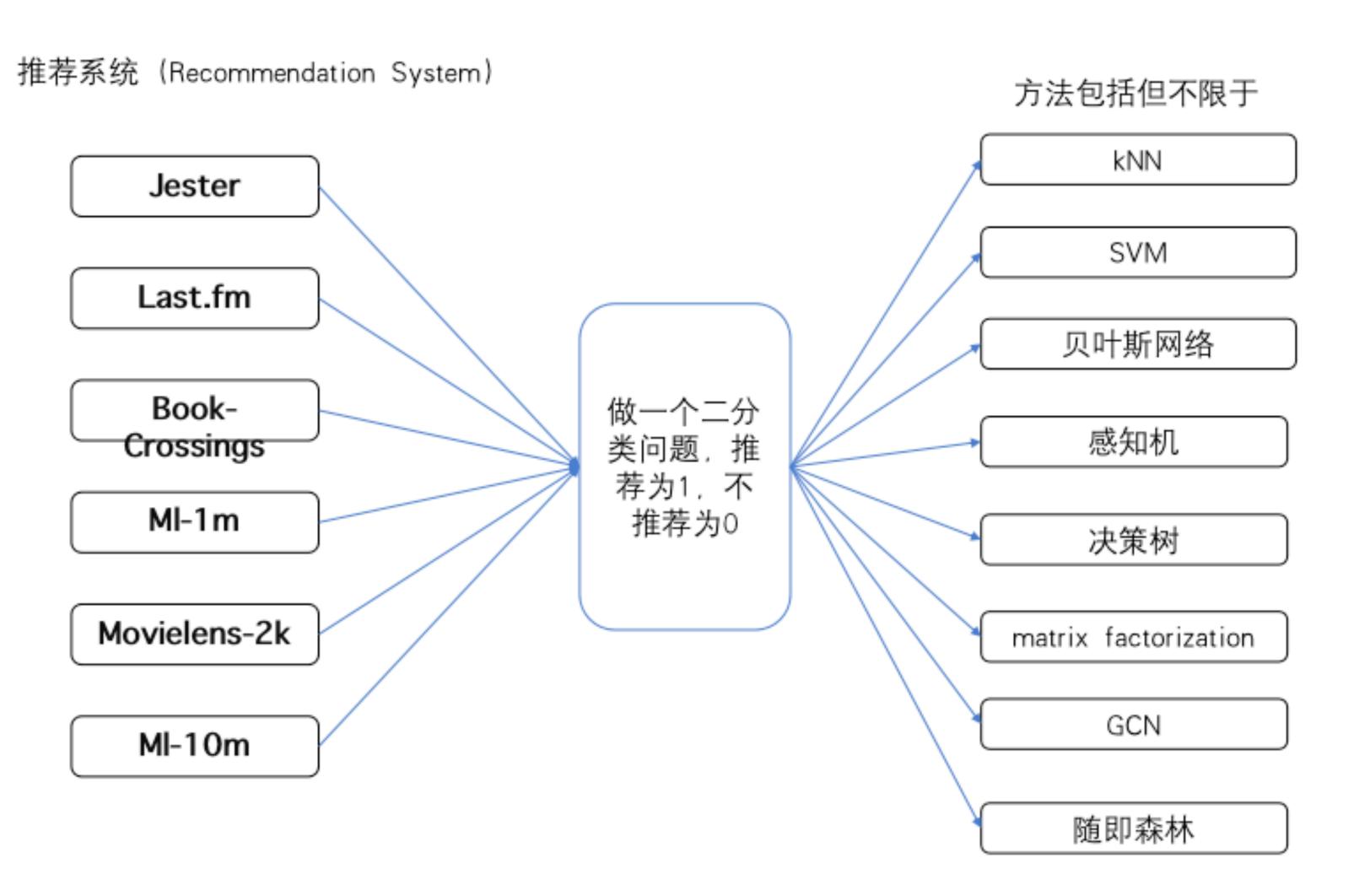
- 1.KNN
- 2.SVM
- 3.逻辑回归
- 4.贝叶斯
- 5.xgboost
- 6.其他
- 深度学习:
- 1.TextCNN
- 2.TextRNN
- 3.FastText
- 4.Transformer
- 5.其他

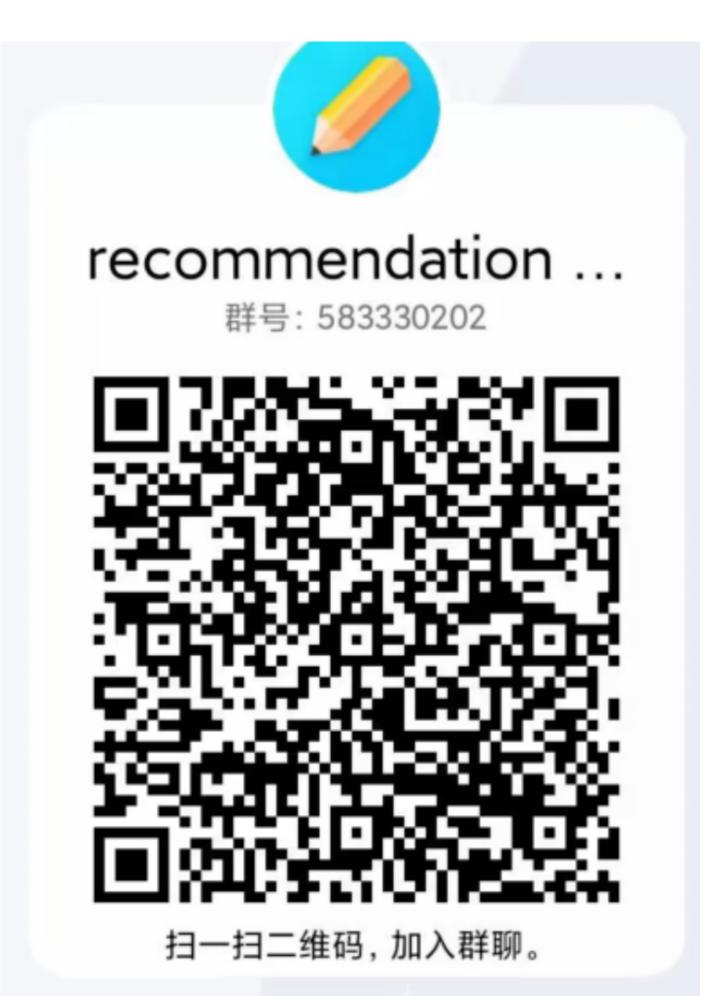


群名称: 文本分类小组

群号: 224822408

☑任务-曲希然: 推荐分析





- ☑提交材料
  - ☑文档报告(长度4页至6页)
  - ☑ 报告格式:《计算机研究与发展》
  - 一音频讲述所做工作,不超过(8分钟)
- ☑提交方式
  - ☑调研和实践报告
  - ☑命名: 工程4+学号+姓名
  - ☑ 课程邮箱:prml\_npu@163.com
  - ☑截止日期: 2023年6月1日, 24:00

- ☑写作要点
  - ☑问题与数据(问题是什么?数据情况)
  - ☑ 所选模型 (模型原理是什么? 在该问题的优势与不足?)
  - ☑实现方法(1.数据预处理, 2.数据集划分, 3.模型实现)
  - ☑ 实验结果与分析(模型至少三种不同参数的指标对比表格、混淆矩阵图、至少1个模型参数变化的曲线图)