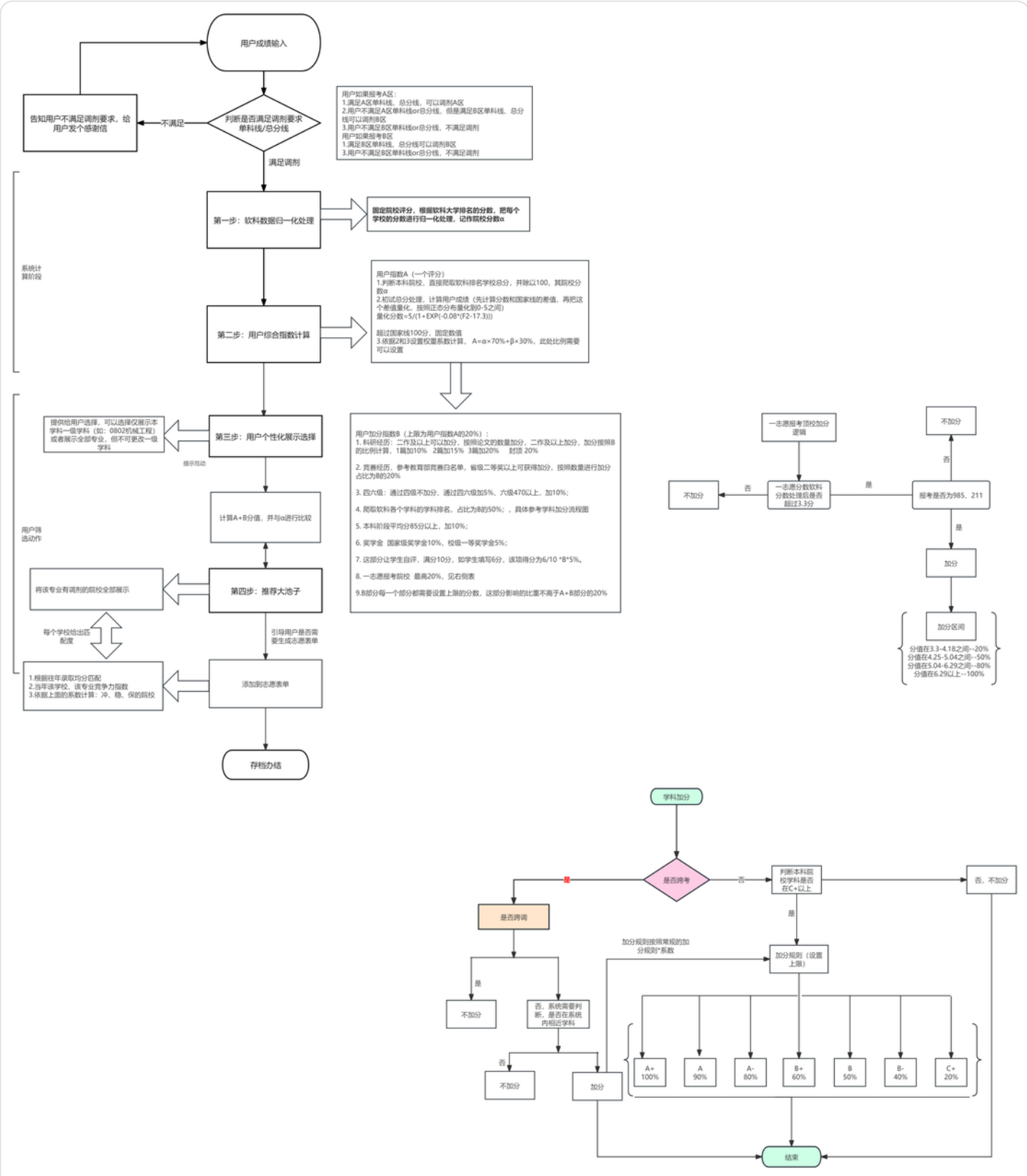


算法推荐逻辑梳理



- 过滤条件
 - 是否满足国家线 分 A B 区
 - 不满足直接无法调剂
 - 是否满足院校线 或 院校单科线
 - 不满足直接无法调剂
- 结果
 - 基础院校列表 （满足分数）

院校协同过滤

- 过滤逻辑
 - 院校排名分 A （大池子）
 - Ranking 学校排名总分/ 100
 - 用户综合分 B
 - 本科院校排名分 B1
 - Ranking 学校排名总分/100
 - 初试总分 B2
 - $f(x) = \frac{L}{1 + e^{-k(x-x_0)}}$ 其中 L（上渐近线）、K（陡峭程度）、x0 中心点
 - $B = B1 * 0.7 + B2 * 0.3$
 - 用户加分 C (加分基数分 C0 = 用户综合分 B * 0.2)
 - 科研经历分 C1
 - 二作及以上 paper_count
 - $C1 = \text{paper_count} \geq 1 : \text{加分基数分} * 0.1$
 - $C1 = \text{paper_count} \geq 2 : \text{加分基数分} * 0.15$
 - $C1 = \text{paper_count} \geq 3 : \text{加分基数分} * 0.2$
 - 竞赛得分 C2
 - 省二级竞赛得奖数量
 - $C2 = \text{count} / 84 * \text{加分基数分} * 0.2$
 - 四六级得分 C3
 - 四级 or 六级分数 ≥ 425
 - $C3 = \text{加分基数分} * 0.05$
 - 六级分数 ≥ 470

- $C3 = \text{加分基数分} \times 0.1$
- 用户所属学科分 C4
 - 是否跨考
 - 是
 - 是否跨调
 - 否
 - 是
 - 否
 - 学院学科在C+以上
 -
- 本科阶段平均分 C5
 - 平均分>85
 - $C5 = \text{加分基数分} \times 0.1$
- 奖学金分数 C6
 - 国家奖学金
 - $C6 = \text{加分基数分} \times 0.1$
 - 学校级以上奖学金
 - $C6 = \text{加分基数分} \times 0.05$
- 自评分 C7
 - $C7 = \text{自评分} / 10 \times \text{加分基数分} \times 0.05$
- 一志愿报考院校 C8
 - 报考院校 ranking 分 > 3.3
 - 是否为985、211
 - 分数区间
 - 3.4~4.18
 - $C8 = \text{加分基数分} \times 0.2 \times 0.2$
 - 4.18 ~ 5.04
 - $C8 = \text{加分基数分} \times 0.2 \times 0.5$
 - 5.04 ~ 6.29
 - $C8 = \text{加分基数分} \times 0.2 \times 0.8$
 - > 6.29

- $C8 = \text{加分基数分} * 0.2 * 1$

- $C = C1 + C2 + C3 + C4 + C5 + C6 + C7 + C8$

- 过滤结果

- 院校备选列表（最终池子）

- $0.75 * A \leq B + C \leq 1.5 * A$

- 概率权重

院校专业匹配度

- 需求表过滤逻辑

- 省份过滤

- 院校备选列表省份 = 目标省份集合

- 院校层次过滤

- 院校备选列表层次 = 目标层次集合

- 专业过滤

- 院校备选列表包含专业 = 目标专业集合

- 学习方式

- 院校备选列表包含专业学习方式 = 目标学习方式

- 学位类型

- 院校备选列表包含专业学位 = 目标学位类型

- 专项计划

- 是

- 专项计划专业

- 匹配逻辑 SimFinal

- 分数匹配 SimA

- 总分 > 去年专业录取平均分

- $\text{SimA} = 1.0 - \min(0.5, (\text{user_score} - \text{avg_score}) / 100)$

- 总分 ≤ 去年专业录取平均分

代码块

```
1 gap = avg_score - user_score
2     if gap <= 10:
3         sim_score = 0.8 - gap * 0.02 # 差10分以内，扣0.2
4     elif gap <= 30:
```

```

5         sim_score = 0.6 - (gap - 10) * 0.01 # 差10-30分，再扣0.2
6     else:
7         sim_score = max(0.1, 0.4 - (gap - 30) * 0.005) # 差30分以上，
        最低0.1

```

- 专业匹配度 SimB

- 本科专业匹配 $\text{SimB1} = 0.5$
- 一志愿报考专业匹配 $\text{SimB2} = 0.5$
- $\text{SimB} = \text{SimB1} + \text{SimB2}$

- 竞争力 SimC

- 冷起展示
 - 用户数 > 3
- 专业竞争力
 - 排名系统专业热度
 - 需要提供数据 @王刚刚
- 院校竞争力
 - 排名系统院校热度
 - 需要提供数据 @王刚刚

最终概率

- $A - (B+C) = -3 \sim 3$
- $\text{SimFinal} = 0 \sim 1$

概率分档

- 冲刺
 - $\text{SimFinal} \leq 0.4$
- 稳妥
 - $\text{SimFinal} \leq 0.8$
- 保底
 - $\text{SimFinal} \leq 0.95$

数据准备

- 国家线
 - 年度
 - 学科门类
 - 以及学科
 - 考生区域
 - 单科满分>100 分数
 - 单科满分=100 分数
 - 总分
 - 特殊计划分
- 自划线
 - 年度
 - 学校
 - 学院
 - 专业
 - 学位类型
 - 科目1
 - 科目2
 - 科目3
 - 科目4
 - 总分
- 院校排名数据
 - 院校编码
 - 院校名称
 - 排名分数
- 学科排名
 - 学科类别
 - 学科编码
 - 学科名称
 - 学院编码
 - 学院名称
 - 排名分数

- 学生基础信息
 - 本科院校
 - 选
 - 本科专业
 - 选
 - 本科绩点
 - 数量
 - 是否过四级及分数
 - 数值
 - 是否过六级及分数
 - 数值
 - 获奖情况
 - 数量
 - 一志愿学校
 - 选
 - 一志愿专业
 - 选
 - 总分及单科分数
 - 数值
 - 一二作论文数量
 - 数量
 - 省二级竞赛获奖数量
 - 数量
 - 自评分 1-10
 - 数值
 - 本科平均分
 - 数值
- 调剂需求表
 - 目标省份
 - 选
 - 目标学校层次

- 选
- 是否跨考
 - 选
- 是否跨调
 - 选
- 可调剂的学科专业
 - 选
- 学位类型
 - 选
- 学习方式
 - 选
- 是否专项计划
 - 选
- 自我介绍
 - 填
- 地区、学校、专业、方向、科目基础表
- 调剂专业表
 - 年份
 - 学校
 - 学院
 - 专业
 - 方向
 - 计划人数
 - 调剂要求
 - 首页展示是否
 - 是否有专项计划
- 调剂录取名单
 - 年份
 - 学校
 - 学院
 - 专业

- 考生编号
- 初试总分