**高考志愿填报小程序概要设计**

**1、应用场景描述与用户画像**

**1.1用户画像**

**1.1.1考生用户画像**

**1.1.2院校用户画像**

**1.2应用场景描述**

**1.2.1考生应用场景**

1． 小红在微信中打开了小程序，输入了自己的基本信息和高考信息，系统根据往年和今年的一些数据向她推荐了建议关注的院校和专业。

2.小红打开程序，查看每个自己关注的院校及专业的当年关注人数以及模拟填报人数

3.小红打开程序，看到了自己的微信好友们的关注院校和模拟志愿填报的情况。当然，她也可以选择不被微信好友发现，但是这样一来她就看不到好友们的信息了。

4.小红根据系统推荐的学校和自己的考虑，输入了平行志愿学校以及志愿专业。系统根据数据库里本省其他考生的信息和学校本省招生情况向她分别反馈了志愿院校被录取的成功率和各志愿专业的成功率。

**1.2.2院校招生应用场景**

1.陈老师打开了微信小程序，选择了所在学校，进行了身份验证。

2.陈老师选择了省份，获得了该省关注本校的考生的分数段及排名段分布报告，并查看了当年本校在该省各专业的热度。

3.陈老师点选了模拟录取按钮，系统在后台自动收集当前所有模拟填报了该学校的考生，并综合学校今年各专业招生计划，给出了反馈报告。

4.陈老师根据反馈报告增加了本校在优质生源省的招生人数，减少了在教育落后地区省的招生人数。

5.陈老师打开了留言板界面，查看了考生留言，并对于其中一些关注度较高的问题给予了回复。

**2、推荐策略设计与分析**

**2.1院校/专业推荐策略设计**

当用户输入考生信息后，首先根据考生成绩与一分一档表查出考生的省内排名，从院校数据库中检索，如果考生排名高于该院校五年录取排名，则为低风险结果；若考生排名高于该院校三至四年录取排名，则为一般风险结果；若考生排名只高于院校三年以下录取排名，则为高风险结果。同时考虑考生排名和院校近五年平均录取排名的比较。若考生排名高于院校近五年平均录取排名，则将风险下调一档。在考虑排名的同时也要综合考虑招生计划的变化带来的影响。

在输入考生信息后选择“偏好专业”或“偏好学校”。偏好专业需选择偏好的专业大类。偏好学校则在推荐结果中按学校排名择优推荐。偏好专业则按所偏好的专业学科评估结果推荐。

从低风险结果中择优选择前N所院校加入推荐结果；从高风险和一般风险院校中选择五年录取排名波动较小的结果中选择M所院校加入推荐结果。将推荐结果进行排序。此为院校初步推荐规则。

考生位次在这里

（推荐院校为院校5、6、7、8、10）

在完成初步推荐之后，根据用户在推荐列表中点选的“关注”确定志愿填报界面的下拉菜单院校。

进入模拟志愿填报界面后，用户可选择下拉菜单中的关注院校，也可手动输入院校。选择院校之后，根据该院校的学科评估结果，综合专业录取排名（与推荐院校类似的策略，但院校推荐的策略相比，因为专业录取人数相对更少，因此要增加招生计划变化所造成影响的比重）给出专业推荐结果。用户也可手动修改志愿专业。

在进行模拟志愿填报之后，程序会给出对当前志愿填报情况的评估结果。如图所示，考生能否成功录取取决于该院校/专业的招生计划和填报该院校/专业的总人数以及考生在同志愿考生排名中的情况。程序统计当前所有用户对当前志愿的填报情况，根据比例关系预估该考生的实际同志愿中的位次，进而估算出报考成功率。

关注该院校的人数

该院校计划招收人数

人数/排名

考生位次（成功率高）

考生位次（成功率低）

可建立模型如下：

南开大学在某省份招收人数 *B*

南开大学某专业在某省份招收人数

南开大学在某省份各专业招收人数 、、…

排名*X*以前的总用户数 *x*

排名在*X*之前且志愿选择了南开大学的用户数 *b*

排名在*X*之前且志愿选择了南开大学某专业的用户数

排名在*X*之前且志愿选择了南开大学各专业的用户数 、、…

则上南开成功率（*w*）

若*w*小于0则令w等于0.01。

上某专业成功率（）

若小于0则令w等于0.01。

各专业成功率类似。

**2.2院校/专业推荐策略分析**

历年院校录取分数线可能会随着各种因素有较大波动，但是录取排名相对稳定。同时为了排除录取排名可能的波动，考虑近五年院校录取排名作为风险评估的依据。择优选择的推荐结果可以大大减少推荐范围，有利于考生选择。对高风险和一般风险院校的操作能一定范围内扩大推荐范围，这样的话推荐结果中“冲一冲”、“稳一稳”和“保一保”的结果就会相对均衡，有利于下一步模拟志愿填报的进一步细化评估。

专业推荐的步骤大致与院校推荐相同，进一步缩小选择范围，同时学科评估权重提高，有利于考生在当前分数内选择最合适、水平最高的专业。

此外，还应考虑因用户行为产生的误差。

1. 存在考生在实际填报时并没有按照模拟填报时的志愿填报的情况。
2. 存在因用户数量不足而程序所收集信息无法反映真实志愿填报的情况。

以上两点均会造成成功率预测的误差。

针对第一点，可在程序界面内加入提示，提示用户尽量按照实际填报情况进行模拟，才能有更为准确的结果；还可采用用户回访的方法，对预测方法进行调整。

针对第二点，应考虑近五年各院校/专业的报录比，根据报录比和当年招生情况预估出当年该院校/专业的关注度期望，用这个期望值对实际收集到的关注度信息进行校正。

院校推荐流程图

否

是

是

否

在一分一档表中检索考生排名

院校近五年录取排名均低于考生

院校三到四年录取排名低于考生线差

低风险

一般风险

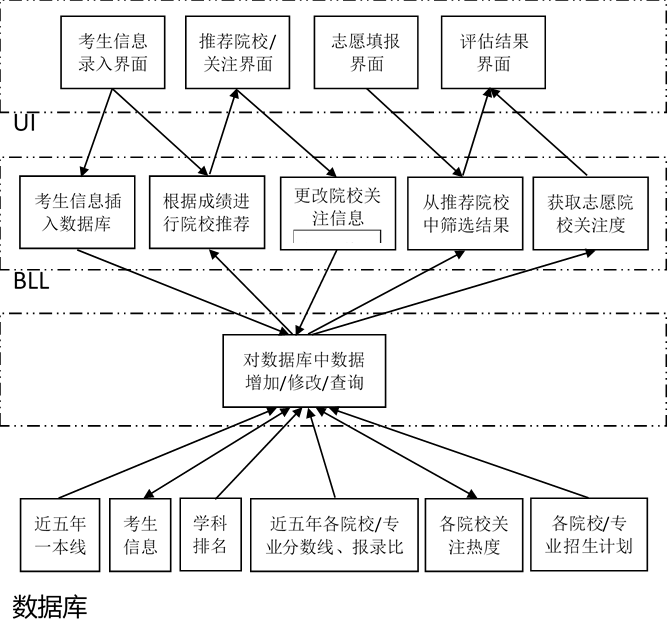
高风险

选择近五年录取排名较稳定的院校

院校推荐列表

**3、系统架构设计**

**3.1三层B/S结构**



三层B/S结构图

**3.2系统功能设计**

对于软件使用者和软件开发者而言，整个系统必须实现以下三大功能：

1.初步推荐学校

（1）实现推荐算法。具体要做的，是实现输入考生的信息，根据往年的分数线等给出初步推荐关注的院校与专业。

2.数据库的建立（往年以及当年）

（1）明确需要收集的当年数据有哪些，建立当年的考生信息和学校招生信息数据库。

（2）建立各校往年分数线数据库。

3.智能志愿评估、个性化风险排查

（1）实现志愿智能评估功能，包括明确从当年数据库中需要抓取的信息，以及推算志愿成功率的算法，反馈报告应包含的内容。

（2）实现核心功能要经过各方研究，关键是搞清楚从学生和学校端都需要获取哪些数据，以达到最准确的评估。

**3.3数据库表结构设计（ER图）**

ER图

院校

专业

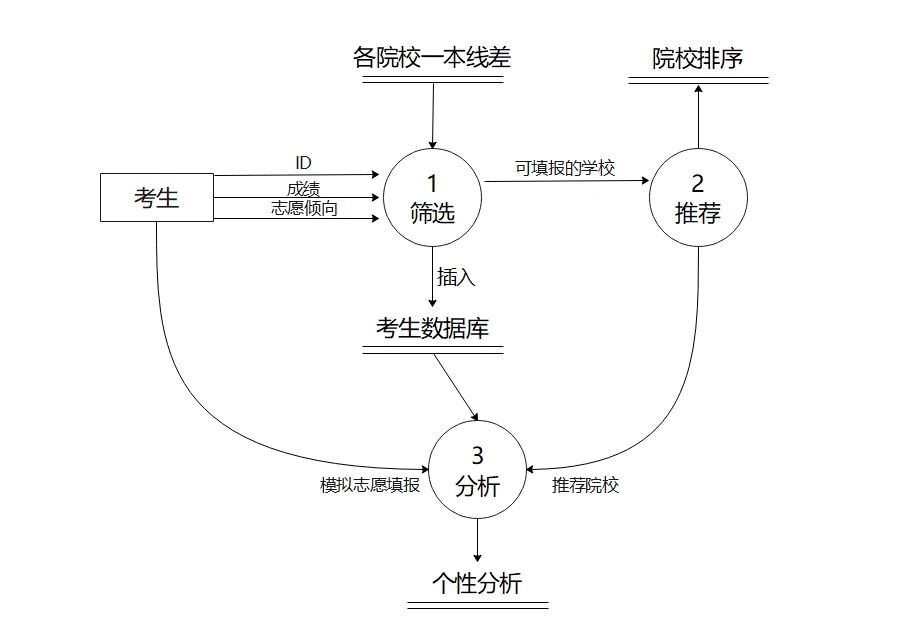
考生

属于

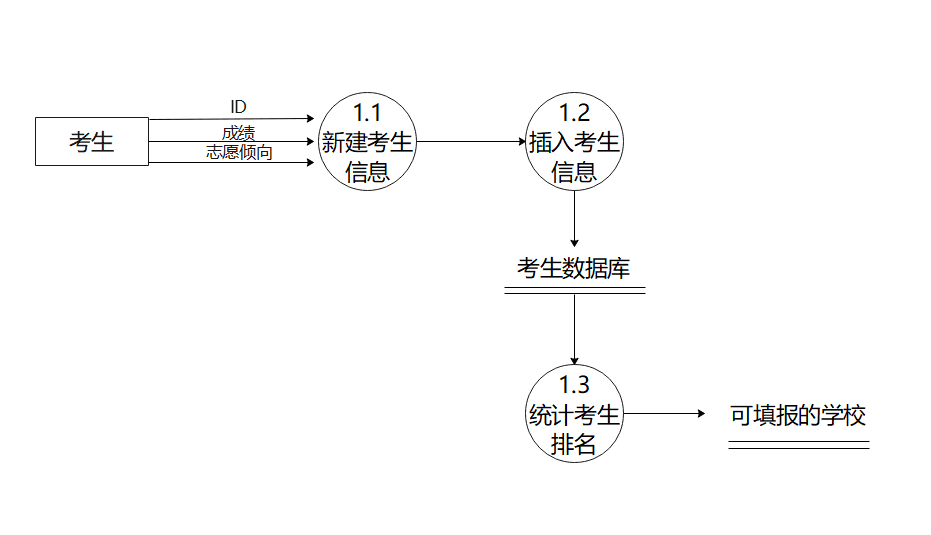
选择

报考

**3.4处理流程设计（流程图）**

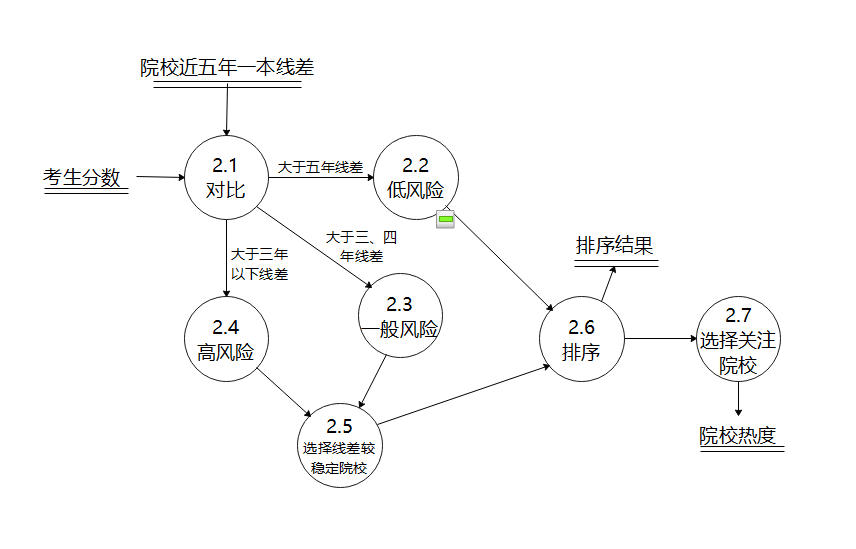


录取排名



数据流图（第二层-1）

数据流图（第一层）



录取排名

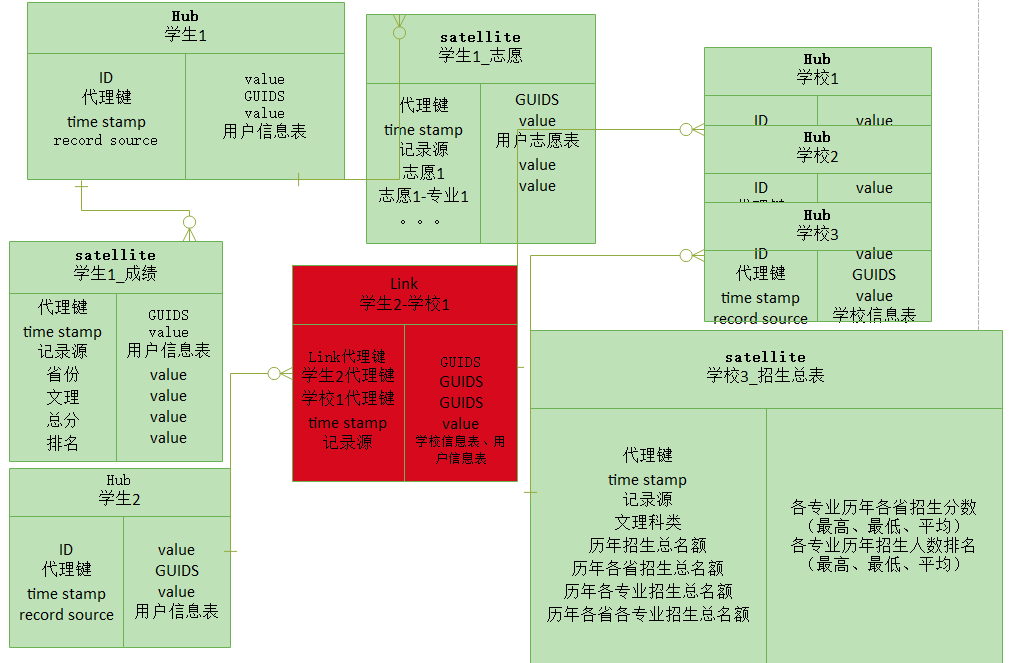
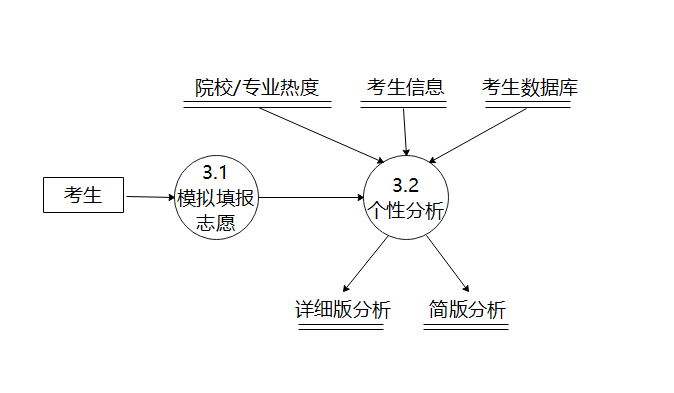
排名

排名

排名

数据流图（第二层-2）

数据流图（第二层-3）

**3.5数据模型设计（数据结构）**

数据流图（第二层-3）

Data Vault模型

**4、评估验证标准**

可利用已知数据进行验证，如利用五年已知数据（2012-2016）预估新一年已知结果（2017）。可虚拟若干考生信息（数据量应大，分布应广泛，考生成绩的分布应近似实际高考分布情况，才能更好地反映程序的真实使用情况），输入程序，根据推荐列表以及推荐专业直接进行填报，对比评估结果和新一年实际招生情况。对同一所院校来说，若某分数段考生评估结果为一较高成功率且这部分考生成绩确实大部分在实际招生范围内，则可初步认为有效。

在初步认定有效之后，统计所有被推荐到该院校的考生人数和实际被录取的考生人数，将比值与成功率比较，若两值相近则可认为结果符合标准。

**5、软件开发环境与开发工具**

**5.1操作系统**

Windows10 64位

**5.2开发环境**

Python3.6

**5.3编程语言**

Python、JS、SQL

**5.4数据库**

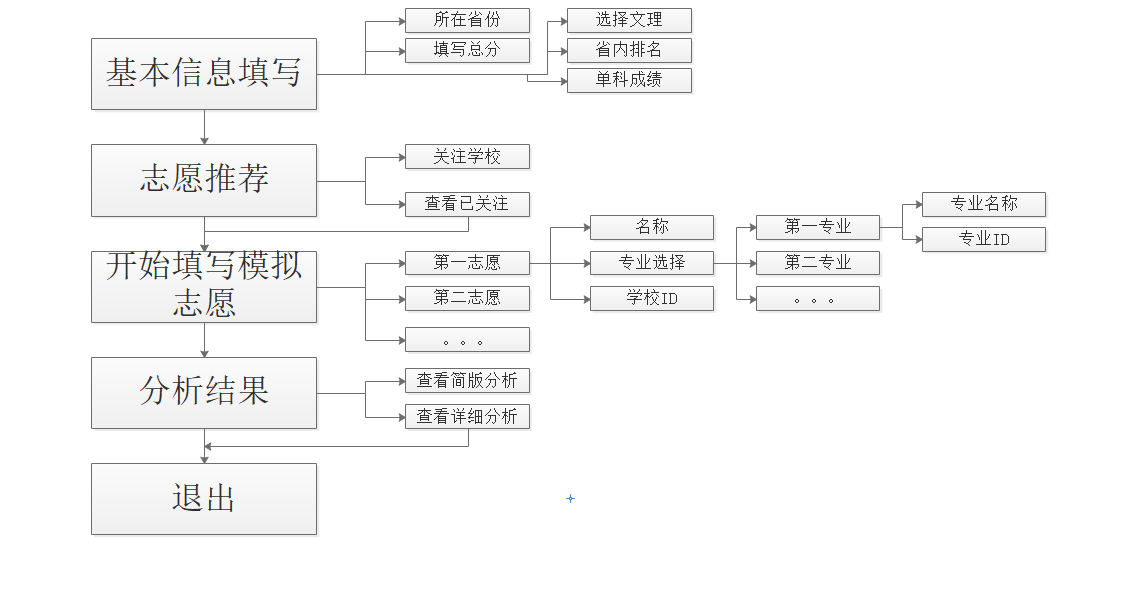
MySQL

**5.5中间件**

**6、数据资源准备**

**6.1静态数据资源**

**6.1.1菜单标题名称**



**6.1.2错误代码**

00000000 XX未填写/XX不能为空

00000001 请填写正确的XX名称/代码！（未在数据库中找到对应学校/专业）

00000010 分数不能超过760！

00000011 排名不能小于0/排名必须是整数

00000100 填写有未识别的字符/数字

**6.2业务数据资源**

1.省份历年信息表：省的一本和二本分数线

2.学校信息表：ID，名称，所在省份

3. 高校招生总表：文理科类、历年招生总名额、历年各省招生总名额、历年

各专业招生总名额、历年各省各专业招生总名额、各专业历年各省招生分数（最高、最低、平均）、各专业历年招生人数排名（最高、最低、平均）

**6.3过程数据资源**

1.用户信息表： ID，省份，文理，总分，省内排名

2.用户志愿表：ID、模拟志愿中的六个学校、每个学校的六个专业

3.统计数据表: 各学校各专业各省关注人数、各学校各省各专业模拟志愿填报人数。