

基于 Web 的智能化问卷调查系统

1. 问卷调查流程

1. 问卷设计阶段：提供问卷设计平台，问卷编写者可以填写需要设计的问题并进行发布给受众填写。
2. 数据收集阶段：提供网页模式的问卷访问，数据被记录进数据库
3. 数据分析阶段：通过数据表、汇总表、折线图、柱状图、饼图、面积图、雷达图等多样报表实时生成分析数据，可更加清晰地了解到数据的占比、分布、趋势等。
4. 发布报告阶段：将数据报表可以容易地被发布为 HTML Web Page

2. 搭建基于 Web 的问卷调查系统的步骤

1. 设计编辑问卷：使用 web 框架设计

一份问卷是由各种各样类型的问题组成的，需要做的就是将期望的问题，选择一种类型，加入到问卷中。

➤ 单响应或多响应的问题

选择类型的问题的响应答案都已经预先设定。例如，询问对某个产品的满意度，可以预先设定为五个可选项，“很满意，满意，一般，不满意，很不满意”；

对于其中的某一个选项，还可以缀一个输入的文本问题。例如，在满意度答案的“很不满意”后面可以放一个输入的文本问题，以得到具体的不满意的原因。

➤ 数值类型的问题

这类的问题，期望响应答案为数值。例如，询问受访者的年龄。

➤ 文本类型的问题

文本类型的问题可以接受任何类型响应回答。例如，询问对本公司产品的建议等。

2. 项目部署：将 web 项目部署到服务器上，可通过 url 进行访问

3. 收集数据：受访者点击收到的问卷调查的 url 就可以开始一次问卷调查，正常结束的一次问卷调查会被视为一次完整的有效数据。意外中断的调查则是不完整的。数据以 json 格式传给后台，进行提取后存入数据库。

4. 分析数据：读取数据库数据，利用 Echarts 框架实现数据可视化

1. 数据预处理：

- 1) 无效样本：剔除无效数据，比如输入不符合规范等
- 2) 处理缺失值数据：取众数/插值/舍弃缺失值
- 3) 异常值：设置正常的区间范围，异常值可用平均值替代
- 4) 数据标签：给特征数据贴上标签，用于有监督学习
- 5) 数据编码：例如以 one-hot 形式将数据进行离散化编码

2. 数据分析：利用 mangodb 存储数据

- 1) 显著性分析：根据某一项的不同属性值将数据分为几部分，增加筛选条件让用户可以进行不同条件的数据筛选并进行可视化展示，如直方图、饼图、折线图等
- 2) 聚类：用户自由选择一类到多类进行聚类，有 k-means、DBSCAN、谱聚类等
- 3) 主成分分析法：在多数情况下，许多变量之间可能存在相关性，通过 PCA，有可能用较少的综合指标分别综合存在于各变量中的各类信息。
- 4) 逻辑回归：二元或多远逻辑回归对数据进行分类
- 5) 交叉分析：分析不同问题的回答之间的相关性以及进行预测



红色方框内，31表示选择了“30~40”和“如果有合适的房源，会出手”的人数。

红色方框内，63.27%表示“30~40”年龄段中选择了“如果有合适房源，会出手”的人群所占比例，即 $31/49=63.27\%$ 。

红色椭圆内，49是“30~40”年龄段的人数。

5. 发布数据：将数据分析得到的报表以 HTML Web Page 或者 Excel 的方式发布出来。

3. 技术重点与难点

1. 使用 Vue + webpack 来构建网站框架
2. 了解 Javascript 跑在服务器上的场景，了解 NodeJS 能够干什么。
3. 进行简单的数据分析后进行进阶的机器学习算法设计。
4. 时间周期
 1. 前端界面设计+实现+功能完善 （2 周）
问卷设计者：设计问卷的界面+问卷数据分析界面
普通用户：填写问卷的界面
 2. 后台读取数据并进行数据分析（2 周）
form 表格提交后将数据写入 mongoDB + 后台从数据库读取数据并进行数据分析后传给前端进行绘图