|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 时间 | 2018-03-20 | 地点 | 微信 | 记录人 | 何铭宜 |
| 参加人员 | 张敏华、林鹏珊、何铭宜、刘彩君、卢程 | | | | |
| 会议主题 | 相似商品功能实现方案讨论 | | | | |
| 会议记录：  会议主要内容：   1. 汇报前两周的任务进展、制订接下来一周的任务计划； 2. 讨论相似商品功能   会议结果：  相似商品的功能方案：  相似性的计算是基于商品名称的文本相似度，因为商品名称既短又精，包含关键词  Tf-idf模型：   1. 将商品名称分词 2. 列出所有的词 3. 计算词频 4. 形成词频向量 5. 计算两个向量的余弦值，余弦值越大，相似度越高     Simhash计算文本相似度:   1. 将商品名称分词 2. 找出词语库中关键词 3. 将每个词语进行哈希操作 4. 计算海明距离   返回相似度高的前n个商品  备注：需要测试两种方案的时间复杂度和准确率来择优选择   |  |  |  |  | | --- | --- | --- | --- | | 编号 | 任务内容 | 参与人员 | 完成时间 | | 1 | 相似商品功能接口实现 | 何铭宜 | 2018.03.25 | | 2 | 相似商品功能接口前后端对接 | 何铭宜、卢程 | 2018.03.28 | | | | | | |