

第十届中国大学生服务外包创新创业大赛企业命题 类赛题相关问题解答 F&Q (6.0)

(截止至 1 月 24 日)

一、【A01】2018 网络零售平台商品分类【浪潮】

1. web 可视化设计的输入是 test.tsv，那么单一和批量查询指的是什么，是文本框的输入形势，还是可以把数据存到数据库，进行指定条目的单一和批量操作，还是自己设计。

答：根据自己设计，最好支持单一输入能返回结果，也能够展示所有的训练集的返回值的结
果

2. 450 万样本的测试时间是指 cmd 的输出时间，还是经过 web 的时间，如何比较不同机器上的时间误差。

答：450 万条数据输出时间，是模型对 450 个商品打标签返回标签值的时间

3. test.tsv 中是否存在 train.tsv 中不存在的标签，如果是，是否给出 test.tsv 中所有的标签类别？

答：test.tsv 的标签都是 train.csv 的标签

4. 训练集中给定的标签可以对它进行更细化的处理么，还是说必须分到给定的标签中

答：分到给定的标签中，否则影响匹配计算正确率

5. 在之前的答疑中看到比赛的分值占比中有 70% 是精确率。由于该问题是一个多标签的分类问题，可以作为评测的指标有很多，比如召回率，准确率，F1 score 等等，请问我们应该采用哪一个？或者命题方是否可以给出一个评价精确率指标的计算公式？

答：准确率公式 = $\text{sum}(\text{测试集模型打上的标签} = \text{测试集标签}(\text{赛题放拥有})) / 450w$

6. 最终的 web 端可视化需要能提供网络服务吗？还是只需要能够在演示的时候能够在本地上进行访问就行了？

答：能够在本地演示即可。

二、【A02】基于华云 chinac 公有云平台，设计云监控管理系统【华云】

1. 请问数据可视化就是用 Grafana 的工具吗？可不可以自己写？

答：数据可视化工具可以是 Grafana，也可以使用别的工具，不做限制。

2. 请问监控管理系统是否仅需包括监控方面，还是说云主机管理、虚拟网络等都要覆盖到？

答：监控管理系统以监控云平台资源为主要考核内容，有能力的同学可以对云资源的管理这块内容进行创意设计。

3. 题目中给出了华云的 API 和使用说明，但是实际开发过程当中不太方便。得知贵公司的云计算平台是基于 OpenStack 开发的，那么可不可以直接基于 OpenStack 的 API（比如 OpenStack 的 JavaAPI:OpenStack4J）来开发监控系统而不采用贵公司提供的 API？还有一个问题就是这个监控系统的用户是贵公司还是客户？

答：企业提供的华云的 API，作为参考，学生也可以使用别的 API 接口，能够实现对华云公有云资源即可。

4. 能否提供一份技术文档包括：样例数据各字段的解释，python SDK 的技术文档以及示例？

答：技术文档请参考华云公有云 OpenAPI 对接指导手册或是华云公有云官网。

三、【A03】大规模资金流入流出的大数据预测【阿里创新创业中心】

1. 在做完视频文档等必需材料之后，能否使用阿里的平台资源（如 DataV）作为解题过程和结果的数据呈现和展示。

答：可以

四、【A07】智能化跨境商品交易平台【东软睿道】

1. 想问一下平台有提供支付的接口和物流信息方面的接口吗？

答：支付接口可以自己对接；物流不会放接口。模拟数据即可。

2. 题目中有说到基于微信小程序实现整体业务，但是供应商在微信端好像没有什么可操作业务。不知道有没有必要添加？

答：微信小程序是供应商用来推广 PC 端业务的，比如说京东商城，有 PC，有 APP，还有小程序。

五、【A08】智能会议室管理系统【虹软】

1. 是否一定要使用虹软提供的安卓人脸识别完成系统开发。由于在官网没看到 php 等后端语言写的人脸识别接口，是否在题目要求的管理端可用其他第三方接口完成人脸识别注册？

答：不同的算法人脸注册是不兼容的。人脸识别还是建议使用虹软的，不同的语言都有能够

调用的机制。

六、【A09】智慧教室学生状态检测系统【慧科集团】

1. face++的 sdk 只有 ios 和安卓是免费的，但是不能适用于开发，而 windows 平台的 sdk 则需要通过商业购买，我们与客服了解到需要 2~3 万资金去购买，主办方是否提供 face++ 的 windows 平台 sdk，或者我们去选择别的 sdk？

答：face++只是一个建议，不是必须的，以完成相应功能为目标。

2. 后台管理系统的服务对象是校方，那么是否需要做两个后台管理系统，一个服务于校方一个服务于公司方，校方的进行设备的管理，而公司方的用于远程系统升级？

答：无需两个管理系统，Demo 能实现相应的功能即可，不需要对接到学校后台。

七、【A10】运用数据分析和可视化技术创造完美的数据大屏交互体验【融创软通】

1. 请题目的要求是做一个 app 用来操控可视化吗？比如选择样式之类的。那网页版的是用来干什么的？

答：题目是需要做一个手机客户端进行页面可视化控制，具体用 Android 和 IOS 原生态开发，还是用 H5 框架来都可以，主要实现小屏控大屏功能就行，当然还是希望客户端界面效果要好。

八、【A11】基于微信的移动智能学习平台【文思海辉】

1. 一般而言，学生只能拿到本校数据接口或者根本没有，但试题要求中提及全网排名功能，是否意味着试题要求面向多个学校而非学生所在的单一学校？另外，试题中要求提供人员管理功能，但正常情况下，任何学校都不会给予这样的权限，那是否可以理解为数据和校方无关，纯粹由作品的后台管理，而不必与校方数据同步？

答：赛题任务要求的开发说明中已说明，如不能获取校方数据的，提供模拟数据用于演示。

九、【A12】用户行为的深度追踪——用户行为分析平台【恒生电子】

1. 提交的程序是否可以为网页形式而不是 PC 端程序？

答：网站和 PC 皆可，取决于学生的熟悉和使用程度。

2. 赛题开发工具里没列出的工具可以使用吗，比如 Python 吗？

答：最好是 java。

3. 请问要做埋点的是网站还是 App？

答：建议是网站。

十、【A13】工业技术类从业者与领域专家的交流平台【海尔】

1. 关于服务器可满足同时在线万人以上的流畅体验，请问服务器是需要小组自己租借还是企业会有服务器平台提供呢？

答：服务器的压力测试，现阶段无需过多考虑，但需在方案里说明方法和措施。产品化初期可采用小型服务器供演示使用，暂无法使用企业服务器

2. 请问企业能否提供现有的专家身份认证标准？

答：目前企业的专家认证，除了必备的个人基础信息外，认证的方式有以下几种：1. 个人简历 2. 名片 3. 发明专利及文献期刊，以人工审核的方式。当然可能有更好的认证方法，此处若有创新的方法会有加分。

十一、【A14】运用文本相似度实现（证券）智能客服【恒生电子】

1. 最终测试集的标准问和训练集的标准问是一样的吗？

答：测试集和训练集多数标准问相同，为测试模型的泛化能力有少量不同的标准问。

2. 赛题中提到的“规则占比过大视为无效”，请问这里的“规则”指什么？

答：规则指的是参赛人员人工制定的硬匹配规则，如“正则表达式”。

3. 技术要求中的短文本相似度大于 0.8 是指客户提问的问题与标准问题的短文本相似度吗？

答：是的。

十二、【A15】景点智能识别语言翻译 APP【文思海辉】

1. 用 Android, 产品无法 Apple store 上线影响大吗？组内成员皆擅长 android, 现学 flutter 成本高，不知用 android 对成绩产生影响大不大。

答：不用 flutter 也可以，但是最好使用 swift，不建议使用 objc。

十三、【A16】AI 数据采集平台【虹软】

1. 题目要求根据需求定制不同的任务并开放给互联网用户，如果收集音频的话这种需求具体指什么呢？

答：比如收集音频的总需求：总共需要 1000 个不同的人，每个人用普通话朗读 10 段特定

文字，并标明每段录音所录制的段落。