

V2.0 A 类赛题答疑（截止至 20191213）

【A01】基于绿色出行场景个人碳排放积分系统开发与运营【八维通】

1、碳积分运行系统。我们是单独做一个 app 或者微信小程序，还是在八维通的 APP 里面内嵌我们的系统模块呢？

答：内嵌的，可以考虑 h5 或者 sdk 的方式

2、在所需数据接口中共有六个接口，请问支付接口中的微信和支付宝代扣需要自己实现吗？

答：不用

【A02】基于地铁出行场景的小游戏设计【八维通】

1、请问这个游戏最终是嵌入现有平台使用呢？还是单独做一个游戏？，如果是嵌入现有平台使用，游戏大小以及风格有没限制？

答：基于接口做独立的 h5 应用。

2、最后测试时，可以拿到 app 来先测试吗？

答：不用集成到 app 测，就基于接口测 demo 就可以。

【A03】工业物联网——高性能边缘计算轻量级容器管理引擎设计【文思海辉】

1、文思海辉赛题讲解视频要求提交的作品是倾向于边缘计算的 web 容器管理工具，而赛题手册中赛题是要求设计一个应用容器引擎，请问以哪个为准提交作品？

答：最终交付以最后发的文档为准，web 的交付耗时，不存在技术上的难点，需要学生实现的是边缘计算应用容器的设计，可以没有 web。

【A04】工夹具全寿命智能管理系统（TMS）【捷普】

1、一个工夹具就是一套还是几个工夹具就是一套？

答：一个就是一套。

2、视频中分为上线夹具和下线夹具，请问有什么区别？

答：上线是夹具从仓库到生产线上（开始使用），下线是从生产线返回仓库（结束使用）。

3、点检是专门有一个人来负责的吗？是定时的吗？

答：点检由专人定期进行。

4、工夹具的定位是简单的用表展示在某个厂区某个区域某人使用即可？请问位置跟踪的具体要求。

答：只需显示厂区，区域，线别，工站。

5、移动端的最终样式是类似独立的功能 app 吗？有无具体要求。

答：移动端 app 是主体的补充，使用同一个数据库。

6、视频中点检记录只有上下线点检，是在出库之后就进行点检，再上生产线；在下生产线时，先点检再入库？那视频中的定时点检又是如何记录到现有的这张纸质表格中？

答：上下线的点检记录是这样的，夹具上线先要点检然后分配到生产线，夹具下线也要点检然后进入到仓库库位保存。至于定时点检的记录是在另外一份定时点检表格中的记录的。

7、现有的这张表格上下线夹具不是同一套，是否可理解为这一条生产线即将下线的一套夹具和即将上线的一套夹具记录在同一张表格中。

答：是的

8、在视频中有提到“需要临时领出去”，记录的表格是产线领退记本，领出去是领到哪里去？

答：领去的地方有很多，有可能去生产线临时需要用一下，有可能去一个工作区域临时做校验验证等，地点是不固定的，临时领取的夹具数量也是单一的，不是产品成套的。

9、另外夹具出入库，这个库是否就是夹具柜，临时领出去和这个出入库的区别？

这个过程在功能综述和流程概述中并没有十分明确的体现，是否需要实现？

答：库就是夹具柜，临时领出去和出入库其实性质是一样的，都是做出入库记录。不一样的地方是正常出入库一定是从仓库领去产线，产线退回仓库，

数量是产品的成套夹具。而临时领出去就如上面讲的出去的地点不确定，数量不确定不是产品的成套夹具。临时出入库其实只要做一个简单的出入库记录功能就可以，便于查询这个夹具的去向。

10、 采购是否会出现一次采购多种夹具，并且夹具数量不为 1 的情况。采购之后夹具上的编号是何时印刷的（这个采购流程是否与我们系统相关，我们是否需要提供编号），另外可否提供一下采购单的照片。采购审核主要审核哪些方面？

答：会出现一次采购多种夹具的情况，每次采购数量不同肯定不为 1. 夹具上的编号在夹具制作的时候就已经生成，系统只要负责录入记录即可。系统可以不要采购审核的内容，系统只要录入已经核实好的信息，比如：夹具名称，夹具本身的编号，夹具的接受编号，夹具的接收日期，夹具属于哪个产品，夹具的 PM 频率，夹具的 PM 类型，夹具的库位信息，夹具的负责人等。

生产线、具体夹具是否只有一个所属的工作部门，对其他工作部门不可见？

答：可编辑和不可编辑，统一由管理员授权，系统通过公司个人的 NTID 登录。

11、 Family（所属大类） 是什么含义，也是要完全独立吗？

答：Family 属于一个产品的家族大类，它下面有不同的型号，型号下面还有不同的频段号。

12、 前端框架是否可以用 vue 框架？

答：可以

13、 关于出入库：产线员工需不需要作为系统用户，如果是是，那么产线员工与仓管员是否都是作为初级用户？仓管员是否作为一个独立于部门存在的职位？

答：产线员工指的是操作这个系统的人员，他们的主要职责是将夹具从仓库中领出然后发放到生产线上，并且在系统上做登记。还有就是临时领用做登记，夹具归还做登记，夹具异常做登记，夹具保养做登记。而仓库管理员的权限在操作员上一个等级，他要夹具的基本信息录入到系统，夹具的库位信息的管理变更等。

14、 系统以 workcell 为单位隔离数据，各个 workcell 之间数据完全独立互不可见，那么请问只有本部门的产线员工能够借所属的工夹具？个人理解是

产线员工都是一个生产部门的

答：Workcell 里面由工程部，生产部，质量部，测试部门，而这个夹具管理的职能部门是工程部是辅助工程部对夹具有系统化可视化的管理，所以对对这个系统有操作权限的都是工程部，其他部门只是可见基本信息即可。这里说的产线员工指的是工程部的实际操作这个系统的在线操作员工。

15、 由于提供的数据缺乏真实性，按数据中推得一个工夹具对应一个采购单号，这与实际情况相悖

答：一个夹具有唯一的自身编号和 buyoff 编号，这里说的 buyoff 编号是经过各个部门验证这个夹具符合功能的接收编号，不是采购编号。

16、 关于申请入库流程：请问贵公司的入库申请实在采购之前，申请之后经过审批再去购买，还是说已经购买回来了，只是需要更新仓库内工夹具信息

答：入库是在 buyoff 之后。

17、 工夹具的哪些基础信息是固定的，比如说同一类的工夹具的共性。

答：固定的有夹具自身编号，夹具 buyoff 编号，夹具的 jbuyoff 日期，ower，使用的产品 family，型号。

18、 有无用于预警预测的数据？预警模型的构建需要有大量的历史记录以及此工具夹在借用之后是否损坏，若损坏则对应的损坏程度如何。

答：目前尚无这方面的具体数据模型，有能力的团队可以充分发挥创造性思维去设计和开发。

【A05】基于人工智能的视觉识别技术【文思海辉】

1、赛题中提到，不能使用商用模块，具体是指什么？

答：只能用 Apache、mit、bsd 授权的开源库。

2、现场测试的图片集是和题目给的数据集类似的吗？

答：是。

3、有数据清洗标准吗？xm 内的 l 字段能否给出详细解释？

答：清洗没有标准，按给出的数据集实现效果就行；xml 是 voc 数据集标注的标准格式，学生团队可自行查询。

【A06】交通时空大数据分析挖掘系统【东软】

1、是否可以使用第三方数据集？是否可以使用第三方接口，或者开源工具？

答：开源工具无限制，第三方数据集及接口如果是为了实现模块任务，在现有数据基础上进行丰富也是可以的。

2、在企业提供的数据外，学生可以自己添加数据吗？

答：数据格式为 csv，查看数据用 Notepad++ 等工具请同学熟悉赛题以及数据后再作决定。我们企业提供的数据可以支撑赛题中所涉及的所有模块。

3、（timestamp）信息记录开始时间为什么比（timestamp1）信息记录结束时间要大？

答：俩个时间取一个既可。

4、企业提供的数据是真实的还是模拟数据？

答：真实业务场景，真实数据。

5、（timestamp）信息记录开始时间为什么比（timestamp1）信息记录结束时间要大？

答：俩个时间取一个既可。

【A08】智·会学习——今目标“智慧校园”场景挖掘与实践应用【今目标】

1、是否允许自己开发应用？

答：关于赛题的延伸，我们可以让有能力的团队在完成基础赛题后进行基于“教师管理场景”的深入思考，开发一套今目标外的小程序用于教师工资的自动核算与发布。

2、我们需要对题目中所列举的几个场景都找到管理难点给出解决方案，还是通过自己调研所总结的难点给出解决方案？

答：关于场景选择，可以根据团队的时间和精力，选择多个场景，可以是题目中罗列的，也可以是其他的场景。题目中的场景只是抛砖引玉，同学们可以先通过调研和走访找到几个具体痛点，然后再通过今目标，搭建解决方案。

3、可以给今目标移动端设计 UI 界面么？

答：设计 UI 界面跟本次赛题关联性不强，不作为评分参考项。

【A09】僵尸企业画像及分类【科创信息】

1、最后放出的用于评估模型的测试集的特征属性？

答：与已发布数据特征一致，只是不同的数据集。可能在建模的时候不同的竞赛单位有特征组合，但是最后的评判还是输入原来的这些特征。

2、题目中判断僵尸企业的条件是.csv 文件中的 flag 列吗？若是 flag，题目叙述的 2000 家僵尸企业为什么在测试集中会有近 9000 个僵尸企业判定，若不是，如何通过训练判断其是僵尸企业？

答：训练集和测试集是分开的，僵尸企业数量不一样没什么问题，主要看学生怎么去用数据。另，每个企业对应多条数据。

3、项目运行中，数据模型提前载入内存会极大提升运行速度，但相对成本较高，如果每次执行任务的时候载入则成本相对较低，但运行速度相对较慢，想问问企业更偏向于成本控制还是运行速度？

答：可以考虑运行速度为主。

4、测试集和验证集必须按赛方给定的比例使用吗？请问在训练模型时必须按照赛方给定的测试集/验证集比例来训练模型吗？是否允许选手自行调整测试集验证集比例，如：将一部分赛方给定的验证集数据用于模型训练。

答：可以。

5、为什么训练集中会有一万多 flag 为 NA 的数据，是给错了吗？

答：提供给参赛队伍的数据集包括：训练数据集（5000 例僵尸企业，10000 例标签不确定的企业，包含僵尸企业数据和非僵尸企业数据）；验证数据集（30000 例企业数据，僵尸企业数据和非僵尸企业数据比例 3：7）。出题方保留的标准测试数据集：10000 例有标签的企业数据。

6、僵尸画像的用户是哪个政府部门？

答：市场监督管理局。

【A10】金融科技服务平台企业数据的无监督分类系统【浪潮】

1、最终模型建完后，一个企业可以有多个标签还是一个标签？

答：多个标签

【A12】基于算法的配送路线优化系统【华云+中诺思】

1、是否应该考虑配送人员的费用问题：题目假设中提到了工作人员每日时长是8h,那么是否可以认为一辆货车无法配送二次,但是人员可以多次参与配送。2、是否派送点只能经过一次：每个派送点只能由一辆车服务一次,是否意味着默认派送点只被经过一次,还是要考虑经过多次,但只派送一次的情况。3、要求中满载率的定义是否为 $\text{单辆车实际所载货物重量}/\text{该辆车最大载重量}$:题目要求满载率最大化,是指所有车辆满载率之和最大化,还是指单辆车满载率最大化?

答: 1、可以不用考虑; 2、派送点只经过一次,经过多次这个属于设计问题,我们只考察一个点由一辆车派送一次,不重复派送; 3、单个车辆的满载率。学生团队在能力允许的范围内,可以做更优化的设计。

2、请问最终,我们最终测试的数据是按照发出来 excel 表中的数据还是题目中给出的数据呢?还有就是题目中的连线代表是只能这样走吗?能不能 A 走到 C、D、E 等等这些点呢?

答: 题目中的数据中根据线的连接走,没有连接的没有给出两点的距离,不能直接走,但是表格中的数据因为是给出的坐标或距离,任何两点都可以走,可以以题目中的做为最终的,表格中的是验证你题目中的设计是否是最优,可以通过不同的数据来验证。

3、题中所提到限定时间是所有货物运输完的时间,还是每辆车的限定工作时间,例如: A 车要在 1 小时完成货物配送,本来的路径需要 1 个半小时,加上限定条件后这个环境就不满足,就要重新规划路线?

答: 可以先不用考虑时间。

4、题目中提到每个员工工作时间为八小时,请问这个八小时是指类似于上下班的固定时间段的八小时,还是其他情况?

答: 暂时不考虑其它情况,这个工作时间也是个不是必要考虑的因素,固定的几个因素配送中心的客户量,每个客户的需求量,每辆车的满载率,车辆一次最大运行的最大距离,这些因素是必须考虑的,学生可以先不考虑工作人员的时间。

5、在内容说明给的样例中,题目让用算法求出车辆数量,而在下面的开发需求中需要用户设置车辆数目,能否给出更加详细的要求?

答: 在没有给定车辆数的情况下,根据车辆的承载量计算需要车辆,还有一种情况就是给出了车辆数和承载,怎么去规划。

6、是只需要考虑车辆的满载率还是需要考虑车辆的耗油量的服务成本？请问关于车辆的服务成本有没有一些参考数据？

答：只考虑车辆的承载量。学生给出的成果中只要考虑了客户数，客户与配送中心的距离，客户需求量，车辆配载这些就可以，工作人员时间和车辆时间做为非必选项。

【A13】运用文本相似度实现主观题自动评阅【华云+中诺思】

1、赛题中所说的测试样例在哪获取？

答：物流专业的简答题基本都是名词解释和概念释义类的题目，语句结构都比较简单，网络上有很多物流相关专业的简答题文档，学生团队可以案例题的样式自己准备数据集。

2、赛题要求将标准答案与学生答案进行对比分析，请问我们所获得的学生答案是否已经经过了 OCR 文本识别？

答：无须经过 OCR 文本识别，目前的场景是在线答题，而非扫描的纸质答卷。

3、题目没有告知评价模型的标准，那我们做完之后，该如何判断自己的模型好不好？

答：题目里给出的标准是对比人工评分，意即模型评分与人工评分的结果一致度越高，模型越好；如学生答案人工评分是 7 分，系统评分为 3 分，就偏差比较大。这里的人工评分标准在于语义是否正确，以及是否正确命中得分点验证的话，可以在网络上找几道物流仓储作业相关的简答题，模拟不同的答题情景让系统评分，然后自己对比参考答案给分，来判断模型是否基本满足需求。基本上物流相关的简答题都是名词解释，释义类的，答案都是比较好理解的。