**经管信息学院2022年大数据技术实训设备采购需求清单**

1.厂家：北京四合天地科技有限公司

2.设备名称：大数据实训管理系统、工业大数据分析实训系统

3.技能指标：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **序号** | **设备名称** | **数量** | **单价** | **总价** | **型号** | **技术指标及要求** |
| **1** | **大数据实训管理系统** | **1** | **170000** | **170000** | **V2.0** | 一、基础平台基础平台以云计算为基础，为教师和学生提供实训管理和实训环境。云实训平台要求操作简单，实训组织方便快捷。教师可快速建立实训课程体系，组织实训内容，创建实训任务，管理实训过程。实训中可以为学生分配独立的、完整的、一致的实训网络环境，进行实训教学，实训结束后可以迅速回收和释放资源，平台要为学校的教学实践活动和实训组织，提供更加高效便捷的服务，主要功能如下：1、本系统在云计算技术的支持下，运行在Linux系统，使用浏览器/服务器模式提供服务,用户可使用最新版本的谷歌浏览器或者火狐浏览器访问系统；2、系统应支持包括系统管理员，教师，学生三种角色。管理员负责系统基础维护配置，教师负责课程维护，实训等教学工作，学生使用本系统创建的实训环境完成实训并提交实训报告；3、管理员可配置包括各项云计算基础参数如镜像服务等资源,管理员可便捷配置构建实训网络环境，快速创建三层网络架构的网络环境池；可维护弹性ip池，系统支持三层网络，可为每个学生的实训环境构建独立的网络结构，使得所有学生的实训环境的每台虚拟主机IP及网络拓扑完全一致；4、管理员可配置系统基础环境，如镜像维护，可文件上传或者URL方式导入，镜像支持Linux和Windows操作系统，设置镜像修改密码属性；支持虚拟化主机计算规格维护；5、管理员可对系统用户进行管理，包括支持学生教师帐号维护，支持用户密码重置、批量导入；学生班级维护；6、管理员可维护实训模板，模板内可配置多台实训云主机的镜像、ip、名称、是否开放学生访问桌面,是否需要弹性ip直连；7、管理员可对实训主机进行管理，可查看和维护平台中所有云主机，对云主机进行重置、停止、启动、修改密码、关闭电源、删除、访问控制台操作；8、管理员可对实训网络环境池管理，可查看环境池使用情况，手动回收环境池资源，并回收环境中的云主机；9、管理员可对异常状态实训任务管理，可查看实训任务运行情况，手动干预实训任务状态；10、管理员可查看日志，如环境分配日志、操作日志、虚拟化日志；11、管理员可查看平台资源如CPU、内存、主存储、环境、云主机使用情况统计信息；12、教师可对课程进行管理，教师可自主创建课程，编排课程下的实训。实训可指定“实训模板”，以MarkDown格式在线编辑实训指导书，上传实训附件。学生可在线查看实训指导书，下载教师上传的实训附件；教师可指定其他授课教师，指定授课班级；13、教师可管理实训任务，教师可在自己的课程下创建实训任务，系统可按照指定的实训模板，为学生自动分配实训环境，创建实训所需的云主机；14、教师可对实训任务过程管理，实训任务启动时，系统需校验当前CPU内存等资源是否满足需求并给出提醒。实训开始后，教师可方便的查看各学生云主机的运行状态。教师可“只读”或“完全控制”两种模式通过浏览器访问学生的云主机桌面。教师可对学生实训中云主机进行重置、停止、启动、修改密码、关闭电源操作；15、系统需支持教师将实训任务挂起，保存学生当前云主机数据。任务继续进行时系统恢复学生云主机实训环境。满足硬件资源有限情况下多批次实训任务交替分多次完成的需求；16、教师可根据实训模板创建自己的实验环境，按照模板创建云主机以及网络环境；教师可把自己实验环境中云主机的桌面VNC方式只读共享给参与实训的学生。学生在进行实训时，可在系统中通过浏览器查看教师共享的云主机桌面，方便教师在教学中演示；17、实训结果管理,学生以MarkDown格式在线编辑实训报告，实训可提交附件，在线编辑实训报告时可上传本机图片或者自动截取云主机桌面图像；教师可查看实训任务中学生提交的实训报告和附件，给出成绩和评价；18、学生登录系统后，可在浏览器中访问自己实训任务云主机的桌面；可自行重置云主机，恢复到云主机初始状态；二、实训配置平台内置餐饮外送统计平台及企业消费服务平台两套用于实训的业务平台，并配备相应的用于进行数据清洗挖掘分析的原始数据以及用于数据可视化的数据，同时平台内置电商行业的实时数据生成器和相关的离线存量数据，用户可基于Web完成大数据相关平台搭建、数据采集与处理（爬虫）、数据清洗与挖掘分析、离线数据抽取、离线数据统计、数据采集（Flume）与实时计算、数据可视化等大数据项目工作过程的训练，涵盖CentOS系统操作、Python程序开发、Java程序开发、Scala程序开发、Hadoop部署管理、Spark部署管理、Hive部署管理、ZooKeeper部署管理、Kafka部署管理、Sqoop部署管理、Flume部署管理、Flink部署管理、Flume操作、MySQL操作、前端vue开发等大数据相关技术技能。内置内容包括：1、内置大数据平台及相关组件安装包的实训集群镜像；2、内置餐饮外送统计平台及企业消费服务平台的实训镜像；3、用于餐饮外送统计平台业务背景下数据清洗与挖掘分析的原始数据不少于4500条；4、用于企业消费服务平台业务背景下数据清洗与挖掘分析的原始数据不少于4500条；5、用于电商业务背景的实时数据生成器脚本和离线存量数据，电商离线存量数据订单表数据量不少于100万条；6、包含已导入用于数据可视化的原始数据的数据库的实训镜像；7、基于镜像创建好相应的实训模板及实训虚拟机或集群，方便用户快速开始相应实训。★要求生产厂家为2021年全国职业院校技能大赛赛项合作企业，能够支持国家级职业院校技能大赛，提供证明材料并加盖原厂公章。★要求提供生产厂家针对本项目的售后服务承诺书并加盖原厂公章 |
| **2** | **工业大数据分析实训系统** | **1** | **200000** | **200000** | **V1.0** | 基于真实的工业大数据分析场景转化，所有内容均直接使用项目研发过程资料作为素材，能够引导学生从零开始完成一个工业大数据分析的研发工作，匹配真实企业工作流程。采集生产设备和制造执行系统数据，根据业务需求完成数据质量评估，制定数据清洗策略，构建工业数据仓库，对设备、产线数据进行分层建模，构建面向数据分析的集成化数据环境，至少包括但不限于数据明细汇总层、数据集市层、数据应用层。接收采集的生产制造执行系统数据，完成数据管理、建模和分析。监控数控设备加工过程，提取状态特征数据，建立设备磨损状态数据模型，预测设备磨损状态，及时发现并进行预测性维护。分析产线整个生产流程，建立物料、订单、工艺、质量、产量和能耗相关模型，在制造执行过程中预测订单加工工时，对产线设备进行预测性维护，减少停机维护时间。依据工业数据仓库数据明细与汇总层的沉淀数据进行多种维度的组合分析，满足特定的查询及挖掘应用，实训系统提供的主要资源及涉及知识技能点如下：1. 提供项目背景类资源，包含项目背景、需求分析、解决方案、数据分析、数据模型、数据流转图等，可指导学生全方位了解项目需求及目标结果；

项目背景资源需包含项目需求、痛点分析、业务介绍、项目的技术选型、整体技术流程及架构等内容；需求分析资源需包含项目需求、项目整体介绍、项目主要模块、工业分析指标、数据需求、技术需求等内容；解决方案资源需包含项目总体架构、技术架构、系统架构等内容；数据分析资源需包含项目数据、逻辑关系、数据示例、预测模块开发示例等内容；数据模型资源需包含业务介绍、关联关系、数据模型等内容；数据流转图资源需包含数据源、数据入口、数据流转等内容；1. 提供项目任务类资源，包含历史数据集、任务说明、任务分析、任务实现，可指导教师进行有序的实训任务安排，并在关键点给予学生必要的实训指导；

任务说明资源需包含任务概述 （报警预测、数据源解析）、实验环境、实验准备（SparkML、 需要引入的依赖）、技术详解（基本程序结构、创建运行环境、随机森林模型创建与训练、建立模型并设置参数、训练模型、 预测数据）等内容；任务分析资源需包含项目搭建，创建Maven项目、 项目框架搭建、声明项目中工具的版本信息、添加项目依赖、数据准备、模块创建和数据准备、基于机器历史数据进行模型训练；常见的机器学习模型， 线性回归（逻辑回归）、贝叶斯网络、 决策树、支持向量机、最近邻等内容；任务实现资源包含基本信息、案例目的、环境准备、案例内容（报警预测、运行代码）等内容；3.提供实训源代码，基于工业大数据分析的源代码可指导学生完成相关实训任务，对接产业真实技术技能需求；4.实训系统涉及Spark、SparkML、Hive SQL分布式计算引擎、工业数据仓库构建相关知识技能，包含分布式计算框架的分治思想、Scala基础操作、Spark Core、Spark SQL、Spark Streaming、SparkML、Hive数据表类型、Hive分区、Hive分桶、Hive拉链表、HiveSQL DDL、HiveSQL DML等。★要求生产厂家为2021年全国职业院校技能大赛赛项合作企业，能够支持国家级职业院校技能大赛，提供证明材料并加盖原厂公章。★要求提供生产厂家针对本项目的售后服务承诺书并加盖原厂公章 |
|  | **合计** | **小写：￥370000.00** | **大写：人民币叁拾柒万元整** |