

政府采购合同书

采购编号：【KRDL】K5-2507085

项目名称：经管院大数据分析平台

西安理工大学经管院大数据分析平台

政府采购合同

甲方：西安理工大学

乙方：北京中软国际教育科技股份有限公司

依据《中华人民共和国民法典》及相关法律法规之规定，甲乙双方在平等、自愿的基础上，经双方协商一致，达成如下协议，以资共同遵照执行，具体条款如下：

一、招标项目及招投标结果描述：

本合同招投标项目：经管院大数据分析平台项目。招标采购编号为：【KRDL】K5-2507085。依法进行招投标程序后，中标供应商为本合同乙方。依据招投标结果，双方签订本合同。

二、采购项目（范围）：

乙方为甲方按招标文件要求以及乙方投标文件提供的全部内容提供服务。

三、期限：

1. 交货期：自合同签订之日起10日历日

四、采购项目付款方式：

采购人自验收合格之日起30日内，支付合同总金额的100%。

五、合同价款

合同总金额：大写：人民币壹佰叁拾叁万元整（¥1330000.00）

本合同总价一次性包死，不受市场价格变化因素的影响，从项目需求调研到设计、开发、实施、调试、验收、人员培训、接口开发和质保期等环节涉及到的一切费用，且包含系统所需对接其他系统的接口费用。按规定要求开具国家正式发票，如有遗漏视为已包含在内。

六、双方权利义务

（一）甲方权利义务

1. 甲方有权对合同规定范围内乙方的行为进行监督和检查，拥有监管权。有权定期核对乙方提供服务所配备的人员数量。对甲方认为不合理的部分有权下达整改通知书，并要求乙方限期整改。

2. 负责检查监督乙方管理工作的实施及制度的执行情况。

3. 根据本合同规定，按时向乙方支付应付费用。

4. 国家法律、法规所规定的甲方承担的其他责任。

5. 甲方负责协助乙方进行需求调研、项目设计和项目调测等，提供有利于项目实施的有关信息、资料及相关职能人员信息等，以便对该项目信息系统进行全面的研究和设计。

6. 其他事宜甲乙双方在合同签订时约定。

(二) 乙方权利义务

1. 对本合同规定的委托服务范围内的项目享有管理及服务义务。

2. 根据本合同的规定向甲方收取相关费用，并有权在本项目管理范围内管理及合理使用。

3. 及时向甲方通告本项目服务范围内有关服务的重大事项。

4. 接受项目行业管理部门及政府有关部门的指导，接受甲方的监督。

5. 国家法律、法规所规定的乙方承担的其他责任。

6. 乙方应按时完成双方商定的项目建设内容，并严格按照实施进度计划，保质、保量、及时、可靠地开展项目建设工作，并负责向甲方提供优质、安全的软件产品。甲方提出整改要求的，乙方应当按照甲方要求整改。

7. 乙方不得将本合同约定的服务事项转包、分包给任何第三方实施，否则甲方有权单方面解除合同并不承担任何责任，乙方除应退还收取的款项后，还应按本合同约定承担违约责任。

8. 乙方未经甲方同意违反上述本合同第六条第(二)项乙方权利义务的，则应及时改进并按甲方要求或由甲方按生产情况合理安排后，继续履行合同，由于乙方故意造成的违约，乙方应承担本合同总金额 %的违约金。

9. 服务要求：

(1) 在系统的服务期内，供应商应确保系统的正常使用。在接到用户服务要求后应立即做出回应，并在承诺的服务时间内实施服务。供应商应具有完善的质保期内及质保期外的售后服务方案。所提供的技术服务为7×24小时支持维护服务，包括邮件、电话、远程维护、现场服务等方式；

(2) 提供固定 2 名技术人员随时提供技术服务。

10. 培训要求：供应商需制定详细的培训计划，对本项目所涉及的系统管理员和使用人员进行操作及相关技术培训，增强维护和使用系统的技能。供应商负

责系统管理员的现场技术培训，包括产品的功能、部署条件、部署步骤和注意事项、产品升级、日常维护事项等方面，使用户达到能独立进行管理、维护和故障处理等工作。对使用人员，除组织系统软件集中培训服务外，要求实施工程师根据学校需求情况进行远程或现场一对一系统软件培训，直至使用人员能熟练掌握使用系统。

七、违约责任

1. 按《中华人民共和国民法典》中的相关条款执行。

2. 未按合同要求提供服务或服务质量不能满足合同要求，采购人有权依据《中华人民共和国民法典》有关条款及合同约定终止合同，并要求中标人承担违约责任。

3. 中标人未按合同约定的交货日期交货的，如中标人在政府采购合同规定的交货期满后 3 天内仍未能交货，则视为中标人不能交货，采购人有权解除采购合同，中标人除退还已收取的货款外，还应向采购人偿付政府采购合同总金额 10% 的违约金，并承担由此给采购人造成的一切损失。

4. 中标人所交付货物及伴随的服务不符合其投标承诺，采购人要求更换一次后仍不符合约定的，采购人有权解除采购合同，并将有关情况上报政府采购监管部门处理。

八、保密

乙方及其工作人员应对技术服务合同签订、履行过程中了解到的涉及到另一方的保密信息承担保密责任，并采取相应的保密措施。双方应承担的保密义务包括但不限于：

1. 未经一方同意，另一方不得将上述保密信息披露给任何第三方；

2. 不得将上述保密信息用于本合同以外的其他目的。

3. 上述保密义务永久有效，不因本合同、本测评项目的终止、中止而终止。

九、不可抗力

1. 不可抗力指在本合同期限内发生的不可预见（或可预见，但其发生或后果不可避免）、非任何一方所能控制且使任何一方无法完全履行本合同的国家政策、法律、法规的限制、地震、台风、火灾、水灾、战争、罢工、暴动、黑客攻击或任何其它社会、政治动荡造成的灾难。

2. 如果出现不可抗力,双方在本合同中的义务在不可抗力影响范围及其持续期间内将中止履行。经另外一方确认的不可抗力影响时间,不计入本合同执行时间。本合同期限可根据中止的期限作相应延长,但须双方协商一致。任何一方均不会因此而承担责任。

3. 声称遭受不可抗力的一方应在不可抗力发生后不迟于 15日内通知另一方,并随附经有关部门确认的不可抗力书面证明,且应尽可能减少不可抗力所产生之影响。

4. 如发生不可抗力,双方应立即协商解决问题的方案。如果不可抗力持续 30 日以上,且致使合同无法继续履约的,经双方书面确认可解除本合同。

5. 在发生不可抗力时,双方对各自控制下的设备、资料负有保管责任,对于未受不可抗力影响并且可以继续履行的合同义务应继续履行。

十、争议解决

执行本合同中产生纠纷,由采购人与中标人双方协商解决;协商不成,向有管辖权的人民法院提起诉讼。

十一、其他

1. 本合同未尽事宜,由双方协商处理。

2. 本合同正本 4 份,具有同等法律效力,甲方 2 份,乙方 2 份。合同自双方签字或盖章之日起即时生效。

3. 招、投标文件为合同的组成部分,与合同具有同等法律效力。

甲方:西安理工大学(单位公章)

地址:西安市金花南路5号

法定代表人:

联系电话:029-82312257

账号:102891574567

开户银行:中国银行西安金花南路支行

签订日期:2025年 月 日

乙方:北京中软国际教育科技股份有限公司(单位公章)

地址:北京市北京经济技术开发区科谷一街8号院8号楼20层2001-1室

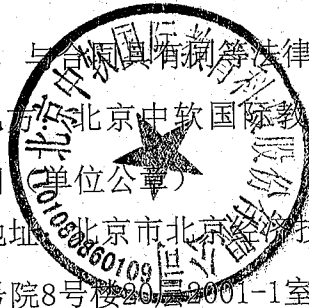
法定代表人:唐振明

联系电话:010-82861666

账号:01090302900120105606242

开户银行:北京银行中关村支行

签订日期:2025年9月5日



附件1：产品技术参数表

投标产品技术参数 标的名称：大数据分析实验教学与科研创新平台（1套）	
一、总体技术架构	
1、系统采用B/S的架构方式，支持学生直接通过浏览器登录使用。兼容Chrome、火狐、360主流浏览器相关版本。	
2、系统采用前后端分离设计，降低前端与后端的耦合度。使得后端通过API接口专注于提供数据，维护系统架构的稳定，保证数据的安全，前端专注于交互，快速响应UI的变化。	
3、系统采用微服务架构，支持Docker容器化部署。	
4、支持在学校本地的服务器上部署和PaaS云平台上部署运行。	
5、大数据教研训资源中心、大数据实践教学与智能管控平台、大数据智慧应用实训与科研平台、数据治理综合实训与科研平台等四大模块化平台进行无缝衔接，具备统一的登录入口和组织架构，支持用户登录平台后可分别进入不同的模块化平台进行操作应用。	
6、平台不限制学校创建学生用户的数量。	
7、平台满足学校信息安全要求。	
二、设备采购清单	
大数据教研训资源中心	
大数据教研训资源中心提供涵盖水利、金融、能源、制造、零售、教育等行业的6000万条行业数据集，支持接入学校可开放的各类业务数据，如教务、教工、学工、人事、财务、资产、校情、科研等；支持多种类型数据源添加和API数据接口配置，提供21个实验库和涵盖多个领域的25个行业大数据项目案例，以全面满足教研训业务的需求。	
（一）数据资源管理及实验资源库模块	
1、为满足智慧水利方向的教研训对数据的需求，需提水务控制系统实时采集的数据信息，数据均需为瞬时测量值，包括时间、原水PH、原水浊度、出水浊度、取水量、PAC 耗等，整体60000条。	
2、支持Influxdb、IoTDB、OpenTSDB、Taosdb等时序数据库接入，并支持Kafka消息队列的数据接入。	
3、为提升学生在智慧水务方向应用大数据知识和技术能力，提供Python语言程序设计开发实验库、面向对象程序设计（JAVA）实验库、数据结构与算法实验库等。	
4、Python语言程序设计开发实验库提供包括关键字和标识符和变量、基本数据类型、运算符和表达式、选择语句、循环语句、跳转语句、列表等82个实验任务，每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套1套教学大纲、60套教学课件（PPT）、60个教学视频、8个课程思政内容、8个教学方案。	
5、面向对象程序设计（JAVA）实验库提供112个实验任务，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测，实验库需配套教学大纲、实验指导书、教学方案、课程思政等资料。	
6、数据结构与算法实验库提供包括但不限于栈的实现及应用、队列的实现及应用、双端队列的实现及应用、列表抽象数据类型及其实现、递归之任意进制转换	

、递归之机器人的运动范围等45个实验任务，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测，实验库需配套教学大纲、实验指导书、教学方案、课程思政等资料。

7、提供城市供水处理投药量建模实践项目案例。支持交互式编程模式和低代码开发模式，其中交互式编程模式提供详细的实训指导手册、交互式编程环境、数据集，支持获取代码答案；低代码开发模式提供详细的实训指导手册、数据集（包括时间、原水PH、原水浊度、出水浊度、取水量、PAC耗等维度，60000个样本）；低代码开发模式提供支持教学演示的示例工程分析模型，包括数据探索模型、城市供水处理投药量预测模型，其中城市供水处理投药量预测模型至少包括数据接入、异常值检测、设置角色、数据拆分、自动回归、模型利用、回归评估等环节。

8、为满足智慧金融方向的教研训对数据的需求，提供增信基金备案详细数据、信用卡欺诈检测数据、银行客户流失预警数据、银行代发薪客户、信贷客户风险预测数据、公募基金购买明细数据等，数量500000条。

9、增信基金备案详细数据包括但不限于所属担保机构编号、担保机构简写、主备案号、批次、业务批次、业务类型、业务方式、借款人、借款人脱敏、首次贷款、借款人性质、证件号码、注册地省、注册地市、注册地县、所属行业、政策扶持领域、上年末从业人数等75个维度，数据量8646条。

10、智慧金融方向支持接入包括但不限于cliclhouse、Gbase 8a、GreenPlum、Teradata、TiDB、Vertica、Guass等MPP数据库。

11、具备数据管理功能，模型授权支持按照组织机构、人员角色或组进行数据集权限的授权分配。数据模型支持数据行级和列级的权限控制，行级权限支持静态规则和动态规则配置，列级权限支持禁止查看列、列数据脱敏功能。

12、为提升学生在智慧金融方向应用大数据知识和技术能力，提供数据采集与预处理实验库、R语言基础实验库、数理与应用统计实验库、金融大数据分析实验库、区块链原理与应用实验库等。

①数据采集与预处理实验库包括但不限于网络爬虫基础知识、Html基本元素、CSS基础、JSON基础、HTTP基本原理、urib的使用、requests的使用、XPath基础、PyMySQL数据库编程、Scrapy爬虫应用实例、Pandas的数据结构、Pandas的基本功能等82个实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲教学课件（PPT）、教学视频、课程思政教学方案等资料。

②R语言基础实验库包括R的字符串操作、R的数据类型、R的条件和循环操作、R的对象和类、R的正则表达式操作、R的数据导入、R语言可视化等20个实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲教学课件（PPT）、教学视频、课程思政教学方案等资料。

③数理与应用统计实验库包括但不限于数字模型与方法、非参数模型、特征工程、线性代数、多元统计分析等74个实验任务，每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲教学课件（PPT）、教学视频、课程思政、教学方案等资料。

④金融大数据分析实验库包括但不限于Python数据类型及金融领域应用、Python控制结构及金融领域应用、Python函数及金融领域应用、python金融领域数据分析应用等101个实验任务。每个实验提供详细的指导手册，实验任务按照关卡进

行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲教学课件（PPT）、教学视频、课程思政、教学方案等资料。
⑤区块链原理与应用实验库包括区块链-数据层、区块链-网络层、区块链-共识层、区块链-合约层、区块链安全性机制等实验任务，实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。
13、提供大型金融投资集团数字化转型项目案例，包括项目信息、业务调研、咨询规划、数据实施、数据应用等5个核心模块，总共12个项目任务，配套10个教学视频、11个教训课件；其咨询规划核心任务包括数据架构设计和数据管理体系规划，数据实施任务包括数据仓库设计、数据集成开发、元数据管理、数据质量管理、数据标准管理、数据安全治理、数据资产目录、数据服务开发等8个项目任务，数据应用核心任务包括报表&驾驶舱开发。支持学生通过进入实训的方式查看详细手册和进入企业级低代码数据治理实训平台进行实战练习。
14、为满足智慧零售方向的教研训对数据的需求，提供零售企业销售部门数据、企业销售业务员数据、企业经营回款计划数据、企业经营分析销售数据等，数量500000条。
15、智慧零售方向支持接入Hive、Impala、Presto、HetuEngine等大数据查询引擎等数据，满足海量数据的研究分析和价值挖掘。
16、为提升学生在智慧零售方向应用大数据知识和技术能力，提供Spark编程基础与应用实验库、大数据存储与处理实验库、大数据管理与应用实验库、商务大数据实训实验库、大数据营销实验库。
①Spark编程基础与应用实验库包括Spark运行基本流程、RDD设计与运行原理、Spark的安装与配置、RDD编程基础、Spark Mllib模型选择和超参数调整等55个实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。
②大数据存储与处理实验库包括HBase环境搭建与常用操作、HBase 开发、HBase中的MapReduce、用户搜索日志分析、成绩查询、用户学历查询、气象数据清洗、招聘数据清洗、电信数据清洗等25个实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。
③大数据管理与应用实验库包括Hadoop的安装和使用、Hadoop的安装和使用(完全分布式)、Spark Scala、实例:Spark初级编程实践、流计算基础知识与应用、Flink安装、Flink编程实践、图计算基础知识与应用、D3可视化库的使用方法、Echarts图表制作等80个实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。
④商务大数据实训实验库包括、某电商平台消费数据分析、商城销售情况分析、某电商货品销售分析案例、巴西电商平台交易数据分析、京东手机销售分析、电商平台化妆品品牌转化率分析、星辰口红流量分析案例、直播电商用户流失行为分析、某淘宝店铺经营分析、电商客户分析等10个实验项目任务，每个实验提供详细的指导手册，并能支持在低代码商业智能可视化实训平台进行实训。
⑤大数据营销实验库包括数据驱动的市场分析、大数据在产品策略中的应用、数据驱动的营销传播、客户关系管理与大数据、数据分析在销售决策中的作用、数据科学在市场营销中的高级应用、全渠道市场营销数据策略开发与执行实战等实

验项目任务，每个实验提供详细的指导手册，并能支持在交互式编程或低代码商业智能可视化实训平台进行实训。

17、提供大型零售集团基于数据治理的应用项目案例。实训任务包括：

①项目目标：支持学生围绕数据治理全过程开展建设规划、数仓建设、数据采集、数据治理和数据应用实施工作。

②建设规划：支持学生模拟通过调研和盘点零售企业业务系统和数据管理的现状和诉求完成紧贴零售企业业务的数据中台建设、数据管理和数据治理方案设计，建立分布式、分层级、分主题的数据仓库，实现数据的统一标准、统一管理和共享服务。

③数仓建设：支持指导学生基于零售企业的数据应用需求，依靠需求分析、数仓分层、数据建模、数据抽取整合等步骤实现数据仓库的建设。其中数仓分层需结合零售企业数据应用需求，仓分三层，分别为ODS层、DW层、ADS层，并设计出各层之间的数据流向关系图。

④要求数据采集实操以零售系统2张表为例展开实战。

⑤元数据管理，支持通过数据源管理->新建数据源->新建MySQL数据源功能操作完成元数据接入；支持使用关系发现算法自动解析业务系统数据表中字段之间的关联关系，为后续构建目录和服务提供依据。

⑥数据标准：支持学生完成新增业务主题、创建审批流程、制定数据标准、标准落地评估等操作；并通过采集工具，将ODS层接入的数据，通过标准化、规范化处理后，加工汇总到DWD层进行数据治理工作。

⑦数据质量：支持学生完成创建业务场景、配置核查对象、配置核查规则、配置核查任务等操作；支持基于治理后形成的标准化、高质量的数据，将DWS层数据按照用户业务需求，处理汇总成相关的分析报表到ADS层。

⑧数据安全：使用数据确权功能将零售销售管理系统确权给“销售中心”，使用数据密级功能完成数据密级分类。

⑨支持学生分析项目治理价值，包括业务价值和教学价值。

18、为满足智慧能源方向的教研训对数据的需求，提供煤炭、风电、光伏、储能、石油、电网等行业数据，数据量5000000条。其中煤炭方向提供煤炭设备泵站出口压力数据、煤炭设备三机电流数据、煤炭设备乳化泵运行数据、煤炭设备总运行数据、煤炭设备位置架上滚筒高度数据、煤炭设备三机开机数据、煤炭设备位置架数据等，均需为煤炭企业真实脱敏数据。

19、智慧能源方向支持接入Oracle、MySQL、Sqlserver、DB2、达梦、Gbase 8t、Postgresql、kyligence、KING BASE、SG-RDB、神通等数据源。

20、为提升学生在智慧能源方向应用大数据知识和技术能力，提供数据库原理与应用实验库、机器学习实验库、自然语言处理与应用实验库、生成式AI与大语言模型实验库。

①数据库原理与应用实验库包括数据库原理、MySQL数据库-安装与配置、库表操作、单表查询、复杂查询、数据库查询-选课系统、数据库开发综合案例-仓库管理系统设计等70个实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。

②机器学习实验库包括模型评估选择与验证、逻辑回归、决策树、kNN算法、朴素贝叶斯分类器、神经网络、随机森林算法、Adaboost、k-means、DBSCAN、AGNES、高斯混合聚类等70个实验任务。每个实验提供详细的指导手册，实验任务按照

<p>关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。</p>
<p>③自然语言处理与应用实验库包括自然语言处理前置技术解析、分词技术、词性标注与命名实体识别、关键词提取算法、句法分析、文本分类与情感分析实战等40个实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。</p>
<p>④生成式AI与大语言模型实验库包括神经网络架构和工作原理、损失函数和优化算法、Transformer架构解析、掩码张量、注意力机制、多头注意力机制、前馈全连接层、规范化层、子层连接结构、编码器、解码器、Transformer模型构建、模型基本测试运行、传统自然语言处理技术回顾-预处理和数据清洗、传统自然语言处理技术回顾-Word2vec、预训练和微调(python)、GPT应用领域-文本生成和生成式任务、GPT应用领域-机器翻译、大模型实战-闲聊机器人等35个实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线编程操作和自动评测。实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。</p>
<p>21、为提升学生在智慧能源方向的大数据实战应用能力和解决问题能力，提供开发油井压裂分析应用实践项目案例。</p>
<p>①项目目标：通过本课程的学习，使学生了解油井压裂前后产油量的变化，并以此为业务背景，学习对油井压裂的相关数据进行挖掘建模及应用，达到精准预测的目的。</p>
<p>②业务理解：支持提供油井压裂业务知识解读。</p>
<p>③数据准备：支持提供油田压裂井的日增油量真实数据集，包括油层组名称、砂岩顶深、砂岩厚度、动液面、油压、套压、流压、日产油、日产水、含水等30个字段维度，数据规模86169条。</p>
<p>④分析建模：支持学生通过交互式编程实训环境和低代码数据挖掘分析实训环境构建日增油量预测模型，分析影响压裂前后日增油量的因素，达到精准预测。在压裂井选取建模时，包含5个步骤：数据读取、数据处理、算法选择、模型评估、执行流程。</p>
<p>⑤可视化分析：建模处理结果数据进行可视化展示，包括数据总览，压裂井日增油分布占比情况，压裂井预测日增油分布占比情况，影响产油因素的重要性，影响产油因素的相关性等。</p>
<p>⑥提供详细的实训指导手册，包括项目背景、业务理解、数据理解、分析思路、分析建模、可视化分析、分析价值等。</p>
<p>22、为满足制造企业生产运营管控方向教研训对数据的需求，提供制造企业生产经营管控数据中心相关业务数据，包括企业经营指标数据、营业成本数据、安全管理、报警数据、工厂销售完成情况、设备停机率、问题列表数据等业务数据表，相关业务数据表核心均来源于业务部门定时填报。</p>
<p>23、支持多种标准协议数据采集，包括OPC-UA、OPC-DA、Modbus、MQTT、CoAP、TCP等。</p>
<p>24、制造企业生产运营管控方向支持将所有的数据源的表、视图按照业务进行分类管理，支持新建文件夹、新建SQL数据集、新建数据表。</p>
<p>25、提供数据治理实验库，包括数据治理与数据管理、数据战略从规划到执行、数据战略规划工具、数据采集的方法、数据存储的概念、数据存储需求、数据存</p>

储模型设计、数据存储架构设计、元数据管理、数据标准管理、主数据管理、数据质量管理、数据安全管理等70个实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线实践和自动评测。实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。

26、为满足智慧教育方向教研训对数据的需求，提供学生信息、师资力量、科研情况、升学、就业等数据，支持RESTFUL的API服务配置方式或脚本编码方式接入数据；并支持接入学校可开放的各类业务数据，如教务、教工、学工、人事、财务、资产、校情、科研等。

27、为提升学生在制造企业生产运营管控方向的大数据实战应用能力和解决问题能力，提供设备生产相关性分析实践项目案例。

①本项目案例支持以轮胎制造企业生产轮胎业务为背景，支持学生结合制造工艺特点，借助大数据分析挖掘技术手段，对于产品成型过程及质量检测数据集进行对应和关联分析挖掘，发现产品制造过程中的质量异常及相关影响因素，精确定位质量问。

②项目案例支持使用交互式编程实训环境和拖拽式的算法建模两种实训环境构建回归预测模型进行产品质量分析。

③项目案例提供产品质量相关数据集，包含规格、时间、花纹、条码、长度测量值、长度伺服测量值、长度设定值、宽度测量值、宽度设定值、角度(料头)测量值、角度(料尾)测量值、裁切温度测量值等 70 个数据维度，数据量 6000 条。

④支持提供详细的项目实训指导手册，内容包括项目背景、业务理解、数据理解、分析思路、知识点、建模步骤、分析价值等。

28、智慧教育方向支持多种数据集成方式，包括离线数据同步、非结构化数据同步、场景化单表数据同步、场景化批量数据同步等场景。离线数据同步支持配置前置SQL、后置SQL、作业管道总数、任务组管道数、脏数据收集器容量、脏数据阈值等相关配置；场景化数据同步支持周期增量、一次全量周期增量、周期全量等同步策略，支持扩展操作，包括：同步注释、数据对账、迁移信息等扩展操作。

29、为提升学生在智慧教育方向应用大数据知识和技术的能力，提供Web开发技术实验库、大数据可视化实验库。

①Web开发技术实验库包括HTML基础、JavaScript中的语句、JavaScript函数、JavaScript与字符串、JavaScript与文档元素等实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线实践和自动评测。实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。

②大数据可视化实验库包括Matplotlib模块和线形图、d3.js可视化、Bokeh进行可视化、图形化安斯库姆四重奏、三维图、地理信息的可视化、Seaborn的可视化、PyEcharts:基本图表、PyEcharts:直角坐标系图表、PyEcharts:树形图表、PyEcharts:组合图表、BI数据可视化、鸢尾花数据集可视化、全国历年GDP可视化分析等实验任务。每个实验提供详细的指导手册，实验任务按照关卡进行排布，并能支持在线实践。实验库整体配套教学大纲、教学课件（PPT）、教学视频、课程思政、教学方案等资料。

30、为提升学生在智慧教育方向的大数据实战应用能力和解决问题能力，提供高校校情管理分析项目案例。实训任务包括业务理解、分析思路、数据理解、可视化页面开发等。并配套实训手册、数据集、分析报告等。

(二) 行业应用实战库
<p>提供涵盖电子商务、金融保险、智慧零售、交通运输、智能电网、移动通信、医疗健康、生产制造等领域基于真实项目实训案例库，25个项目实战案例，应用于大数据、信管、经济、会计、金融、国贸及其他交叉学科的实验教学和科学研究。</p>
<p>1、电子商务行业-电商客户分析项目实战：本项目案例指导学生根据淘宝用户从点击、收藏、加购到购买的行为数据，通过商业智能工具探索不同种类商品的需求和客户转化的关系，分析客户需求与实际购买的情况，了解平台在商品推荐、广告投放及活动运营策略方面可能存在的问题，并对后续工作提升提供改善建议。项目支持低代码可视化分析实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>2、电子商务-互联网平台采购意向预测项目实战：本实验以某互联网平台的用户、商品和行为数据（脱敏后）为基础，通过数据挖掘的技术和机器学习的算法，构建用户购买商品的预测模型，输出高潜用户和目标商品的匹配结果，为精准营销提供高质量的目标群体。项目支持交互式编程实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>3、智慧营销-公募基金精准营销项目实战：本项目以金融行业的公募基金产品营销为业务背景，让学生运用机器学习知识对基于公募基金产品历史营销数据进行挖掘分析，一方面让学生了解金融行业公募基金产品营销的业务背景和传统营销方式存在的问题，另一方面让学生掌握通过机器学习技术如何建立客户细分模型和潜在客户预测模型来实现精准营销的目的，从而优化外呼营销策略，提升公募基金产品营销的意向成功率，降低成本。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>4、智慧金融-银行市场营销活动分析：本项目指导学生利用随机森林算法构建分类模型，通过模型分析影响定存业务的因素。利用分类模型预测目标客群，为电话销售经理提供有价值的信息，降低营销成本，提升营销效果。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>5、智慧零售-某零售企业经营分析项目实战：本项目指导学生基于零售门店经营数据，分析企业经营管理、人员管理、运营管理等基本情况，使用数据可视化方法找到业务规律和问题，为企业经营管理提供决策支持和依据。本案例涉及到的知识点主要有可视化分析的一般步骤、折线图、柱状图、条形图及气泡图的基本知识，以及可视化分析方法等。项目支持低代码可视化分析实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>6、智慧零售-企业运营综合分析项目实战：本项目以某零售公司的经营分析为背景，以培养学生商业分析能力为目标，以提高学生的就业竞争能力为重点，指导学生基于企业经营数据梳理并搭建公司管理层核心关注的指标体系、业务分析主题，并借助可视化分析工具进行企业经营数据处理与数据可视化，建立数据分析框架，提高商业业务分析能力，完成企业经营分析报告。项目支持低代码可视化分析实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>7、智慧人力-企业人力薪酬分析实战项目实战：随着公司的发展，公司的职员越多，则人员的分配及更替越是难题。本实训指导学生将对人员信息进行数据分析，快速分析、识别公司的人员特征，比如人员在各部门的分配情况、是否达到退休年龄等。从而指导企业对人力资源进行调整。项目支持低代码可视化分析实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>

<p>8、智慧财务-公司财务报表分析项目实战：本项目指导学生基于某公司的财务报表相关数据，通过可视化技术手段分析其主要财务指标，包括盈利能力、偿债能力、成长能力和运营能力等情况，帮助决策者了解企业过去的财务情况和财务现状，有效的检验管理人员的工作业绩，为未来如何实现企业的利益最大化，以及股东利益的最大化的最终目标决策打下有效基础，对企业的经营成果、财务状况以及管理者的未来决策提供真实有效的数据依据。项目支持低代码可视化分析实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>9、交通运输-地铁客流量分析预测项目实战：本实训以地铁客流量数据为基础，让学生将所学习的数据挖掘理论知识应用到实际的业务当中，建立相应的时间序列预测模型，帮助地铁运营优化；同时加深学生对ARIMA的理论知识理解，增强学生数据分析的能力以及养成良好的建模分析思维。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>10、交通运输-基于卡车缸温监测的故障预警分析实战：项目支持数据挖掘实训环境模块建立卡车发动机缸温预测模型。提供卡车发动机缸温数据集，包括环境温度、环境压力、风速、车速、马力发动机转速、加速踏板百分比、制动踏板百分比、k#缸温、k+1#缸温等50个维度，总体30000条数据。支持提供详细的项目实训指导手册，内容包括课程目标、业务背景、数据理解、数据准备、模型构建、缸温预测可视化、应用价值等。</p>
<p>11、智能电网-基于设备运行状态的重过载精准预测实践项目实战：项目支持拖拽式的算法建模模块建立设备运行状态智能感知模型和设备重过载预测模型。支持提供配电网设备相关数据集，包括描述设备编号、设备档案等相关信息，包括设备基本信息、过载信息等字段信息，每一个样本是每一台设备的相关信息，20000个样本，支持提供详细的项目实训指导手册，内容包括项目背景、业务理解、数据理解、分析思路、知识点、建模步骤、分析价值等。</p>
<p>12、智能电网-企业用能环保监测分析项目实战：本项目针对生态环境局重点企业无法监测、污染行为无法制止、政策执行无法把控的痛点难点，指导学生充分利用电力数据采样频度高、准确性好、可追溯性强等优势，聚焦生态环境部门微观出击、靶向治污、宏观把控三方面需求，构建重点企业污染防治大数据监测分析模型。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>13、智能电网-风电设备健康预警模型项目实战：由于风电机组受风速、风向的影响，长期在变转速、变负荷状态下运行，使得机组运行工况具有波动性、间歇性的特点。本项目指导学生利用大数据分析方法和数据挖掘技术迅速、准确地对齿轮箱故障进行预警，降低风电场的运维成本、提高风电场的经济效益、提高风电机组运行的可靠性，并提供实训数据集、实训指导手册。</p>
<p>14、移动通信-移动通信客户细分项目实战：本项目指导学生基于聚类算法利用消费行为数据对用户进行细分，了解不同群体的手机消费习惯，识别有价值的潜在客户，指导企业进行有针对性的产品或服务设计，提高客户满意度。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>15、移动通信-基于基站定位数据的商圈分析项目实战：本项目指导学生利用基站用户的历史定位数据进行科学的分析（聚类分群），归纳出商圈的人流特征和规律，识别出不同类别的商圈，为区域性差异化营销提供依据。</p>
<p>16、移动通信-手机客户流失项目实战：本项目指导学生基于运营商的真实数据，利用贝叶斯网络分类算法构建预测模型，分析影响手机客户流失的因素，帮助</p>

<p>通信行业从相关因素入手，留住高价值客户。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>17、医疗健康-糖尿病病情预测：本项目指导学生基于皮马印第安人糖尿病患病信息，利用随机森林算法构建糖尿病病情预测模型，预估发病风险，发现高危人群，帮助相关机构通过科学预防及早期治疗实现疾病健康管理。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>18、医疗健康-中医证型关联规则挖掘项目实战：本项目指导学生借助三阴乳腺癌患者的病理信息，挖掘患者的症状与中医证型之间的关联关系，对截断治疗提供依据，挖掘潜在证素，帮助患者建立合理的认知，促进恢复健康。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>19、医疗健康-乙型病毒性肝炎诊断项目实战：本项目指导学生基于历史数据分析乙肝患者的病状表现，构建分类预测模型，根据患者的症状表现预测患者是否患有乙肝，为医生诊断乙型病毒性肝炎提供科学依据。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>20、生产制造-制造企业生产经营管控：本项目以某企业生产经营管理过程为背景，指导学生通过可视化手段与方法，基于实际生产经营业务数据，进行生产过程中生产基本情况、质量、效率等多个维度的分析和展示，搭建企业生产经营管控分析看板，实时展示企业生产经营的现状，为相关的业务管理人员与一线生产人员提供管理的决策依据。项目支持低代码可视化分析实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>21、生产制造-设备质量相关性分析：本项目指导学生通过企业级机器学习工具将轮胎生产及质量数据有效融合，并结合轮胎成型制造工艺特点，进一步分析影响轮胎质量的影响因素及其影响程度，从而构建可用于分析轮胎质量影响因素、预测对应轮胎质量品级的智能模型，进一步为生产过程精细化管理、工艺过程优化和质量提升提供服务。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>22、基于电商平台商品评论的观点情感分析和话题分析：本课程通过学习和了解文本分析技术在解决电商行业中的面对海量用户评论数据如何实现精准洞悉的实际应用问题，使学生能够掌握文本分析的一些技术和分析思路，能够利用文本分析的技术解决实际案例中的一些非结构化数据的案例。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>23、基于向量空间和分类算法的新闻资讯文本分类分析项目实战：随着互联网资讯的激增，新闻资讯的自动化分类管理显得尤为重要。本项目指导学生基于向量空间和分类算法对新闻资讯文本数据分类建模，达到对新闻文本实现自动化分类和基于分类进行用户个性化推荐的目的。项目支持低代码数据挖掘实训环境进行实战演练，并提供实训数据集、实训指导手册。</p>
<p>24、智慧石油-稠油掺稀精准预测分析实践项目实战：核心实训项目任务包括：</p>
<p>①业务背景：以稠油掺稀业务为背景，具体包括油田生产的原油有重质原油、中质原油、轻质原油及少量的凝析油，其中生产的重质原油约占57%。在地层温度、压力条件下具有较好的流动性，但在举升过程中随着井筒压力、温度降低，黏度迅速增大，导致井筒流动阻力大幅度增加，增加了举升难度。</p>
<p>②课程目标：本案例以掺稀相关的生产数据为基础，指导学生建立稠油掺稀模型，帮助工程师在满足产量要求的情况下进一步降低稀油的使用量，降低稠油的开采成本。</p>

③数据准备：支持提供相关数据集，包括采油区块、产时、泵径、泵深、排量、油嘴、冲程、油压、套压井温、掺稀温度、产液密度及粘度等30个字段，数据量300000条。
④数据探索：支持对采集到的数据进行查看与探索。例如：各个数值型变量的最大值、最小值、均值、标准差、完整个数、缺失个数等。
⑤稠油掺稀精准预测模型：支持学生利用交互式编程实训环境和低代码数据挖掘分析实训环境，基于历史掺稀数据及生产数据，构建出掺稀量的预测模型。。
⑥模型评估：支持通过大数据分析实训环境提供的回归评估节点，检验回归模型的准确性和可靠性。
⑦提供详细的实训指导手册，包括项目背景、业务理解、数据理解、分析思路、分析建模、分析价值等。
25、智慧石油-勘探设备需求预测分析实践项目实战：核心实训任务包括：
①项目背景：本案例以石油地质勘探业务为背景，通过设备预测分析帮助企业进行设备管理。
②课程目标：本案例支持基于石油地质勘探项目设备与项目设计参数数据，建立设备选型模型和设备数量预测模型，帮助项目人员根据不同地表条件，判断项目需要采用的设备类型以及数量。
③数据准备：支持提供真实的项目勘探设备数据，包括项目名称、地表类型、作物区类型、震源、钻机、仪器车、地震仪器、检波器炮数、炮密度、检波器总数、震源类型、井深、前放增益、束线滚动距离、最大炮检距、组合井数震动次数等字段维度。
④设备选型数据探索：支持指导学生基于项目的地形和设备数量，寻找施工地形和设备类型的关联关系。支持对采集到的数据在建模前期进行查看与探索。
⑤设备数量预测数据探索：支持对不同地形所需的设备数量进行分析，将时长、日效、最大炮检距、前放增益、炮线距、炮密度、道间距等作为影响因素，将数据集拆分成训练集和测试集，采用相关算法对训练集进行建模，寻找影响因素和设备数量之间的数学模型，然后通过测试集来验证该数学模型的准确率。
⑥设备选型模型：支持根据历史的施工地形与设备类型数据，构建设备选型模型，寻找它们的关联关系，以针对不同的地形特征，判断所需要的设备类型。
⑦设备数量预测模型：支持根据前期相关性分析结果数据构建设备数量预测模型。
⑧提供详细的实训指导手册，包括项目背景、业务理解、数据理解、分析思路、分析建模、分析价值等。
（三）科研文献数据检索模块
1、具备全域语义检索功能，接入全球1.5亿+外文期刊资源池，支持跨百余种学科分类的联合检索。
2、具备文献图谱解析功能，支持基于主题聚类的智能语义聚合技术，一键生成全景式研究演进文献图谱，动态标注领域研究热点与学术争议域。
3、具备综述生成模组功能，支持基于知识密度压缩技术，自动提取多篇文献核心论点，一键生成文献综述。
4、提供覆盖科研全周期文档体系，包括国家社科基金申报书、国家自然科学基金申报书、教育部人文社科申请书、科研立项申请书、科研结项报告、调研报告、决策咨询报告、可行性研究报告、行业发展报告等12种科研全场景，内置学科领域知识包，自动匹配学术范式。

5、具备多模型协同能力，采用主流大模型引擎协同架构，智能规避单一模型幻觉风险，支持多版本内容衍生与溯源对比，并能实现动态润色，实现学术语言风格校准、逻辑链强化及图表数据一致性校验，提升报告严谨度。
大数据实践教学与智能管控平台
(一) 行业大数据技能实践环境模块
1、平台在大数据方向提供如下实验镜像环境：Hadoop, Hbase, Hive, Spark, Flink, Storm, Zookeeper, Scala, Kafka。
2、平台在Python方向提供包括实验镜像环境：numpy, scikit-learn, pandas, opencv-python, matplotlib, scipy, sklearn, scrapy, pytorch, python_utils, rabin, miller, crypto等；
3、在JAVA程序设计方向提供包括镜像环境：Java+maven+ant等。
4、平台在数据采集预处理方向提供包括镜像环境：Python、JAVA爬虫相关镜像等；在R语言方向提供包括实验镜像环境：R4、vim、man。
5、平台支持多种环境之间可以相互组合，教师可根据自身教学需求进行选择，如MySQL+Python、Java+Hive+Hadoop等，组建个性化实验环境。
6、支持按照实践环境名称、实践镜像名称、镜像类型等条件查询平台内所有的实践环境，支持对实践环境进行删除、编辑、查看版本记录的操作。
7、支持查看以环形图方式查看应用服务器、计算节点服务器资源的使用情况，包括CPU核数、内存(GB)、硬盘(GB)等信息。
8、支持设置同一用户同时开启多个实验，设置完成后同一用户可同时启动多个实验环境将持续使用资源。
9、支持以列表方式查看用户的K8S资源使用情况，包括姓名、账号、课程名称、实验类型、开启时间、运行时长、CPU、内存(GB)等信息，支持对K8S资源进行强制停止。
10、支持以列表方式查看用户的Docker资源使用情况，包括课程名称、课程类型、实验类型、开启时间、运行时长、CPU、内存(GB)、状态等信息，支持对Docker资源进行强制停止和删除。
11、提供课程统计分析功能。总体统计包括课程总数、内置课程数、教师公开课程数等信息；支持以列表方式统计包括课程名称、访问次数、访问人数、人均访问次数、创建课堂次数等信息，并支持以EXCLE形式导出至本地。
12、提供实践统计分析功能。总体统计包括实践总数、内置实践数、教师公开实践数等信息；支持以列表方式统计包括实践名称、访问次数、访问人数、人均访问次数、添加至课堂次数等信息，并支持以EXCLE形式导出至本地。
13、提供教师使用统计功能，支持统计包括姓名、工号、学院、专业、创建课程数、创建实践数、个人发布实践数、公开发布实践数等信息，并支持以EXCLE形式导出至本地。
14、提供学生使用统计功能，支持统计包括姓名、学号、学院、专业、年级、班级、登录次数、开始实践次数、实践学习时长、教学资源学习时长等信息，并支持以EXCLE形式导出至本地。
(二) 行业大数据在线实验模块：
1、为满足学生在智慧水利、智慧金融、智慧零售、智慧能源、制造企业生产运营管控、智慧教育等方向的大数据应用知识学习和技术在线实践需求，所有实验内容均基于浏览器的B/S模式，学生机器无需安装任何软件即可开始实验。
2、大数据技术在线实践提供包括在线编码、命令行、云桌面等多种实验方式，

支持根据不同实验目标、实验内容和实验难易度，匹配不同实验方式。
3、在线实践支持Python、R、SQL、Java多种编程语言环境。
5、大数据技术实验环境底层支持基于容器化相关技术，在后台无感知的实现容器的创建、销毁、资源回收。支持修改配置每个课程、每个学生所使用的cpu、内存、硬盘资源，配置每个学生、每个课程的硬盘持久化路径。每个学生的容器实验环境不会互相影响，保证每个学生的资源独立。
6、平台支持采用闯关的实验模式，将知识点融汇到各个关卡之中，学生通过答题或编程进行闯关，闯关成功可获得金币等形式的虚拟奖励。
7、支持代码实验自动评测功能，学生在线提交代码评测后，系统会对学生编写的代码的准确性进行自动化评测，实时获取评测结果。
8、支持云端编程实践模式设置多组测试集对学生代码运行结果进行校验，保障学生代码的可靠性。
9、支持根据实践课程内容对不同题目设置不同评测时长，防止代码出现死循环造成环境崩溃，同时也对学生代码性能作出要求，评测时间超过评测时长即视为代码不通过。
10、支持学生进行实践代码评测时，一键点击即可启用题目所需实验环境，启动评测时占用资源，评测完成后资源自动回收，快速释放实验资源，实现实验资源即启即用，自动分配利用。
11、提供重置单页代码、重置代码仓库、返回通关时代码等功能，让学生可将当前或全部的代码恢复到初始状态，重新开始学习，防止代码环境发生改变造成实践无法继续进行。
12、支持学生自行修改云端编程实践环境中代码字体大小。
13、支持对学生在云端编程实践环境中的学习行为进行监控与统计，记录学生学习时长，通关时间，通关答案。
14、远程命令行实践模式基于容器化方式实现，学生可在线根据题目要求在命令行上进行操作，提交作业。
15、学生在线提交评测后，系统会对学生作业的准确性进行自动化评测。若未通过，需返回失败原因。
16、支持远程命令行模式中记录学生的学习时长，通关时间。
17、支持远程命令行重置环境功能，让学生可将当前容器重置到初始状态，重新开始学习。
18、支持限制每个学生实践时使用的CPU、内存，可指定持久化保存容器内多个路径下学生的文件。
19、自动化创建学生的学习容器，在一定时间不操作后，资源自动回收，实现无感知的实验资源管理。
(三) 创建实践课程
1、支持教师自行创建或者通过一键复制方式开发创建实验课程，可自定义实验名称、实验分类、实验介绍、难易度、方向分类并根据所建课程需求自行选择实践环境。
2、支持对实验环境占用的资源进行自定义配置，配置项包括存储空间、内存、CPU等。
3、支持教师设置持久化路径，保存容器内多个路径下学生的文件。
4、支持教师在实验课程中创建选择题、判断题、实践题等多种类型实践关卡。
5、支持教师对创建的实验关卡进行编辑、删除、拖拽排序。

6、集成gitlab, 提供在线代码仓库, 支持教师在线编辑学生任务文件及代码评测脚本或将本地代码文件上传至在线代码仓库。
7、在线代码仓库支持创建文件夹对全部代码文件进行分类管理, 文件及文件夹支持拖拽移动。
8、在线代码仓库中需标识出已编辑但未保存的文件, 并统计全部未保存文件数量。
9、提供实验课程数据集, 支持教师将本地数据文件上传至数据集中, 通过复制容器地址的方式在代码文件中调用。
10、支持教师自主配置实验任务使用云端代码编辑器或在线命令行作为实践练习工具。
11、支持对有强关联性的实验课程设置不允许跳关。
12、支持教师将创建好的实验课程进行个人发布和公开发布。
(四) 课堂管理模块
1、支持教师进行课堂的创建, 根据自身需求设置课堂名称、学分、课堂开始与结束时间。
2、支持教师自行选择相关专业学生, 并将其添加至自己的课堂中、进行编辑、批量移除。
3、支持教师将公共库中的实验、基础实验、实训项目案例添加至自己创建的课堂中。
4、添加至课堂内的专业实践课程支持自动分配章节、排序、标号, 支持拖拽快速排序, 可快速拖拽实践课程重新分配章节并自动生成课程章节编号。
5、支持采用闯关的实验模式进行课程实践, 支持自定义实验课程关卡分值, 可按照均分比例、关卡难易度自动分配实验内的各关卡对应分值, 同时也支持教师手动输入各关卡分值。
6、支持一键发布全部添加至课堂内的实验。
7、对添加至课堂内的实验与实训项目可自由定制, 支持设置其必修/拓展属性, 设置为拓展将单独统计成绩, 不会计入课程总成绩中。
8、支持设置实验课程的补交扣分, 对于晚于课程规定时间提交的作业, 将默认扣除相应分值, 对学生提交作业的准时性进行区分。
9、支持根据教师发布状态、学生学习状态和时间进度自动分配课程状态, 区分未发布、学习中、补交中、已完成的课程。
10、支持统计课堂内全部学生对于课程的学习情况, 并将学习中、已完成、未开始状态下的学生人数统计展示在课程下方。
11、支持自定义实训项目作业要求, 教师可自行设置是否需要提交设计文件、实验报告、选择报告模板, 报告模板包括数据分析实验报告模板和开发实验报告模板。
12、支持教师对课堂内学生提交的项目作业统一查看, 提供AI建模流程/Word/pdf/Jupyter Notebook等多种形式作业的在线预览以及pdf下载。
13、支持教师可对学生项目作业进行主观评分, 支持教师对已评分的项目作业进行评分及评语修改, 支持教师将完成度高的项目作业评选为优秀作业。
14、支持按课堂、实训项目归类展示相关优秀作业, 并标注出对应课堂、实训项目中被评选为优秀作业的数量。
15、对于课程实验, 支持在线自动评分, 学生完成对应关卡后添加通关标识, 并对实践课程内全部关卡学习进度进行统计。

16、支持教师查看课堂内课程的全部学生学习情况及成绩，展示内容包括学生姓名、学号、作业完成状态、提交时间、作业总耗时、完成关卡数、关卡得分、补交扣分及课程总成绩。
17、支持以表格的形式分别统计课堂内全部学生的必修课程学习情况与拓展课程学习情况，展示内容包括学生姓名、学号、在线状态、年级、班级、实验完成数、实验平均分、实训完成数、实训平均分、课程平均分、总学习时长；并且支持通过点击表头中学号、年级、班级、实验完成数、实验平均分、实训完成数、实训平均分、课程平均分、总学习时长对列表内容进行排序。
18、支持教师上传课堂教学过程中所需的全部文档、图片、代码包等课程文件资源，实时共享到课程共享云盘中，方便学生进行查看、下载。
19、支持选择历史课堂名称，填写新建课堂名称、学分、起止时间及课堂内学生等，复用历史课堂资源组建新课。
20、支持修改历史已结束课堂的结束时间重启上课。
21、支持一键智能生成课堂章节，支持基于章节名称智能推荐元子实验和基础实验，支持用户自定义编辑。
22、支持一键智能推荐课堂相关的应用实训。
23、支持用户输入试卷名称、方向分类、题目类型、题目数量及对应分数和试卷难易程度等参数结合本课堂内容一键智能出题，支持在线预览试卷、自定义试卷名称。
24、教学管理模块具备智能生成前景描述、智能生成课程引导、教学总结、智能推荐、课堂图谱功能，其中课堂图谱提供课程目录、类思维导图的知识树、知识图谱，支持用户在目录中新增标签知识点。
25、支持智能生成课堂前景描述以及课程引导，教师端预览后学生端可查看。
27、支持在线预览及下载教学总结报告，包含课堂图谱、学情分析、考情分析及试卷详情分析。
(五) 学习探索模块
1、支持学生自定义自己的个人信息，包含头像、个性签名、标签等。
2、支持学生在课堂主页查看到教师发布出来的实验课程与实训项目，并且以标签的形式提示当前课程的学习状态，可通过标签筛选出未开始、学习中、待补交、已完成课程，快速查看到当前还有哪些课程待完成
3、支持学生打开浏览器点击课堂内的课程实验即可进入在线编码环境进行实验训练或使用编码式、拖拽式操作工具进行在线实训。
4、对于课程实验，支持在线自动评分，学生完成对应关卡后添加通关标识，并对实践课程内全部关卡学习进度进行统计。
5、支持学生在完成课程内容后可查看课程成绩，根据教师用户的设置，若课程设置了公开成绩则可以看到其他学生成绩；若设置不公开成绩则学生在列表中仅能查看到自己。
6、支持学生查看教师共享至课堂中的全部课件、文档、图片、视频、代码包等教学资源，并可将其下载至本地。
7、支持学生自行数据挖掘实训项目，在拖拽式、低代码的操作环境中，专注于大数据算法建模和洞察分析相关能力练习，自助进行项目实践。
8、支持学生查看被教师评选为优秀的实训项目作业，包括数据分析设计文件图、Jupyter Notebook代码文件以及pdf实验报告。
9、支持学生进入课堂后自动签到。

10、支持以日历图形式展示当天已签到人数、未签到人数以及未签到学生姓名。
11、支持导出本课堂考勤记录excel表格，包括自开课以来的所有学生签到情况。
12、支持学生查看课堂的前景描绘、课程引导，便于理解课堂内容。
13、支持学生查看课堂中实验、实训、实战及考试等学习路径甘特图。
(六) 大数据考试中心模块
平台内置了试卷、试题库，支持自定义新建试题或批量导入试题，以及通过模板自动组卷和手动选题组卷两种方式创建试卷，积累自己的教学成果。可通过试卷创建考试并发送至课堂，以课堂为单位进行在线考试，可在线评阅试卷并查阅考试成绩、成绩分布及详细答题情况分析，帮助教师快速了解学生对课程知识的掌握情况。
1、提供试卷、试题库，支持按照大数据技术、人工智能、数据库、编程语言、计算机技术、Web开发等方向进行筛选查询。
2、支持按照数据采集、数据存储与管理、数据分析处理、数据可视化、机器学习、深度学习、自然语言处理、语音识别、集成学习、SQL Server、Mongo DB、MySQL、Java、Python、操作系统、数据结构与算法JavaScript、HTML、CSS、Vue等分类进行筛选查询。
3、支持按照初级、中级、高级等难易度进行筛选查询
4、支持教师在线创建试卷，包括自定义试卷名称、方向分类、方向分类、组卷方式等。
5、支持教师自定义新建试题或批量导入试题，以及通过模板自动组卷和手动选题组卷两种方式创建试卷，积累自己的教学成果。
6、提供并系统预置大数据配套题库：配套习题集超过3000道，包括相关的单选题、多选题、判断题、简答题等。
7、支持教师通过试卷创建考试并发送至课堂，以课堂为单位进行在线考试。
8、支持在线预览试卷，并支持以word形式导出试卷至本地。
9、支持学生在线考试，可实时查看考试进度、考试剩余时间等信息。
10、支持客观题自动评分和主观题教师自主评阅两种评分模式。
11、支持成绩自动统计，包括平均分、最高分、最低分、及格率、平均答题时间、完成情况、待评阅试卷，并支持考试成绩以Excel形式导出。
12、支持以图表看板形式，包括以柱状图和饼图等形式展示成绩分布情况。
13、支持试题智能分析，分析结果包括满分人数、有效填写数、考试人数、满分率等，帮助教师快速了解学生对课程知识的掌握情况。
14、支持用户输入试卷名称、方向分类、题目类型、题目数量及对应分数和试卷难易程度等内容一键智能出题，支持用户自定义编辑。
(七) 课程创建模块
1、提供公共课程和我的课程两个资源库，支持按照大数据技术、人工智能、数据库、编程语言、计算机技术、Web开发、深度学习等维度进行筛选查询。
2、支持教师通过复制课程模版的方式进行课程创建，复制课程时可自定义课程名称，并支持是否选择现有课程自带的课程实验、教学课件、教学视频、课程考核、教学方案、课程思政等内容。
3、支持将当前专业课程复制至我的课程模块进行编辑、复制、发布、删除等操作。
4、支持自定义课程的基本信息，包括课程名称、课程封面、开发团队、合作团

队、课程标签、参考教材、课程简介等信息。
5、支持自定义课程教学大纲，并支持教学大纲预览和发布。
6、支持自定义课程内容，包括课程实验、教学课件、教学视频、教学方案、课程思政等内容，
7、课程实验支持添加、查看、重命名、删除等操作。
8、教学课件、教学视频、教学方案、课程思政等内容支持上传、查看删除等操作。
9、课程考核支持添加考试、预览、发布等功能，其中添加考试可自定义考试名称和选择试卷。
10、支持通过用户输入课程名称，智能生成课程简介或者基于用户输入的课程简介智能优化润色；
11、支持一键智能生成教学大纲，包括课程简介、基础信息、教学目标及要求以及章节课程大纲，支持用户自定义编辑；
12、支持根据教学大纲一键智能生成对应章节，支持基于章节名称智能推荐元子实验和基础实验，智能生成对应章节的教学课件、教学视频、课程思政及教学方案，支持用户在线预览及删除；
13、支持用户输入试卷名称、方向分类、题目类型、题目数量及对应分数和试卷难易程度等参数结合本课程内容一键智能出卷，支持自定义试卷名称；14、围绕课程章节和知识技能点智能自动生成课程图谱，包含课程目录、类思维导图的知识树以及课程知识图谱，支持用户新增章节知识点。
大数据智慧应用实训与科研平台
大数据智慧应用实训与科研平台支持以智慧水利、智慧金融、智慧零售、智慧能源、制造企业生产运营管控、智慧教育等行业应用为引导，提供企业级开发实战环境，通过实战演练让学生了解行业最新实践与应用场景，提升学生解决实际问题的能力，使知识传授、能力培养和素质拓展融为一体。另一方面支持教师将科研课题成果转化实训项目案例，让科研反哺教学，不断优化教学成果。
(一) 总体设计
1、提供公共实训和个人实训两部分内容，支持教师将公共实训和个人实训的项目案例资源添置个人课程中进行编排和发布应用于教学中。
2、大数据项目实训内容均支持基于浏览器的B/S模式，学生机器无需安装任何软件即可访问实训环境开始实训练习。
3、平台提供包括低代码开发、交互式编程、云桌面、智能体开发在内的四类实训。
4、每个实训项目提供详细实训指导手册，包括不限于项目背景、业务要求、分析思路、知识点讲解、应用价值等方面。
5、实训指导手册中支持插入图片、表格、代码块等。
(二) 交互式编程实训
1、提供以Jupyter Notebook为工具的交互式编程实训环境，支持学生在线根据题目要求在Jupyter Notebook上进行编码式大数据分析相关实训练习与学习。
2、提供代码训练功能，教师可通过自由设置，将代码模块留空，交予学生进行编程训练。
3、所提供实验环境文本编辑器支持多种编辑器风格，包括但不限于：markdown。
4、提供代码片段库功能，可直接插入Notebook内容，并运行；

5、支持学生提交.ipynb文件及实验报告作为实训项目作业。
6、提供重置环境功能，让学生可将当前容器重置到初始状态，重新开始学习。
7、支持自动化创建学生的学习容器，并进行资源回收计时。
8、支持在容器使用倒计时临近时对用户提醒，用户可根据自身需要进行续时，若使用时间到期且未续时，资源将自动回收，防止出现服务器资源被大量无效占用导致学生实验启动失败或卡顿问题。
9、实训任务提供实训数据集，支持教师将本地数据文件上传至数据集中，通过复制容器地址的方式在Jupyter Notebook中调用。
(三) 低代码实训
1、支持学生基于低代码数据挖掘分析实训环境，通过拖拽式布局、连线式流程编排和指导式流程配置的操作方式快速构建完成挖掘分析流程。
2、支持流程断点缓存功能，包括开启缓存、关闭缓存、清除缓存、从缓存处执行、执行到当前节点、从下一个节点开始执行等功能。支持配置流程运行资源，同时支持资源使用情况进行监控。
3、为方便学生能快速查询了解智慧水利方向的应用算法，支持常用算子功能，学生可将常用的节点进行收藏，方便使用。
4、提供24种行、列、高级节点的数据预处理算子，包括自动数据处理、数据去重、数据过滤、随机抽样、属性过滤、缺省值处理、异常值检测、数据平滑等。
5、提供5种数据融合算法，包括数据连接、数据拆分、数据分解等。
6、提供11种特征工程算法，包括属性生成、特征编码、主成分分析、因子分析、奇异值分析等。
7、提供7种统计分析算法，包括方差分析、相关系数、典型相关分析、偏相关分析等。
8、支持平行坐标图、散点图、箱线图、P-P图、Q-Q图等10种图形节点。
9、提供分类、聚类、回归、关联、时间序列、综合评价、推荐等7大类51种机器学习算法组件。其中分类算法包括C4.5+决策树分类、Xgboost分类、KNN、朴素贝叶斯、BP神经网络分类、L1/2稀疏迭代分类等，聚类算法包括KMeans、EM聚类、两步聚类、模糊C均值、视觉聚类等；回归算法包括线性回归、SVM回归、梯度提升树回归、BP神经网络回归、保序回归、L1/2稀疏迭代回归等；关联算法包括Apriori、FPGrowth等；时间序列算法包括ARIMA、稀疏时间序列、指数平滑、灰色预测、回声状态网络等；综合评价算法包括熵值法、层次分析法、模糊综合评价法；推荐算法包括协同过滤。
10、提供4种集成学习框架，包括Bagging分类、Bagging回归、Voting分类、Voting回归等。
11、提供5种深度学习算法，包括DNN回归、DNN分类、RNN分类、RNN回归、LSTM等。
12、提供13种文本分析算法，包括分词、信息抽取、文本过滤、向量空间、关键词提取、命名实体识别、文本相似度、观点情感分析等。
13、提供5种自动学习算法，包括自动择参、自动分类、自动回归、自动聚类等。
14、支持一键式建模功能，用户只需输入数据，通过该功能可以自动完成数据处理、特征工程、算法及参数选择及模型评估等环节。
15、提供6种扩展编程算法，支持用户编制SQL\R\Python\Java\Scala\Matlab等脚本语言实现个性化的算法扩展。

16、平台提供6种评估算法，平台支持对构建的挖掘模型进行K-S、Pr、Roc等评估，给出最优模型，同时能够在建模过程中，对模型进行输出、读取及利用。支持模型评估、模型读取、模型输出、模型利用等。
17、平台提供自定义算法功能，支持学生通过R\Python\Java\Scala\PySpark基于平台规范封装自主算法并发布形成平台节点。
18、平台提供30种工业信号分析算法组件，包括信号输入、信号输出、信号预处理、信号特征工程、信号变换、谱分析、信号滤波等7大类信号分析功能；其中信号变化功能支持模糊函数、希尔伯特变换、傅里叶变换、逆傅里叶变换、变分模态分解、小波变换等相关组件，每一个信号组件提供详细节点描述，包括功能、输入端口、输出端口、参数等信息；支持通过拖拽、连线、参数配置的方式，搭建包括IoTDB信号输入、信号分割、信号特征提取、数据转信号、信号文件输出等组件的建模流程。
19、平台提供建模洞察功能，支持包括数据接入、数据处理、算法模型、模型评估等在内的节点，均可以在平台的洞察中查看中间结果。并支持洞察报告预览功能及将洞察内容导出到WORD。
20、支持流程和模型的多版本管理机制，便于用户进行历史流程的回溯，或不同模型版本的引用。
21、进程管理：支持对本地、集群及当前用户所在队列进程的资源使用情况进行监控，用户可以通过查看CPU核数、内存的使用及剩余情况，设置服务的环境参数，支持修改CPU个数、内存大小。
22、支持学生在智慧水利、智慧金融、智慧零售、智慧能源、制造企业生产运营管控、智慧教育等方向的项目实训中提交机器学习实训成果及实验报告作为实训作业。
23、制造企业生产运营管控方向的大数据项目实训任务采用拖曳式的数图映射模式，支持学生基于低代码可视化分析实训环境，通过拖曳数据字段即可生成相关图形及构建可视化场景。
24、平台提供表关联、表结构同步、列重命名、列隐藏、列合并、计算列、地理分析、数据权限、列分析、类型转换、替换值、日期格式、按范围分段、数据权限分配等数据准备功能。
25、平台的图形组件支持常用、高级、组件、时序、实时等5类62余种，其中常用图形包括柱形图、条形图、柱线组合图、纵向组合图、面积图、饼图、玫瑰图、仪表盘、文字KPI、水球图（KPI）、行政地图、标记地图、迁徙地图、列表、交叉表、自由式报表，高级图形包括旭日图、玉玦图、矩形树图、漏斗图、关系图、词云图、瀑布图、箱线图、桑基图、力向导图、热力图、甘特图等，时序图形包括时序线图、时序柱图、时序面积图、时序仪表盘、时序水球图等，实时图形包括实时线图、实时标签图。
26、支持上卷、下钻、联动、链接、保留、排除等6种图形交互方式。
27、支持强大的分析计算能力：包括聚合计算，包括合计、计数、总体标准差、总体方差、平均值等；支持丰富的计算函数；支持复杂的分析场景：包括包括同比、环比、累计占比等；支持数据的预警分析，能够实现多样的条件格式；支持趋势线拟合、参考线、时序预测、聚类分析等功能。
28、平台提供3D图形组件，内置3D渲染引擎与3D效果设计器，支持上传OBJ类型的3D模型与数据指标进行映射展示，支持3D模型管理和3D脚本编码能力。
29、平台提供“中国式”复杂报表在内的多种报表样式，支持多表头合并、行列

转置、分组交叉、混合报表、同比、环比等，轻松实现任意形态的报表展现。
30、平台支持R语言组件、Python语言组件、JS脚本功能等图形展示效果扩展功能。
31、平台内置丰富的设计与美化素材，提供场景配置的快速入口并可一键完成场景优化，节省学生页面设计时间，提高效率。同时，学生可将场景页保存为个人模板，方便再次使用。
32、平台提供word报告生成模式，支持学生通过在word中插入平台的数据指标、图形报表、函数计算规则等，并支持word模板的上传与下载。
33、平台支持将多个可视化场景组装成一个专题报告。
34、平台支持学生提交可视化分析实训成果及实验报告作为实训作业。
35、支持在可视化分析实训模式下记录学生的学习时长。
36、支持机器学习与可视化分析相融合的实训模式，实现在一个实训中设置多个作业节点分别进行机器学习与可视化分析，利用其结果组合完成复杂度高的大型数据分析实训项目。
37、支持学生组合提交机器学习与可视化分析实训成果及实验报告作为实训作业。
38、支持分别记录学生在多个作业节点中的学习时长及学习的总时长。
(四) 云桌面实训
1、在行业大数据平台搭建与运维项目实训中，平台提供云桌面实训环境，利用虚拟化技术并使用虚拟机作为实训工具，支持创建多节点的集群，帮助学生进行大数据集群搭建、集群操作、数据处理等练习实训。
2、云桌面实训内容基于浏览器的B/S模式，学生机器无需安装任何软件即可访问云桌面实训环境开始实训练习。
3、支持复制外部内容到实验环境内，可将本文内容粘贴到复制功能文本框，点击保存，在实验环境中点击右键进行粘贴
4、支持复制实验环境中内容到外部，可在云桌面环境中复制内容，复制成功后内容会显示在文本框中，再次复制文本框中内容即可粘贴至本地。
5、支持实验手册和实验环境分屏显示，并支持实验环境全屏显示。
6、支持云桌面环境重置，开始重置功能后实现环境将被重置为初始状态，学生在云桌面中所做的所有修改将被撤销。
7、支持实验结束后在线提交PDF格式的实训作业。
(五) 大模型应用实训内容
(1) 支持为学生提供基于大模型技术的智能聊天助手开发项目的业务场景、实战数据、实战环境等。
(2) 基于大模型技术的智能聊天助手开发项目包括项目信息、数据准备、模型训练、大模型对话等4个部分。
(3) 为满足学生训练大模型对数据的需求，支持提供丰富的语料库，主要为对话式的文本记录，如社交媒体对话、客服问答、论坛讨论等。
(4) 实战环境直观展示项目实训镜像、存储空间、内存限制、算力限制等信息。
(5) 数据准备的核心目标包括：①理解数据源：深入理解txt聊天文件的内容、结构以及可能存在的数据格式问题。②数据清洗：学习并实践如何对txt文件中的聊天数据进行清洗，包括去除无关信息（如时间戳、用户标识中的非必要部分）、修正格式错误、以及处理可能的噪声数据。

(6) 模型训练阶段核心目标包括：让学生掌握大模型基本原理与训练流程，通过实操训练掌握数据准备、模型配置、训练、评估与调优的技能，同时培养解决问题的能力、团队协作精神和持续学习的能力，以确保学生具备将训练好的模型成功部署并应用于实际场景的综合能力。

(7) 大模型对话阶段目标：实现一个能够与用户进行问答式智能聊天的系统，通过编写和执行交互代码，使聊天助手能够理解用户输入的问题并给出相关、准确的回答，从而展现智能交互的能力。

(六) 大模型应用实训环境

1. 智能体开发环境

(1) 支持按照关键词进行应用工程的检索。

(2) 支持按照聊天助手、Agent、工作流分类展示应用。

(3) 支持按照应用创建角色不同进行团队和个人的应用筛选。

(4) 按照卡片样式展示应用的名称、图标、类型以及描述。

(5) 支持按照关键词或助手、Agent、工作流等分类检索。

(6) 支持基于大模型智能体模板一键创建应用

(7) 支持通过上传DSL文件或URL一键导入应用。

(8) 支持创建包括聊天助手、Agent、文本生成应用、Chatflow、工作流等智能应用。

(9) 支持自定义应用名称、图标以及描述创建应用。

(10) 支持通过设置提示词、变量、添加上下文、对话开场白、下一步问题及建议、语音文字互转、引用和归属、内容审查、标注回复、系统变量、环境变量、会话变量等进行应用编排。

(11) 针对工作流类支持通过拖拽连接及低代码方式配置工作流节点进行编排，如LLM、知识检索、结束、问题分类器、条件分支、迭代、代码执行、模板转换、变量聚合器、文档提取器、变量赋值、参数提取器、Http请求、列表操作等进行应用编排。

(12) 支持可视化实时调试与预览，可同时最多选择四个模型调试。

(13) 支持直接运行、嵌入网站、在应用中心打开三种应用发布方式，其中网站支持三种嵌入样式。

(14) 提供丰富全面的API文档供开发者查看，并管理可访问API的凭据。

(15) 支持在日志列表中查看应用的运行情况，包括用户的输入和AI的回复，包括对话标题、用户名、消息数、创建及更新时间等。

(16) 支持在标注列表中查看历史对话标注详情，包括问题、回答、命中来源、匹配相似分数、创建时间以及命中次数，支持开启/关闭/删除/编辑/新添加标注、设置标注阈值和模型以及批量导入导出标注问答对。

(17) 支持监控、跟踪应用程序在生产环境中的性能，在数据分析仪表盘内分析应用的全部消息数、活跃用户数、费用消耗、平均用户调用次数、Token输出速度、用户满意度等指标。

(18) 支持自定义标签进行应用工程的分类管理和检索。

(19) 支持编辑已创建应用的名称、图标以及描述。

(20) 支持复制、删除、导出DSL类型以及在应用中心打开应用。

2. 开发与微调

(1) 支持接入Deepseek、Qwen、Llama等主流大模型通过参数配置的方式进行开

发与微调，无需编写代码，支持配置模型路径、微调方法、检查点路径、量化等级、量化方法、对话模板、RoPE 插值方法、加速方式等，支持上传数据集、选择训练阶段目前采用的训练方式、数据路径、选择数据集、学习率、训练轮数、最大梯度范数、最大样本数、截断长度、梯度累积等模型训练参数配置，支持在评估预测阶段对最大生成长度、Top-p 采样值、温度系数等参数配置以及可进行预览命令、开始、中断等操作。
(2) 本地模型的文件路径或 Hugging Face 的模型标识符。
(3) 支持多种微调方法，包括lora、freeze、full等。
(4) 支持配置启用量化等级。
(5) 支持多种量化方法，包括bitsandbytes、hqq、eetq等。
(6) 提供丰富的模板用于构建提示词，包括qwen、deepseek3等。
(7) 支持多种RoPE 插值方法，包括linear、dynamic、yarn、llama3等。
(8) 支持多种加速方式，包括auto、flashattn2、unsloth、liger_kernel等。
(9) 支持上传训练的json数据集。
(10) 支持选择目前的训练阶段，包括Supervised Fine-Tuning、PPO、DPO等。
(11) 支持选择对应阶段的数据集，支持数据集的在线预览。
(12) 支持设置AdamW 优化器的初始学习率。
(13) 支持设置需要执行的训练总轮数。
(14) 支持设置用于梯度裁剪的范数。
(15) 支持设置每个数据集的最大样本数。
(16) 支持选择是否使用混合精度训练，如bf16、fp16等。
(17) 支持设置输入序列分词后的最大长度，不超过131072。
(18) 支持设置每个 GPU 处理的样本数量，不超过1024。
(19) 支持设置梯度累积的步数，不超过1024。
(20) 支持设置验证集占全部样本的百分比。
(21) 支持选择学习率调度器的名称。
(22) 支持日志间隔、保存间隔、预热步数、NEFTune 噪声参数、额外参数。
(23) 支持序列打包、学习提示词、无污染打包、学习历史对话、更改词表大小、使用 LLaMA Pro等参数设置。
(24) 支持可训练层数、可训练模块、额外模块等参数设置。
(25) 支持LoRA秩、缩放系数、随机丢弃、学习率比例、作用模块、使用 DoRA、使用 rslora、使用 PiSSA、附加模块等参数设置。
(26) 支持Beta 参数、Ftx gamma、损失类型、奖励模型等参数设置。
(27) 支持GaLore 秩、使用 GaLore、更新间隔、缩放系数、作用模块等参数设置。
(28) 支持 APOLLO 秩、使用APOLLO、更新间隔、缩放系数、作用模块等参数设置。
(29) 支持BAdam 模式、使用 BAdam、切换策略、切换频率、Block 更新比例等参数设置。
(30) 支持SwanLab 项目名、使用 SwanLab、模式、实验名、工作区、密钥等参数设置。
(31) 支持设置输入序列分词后的最大长度，不超过131072。
(32) 支持设置每个数据集的最大样本数。

(33) 支持设置每个 GPU 处理的样本数量，不超过1024。
(34) 支持设置最大生成长度，不超过4096。
(35) 支持设置Top-p采样值。
(36) 支持设置温度系数。
(37) 支持设置多种推理引擎，如huggingface、vllm、sglang。
(38) 支持设置多种推理数据类型，如auto、float16、bfloat16、float32。
(39) 支持设置单个模型文件的最大大小，不超过100GB。
(40) 支持设置量化导出模型。
(41) 支持设置量化过程中使用的校准数据集。
(42) 支持设置导出模型使用的设备类型，如cpu、auto。
(七) 大数据实训创建
1、平台支持教师选择低代码开发、云桌面、交互式编程、智能体开发等四种类型的实训环境开发实训项目课程，其中智能体开发支持自定义实训名称、实训简介、所属行业、实训难易度、实验课时、作业名称等信息，实训作业的实训环境模块选择范围包括聊天助手、Agent、文本自动生成、Chatflow、工作流等五种智能体开发模式，支持编辑实训手册，支持添加实训资料，包括上传pdf、ppt、excel、word、MP4等文件等。
2、支持教师组合使用可视化分析与机器学习工具，制作多作业节点的融合型拖拽式实训。可自定义拖拽式实训的实训名称、实训简介、所属行业、实训难易度、实验课时、设置项目作业节点以及是否提交实验报告、选择实训数据、创建示例工程、编辑手册、添加文件。
3、支持教师自行创建基于Jupyter Notebook工具的编码式实训课程。可自定义交互式编程实训的实训名称、实训简介、所属行业、GPU是否支持、实训难易度、实验课时、实训手册、作业要求、是否提交实验报告。支持教师根据所建课程需求自行选择实训环境，并支持对实训环境占用的资源进行自定义配置，配置项包括存储空间、内存、CPU，支持自定义持久化路径。
4、支持教师在线编辑实训作业中的示例代码文件或上传本地代码文件至Jupyter Notebook环境中，在完成编码式实训创建后，可同步至学生端，在学生实训过程中给予文字或代码提示。
5、提供实训数据集，支持教师将本地数据文件上传至数据集中，通过复制文件在容器中地址的方式在Jupyter Notebook中调用。
6、支持教师选择平台云桌面环境创建实训课程，支持教师云桌面课程的实训名称、实训简介、所属行业、实训难易度、实验课时、实训手册、作业要求、是否提交实验报告。
7、支持教师选择多节点实训环境类型，并支持通过高级配置功能配置实训环境所需要的运行资源，包括存储空间、内存限制、CPU限制等。
8、支持教师将创建好的实训课程进行个人发布，快速使用至自己的课堂中。
9、支持教师将创建好的实训课程进行公开发布，添加至实践课程资源库，供全平台教师用户复用。
10、支持教师对自己创建的实训课程进行统一管理，包括编辑、发布、下架、删除等操作。
数据治理综合实训与科研平台
为提升学生在智慧水利、智慧金融、智慧零售、智慧能源、制造企业生产运营管控、智慧教育等方向数据治理应用能力，需建设集数据采集、处理、分析、管理

<p>和服务于一体的数据治理综合实训与科研平台，以数据规划与管理咨询方法为指导，开展数据架构设计及模型设计，实现多源异构数据的统一归集与管理，践行数据仓库建设，辅以元数据管理、数据标准、数据质量及数据安全等治理活动，沉淀数据资产，并支持基于资产目录的数据订阅与服务发布，面向数据可视化展现和模型构建提供数据自助治理探索，从而实现数据价值释放。</p>
<p>(一) 数据标准管理</p>
<p>1、为确保智慧水利数据治理综合实训与科研中的数据的一致性、准确性和互操作性，以支持业务决策、降低管理成本并符合法规要求，平台具备多种类型数据的标准管理能力，包括主数据标准、参考数据标准和文档类标准；</p>
<p>2、为确保学生在智慧水利数据治理综合实训与科研方向能优化数据管理流程并提升数据应用价值，平台支持对主数据标准/参考数据标准的新增、修改、删除、智能识别、检索、导出及分类管理等；</p>
<p>3、平台支持对标准文档进行在线管理，包括但不限于上传、发布、下载、搜索、查看、预览和分类管理；</p>
<p>4、平台支持通过算法智能识别数据模型、数据库中的主数据表及属性，对主数据标准、参考数据标准进行快速提取和推荐；</p>
<p>5、平台支持对业务系统中的数据表一键快速生成数据标准；</p>
<p>6、平台支持对数据模型、数据标准及数据库之间的关系进行自动映射，并对映射关系进行编辑、查看和管理；</p>
<p>7、为支持学生在智慧水利数据治理综合实训与科研中能优化数据治理策略和提升数据管理水平，平台支持对主数据标准、参考数据标准落地情况、系统贯标情况进行整体评估分析，支持导出评估结果；</p>
<p>8、在智慧水利数据治理综合实训与科研方向，平台支持对数据标准的申请、审批、使用、监测等全流程管理。</p>
<p>(二) 数据模型管理</p>
<p>1、平台支持对数据模型（逻辑模型和物理模型）进行可视化展现，包括但不限于目录树、关系图谱等形式；支持对模型进行搜索、查看及分类管理；</p>
<p>2、平台支持对数据模型不同版本内容进行智能差异化对比分析，支持差异结果SQL脚本自动生成并导出；</p>
<p>3、平台支持对数据模型的基线管理和版本管理，提供数据模型的编辑、发布、扩充和审核等功能；</p>
<p>4、平台支持数据模型的逆向设计、在线设计能力</p>
<p>5、平台支持物理模型导出sql脚本，模型入库能力；</p>
<p>6、平台支持物理模型导出word和excel文档，作为模型设计的成果物进行管理；</p>
<p>7、平台支持对设计态数据模型和运行态数据模型进行智能差异比对，可自由设置调度规则，对数据表结构变更进行说明，实现数据模型的两态差异感知。</p>
<p>(三) 数据采集开发</p>
<p>1、为保障学生在智慧能源数据治理和深入科研中能获取全面、准确的数据，平台支持通过多种数据采集策略，对多源异构的各类数据进行统一归集。</p>
<p>2、平台支持但不限于达梦、Mysql、oracle、SQLserver、postgresql、hive、greenplum、Gbase 8a、神通、人大金仓、DB2、虚谷、ClickHouse等类型数据库之间的单表、批量同步。</p>
<p>3、同步时应支持对源端数据进行过滤筛选，只同步符合条件的数据；</p>
<p>4、目标端应支持新建表和选择已有表两种方式；</p>

5、同步时应支持点选式快速选择周期增量、一次全量周期增量、周期全量同步策略；
6、周期增量/一次全量周期增量场景下，应支持配置时间限定，应支持分钟增量、小时增量、日增量、月增量；应支持配置增量区间大小；应支持配置向前偏移时间；
7、应支持但不限于全删全增、主键更新、增量更新、版本表、缓存表等多种数据写入策略，应支持点选方式快速配置数据写入策略；
8、同步时应支持进行数据预览，查看源端与目标端数据映射关系，支持对字段映射进行编辑。建表应支持在迁移过程中添加新字段，新增字段应支持参数、函数、常量三种类型；应支持点选方式快速设置增量字段及唯一键；
9、应支持批量数据同步，一次性配置多张表的同步；
10、应支持配置分批读取行数、批量写入条数；
(四) 元数据管理
1、在大型零售集团基于数据治理的应用综合实训与科研方向，平台支持多源数据的快捷接入，支持数据自动抽取，一键式完成源数据同步入库；支持多数据源的统一管理；
2、为提升大型零售集团数据治理实训中的数据管理和分析效率，平台对接源业务数据库后，系统支持自动创建数据库模型，统一管理并可视化展示数据库表结构、关联关系等；
3、支持对源业务系统表结构的动态感知，并能够设置进行自动同步；
4、支持元数据变化版本的记录和管理，支持对比元数据不同版本变化内容；
5、支持基于业务数据对数据模型进行关系发现，可对关系发现算法进行配置维护。
6、支持基于元数据智能探查算法的数据探查，标记元数据的业务信息，比如身份证号、姓名、地址、电话号码等信息。支持智能分析字段取值，最小值、最大值、取值的频数和频率、模式频数频率、时间格式的频数和频率、最早时间、最晚时间、长度频数频率的统计分析。
(五) 数据血缘分析
1、在智慧水利、智慧金融、智慧零售、智慧能源、制造企业生产运营管控、智慧教育等方向，平台支持自动解析数据存储、数据同步、加工计算流程，自动解析数据之间的血缘与影响关系，直观展现数据流向；
2、支持手动创建维护数据的血缘、影响关系；
3、支持将自动解析的血缘和手动维护的血缘关系的自动合并、统一展现，展示数据全链关系；
4、支持展示数据表级、字段级的血缘关系图、影响关系图和全链关系图；
5、支持对数据血缘分析结果、完成情况、进展情况进行统计分析。
(六) 数据标签管理
1、支持对数据标签进行分类管理，支持标签的新建，绑定规则、删除标签、启用/停用标签等；
2、支持对标签设置规则，并构建统一的标签库进行管理；
3、支持自动打标签、手动打标签的方式对数据进行标签化管理；
4、支持按照标签维度对数据进行快速检索、统计分析和列表呈现。
(七) 数据质量管理

1、在智慧能源-某物资集团数字化转型项目实训和研究方向，平台支持数据质量场景管理，包括场景新增、删除、查看、统计等，能够配置对应的核查规则；
2、提供数据对象管理，支持配置数据表进行质量核查；
3、支持数据质量核查规则管理，包括规则的导入、删除、可视化配置规则、自定义SQL配置规则、智能化算法核查等；支持核查结果导出、入库、配置及生成报告；
4、支持对核查任务进行编辑、删除、配置、任务调度、异常监控、任务管理等；
5、支持业务核查和数据核查的双重质量核查，数据核查能够基于质量规则以评估任务调度的方式，对相关数据定时进行自动化的质量评估，同时支持手动实时进行质量评估；同一个评估任务支持同时评估多个数据模型；业务核查，可自由定义核查结果表，灵活按照业务需求输出核查清单。
6、支持设定不同规则的权重值，并基于权重值计算模型质量评估得分；
7、支持构建数据质量看板，对数据质量看板进行自定义配置、发布、查看等；支持输出质量规则、模型属性、质量改进趋势等维度的多维数据质量评估报告和脏数据明细，并提供查看和下载导出能力；
8、支持构建数据质量工单，通过脏数据占比等阈值触发生成工单，设置相关的责任人进行跟进、处理及反馈；
9、支持通过邮件等方式进行数据质量告警提示；
(八) 数据资源目录
1、在智慧水利、智慧金融、智慧零售、智慧能源、制造企业生产运营管控、智慧教育等方向的数据治理项目实训和科研中，平台支持对数据资源目录、数据资源目录元数据进行新增、删除、修改和管理；
2、支持对数据资源目录与数据表进行关系维护，包括目录加入、删除、清空及移动；
3、支持发布数据资源目录，支持共享数据资源目录；
4、支持数据目录按照图谱可视化方式展示数据资源以及资源间的关联关系；
5、支持对数据资源进行检索，包括关键字检索、目录检索、高级检索、检索结果查看等。
(九) 数据服务
1、在智慧水利、智慧金融、智慧零售、智慧能源、制造企业生产运营管控、智慧教育等方向的数据治理项目实训和科研中，平台支持零代码方式、SQL脚本方式灵活创建数据服务，并可对服务接口进行灵活配置管理，支持参数配置、字段配置、值映射配置等配置能力
2、支持数据脱敏规则配置、敏感数据处理等功能，确保数据的安全；
3、支持新增、删除、修改、维护服务目录，并可对服务与目录关系进行维护管理，支持对服务进行检索查询；
4、支持通过线上提需求的流程进行服务构建；
5、支持数据服务接口发布到服务超市上，供数据消费者获取。
6、支持服务的申请、审批、使用、监测等全流程管理；
7、支持白名单、黑名单管控，支持流控策略管控，确保数据服务稳定安全。
8、支持对服务访问、服务申请等服务运营情况进行监测统计。
(十) 数据生命周期管理

1、在智慧水利、智慧金融、智慧零售、智慧能源、制造企业生产运营管控、智慧教育等方向的数据治理项目实训和科研中，平台支持对数据资产进行归档管理，可根据数据热度进行数据在线、近线或离线区等自动归档，同时支持用户手动归档；
2、支持根据数据的使用情况进行销毁，提供相应的审批流程，只有通过审批的数据才能被销毁，临时销毁的数据存放在回收站内，支持一键还原；
3、支持对数据生命周期进行统计，用户可通过数据生命周期统计了解各系统中在线、近线、离线、销毁的数据表量，从业务系统角度统计各归档区数据量占比并统计各归档区数据变化走势。
(十一) 数据权限与安全
1、在智慧水利、智慧金融、智慧零售、智慧能源、制造企业生产运营管控、智慧教育等方向的数据治理项目实训和科研中，平台支持按照数据库、数据表、数据字段粒度进行数据确权管理，明确数据资产的权属关系。
2、支持按照数据的安全分级分类，根据不同的数据分级，与系统的人员密级进行匹配，设置不同的数据访问策略。
3、支持按照组织机构、按照角色、按照用户进行数据授权管理；
4、支持数据目录、数据服务、数据共享环节的脱敏处理。
5、支持自定义构建脱敏规则，内置脱敏规则库。
6、支持完善的权限与安全管控机制，从数据库、数据、文件等维度进行逐级的权限管理；支持对数据、文件及人员密级进行分级管理及匹配；
7、支持给不同角色配置数据源、数据分类的应用权限，具备权限后，用户可在具体的数据源、数据分类下进行模型管理；
8、支持给不同用户配置一个表的数据行具有不同的查看和维护权限；支持不同角色用户对同一表的列具有不同的查看和维护权限；
9、支持对敏感数据（如用户表等）采用可靠的加密算法，加密后存储，支持国密算法加密；
10、支持对数据资源目录进行分级权限控制，支持负面清单的设置与管理。

“数据安全”

附件2：售后服务

售后服务

我司针对该项目提供的售后服务内容如下：

1、质量保修范围和保修期

(1) 我司针对本项目所提供的产品为原厂原装、全新的、符合国家有关质量标准的产品，无版权纠纷。

(2) 保修期：在完成产品的最终验收后，提供 3 年免费质保期服务，产品保修期自验收合格之日起计算。

(3) 我司质保期满后，仍提供原厂专业维修服务，软件系统提供终身免费升级服务。

2、故障响应时间及维修时限

我司在质保期内提供原厂售后服务，包括提供 7*24 小时电话免费服务；如果产品在保修期内出现问题，在 1 小时内响应，最迟响应时间不晚于 2 小时；遇有严重技术问题、重大故障，需要现场维护，将在 8 小时内到达用户现场，24 小时内解决问题。

3、其他服务

我司在保期内由原厂专业人员对学生进行大数据项目集中特训服务，特训内容包括大数据分析项目、数据治理项目等，原厂专业人员将在进行线下教学，特训服务6次，每次时长为一周。原厂专业人员将具备工信部大数据分析证书或 DAMA 数据治理工程师证书。

4、产品对接与共享

1、我司根据客户要求提供软件产品与用户单位数据中台及统一身份认证系统免费对接的服务，以保证该产品与用户单位教务教学管理系统、实验室管理系统、科研管理系统、学生信息管理系统等系统的无缝对接，由此产生的费用乙方承担。

2、我司针对本项目所提供产品能满足甲方的仪器设备共享的要求。