

关于举办全国数字人才技能提升师资培训班的通知

(2024) 第二期

各院校：

随着数字经济的快速发展，社会对数字人才的需求与日俱增。为贯彻落实党中央、国务院关于发展数字经济的决策部署，发挥数字人才支撑数字经济的基础性作用，加快推动形成新质生产力，为高质量发展赋能蓄力，数字人才培养成为当务之急。根据《加快数字人才培养支撑数字经济发展行动方案（2024-2026年）》通知，“泰迪杯”数据挖掘挑战赛组织委员会将于2024年7月18日-8月3日举办2024年第二期全国数字人才技能提升师资研修班，旨在提升高校教师在大数据、人工智能、大模型等领域的专业知识和技能，以更好地支持数字经济的发展。本次研修班将开设三个线下专题和九个线上专题，涵盖了数字技术核心领域的最新进展和应用实践，具体安排如下文，期待您的参与。

一、培训组织

主办单位：泰迪杯数据挖掘挑战赛组织委员会

协办单位：广东省工业与应用数学学会

人民邮电出版社有限公司

北京泰迪云智信息技术研究院

华南师范大学人工智能机器人教育产业学院（广东省示范性产业学院）

承办单位：广东泰迪智能科技股份有限公司

二、课程安排

本次培训分为线下和线上两种班型，详细课程内容见文末附件课程大纲。

1 线下班（广州&青岛&北京）

混合模式，技能学习课程放置在线上云课堂以便学员系统学习，核心案例实战采用线下形式由讲师亲授同步操作演练。

线下广州班 [AIGC 技术与大模型应用开发实战](#)

基本信息	技能学习(线上云课堂)	现场案例实战, 7.19-7.23
时间: 7月18日-23日 线下报到时间: 7月18日 地点: 广州 学时: 共计56学时 证书: CBDA大模型应用工程师 费用: 3980元/人 学习环境要求: Windows10或以上操作系统(64位), 8G+内存 【专题特色: 从大模型工具使用到应用开发、从全新的提示工程到大模型对各行业的应用赋能、从RAG到AI Agent开发, 带你掌握大模型时代全面的实战技能。】	1. Python编程基础 2. 大模型赋能科研(拓展) 3. AIGC与图像生成(拓展) 4. 大模型在教育中的应用(拓展) 5. 大模型辅助金融数据分析(拓展) 6. 基于大模型的图像生成艺术(拓展)	1. 大模型与AIGC概述 2. 提示工程 3. AIGC的教学应用 4. AIGC生产力提升 5. RAG应用: 构建一个答疑小助手 6. Agent应用: 构建你的私人助理 7. LangChain基础使用 8. LangChain实战-RAG应用 9. 实战演练-产品手册智能问答助手 10. 实战演练-搭建一个AI Agent

线下青岛班 [人工智能技术与应用实战](#)

基本信息	技能学习(线上云课堂), 7.19-7.22	现场案例实战, 7.24-7.28
时间: 7月19日-28日 线下报到时间: 7月23日 地点: 青岛 学时: 共计80学时 证书: 高级人工智能应用工程师职业技术证书 费用: 3980元/人 学习环境要求: Windows10或以上操作系统(64位), 8G+内存 【专题特色: 贯穿机器学习与深度学习的核心内容, 深度探索计算机视觉和自然语言处理任务, 全面剖析深度学习原理, 并结合实际项目进行教学实践】	1. Python编程基础 2. Python数据分析与应用 3. Python机器学习实战 4. PyTorch框架基础实践 5. PyTorch深度学习原理与实现 6. 特别内容: ① ChatGPT教学应用 ② 金牌助手ChatGPT, 让应用开发更高效 ③ 大模型(ChatGLM)产业应用漫步与畅想	1. 核心技能串讲与巩固 2. 基于大模型微调的命名实体识别 3. 基于YOLOv8的岩石样本图像分割 4. 实战演练-基于YOLOv8的岩石样本图像识别

线下北京班 [大模型微调与私有化部署实战](#)

基本信息	技能学习(线上云课堂), 7.26-7.28	现场案例实战, 7.30-8.3
时间: 7月26日-8月3日 线下报到时间: 7月29日 地点: 北京 学时: 共计72学时 证书: CBDA大模型应用工程师 费用: 3980元/人 学习环境要求: Windows10或以上操作系统(64位), 8G+内存 【专题特色: 从基础神经网络到Transformer、从传统机器学习到预训练和微调, 彻底讲清楚大模型的工作原理与机制; 手把手带你在本地图部署属于自己的大模型 】	1. Python编程基础 2. Python数据分析与应用 3. Python机器学习实战 4. 大模型原理与实现 5. PyTorch框架基础实践(拓展) 6. PyTorch深度学习原理与实现(拓展)	1. 大模型与AIGC概述 2. 核心技能串讲与巩固 3. 大模型微调技术应用-基于大模型微调的新闻文本分类 4. 大模型部署 5. 大模型微调与部署实战: 基于通义千问(Qwen)的创意广告生成 6. 大模型微调与部署演练: 基于大模型的校内问答系统

2 线上班

本次培训教学视频六个月有效期可反复观看学习, 同步提供配套资源, 线上学习无需脱产, 可灵活安排时间。

专题一 [AIGC技术与应用实战](#)

基本信息	技能学习	案例实战
时间: 7月20日-30日 学时: 共计88学时 证书: CBDA大模型应用工程师 费用: 1980元/人 【专题特色: 紧贴大模型前沿应用、切实提升工作效率、全程无代码、真正零门槛学习最新技术 】	1. 大模型与AIGC概述 2. 提示工程 3. AIGC的教学应用 4. 大模型其它教育应用 5. AIGC生产力提升 6. 大模型辅助金融数据分析 7. 基于大模型的图像生成艺术	1. 大模型赋能科研 2. RAG应用: 构建一个答疑小助手 3. Agent应用: 构建你的私人助理 4. 基于大模型的小红书文案写作(拓展)

专题二 [大模型应用开发实战\(LangChain\)](#)

基本信息	技能学习	案例实战
时间: 7月20日-29日 学时: 共计80学时 证书: CBDA大模型应用工程师	1. Python编程基础 2. Python数据分析与应用 3. 大模型与AIGC概述 4. 提示工程	1. 大模型赋能科研 2. RAG实战-论文阅读助手 3. Agent实战-在线教育课程订单智能助手

费用：1980 元/人	5. AIGC 的教学应用	4. AIGC 与图像生成（拓展）
学习环境要求：Windows10 或以上操作系统（64 位），8G+内存	6. LangChain 实战-基础使用	
	7. LangChain 实战-RAG 应用	
	8. LangChain 实战-Agent 应用	

【专题特色】从大模型工具的基本使用开始，逐步深入到 LangChain 应用，以实战案例演练来了解大模型应用方向 RAG(检索增强生成)技术与 Agent(代理)技术。

专题三 [大模型微调应用实战](#)

基本信息	技能学习	案例实战
时间：7 月 20 日-30 日 学时：共计 88 学时 证书：CBDA 大模型应用工程师 费用：2480 元/人 学习环境要求：Windows10 或以上操作系统（64 位），8G+内存 【专题特色】 从基础神经网络到 Transformer、从传统机器学习到预训练和微调，彻底讲清楚大模型的工作原理与机制；手把手带你搭建属于自己的大模型	1. Python 编程基础 2. Python 数据分析与应用 3. 大模型与 AIGC 概述 4. Python 机器学习实战 5. PyTorch 框架基础实践 6. 大模型原理与实现 7. 大模型微调技术应用	1. 基于通义千问（Qwen）的创意广告生成 2. 基于大模型微调的命名实体识别 3. 基于大模型的金融问句语义相似度计算 4. 基于 Transformer 的疫情问诊系统自动翻译实现（拓展）

专题四 [大模型私有化部署和应用开发实战](#)

基本信息	技能学习	案例实战
时间：7 月 20 日-29 日 学时：共计 80 学时 证书：CBDA 大模型应用工程师 费用：2480 元/人 学习环境要求：Windows10 或以上操作系统（64 位），8G+内存 【专题特色】 本专题重点介绍了多种部署框架如 Langchain、Ollama 及多样应用案例，教导学员本地部署属于自己的大模型。学习过程融入实战和前端开发，提升互动性和实用性，促进深	1. Python 编程基础 2. Python 数据分析与应用 3. 大模型与 AIGC 概述 4. 开源社区 Huggingface 和 ModelScope 5. 基于开源代码的 Qwen 大模型部署 6. 基于 Langchain 的本地大模型部署和应用开发	1. Qwen 大模型部署与电商客服模型微调 2. 基于 ChatGLM 的校内问答系统 3. 基于 Langchain 部署本地大模型的知识库助手应用开发 4. 基于 Ollama 和 Dify 的本地大模型的部署和应用开发

专题五 [数据分析与挖掘实战（泰迪杯竞赛方向）](#)

基本信息	技能学习	案例实战
时间：7月20日-30日 学时：共计88学时 证书：高级机器学习工程师 职业技术证书 费用：1980元/人 学习环境要求：Windows10 或以上操作系统（64位）， 8G+内存 【专题特色】 ：课程融合泰迪杯经典竞赛，涵盖机器学习、爬虫和深度学习基础。分析历年竞赛案例，深入实战，理论与实践相结合，带你领略竞赛魅力，掌握实战技巧。】	1. Python 编程基础 2. Python 数据分析基础 3. Python 数据分析与应用 4. Python 机器学习实战 5. Python 网络爬虫实战 6. PyTorch 框架基础实践 7. PyTorch 深度学习原理与实现（拓展） 8. 特别内容： ① ChatGPT 教学应用 ② 金牌助手 ChatGPT，让应用开发更高效 ③ 大模型(ChatGLM)产业应用漫步与畅想	1. 学生校园卡消费行为分析（2019年技能赛B题相关） 2. 泰迪内推平台招聘信息采集与分析（2023年挑战赛C题相关） 3. 基于YOLOv8的岩石样本图像分割（2021年挑战赛B题相关）

专题六 [大数据技术应用实战\(Hadoop+Spark\)](#)

基本信息	技能学习	案例实战
时间：7月20日-29日 学时：共计80学时 证书：高级大数据技术应用 职业技术证书 费用：2980元/人 学习环境要求：Windows10 或以上操作系统（64位）， 8G+内存 【专题特色】 ：本专题课程涵盖了Hadoop生态圈的多种组件，如Hadoop、Hive、Spark、HBase、ZooKeeper等，这种全面的知识体系有助于学员从整体上把握Hadoop生态圈的架构和工作原理。课程注重理论与实践的结合，让学员了解最新的大数据处理技术，并在实践中掌握相关的操作技能。】	1. Linux 操作系统基础 2. Java 编程基础 3. Scala 编程基础 4. Hadoop 大数据基础 5. Hive 大数据仓库 6. Spark 大数据技术基础 7. Zookeeper 大数据分布式消息（拓展） 8. HBase 非关系型数据库（拓展） 9. Flume 数据采集（拓展） 10. Kafka 消息系统（拓展） 11. Flink 大数据实时处理（拓展） 12. 特别内容： ① ChatGPT 教学应用 ② 金牌助手 ChatGPT，让应用开发更高效 ③ 大模型(ChatGLM-6B)产业应用漫步与畅想	1. 航空客户价值分析 2. 广电大数据用户画像（Hadoop + Spark + Hive） 3. 大数据分布式消息 Zookeeper（拓展） 4. 商品实时推荐系统（拓展）

专题七 [数据采集与处理实战（Python&八爪鱼）](#)

基本信息	技能学习	案例实战
------	------	------

<p>时间：7月20日-30日</p> <p>学时：共计88学时</p> <p>证书：高级Python技术应用工程师职业资格证书</p> <p>费用：1980元/人</p> <p>学习环境要求：Windows10或以上操作系统（64位），8G+内存</p> <p>【专题特色】课程以Python编程实现与八爪鱼工具应用双轨并行。从零开始讲解数据采集流程Python代码实现，以及使用无编程门槛的八爪鱼工具高效实现数据采集工作，实现工具与编程技能的双重掌握与优化应用。</p>	<ol style="list-style-type: none"> 1. Python 编程基础 2. Python 数据分析基础 3. Python 数据分析与应用 4. Python 网络爬虫实战 5. 特别内容： <ol style="list-style-type: none"> ① ChatGPT 教学应用 ② 金牌助手 ChatGPT，让应用开发更高效 ③ 大模型(ChatGLM)产业应用漫步与畅想 	<ol style="list-style-type: none"> 1. 泰迪内推平台招聘信息采集与分析（Python） 2. 新浪财经数据采集（八爪鱼工具） 3. 汽车质量投诉平台数据采集（八爪鱼工具） 4. 微博疫情评论数据采集（Python） 5. 网站图像素材采集实战（拓展）
--	---	--

专题八 [商务数据分析实战（Python）](#)

基本信息	技能学习	案例实战
<p>时间：7月20日-30日</p> <p>学时：共计88学时</p> <p>证书：高级大数据分析师职业资格证书</p> <p>费用：1980元/人</p> <p>学习环境要求：Windows10或以上操作系统（64位），8G+内存</p> <p>【专题特色】课程以六大商务数据分析场景为切入点，以数据分析为导向，逐步深入分析案例。学员们将能够更容易地理解商务数据分析的核心概念与实践应用。</p>	<ol style="list-style-type: none"> 1. Python 编程基础 2. Python 数据分析基础 3. Python 数据分析与应用 4. Python 数据分析实训 5. 商务数据分析概述 6. 特别内容： <ol style="list-style-type: none"> ① ChatGPT 教学应用 ② 金牌助手 ChatGPT，让应用开发更高效 ③ 大模型(ChatGLM)产业应用漫步与畅想 	<ol style="list-style-type: none"> 1. 牛刀小试-T公司运营数据分析与业务决策 2. 回归分析-某市财政收入预测 3. 聚类分析-对某航空公司客户分群 4. 分类分析-运营商用户流失判别 5. 关联规则-购物篮商品推荐 6. 协同过滤-泰迪内推平台信息精准推荐应用

专题九 [PyTorch 深度学习实战](#)

基本信息	技能学习	案例实战
<p>时间：7月20日-30日</p> <p>学时：共计88学时</p> <p>证书：高级人工智能应用工程师职业资格证书</p> <p>费用：1980元/人</p> <p>学习环境要求：Windows10或以上操作系统（64位），8G+内存</p> <p>【专题特色】重点专注于计</p>	<ol style="list-style-type: none"> 1. Python 编程基础 2. Python 数据分析与应用 3. Python 机器学习实战 4. PyTorch 框架基础实践 5. PyTorch 深度学习原理与实现 6. 特别内容： <ol style="list-style-type: none"> ① ChatGPT 教学应用 ② 金牌助手 ChatGPT，让应用开发更高效 	<ol style="list-style-type: none"> 1. 脑PET图像分析与疾病预测 2. 基于textCNN的公众健康问句分类 3. 基于大模型微调的命名实体识别 4. 基于YOLOv8的岩石样本图像分割与识别

三、课程特色

1、数字人才赋能产业+落地实操授课

课程全程强调动手实操，以实际编码为核心，通过解析企业案例，培养学员在产业技术和数字技术方面的综合能力，助力产业数字化转型和高质量发展。

2、全方位答疑辅导学习

课程设有答疑交流讨论群，培训期间助教全程辅助教学。线上学习部分每天提供 10 小时的实时在线答疑辅导。

3、内容从浅及深更易入门

本课程配套有基础知识内容，即使零基础学员也能找到适合自己的学习内容和节奏，快速掌握课程知识和技能，助力学员入门数字经济领域。

4、大模型应用：畅想数据智能新未来

紧贴产业前沿，特邀三位专家畅谈大模型应用，分享“ChatGPT 教学应用”，“金牌助手 ChatGPT，让应用开发更高效”，“大模型(ChatGLM) 产业应用漫步与畅想”等专题，一起畅想数据智能新未来。

5、提供课程资源和回看功能

所有课程相关源代码、数据、PPT、案例素材全部提供下载，即学即用，教学更轻松！线上课程内容支持六个月内免费回看，以便复习和参考。

6、满足教学和科研需求

通过技能学习和案例实战，学员将在具体应用场景中全面掌握相关技能，提升实训教学工作的实际动手能力并为后续科研打下坚实基础。

四、证书颁发

1、CBDA 大模型应用工程师证书

线下广州班、线下北京班、线上专题一、线上专题二、线上专题三和线上专题四的学员完成培训并经考核合格后，可以获得由泰迪智能研究院国际培训中心颁发的“大模型应用工程师”证书。



2、职业技术证书

其他专题学员经培训并考试合格后，可以获得由工业和信息化部教育与考试中心颁发的相应职业技术证书，证书可登录工业和信息化部教育与考试中心官网查询。



五、报名及联系方式

1、报名材料：报名申请表、身份证复印件、两寸近期正面免冠彩色半身证件照电子版(要求：背景：白色，格式：JPG，大小：14-20K)。

2、本期研修班由广东泰迪智能科技股份有限公司收取费用并开具发票。

3、联系方式

联系人：曾老师

电 话：13246821827



全国数字人才技能提升师资研修班报名申请表

单位名称									
部门/院系									
通讯地址									
发票抬头									
纳税号				电子发票接收邮箱					
联系人		电话		发票内容					
以下报名信息表请认真完整填写									
姓名	手机号码	性别	身份证号	职称	主要任课专业	毕业院校	最高学历	专题选择	是否住宿 (单间/标间)

费用支付方式	1、汇款到指定账号。 2、扫码支付（报名后联系工作人员索要支付码）。 注： 付款时请注明“数字人才技能提升研修班+单位或姓名”字样，方便查账备案。								
账户信息	账户名：广东泰迪智能科技股份有限公司 开户行：中国工商银行广州花城支行 账户号：3602 0285 0920 1663 221								
备注	请将报名表发送至邮箱： zengaizhi@tipdm.com 联系人：曾老师 13246821827 微信：antonia602501								

附件一 线下班课程大纲

<线下广州班> AIGC 技术与大模型应用开发实战课程大纲

技能学习（线上云课堂）课程安排

时间：报名成功后即可开始学习

Python 编程基础

- 1 认识 Python
- 2 编写 Python 程序
- 3 认识 Python 数据结构
- 4 条件判断及分支语句
- 5 使用 def 定义函数
- 6 认识面向对象
- 7 读取文件数据
- 8 模块和第三方库

拓展课程

大模型赋能科研

- 1 抛出目标-目标路径-实际场景
- 2 确定发明专利的关键点
- 3 完成发明专利的基本内容
- 4 完成发明专利的详细阐述
- 5 赋能科研-科研小助手
- 6 赋能科研-kimi

AIGC 与图像生成

- 1.1 AI 绘画简介
- 1.2 绘图提示词原则与框架
- 1.3 绘图工具介绍
- 2.1 绘图流程
- 2.2 文生图与图生图
- 2.3 图像连续创作

大模型在教育中的应用

- 1 学术版 ChatGPT 应用指南
 - 1.1 ChatGPT 学术版简介
 - 1.2 ChatGPT 学术版的安装
 - 1.3 ChatGPT 学术版的基础使用
 - 1.4 ChatGPT 学术版的高级功能
 - 1.5 ChatGPT 学术版在学术研究中的应用
- 2 大模型课堂教学应用

大模型辅助金融数据分析

- 1 大模型+金融数据分析介绍
- 2 智谱清言大模型介绍
- 3 应用大模型进行金融数据分析
 - 3.1 金融数据清洗和数据处理
 - 3.1.1 数据排重
 - 3.1.2 缺失数据处理
 - 3.1.3 数据格式检查
 - 3.1.4 异常值处理
 - 3.2 股票数据分析
 - 3.2.1 股票的收益率计算
 - 3.2.2 绘制 K 线图和成交量示意图
 - 3.2.3 绘制投资组合 K 线图和成交量示意图
 - 3.3 量化投资
 - 3.3.1 绘制股票价格走势
 - 3.3.2 双均线投资策略的制定
 - 3.3.3 收益情况分析
- 4 小结

基于大模型的图像生成艺术

- 1 大模型与计算机视觉简介
 - 1.1 计算机视觉的定义与应用
 - 1.2 大模型在计算机视觉中的应用
- 2 利用大模型对图像进行描述
 - 2.1 图像标注与描述技术概览
 - 2.2 图像描述的实际操作
 - 2.3 图像描述的应用案例
- 3 利用大模型生成图像
 - 3.1 文本到图像的转换
 - 3.2 大语言模型在图像生成的应用
 - 3.3 面临的挑战与未来发展
- 4 利用大模型绘制数据图
 - 4.1 利用大模型绘制柱状、折线图
 - 4.2 利用大模型绘制饼图、雷达图
 - 4.3 利用大模型绘制散点图、漏斗图、思维导图
5. 开源图像小模型
 - 5.1 魔搭与 PaddleHUB
 - 5.2 图像小模型
 - 5.3 图片文字识别
 - 5.4 图像去水印
 - 5.5 图像清晰化
 - 5.6 色彩增强
 - 5.7 风格迁移
- 6 使用开源模型进行艺术绘画
 - 6.1 Stable Diffusion-webUI
 - 6.2 其它开源绘画模型
7. 利用大模型生成视频
 - 7.1 基于文心一言一键流影
 - 7.2 stable-video-diffusion-webui
 - 7.3 魔搭平台的视频生成模型

<p>2.1 利用大语言模型对老师的课堂内容总结</p> <p>2.2 利用大语言模型对老师的课堂内容问答</p> <p>2.3 利用大语言模型随机出题与批改作业</p> <p>3 人工智能行业法规</p>		
现场案例实战课程安排		
<p>时间：7.19 (8:30-11:30)</p> <p>大模型与 AIGC 概述</p> <p>1 大模型是什么</p> <p>1.1 大模型的定义</p> <p>1.2 大模型的特性与应用</p> <p>2 大模型的前世今生</p> <p>3 大模型爆火的原因</p> <p>4 大模型原理介绍</p> <p>4.1 ChatGPT 工作机制与训练流程</p> <p>4.2 自监督预训练解读</p> <p>4.3 有监督微调介绍</p> <p>4.4 奖励建模与强化学习</p> <p>4.5 token 是什么</p> <p>5 AIGC 简介</p> <p>提示工程</p> <p>1 提示工程 (Prompt Engineering) 简介</p> <p>1.1 提示 (Prompt) 是什么</p> <p>1.2 提示的发展历程</p> <p>1.3 提示工程 (Prompt Engineering) 是什么</p> <p>1.4 提示词编写原则与策略介绍</p> <p>2 提示编写原则与策略</p> <p>2.1 原则与策略 1: 编写清晰的提示</p> <p>2.2 原则与策略 2: 提供参考示例</p> <p>2.3 原则与策略 3: 让模型一步步思考</p> <p>2.4 原则与策略 4: 调用外部工具</p> <p>2.5 原则与策略 5: 将复杂任务分解成子任务</p> <p>2.6 原则与策略 6: 采用系统的提示框架</p> <p>2.7 原则与策略 7: 用结构化方式进行提示</p> <p>2.8 自动生成 Prompt</p>	<p>时间：7.20 (8:30-11:30)</p> <p>AIGC 生产力提升</p> <p>1 学习准备与声明</p> <p>2 大模型辅助生成 PPT</p> <p>2.1.1 一步到位生成自我介绍 PPT</p> <p>2.1.2 基于文档制作 PPT</p> <p>2.2.1 提炼并梳理文档中的核心内容</p> <p>2.2.2 将 markdown 内容转为 PPT</p> <p>3 大模型辅助生成思维导图</p> <p>3.1 利用大模型生成思维导图</p> <p>3.2 将已有内容提炼成思维导图</p> <p>4 大模型辅助绘制流程图</p> <p>RAG 应用: 构建一个答疑小助手</p> <p>1 目标与分析</p> <p>2 RAG 介绍及其工作流程</p> <p>3 任务实现</p> <p>3.1 RAG 实现 1: 将文档直接上传至大模型</p> <p>3.2 RAG 实现 2: 创建一个 Bot</p> <p>3.2.1 设定 Bot 角色及回复逻辑</p> <p>3.2.2 上传本地知识库并调试 Bot</p> <p>3.2.3 发布并使用 Bot</p> <p>4 小结</p> <p>时间：7.20 (14:00-17:00)</p> <p>Agent 应用: 构建你的私人助理</p> <p>1 Agent 是什么</p> <p>1.1 引言: 大模型与通用人工智能的差距</p> <p>1.2 Agent 的概念和示例</p> <p>2 第一个 Agent: 天气播报员</p> <p>2.1 利用扣子平台创建一个 Bot</p>	<p>时间：7.21 (8:30-11:30)</p> <p>LangChain 实战-基础使用</p> <p>1.1 LangChain 简介</p> <p>1.2 LangChain 环境安装</p> <p>2.1 申请阿里云的 API KEY</p> <p>2.2 LangChain 示例</p> <p>2.3 链的使用</p> <p>3.1 提示模板</p> <p>3.2 模型调用</p> <p>3.3 输出解析</p> <p>3.4 基于 LangChain 开发情感极性分析应用</p> <p>3.5 课堂练习: 基于 LangChian 开发命名实体识别应用</p> <p>时间：7.21 (14:00) -7.22(11:30)</p> <p>LangChain 实战-RAG 案例应用</p> <p>4.1 RAG 的概念</p> <p>4.2 文档加载</p> <p>4.3 文档分块</p> <p>4.4 向量化存储</p> <p>4.5 检索器</p> <p>4.6 函数式链构造 RAG</p> <p>4.7 记忆机制</p> <p>4.8 为 RAG 添加记忆功能</p> <p>4.9 使用 streamlit 开发界面应用</p> <p>时间：7.22 (14:00-17:00)</p> <p>实战演练-产品手册智能问答助手</p> <p>1. 任务提出</p> <p>2. 任务实现路径分析</p> <p>3. 本地知识库准备</p> <p>4. RAG 实践</p> <p>5. 效果检测</p> <p>实战演练-搭建一个 AI Agent</p>

<p>3 总结</p> <p>时间：7.19 (14:00-17:00)</p> <p>AIGC 的教学应用</p> <ol style="list-style-type: none"> 1 学习准备与声明 2 大模型辅助教案撰写 <ol style="list-style-type: none"> 2.1 编写一份教案模板 2.2 撰写具体教案内容 3 题库题目生成 4 大模型辅助文献阅读 5 大模型辅助编程 6 大模型辅助数据分析 7 总结 	<p>(Agent)</p> <ol style="list-style-type: none"> 2.2 为 Agent 配置人设并添加天气插件 2.3 发布并使用第一个 Agent 3 创建工作流 <ol style="list-style-type: none"> 3.1 工作流介绍 3.2 创建一个工作流 3.3 在 Agent 中关联使用工作流 4 第二个 Agent: AI 资讯早知道 <ol style="list-style-type: none"> 4.1 目标与实现思路介绍 4.2 创建相应工作流 4.3 完成 Agent 配置及发布 5 小结 	<ol style="list-style-type: none"> 1. 任务提出 2. 任务实现路径分析 3. 创建 Agent 4. 创建工作流 5. 在 Agent 中关联使用工作流 <p>时间：7.23 (9:00-17:00)</p> <ol style="list-style-type: none"> 1. 结课考核 2. 企业参观 <p>跳转至线下课程安排</p>
--	--	---

<线下青岛班> 人工智能技术与应用实战课程大纲

技能学习课程安排		
<p>时间：报名成功后即可开始学习</p> <p>Python 编程基础</p> <ol style="list-style-type: none"> 1 认识 Python 2 编写 Python 程序 3 认识 Python 数据结构 4 条件判断及分支语句 5 使用 def 定义函数 6 认识面向对象 7 读取文件数据 8 模块和第三方库 <p>Python 数据分析与应用</p> <ol style="list-style-type: none"> 1 Python 数据分析概述 2 NumPy 数值计算基础 3 Matplotlib 数据可视化基础 4 Pandas 统计分析基础 5 使用 Pandas 进行数据预处理 	<p>时间：7.20 (9:00-18:00)</p> <p>PyTorch 框架基础实践</p> <ol style="list-style-type: none"> 1 PyTorch 简介 2 张量操作 <ol style="list-style-type: none"> 2.1 创建张量 2.2 张量与数组相互转化 3 构建一个线性模型 <ol style="list-style-type: none"> 3.1 任务描述构建一个线性模型 3.2 读取数据 3.3 构建初始模型及损失函数 3.4 test-构建优化器 3.4 构建优化器 3.5 最小化方差 (训练) 3.6 执行多轮训练 3.7 训练过程可视化 4 识别手写数字 <ol style="list-style-type: none"> 4.1 案例目标与流程 4.2 加载数据 4.3 加工数据 4.4 模型结构介绍 4.5 构建模型 4.6 模型配置 4.7 模型训练 4.8 执行多轮训练 4.9 模型性能评估 4.10 保存模型 	<p>时间：7.21 (9:00-18:00)</p> <p>PyTorch 深度学习原理与实现</p> <ol style="list-style-type: none"> 1 引言 2 卷积神经网络 CNN <ol style="list-style-type: none"> 2.1 浅层神经网络的局限 2.2 卷积操作 2.3 卷积操作的优势 2.4 池化及全连接 2.5 高维输入及多 filter 卷积 2.6 实现卷积操作 2.7 将卷积结果可视化 2.8 实现池化操作 <p>时间：7.22 (9:00-18:00)</p> <ol style="list-style-type: none"> 3 循环神经网络 RNN <ol style="list-style-type: none"> 3.1 循环神经网络简介 3.2 循环神经网络的常见结构 4 长短时记忆网络 LSTM <ol style="list-style-type: none"> 4.1 LSTM 的三个门 4.2 LSTM 三个门的计算示例 4.3 实现 LSTM 操作 4.4 LSTM 返回值解读 5 利用 LSTM 实现手写数字识别 <ol style="list-style-type: none"> 5.1 加载数据 5.2 数据加工 5.3 搭建循环神经网络 5.4 模型配置
<p>时间：7.19 (9:00-18:00)</p> <p>Python 机器学习实战</p> <ol style="list-style-type: none"> 1 机器学习绪论 <ol style="list-style-type: none"> 1.1 引言 1.2 基本术语 1.3 假设空间&归纳偏好 2 模型评估与选择 <ol style="list-style-type: none"> 2.1 经验误差与过拟合 2.2 评估方法 2.3 性能度量 		

<p>2.4 性能度量 Python 实现</p> <p>3 回归分析</p> <p>3.1 线性回归基本形式</p> <p>3.2 线性回归模型的 Python 实现</p> <p>4 神经网络</p> <p>4.1 单个神经元介绍</p> <p>4.2 经典网络结构介绍</p> <p>4.3 神经网络工作流程演示</p> <p>4.4 如何修正网络参数-梯度下降</p> <p>4.5 网络工作原理推导</p> <p>4.6 网络搭建准备</p> <p>4.7 样本从输入层到隐层传输的 Python 实现</p> <p>4.8 网络输出的 Python 实现</p> <p>4.9 单样本网络训练实现</p> <p>4.10 全样本网络训练实现</p> <p>4.11 网络性能评价</p> <p>4.12 实现神经网络算法</p>	<p>4.11 加载模型</p> <p>4.12 模型应用</p>	<p>5.5 模型训练</p> <p>5.6 模型性能验证</p> <p>特别内容:</p> <p>① ChatGPT 教学应用</p> <p>② 金牌助手 ChatGPT, 让应用开发更高效</p> <p>③ 大模型 (ChatGLM-6B) 产业应用漫步与畅想</p>
--	-----------------------------------	---

案例实战课程安排

<p>时间: 7.24 (8:30-17:00)</p> <p>核心技能串讲与巩固</p> <p>1 PyTorch 简介</p> <p>2 张量操作</p> <p>3 构建一个线性模型</p> <p>4 识别手写数字</p> <p>5 深度学习引言</p> <p>6 卷积神经网络 CNN</p> <p>7 循环神经网络 RNN</p> <p>8 长短时记忆网络 LSTM</p> <p>9 利用 LSTM 实现手写数字识别</p>	<p>时间: 7.26 (8:30-17:00)</p> <p>基于 YOLOv8 的岩石样本分割</p> <p>1 背景与目标</p> <p>2 目标分析</p> <p>3.1 YOLOv5 模型概况</p> <p>3.2 YOLOv8 网络结构</p> <p>3.3 数据增强</p> <p>3.4 Backbone 和 Neck 结构</p> <p>3.5 Head 结构和非极大值抑制</p> <p>4.1 环境准备</p> <p>4.10 模型训练方式说明</p> <p>4.11 模型训练</p> <p>4.12 模型验证</p> <p>4.13 模型预测</p> <p>4.14 含油面积计算</p> <p>4.15 图像分割过程总结</p>	<p>时间: 7.27 (8:30-17:00)</p> <p>实战演练-基于 YOLOv8 的岩石样本识别</p> <p>1.1 目标分析</p> <p>1.2 工程准备</p> <p>1.3 数据格式整理</p> <p>1.4 图像分类模型训练</p> <p>1.5 图像分类模型预测</p>
<p>时间: 7.25 (8:30-17:00)</p> <p>基于大模型微调的命名实体识别</p> <p>1 背景与目标</p> <p>2.1 命名实体识别的概念</p> <p>2.2 命名实体识别的方法</p> <p>3.1 数据格式</p> <p>3.10 NERDataset 结果返回</p> <p>3.11 NERDataset 代码调试</p> <p>3.12 构建最终的训练数据</p> <p>3.2 环境配置</p> <p>3.3 数据格式转换</p> <p>3.4 Label-Studio 数据标注工具使用</p>	<p>4.2 模型和预训练权重下载</p> <p>4.3 lableme 安装</p> <p>4.4 数据标注</p> <p>4.5 认识不同的数据格式</p> <p>4.6 数据格式转化</p> <p>4.7 数据可视化</p> <p>4.8 数据划分</p> <p>4.9 数据配置文件修改</p>	<p>时间: 7.28 (9:00-17:00)</p> <p>1. 职业技术考试</p> <p>2. 企业参观</p> <p>跳转至线下班课程安排</p>

3.5 Label-Studio 数据格式转换 3.6 Dataset 数据格式说明 3.7 __init__ 函数配置 3.8 __get_item__ 函数 3.9 分词器 4.1 模型介绍 4.2 模型下载使用 4.3 Pytoch Lightning 框架 4.4 模型训练 4.5 指标评价 4.6 模型验证 5 小结		
---	--	--

<线下北京班> 大模型微调与私有化部署实战课程大纲

技能学习课程安排		
时间：报名成功后即可开始学习 Python 编程基础 1 认识 Python 2 编写 Python 程序 3 认识 Python 数据结构 4 条件判断及分支语句 5 使用 def 定义函数 6 认识面向对象 7 读取文件数据 8 模块和第三方库 Python 数据分析与应用 1 Python 数据分析概述 2 NumPy 数值计算基础 3 Matplotlib 数据可视化基础 4 Pandas 统计分析基础 5 使用 Pandas 进行数据预处理 时间：7.26 (9:00-18:00) Python 机器学习实战 1 机器学习绪论 1.1 引言 1.2 基本术语 1.3 假设空间&归纳偏好 2 模型评估与选择 2.1 经验误差与过拟合 2.2 评估方法 2.3 性能度量 2.4 性能度量 Python 实现	时间：7.27 (9:00-18:00) 大模型原理与实现 1. 大模型简介 2.1 Transformer 引入 2.2 Transformer 训练阶段 2.3 Transformer 推理阶段 2.4 Input Embedding 操作 2.5 注意力机制介绍 2.6 attention 层的计算过程 01 2.7 attention 层的计算过程 02 2.8 从单头到多头注意力 2.9 Encoder 整体计算流程 2.10 encoder 代码实现 2.11 Decoder 整体计算过程 2.12 Masked Attention 及 Decoder 输出 2.13 decoder 代码实现 2.14 Transformer 代码实现 2.15 Transformer 的并行计算能力 时间：7.28 (9:00-18:00) 3 生成式预训练模型 GPT 4 双向编码模型 BERT 5 Transformer 应用 6 小结	拓展内容： PyTorch 框架基础实践 1 PyTorch 简介 2 张量操作 2.1 创建张量 2.2 张量与数组相互转化 3 构建一个线性模型 3.1 任务描述：构建线性模型 3.2 读取数据 3.3 构建初始模型及损失函数 3.4 test-构建优化器 3.4 构建优化器 3.5 最小化方差（训练） 3.6 执行多轮训练 3.7 训练过程可视化 4 识别手写数字 4.1 案例目标与流程 4.2 加载数据 4.3 加工数据 4.4 模型结构介绍 4.5 构建模型 4.6 模型配置 4.7 模型训练 4.8 执行多轮训练 4.9 模型性能评估 4.10 保存模型 4.11 加载模型

<p>3 回归分析</p> <p>3.1 线性回归基本形式</p> <p>3.2 线性回归模型的 Python 实现</p> <p>4 神经网络</p> <p>4.1 单个神经元介绍</p> <p>4.2 经典网络结构介绍</p> <p>4.3 神经网络工作流程演示</p> <p>4.4 如何修正网络参数-梯度下降</p> <p>4.5 网络工作原理推导</p> <p>4.6 网络搭建准备</p> <p>4.7 样本从输入层到隐层传输的 Python 实现</p> <p>4.8 网络输出的 Python 实现</p> <p>4.9 单样本网络训练实现</p> <p>4.10 全样本网络训练实现</p> <p>4.11 网络性能评价</p> <p>4.12 实现神经网络算法</p>		<p>4.12 模型应用</p> <p>PyTorch 深度学习原理与实现</p> <p>1 引言</p> <p>2 循环神经网络 RNN</p> <p>3 长短时记忆网络 LSTM</p>
--	--	--

案例实战课程安排

<p>时间：7.30 (8:30-17:00)</p> <p>大模型与 AIGC 概述</p> <p>1 大模型是什么</p> <p>1.1 大模型的定义</p> <p>1.2 大模型的特性与应用</p> <p>2 大模型的前世今生</p> <p>3 大模型爆火的原因</p> <p>4 大模型原理介绍</p> <p>4.1 ChatGPT 工作机制与训练流程</p> <p>4.2 自监督预训练解读</p> <p>4.3 有监督微调介绍</p> <p>4.4 奖励建模与强化学习</p> <p>4.5 token 是什么</p> <p>5 AIGC 简介</p> <p>核心技能串讲与巩固</p> <p>1 机器学习模型构建</p> <p>2 PyTorch 框架基础实践</p> <p>3 PyTorch 深度学习原理与实现</p> <p>4 万模基座 Transformer</p>	<p>时间：7.31 (8:30-17:00)</p> <p>大模型微调技术应用-基于大模型微调的新闻文本分类</p> <p>1. 大模型微调的概念</p> <p>2.1 大模型微调方法</p> <p>2.2 硬件配置要求说明</p> <p>2.3 环境准备</p> <p>2.4 模型准备</p> <p>2.5 数据准备</p> <p>3.1 全量微调 FFT-原理介绍</p> <p>3.2 全量微调 FFT-配置训练参数</p> <p>3.3 全量微调 FFT-构建训练器</p> <p>3.4 全量微调 FFT-模型训练</p> <p>3.5 全量微调 FFT-模型效果验证</p> <p>4.1 参数高效微调 PEFT</p> <p>4.2 Prompt Tuning 微调原理</p> <p>4.3 Prompt Tuning 实现</p> <p>4.4 Prompt Tuning 结果分析</p> <p>4.5 Prompt Tuning 模型验证</p> <p>4.6 P-Tuning 微调原理</p> <p>4.7 P-Tuning 实现</p> <p>4.8 LoRA 微调原理</p> <p>4.9 LoRA 实现</p> <p>时间：8.1 (8:30-11:30)</p>	<p>时间：8.1 (14:00-17:00)</p> <p>大模型部署</p> <p>1 基于 Gradio 的页面部署</p> <p>1.1 Gradio 库介绍与安装指南</p> <p>1.2 使用 Gradio 构建 Web 界面</p> <p>1.3 整合后端模型与前端界面</p> <p>1.4 广告创意生成的页面部署</p> <p>时间：8.2 (8:30-17:00)</p> <p>实战演练:基于大模型的校内问答系统</p> <p>1 制作学校相关语料库</p> <p>2 利用文心千帆平台开发问答系统</p> <p>3 部署开源的 ChatGLM 模型</p> <p>4 利用语料库进行模型微调</p> <p>5 基于 Gradio 的前端开发</p> <p>时间：8.3 (9:00-17:00)</p> <p>1. 结课考核</p> <p>2. 企业参观</p> <p>跳转至线下班课程安排</p>
---	--	--

	基于大模型微调的命名实体识别 1 背景与目标 2 目标分析 3 模型准备 3.1 环境准备 3.2 权重下载 4 数据准备 4.1 数据格式说明 4.2 数据格式转化 4.3 构建 Dataset 5 模型微调训练 5.1 构建模型训练函数 5.2 构建模型测试函数 5.3 模型训练 5.4 模型测试 6 小结	
--	--	--

附件二 线上班课程大纲

专题一 AIGC 技术与应用实战课程大纲

技能学习课程安排		
时间：7.20 (9:00-18:00) 大模型与 AIGC 概述 1 大模型是什么 1.1 大模型的定义 1.2 大模型的特性与应用 2 大模型的前世今生 3 大模型爆火的原因 4 大模型原理介绍 4.1 ChatGPT 工作机制与训练流程 4.2 自监督预训练解读 4.3 有监督微调介绍 4.4 奖励建模与强化学习 4.5 token 是什么 5 AIGC 简介 提示工程 1 提示工程简介 1.1 提示 (Prompt) 是什么 1.2 提示的发展历程 1.3 提示工程是什么 1.4 提示词编写原则与策略介绍	时间：7.22 (9:00-18:00) 大模型其它教育应用 1 学术版 ChatGPT 应用指南 1.1 ChatGPT 学术版简介 1.2 ChatGPT 学术版的安装 1.3 ChatGPT 学术版的基础使用 1.4 ChatGPT 学术版的高级功能 1.5 ChatGPT 学术版在学术研究中的应用 2 大模型课堂教学应用 2.1 利用大语言模型对老师的课堂内容总结 2.2 利用大语言模型对老师的课堂内容问答 2.3 利用大语言模型随机出题与批改作业 3 人工智能行业法规 时间：7.23 (9:00-18:00) AIGC 生产力提升 1 学习准备与声明	时间：7.25 (9:00-18:00) 基于大模型的图像生成艺术 1 大模型与计算机视觉简介 1.1 计算机视觉的定义与应用 1.2 大模型在计算机视觉中的应用 2 利用大模型对图像进行描述 2.1 图像标注与描述技术概览 2.2 图像描述的实际操作 2.3 图像描述的应用案例 3 利用大模型生成图像 3.1 文本到图像的转换 3.2 大语言模型在图像生成中的实际应用 3.3 面临的挑战与未来发展 4 利用大模型绘制数据图 4.1 利用大模型绘制柱状图、折线图 4.2 利用大模型绘制饼图、雷达图 4.3 利用大模型绘制散点图、漏斗图、思维导图

<p>2 提示编写原则与策略</p> <p>2.1 原则与策略 1: 清晰的提示</p> <p>2.2 原则与策略 2: 参考示例</p> <p>2.3 原则与策略 3: 让模型一步步思考</p> <p>2.4 原则与策略 4: 调用工具</p> <p>2.5 原则与策略 5: 将复杂任务分解成子任务</p> <p>2.6 原则与策略 6: 采用系统的提示框架</p> <p>2.7 原则与策略 7: 用结构化方式进行提示</p> <p>2.8 自动生成 Prompt</p> <p>3 总结</p> <p>时间: 7.21 (9:00-18:00)</p> <p>AIGC 的教学应用</p> <p>1 学习准备与声明</p> <p>2 大模型辅助教案撰写</p> <p>2.1 编写一份教案模板</p> <p>2.2 撰写具体教案内容</p> <p>3 题库题目生成</p> <p>4 大模型辅助文献阅读</p> <p>5 大模型辅助编程</p> <p>6 大模型辅助数据分析</p> <p>7 总结</p>	<p>2 大模型辅助生成 PPT</p> <p>2.1.1 一步到位生成自我介绍 PPT</p> <p>2.1.2 基于文档制作 PPT</p> <p>2.2.1 提炼并梳理文档中的核心内容</p> <p>2.2.2 将 markdown 内容转为 PPT</p> <p>3 大模型辅助生成思维导图</p> <p>3.1 利用大模型生成思维导图</p> <p>3.2 将已有内容提炼成思维导图</p> <p>4 大模型辅助绘制流程图</p> <p>时间: 7.24 (9:00-18:00)</p> <p>大模型辅助金融数据分析</p> <p>1 大模型+金融数据分析介绍</p> <p>2 智谱清言大模型介绍</p> <p>3 应用大模型进行金融数据分析</p> <p>3.1 金融数据清洗和数据处理</p> <p>3.1.1 数据排重</p> <p>3.1.2 缺失数据处理</p> <p>3.1.3 数据格式检查</p> <p>3.1.4 异常值处理</p> <p>3.2 股票数据分析</p> <p>3.2.1 股票的收益率计算</p> <p>3.2.2 绘制 K 线图和成交量示意图</p> <p>3.2.3 绘制投资组合 K 线图和成交量示意图</p> <p>3.3 量化投资</p> <p>3.3.1 绘制股票价格走势</p> <p>3.3.2 双均线投资策略的制定</p> <p>3.3.3 收益情况分析</p> <p>4 小结</p>	<p>时间: 7.26 (9:00-18:00)</p> <p>5. 开源图像小模型</p> <p>5.1 魔搭与 PaddleHUB</p> <p>5.2 图像小模型</p> <p>5.3 图片文字识别</p> <p>5.4 图像去水印</p> <p>5.5 图像清晰化</p> <p>5.6 色彩增强</p> <p>5.7 风格迁移</p> <p>6 使用开源模型进行艺术绘画</p> <p>6.1 Stable Diffusion-webUI</p> <p>6.2 其它开源绘画模型</p> <p>7. 利用大模型生成视频</p> <p>7.1 基于文心一言一键流影</p> <p>7.2</p> <p>stable-video-diffusion-webui</p> <p>7.3 魔搭平台的视频生成模型</p>
案例实战课程安排		
<p>时间: 7.27 (9:00-18:00)</p> <p>大模型赋能科研</p> <p>1 抛出目标-目标路径-实际场景</p> <p>2 确定发明专利的关键点</p> <p>3 完成发明专利的基本内容</p> <p>4 完成发明专利的详细阐述</p> <p>5 赋能科研-科研小助手</p> <p>6 赋能科研-kimi</p>	<p>时间: 7.29 (9:00-18:00)</p> <p>Agent 应用: 构建你的私人助理</p> <p>1 Agent 是什么</p> <p>1.1 引言: 大模型与通用人工智能的差距</p> <p>1.2 Agent 的概念和示例</p> <p>2 第一个 Agent: 天气播报员</p> <p>2.1 利用扣子平台创建一个 Bot</p>	<p>时间: 7.30 (9:00-18:00)</p> <p>结课考核</p> <p>跳转至线上课程安排</p> <p>拓展内容:</p> <p>基于大模型的小红书文案写作</p> <p>1 小红书文案介绍</p> <p>2 基于大模型撰写小红书文案并整</p>

<p>时间：7.28 (9:00-18:00)</p> <p>RAG 应用: 构建一个答疑小助手</p> <ol style="list-style-type: none"> 1 目标与分析 2 RAG 介绍及其工作流程 3 任务实现 <ol style="list-style-type: none"> 3.1 RAG 实现 1: 将文档直接上传至大模型 3.2 RAG 实现 2: 创建一个 Bot <ol style="list-style-type: none"> 3.2.1 设定 Bot 角色及回复逻辑 3.2.2 上传知识库并调试 Bot 3.2.3 发布并使用 Bot 4 小结 	<p>(Agent)</p> <ol style="list-style-type: none"> 2.2 为 Agent 配置人设并添加天气插件 2.3 发布并使用第一个 Agent 3 创建 workflow <ol style="list-style-type: none"> 3.1 workflow 介绍 3.2 创建一个 workflow 3.3 在 Agent 中关联使用 workflow 4 第二个 Agent: AI 资讯早知道 <ol style="list-style-type: none"> 4.1 目标与实现思路介绍 4.2 创建相应 workflow 4.3 完成 Agent 配置及发布 5 小结 	<p>理文案排版</p> <ol style="list-style-type: none"> 3 基于大模型设置文案标题 4 基于大模型增加文案内容 5 基于大模型润色整体文案并做出总结 6 实训 小红书文案写作
--	---	---

专题二 大模型应用开发实战 (LangChain) 课程大纲

技能学习课程安排		
<p>时间：报名成功后即可开始学习</p> <p>Python 编程基础</p> <ol style="list-style-type: none"> 1 认识 Python 2 编写 Python 程序 3 认识 Python 数据结构 4 条件判断及分支语句 5 使用 def 定义函数 6 认识面向对象 7 读取文件数据 8 模块和第三方库 <p>Python 数据分析与应用</p> <ol style="list-style-type: none"> 1 Python 数据分析概述 2 NumPy 数值计算基础 3 Matplotlib 数据可视化基础 4 Pandas 统计分析基础 5 使用 Pandas 进行数据预处理 	<p>时间：7.21 (9:00-18:00)</p> <ol style="list-style-type: none"> 2 提示编写原则与策略 <ol style="list-style-type: none"> 2.1 原则与策略 1: 编写清晰的提示 2.2 原则与策略 2: 参考示例 2.3 原则与策略 3: 让模型一步步思考 2.4 原则与策略 4: 调用工具 2.5 原则与策略 5: 将复杂任务分解成子任务 2.6 原则与策略 6: 采用系统的提示框架 2.7 原则与策略 7: 用结构化方式进行提示 2.8 自动生成 Prompt 3 总结 	<p>时间：7.24 (9:00-18:00)</p> <p>LangChain 实战-RAG 案例应用</p> <ol style="list-style-type: none"> 4.1 RAG 的概念 4.2 文档加载 4.3 文档分块 4.4 向量化存储 4.5 检索器 4.6 函数式链构造 RAG 4.7 记忆机制 4.8 为 RAG 添加记忆功能 4.9 使用 streamlit 开发界面应用
<p>时间：7.20 (9:00-18:00)</p> <p>大模型与 AIGC 概述</p> <ol style="list-style-type: none"> 1 大模型是什么 <ol style="list-style-type: none"> 1.1 大模型的定义 1.2 大模型的特性与应用 2 大模型的前世今生 3 大模型爆火的原因 4 大模型原理介绍 <ol style="list-style-type: none"> 4.1 ChatGPT 工作机制与训练流程 4.2 自监督预训练解读 	<p>时间：7.22 (9:00-18:00)</p> <p>AIGC 的教学应用</p> <ol style="list-style-type: none"> 1 学习准备与声明 2 大模型辅助教案撰写 <ol style="list-style-type: none"> 2.1 编写一份教案模板 2.2 撰写具体教案内容 3 题库题目生成 4 大模型辅助文献阅读 5 大模型辅助编程 6 大模型辅助数据分析 7 总结 	<p>时间：7.25 (9:00-18:00)</p> <p>LangChain 实战-Agent 应用</p> <ol style="list-style-type: none"> 5.1 Agent 的概念 5.2 load_tools 工具导入 5.3 工具类函数调用 5.4 Tool 工具类 5.5 Agent 其他组件 5.6 initialize_agent 创建 Agent 5.7 Toolkits 创建 Agent 5.8 界面化开发 Agent

<p>4.3 有监督微调介绍</p> <p>4.4 奖励建模与强化学习</p> <p>4.5 token 是什么</p> <p>5 AIGC 简介</p> <p>提示工程</p> <p>1 提示工程 (Prompt Engineering) 简介</p> <p>1.1 提示 (Prompt) 是什么</p> <p>1.2 提示的发展历程</p> <p>1.3 提示工程 (Prompt Engineering) 是什么</p> <p>1.4 提示词编写原则与策略介绍</p>	<p>时间：7.23 (9:00-18:00)</p> <p>LangChain 实战-基础使用</p> <p>1.1 LangChain 简介</p> <p>1.2 LangChain 环境安装</p> <p>2.1 申请阿里云的 API KEY</p> <p>2.2 LangChain 示例</p> <p>2.3 链的使用</p> <p>3.1 提示模板</p> <p>3.2 模型调用</p> <p>3.3 输出解析</p> <p>3.4 基于 LangChain 开发情感极性分析应用</p> <p>3.5 课堂练习：基于 LangChian 开发命名实体识别应用</p>	
---	---	--

案例实战课程安排

<p>时间：7.26 (9:00-18:00)</p> <p>大模型赋能科研</p> <p>1 抛出目标-目标路径-实际场景</p> <p>2 确定发明专利的关键点</p> <p>3 完成发明专利的基本内容</p> <p>4 完成发明专利的详细阐述</p> <p>5 赋能科研-科研小助手</p> <p>6 赋能科研-kimi</p>	<p>时间：7.28 (9:00-18:00)</p> <p>Agent 实战-在线教育课程订单智能助手</p> <p>1 背景与目标</p> <p>2 目标分析</p> <p>3.1 模型准备</p> <p>3.2 创建表格智能体</p> <p>3.3 工具准备</p> <p>3.4 Agent 创建和使用</p> <p>4 思维拓展：RAG 结合 Agent</p> <p>5 小结</p>	<p>时间：7.29 (16:00-18:00)</p> <p>结课考核</p> <p>跳转至线上班课程安排</p>
<p>时间：7.27 (9:00-18:00)</p> <p>RAG 实战-论文阅读助手</p> <p>1 背景与目标</p> <p>2.1 模型准备</p> <p>2.2 文档读取和预处理</p> <p>2.3 文档拆分</p> <p>2.4 向量存储</p> <p>2.5 构建阅读问答链</p> <p>2.6 阅读应用</p> <p>2.7 RAG 增强技术</p> <p>3 阅读机器人界面开发</p> <p>4 小结</p>	<p>拓展内容：</p> <p>AIGC 与图像生成</p> <p>1.1 AI 绘画简介</p> <p>1.2 绘图提示词原则与框架</p> <p>1.3 绘图工具介绍</p> <p>2.1 绘图流程</p> <p>2.2 文生图与图生图</p> <p>2.3 图像连续创作</p>	

专题三 大模型微调应用实战课程大纲

技能学习课程安排		
<p>时间：报名成功后即可开始学习</p> <p>Python 编程基础</p> <p>1 认识 Python</p>	<p>时间：7.22 (9:00-18:00)</p> <p>PyTorch 框架基础实践</p> <p>1 PyTorch 简介</p>	<p>时间：7.25 (9:00-18:00)</p> <p>大模型微调技术应用</p> <p>1. 大模型微调的概念</p>

2 编写 Python 程序	2 张量操作	2.1 大模型微调方法
3 认识 Python 数据结构	2.1 创建张量	2.2 硬件配置要求说明
4 条件判断及分支语句	2.2 张量与数组相互转化	2.3 环境准备
5 使用 def 定义函数	3 构建一个线性模型	2.4 模型准备
6 认识面向对象	3.1 任务描述：构建线性模型	2.5 数据准备
7 读取文件数据	3.2 读取数据	3.1 全量微调 FFT-原理介绍
8 模块和第三方库	3.3 构建初始模型及损失函数	3.2 全量微调 FFT-配置训练参数
Python 数据分析与应用	3.4 test-构建优化器	3.3 全量微调 FFT-构建训练器
1 Python 数据分析概述	3.4 构建优化器	3.4 全量微调 FFT-模型训练
2 NumPy 数值计算基础	3.5 最小化方差（训练）	3.5 全量微调 FFT-模型效果验证
3 Matplotlib 数据可视化基础	3.6 执行多轮训练	4.1 参数高效微调 PEFT
4 Pandas 统计分析基础	3.7 训练过程可视化	4.2 Prompt Tuning 微调原理
5 使用 Pandas 进行数据预处理	4 识别手写数字	4.3 Prompt Tuning 实现
时间：7.20 (9:00-18:00)	4.1 案例目标与流程	4.4 Prompt Tuning 结果分析
大模型与 AIGC 概述	4.2 加载数据	4.5 Prompt Tuning 模型验证
1 大模型是什么	4.3 加工数据	4.6 P-Tuning 微调原理
1.1 大模型的定义	4.4 模型结构介绍	4.7 P-Tuning 实现
1.2 大模型的特性与应用	4.5 构建模型	4.8 P-Tuning v2 微调原理
2 大模型的前世今生	4.6 模型配置	4.9 P-Tuning v2 实现
3 大模型爆火的原因	4.7 模型训练	时间：7.26 (9:00-18:00)
4 大模型原理介绍	4.8 执行多轮训练	4.10 LoRA 微调原理
4.1 ChatGPT 工作机制与训练流程	4.9 模型性能评估	4.11 LoRA 实现
4.2 自监督预训练解读	4.10 保存模型	4.12 AdaLoRA 原理
4.3 有监督微调介绍	4.11 加载模型	4.13 AdaLoRA 实现
4.4 奖励建模与强化学习	4.12 模型应用	4.14 IA3 原理与实现
4.5 token 是什么	时间：7.23 (9:00-18:00)	4.15 BitFit 原理与实现
5 AIGC 简介	大模型原理与实现	5.1 如何降低大模型显存占用
Python 机器学习实战	1. 大模型简介	5.2 半精度 F16 与 BF16
1 机器学习绪论	2.1 Transformer 引入	5.3 半精度微调实现
1.1 引言	2.2 Transformer 训练阶段	5.4 INT8 量化原理
1.2 基本术语	2.3 Transformer 推理阶段	5.5 INT8 量化实现
1.3 假设空间&归纳偏好	2.4 Input Embedding 操作	5.6 QLoRA 原理
2 模型评估与选择	2.5 注意力机制介绍	5.7 QLoRA 实现
2.1 经验误差与过拟合	2.6 attention 层的计算过程 01	6 小结
2.2 评估方法	2.7 attention 层的计算过程 02	7.1 拓展 1: modelscope 平台启动
2.3 性能度量	2.8 从单头到多头注意力	7.2 拓展 1: modelscope 微调步骤实现
2.4 性能度量 Python 实现	2.9 Encoder 整体计算流程	8.1 拓展 2: LLaMA-Factory 界面微调工具
3 回归分析	2.10 encoder 代码实现	8.2 拓展 2: LLaMA-Factory 环境配置
3.1 线性回归基本形式	2.11 Decoder 整体计算过程	8.3 拓展 2: LLaMA-Factory 配置和训练
3.2 线性回归模型的 Python 实现	2.12 Masked Attention 及 Decoder 输出	
时间：7.21 (9:00-18:00)	2.13 decoder 代码实现	
	2.14 Transformer 代码实现	

<p>4 神经网络</p> <p>4.1 单个神经元介绍</p> <p>4.2 经典网络结构介绍</p> <p>4.3 神经网络工作流程演示</p> <p>4.4 如何修正网络参数-梯度下降法</p> <p>4.5 网络工作原理推导</p> <p>4.6 网络搭建准备</p> <p>4.7 样本从输入层到隐层传输的Python实现</p> <p>4.8 网络输出的Python实现</p> <p>4.9 单样本网络训练实现</p> <p>4.10 全样本网络训练实现</p> <p>4.11 网络性能评价</p> <p>4.12 实现神经网络算法</p>	<p>2.15 Transformer的并行计算能力</p> <p>时间：7.24 (9:00-18:00)</p> <p>3 生成式预训练模型 GPT</p> <p>4 双向编码模型 BERT</p> <p>5 Transformer 应用</p> <p>6 小结</p>	<p>8.4 拓展2: LLaMA-Factory 其他功能使用</p> <p>拓展内容：</p> <p>PyTorch 深度学习原理与实现</p> <p>1 引言</p> <p>2 循环神经网络 RNN</p> <p>3 长短时记忆网络 LSTM</p>
--	--	--

案例实战课程安排

<p>时间：7.27 (9:00-18:00)</p> <p>基于通义千问 (Qwen) 的创意广告生成</p> <p>1 背景与目标</p> <p>2.1 千问大模型介绍</p> <p>2.2 modelscope 免费实例</p> <p>2.3 创建虚拟环境</p> <p>2.4 模块安装</p> <p>2.5 模型下载和使用</p> <p>2.6 模型文件路径说明</p> <p>3.1 数据格式要求</p> <p>3.2 数据格式转换</p> <p>4.1 LoRA 技术</p> <p>4.2 finetune 代码解读</p> <p>4.3 LoRA 微调和结果</p> <p>4.4 模型验证</p> <p>4.5 模型验证的注意事项</p> <p>5 小结</p> <p>6 拓展：modelscope 环境的使用</p>	<p>时间：7.28 (9:00-18:00)</p> <p>基于大模型微调的命名实体识别</p> <p>1 背景与目标</p> <p>2.1 命名实体识别的概念</p> <p>2.2 命名实体识别的方法</p> <p>3.1 数据格式</p> <p>3.10 NERDataset 结果返回</p> <p>3.11 NERDataset 代码调试</p> <p>3.12 构建最终的训练数据</p> <p>3.2 环境配置</p> <p>3.3 数据格式转换</p> <p>3.4 Label-Studio 数据标注工具使用</p> <p>3.5 Label-Studio 数据格式转换</p> <p>3.6 Dataset 数据格式说明</p> <p>3.7 __init__ 函数配置</p> <p>3.8 __get_item__ 函数</p> <p>3.9 分词器</p> <p>4.1 模型介绍</p> <p>4.2 模型下载使用</p> <p>4.3 Pytoch Lightning 框架</p> <p>4.4 模型训练</p> <p>4.5 指标评价</p> <p>4.6 模型验证</p> <p>5 小结</p>	<p>时间：7.29 (16:00-18:00)</p> <p>基于大模型的金融问句语义相似度计算</p> <p>1 案例背景及挖掘目标</p> <p>2 语义文本相似度</p> <p>3.1 预处理：文件加载</p> <p>3.2 预处理：错别字纠正</p> <p>3.3 预处理：数据类型转化与长度统计</p> <p>3.4 预处理：数据准备</p> <p>3.5 加载预训练模型</p> <p>3.6 定义损失函数</p> <p>3.7 模型训练</p> <p>3.8 模型预测</p> <p>3.9 模型评估</p> <p>4 小结</p> <p>拓展内容：</p> <p>基于 Transformer 的疫情问诊系统自动翻译实现</p> <p>1.1 机器翻译发展历程</p> <p>1.2 任务目标及实现流程</p> <p>2.1 数据预处理</p> <p>3 预训练模型加载</p> <p>4 模型训练</p> <p>5 模型推理</p> <p>时间：7.30 (16:00-18:00)</p>
--	--	--

专题四 大模型私有化部署和应用开发实战课程大纲

课程安排

<p>时间：报名成功后即可开始学习</p> <p>Python 编程基础</p> <ol style="list-style-type: none"> 1 认识 Python 2 编写 Python 程序 3 认识 Python 数据结构 4 条件判断及分支语句 5 使用 def 定义函数 6 认识面向对象 7 读取文件数据 8 模块和第三方库 <p>Python 数据分析与应用</p> <ol style="list-style-type: none"> 1 Python 数据分析概述 2 NumPy 数值计算基础 3 Matplotlib 数据可视化基础 4 Pandas 统计分析基础 5 使用 Pandas 进行数据预处理 	<p>时间：7.22 (9:00-18:00)</p> <p>Qwen 大模型部署与电商客服模型微调</p> <ol style="list-style-type: none"> 1.1 目标分析 1.2 模型部署测试 1.3 电商客服数据准备 1.4 源码微调 1.5 效果检测 2 高效微调方法介绍 <ol style="list-style-type: none"> 2.1 高效微调的概念 2.2 LoRA 微调原理与实现 2.3 QLoRA 微调原理与实现 3 界面化微调 LLaMa-Factory 	<p>时间：7.26 (9:00-18:00)</p> <p>基于 Langchain 部署本地大模型的知识库助手应用开发</p> <ol style="list-style-type: none"> 3.1 RAG 概述 3.2 文档读取 3.3 文档处理和分段 3.4 向量化与向量数据库 3.5 检索器 3.6 构建提示词 3.7 构建问答链 3.8 问答机器人测试 4 基于 Gradio 的知识库助手界面开发 <ol style="list-style-type: none"> 4.1 Gradio 简介 4.2 构建检索问答链 4.3 调用检索问答链 4.4 Web 界面开发 4.5 Web 界面的外部调用 5 小结
<p>时间：7.20 (9:00-18:00)</p> <p>大模型与 AIGC 概述</p> <ol style="list-style-type: none"> 1 大模型是什么 <ol style="list-style-type: none"> 1.1 大模型的定义 1.2 大模型的特性与应用 2 大模型的前世今生 3 大模型爆火的原因 4 大模型原理介绍 <ol style="list-style-type: none"> 4.1 ChatGPT 工作机制与训练流程 4.2 自监督预训练解读 4.3 有监督微调介绍 4.4 奖励建模与强化学习 4.5 token 是什么 5 AIGC 简介 <p>开源社区 Huggingface 和 ModelScope</p> <ol style="list-style-type: none"> 1 Huggingface 社区介绍 <ol style="list-style-type: none"> 1.1 开源模型下载 1.2 模型使用 	<p>时间：7.23 (9:00-18:00)</p> <p>基于 ChatGLM 的校内问答系统</p> <ol style="list-style-type: none"> 1 制作学校相关语料库 <ol style="list-style-type: none"> 1.1 语料库的重要性与构建方法 1.2 收集与整理有监督数据（对话数据） 1.3 收集无监督数据（纯文本数据） 1.4 数据清洗、标注与存储 2 部署开源的 ChatGLM 模型 <ol style="list-style-type: none"> 2.1 ChatGLM 模型与特点分析 2.2 模型部署的环境与硬件要求 2.3 部署与常见问题解决方案 2.4 实战演练：成功部署并测试 ChatGLM 模型 3 利用语料库进行模型微调 <ol style="list-style-type: none"> 3.1 模型微调的概念与意义 3.2 利用文心千帆对大模型进行微调 3.3 本地对 ChatGLM 模型微调 3.4 本地微调过程中的技巧与注意事项 <p>时间：7.24 (9:00-18:00)</p>	<p>时间：7.27 (9:00-18:00)</p> <p>基于 Ollama 和 Dify 的本地大模型的部署和应用开发</p> <ol style="list-style-type: none"> 1 目标分析 <ol style="list-style-type: none"> 1.1 任务分析 1.2 实现方式 2 基于 Ollama 的本地大模型部署 <ol style="list-style-type: none"> 2.1 Ollama 使用简介 2.2 Ollama 环境部署 2.3 使用模型文件部署 Qwen 大模型 2.4 Langchain 对接 Ollama 本地 Qwen 大模型 2.5 基于 GGUF 模型文件的本地部署 2.6 基于 bge-large-zh 向量大模型文件的本地部署 2.7 Langchain 对接 Ollama 本地向量大模型 2.8 基于 Ollama 部署本地大模型的

<p>1.3 数据集使用</p> <p>1.4 Space 空间体验</p> <p>2 ModelScope 社区介绍</p> <p>2.1 开源模型库</p> <p>2.2 模型使用</p> <p>2.3 开源数据集</p> <p>2.4 ModelScope 算力平台</p> <p>2.5 OpenCSG 平台</p> <p>3 大模型部署推理</p> <p>3.1 开源大模型的文件类型</p> <p>3.2 开源大模型部署</p> <p>3.3 使用 Transformer 进行模型推理</p>	<p>3.5 实战演练 1:使用自己的语料库对云端 ChatGLM 进行微调</p> <p>3.6 实战演练 2:使用自己的语料库对本地 ChatGLM 进行微调</p> <p>4 基于 Gradio 的前端开发</p> <p>4.1 Gradio 库介绍与安装指南</p> <p>4.2 使用 Gradio 构建 Web 界面</p> <p>4.3 整合后端模型与前端界面</p> <p>4.4 实战演练: 为校内问答系统开发一个用户友好的前端界面</p> <p>5 总结与展望</p> <p>5.1 回顾课程重点内容</p> <p>5.2 分享学习心得与项目经验</p> <p>5.3 探讨自然语言处理的未来发展趋势</p>	<p>知识库助手应用开发</p> <p>3 Dify 框架理解与安装</p> <p>3.1 开源应用 Dify 简介</p> <p>3.2 Docker 安装</p> <p>3.3 Dify 安装</p> <p>3.4 大语言模型配置——本地私有大型模型</p> <p>3.5 大语言模型配置——公有大型模型 API</p> <p>3.6 简单使用: 创建文章生成器</p>
<p>时间: 7.21 (9:00-18:00)</p> <p>基于开源代码的 Qwen 大模型部署</p> <p>1 Github 与开源 Qwen 大模型项目</p> <p>1.1 Github 介绍</p> <p>1.2 开源项目准备</p> <p>1.3 模型权重下载</p> <p>2 Qwen 大模型开源项目部署使用</p> <p>2.1 大模型推理调用</p> <p>2.2 命令行调用</p> <p>2.3 网页 demo 调用</p> <p>2.4 部署自己的大模型 API</p> <p>2.5 通过 curl 方法调用大模型 API</p> <p>2.6 通过 requests 方法调用大模型 API</p> <p>2.7 量化模型部署和使用</p> <p>2.8 其他部署方式</p> <p>3 其他模型部署</p>	<p>时间: 7.25 (9:00-18:00)</p> <p>基于 Langchain 的本地大模型部署和应用开发</p> <p>1 Langchain 框架介绍与安装</p> <p>1.1 Langchain 框架介绍</p> <p>1.2 框架安装</p> <p>1.3 Langchain 的快速使用</p> <p>2 LangChain 使用实例</p> <p>2.1 基于 Langchain 的公有大模型 API 调用</p> <p>2.2 基于 Langchain 的私有大型模型 API 调用</p> <p>2.3 开发实例: 通过本地大模型进行情感分析</p> <p>2.4 开发实例: 调用本地大模型进行实体提取</p>	<p>时间: 7.28 (9:00-18:00)</p> <p>4 基于 Dify 的 RAG 应用开发</p> <p>4.1 知识库上传和分段</p> <p>4.2 文本向量化与存储</p> <p>4.3 创建 RAG 应用</p> <p>4.4 提示词</p> <p>4.5 关联知识库以及其他配置</p> <p>4.6 Web 站点应用发布</p> <p>4.7 基于 APIs 的应用开发</p> <p>4.8 嵌入式网站机器人开发</p> <p>5 其他功能介绍</p> <p>5.1 工作流 (Workflow)</p> <p>5.2 第三方工具</p> <p>5.3 构建一个简单的 AI Agent</p> <p>5.4 Coze 功能概览</p> <p>6 小结</p>
		<p>时间: 7.29 (16:00-18:00)</p> <p>结课考核</p> <p>跳转至线上上班课程安排</p>

专题五 数据分析与挖掘实战（泰迪杯竞赛方向）课程大纲

技能学习课程安排		
<p>时间: 报名成功后即可开始学习</p> <p>Python 编程基础</p> <p>1 认识 Python</p> <p>2 编写 Python 程序</p> <p>3 认识 Python 数据结构</p> <p>4 条件判断及分支语句</p>	<p>时间: 7.22 (9:00-18:00)</p> <p>5 神经网络</p> <p>5.1 单个神经元介绍</p> <p>5.2 经典网络结构介绍</p> <p>5.3 神经网络工作流程演示</p> <p>5.4 修正网络参数-梯度下降法</p>	<p>时间: 7.27 (9:00-18:00)</p> <p>PyTorch 框架基础实践</p> <p>1 PyTorch 简介</p> <p>2 张量操作</p> <p>2.1 创建张量</p> <p>2.2 张量与数组相互转化</p>

<p>5 使用 def 定义函数</p> <p>6 认识面向对象</p> <p>7 读取文件数据</p> <p>8 模块和第三方库</p> <p>Python 数据分析基础</p> <p>1 Python 数据分析概述</p> <p>2 NumPy 数值计算基础</p> <p>3 Matplotlib 数据可视化基础</p> <p>4 Pandas 统计分析基础</p> <p>时间：7.20 (9:00-18:00)</p> <p>Python 数据分析与应用</p> <p>1 使用 Pandas 进行数据预处理</p> <p>1.1 合并数据</p> <p>1.2 清洗数据</p> <p>1.3 标准化数据</p> <p>1.4 转换数据</p> <p>时间：7.21 (9:00-18:00)</p> <p>Python 机器学习实战</p> <p>1 机器学习绪论</p> <p>1.1 引言</p> <p>1.2 基本术语</p> <p>1.3 假设空间&归纳偏好</p> <p>2 模型评估与选择</p> <p>2.1 经验误差与过拟合</p> <p>2.2 评估方法</p> <p>2.3 性能度量</p> <p>2.4 性能度量 Python 实现</p> <p>3 回归分析</p> <p>3.1 线性回归基本形式</p> <p>3.2 线性回归模型实现</p> <p>4 决策树</p> <p>4.1 从女生相亲到决策树</p> <p>4.2 明天适合打球吗</p> <p>4.3 决策树拆分属性选择</p> <p>4.4 决策树算法家族</p> <p>4.5 泰坦尼克号生还者预测—数据预处理</p> <p>4.6 泰坦尼克号生还者预测—模型构建与预测</p>	<p>5.5 网络工作原理推导</p> <p>5.6 网络搭建准备</p> <p>5.7 样本从输入层到隐层传输的 Python 实现</p> <p>5.8 网络输出的 Python 实现</p> <p>5.9 单样本网络训练实现</p> <p>5.10 全样本网络训练实现</p> <p>5.11 网络性能评价</p> <p>5.12 sklearn 实现神经网络</p> <p>时间：7.24 (9:00-18:00)</p> <p>Python 网络爬虫实战</p> <p>1 Python 爬虫环境与爬虫简介</p> <p>1.1 认识爬虫</p> <p>1.2 认识反爬虫</p> <p>1.3 配置 Python 爬虫环境</p> <p>2 网页前端基础</p> <p>2.1 认识网络信息传输过程</p> <p>2.2 认识 HTTP</p> <p>3 简单静态网页爬取</p> <p>3.1 认识静态网页</p> <p>3.2 实现 HTTP 请求</p> <p>3.3 解析网页</p> <p>3.4 存储数据</p> <p>时间：7.25 (9:00-18:00)</p> <p>4 认识动态网页</p> <p>4.1 认识动态网页</p> <p>4.2 逆向分析爬取动态网页</p> <p>4.3 使用 Selenium 爬取动态网页</p>	<p>3 构建一个线性模型</p> <p>3.1 任务描述：构建线性模型</p> <p>3.2 读取数据</p> <p>3.3 构建初始模型及损失函数</p> <p>3.4 test-构建优化器</p> <p>3.4 构建优化器</p> <p>3.5 最小化方差（训练）</p> <p>3.6 执行多轮训练</p> <p>3.7 训练过程可视化</p> <p>4 识别手写数字</p> <p>4.1 案例目标与流程</p> <p>4.2 加载数据</p> <p>4.3 加工数据</p> <p>4.4 模型结构介绍</p> <p>4.5 构建模型</p> <p>4.6 模型配置</p> <p>4.7 模型训练</p> <p>4.8 执行多轮训练</p> <p>4.9 模型性能评估</p> <p>4.10 保存模型</p> <p>4.11 加载模型</p> <p>4.12 模型应用</p> <p>拓展内容：</p> <p>PyTorch 深度学习原理与实现</p> <p>1 引言</p> <p>2 循环神经网络 RNN</p> <p>3 长短时记忆网络 LSTM</p> <p>特别内容：</p> <p>① ChatGPT 教学应用</p> <p>② 金牌助手 ChatGPT，让应用开发更高效</p> <p>③ 大模型(ChatGLM)产业应用漫步与畅想</p>
案例实战课程安排		
<p>时间：7.23 (9:00-18:00)</p> <p>学生校园卡消费行为分析（2020</p>	<p>时间：7.26 (9:00-18:00)</p> <p>泰迪内推平台招聘信息采集与</p>	<p>时间：7.28 (9:00-18:00)</p> <p>基于 YOLOv8 的岩石样本图像分割</p>

年技能赛 A 题)	分析 (2023 年挑战赛 C 题)	(2021 年挑战赛 B 题)
1 案例背景和目标 2.1 数据预处理 2.2 学生食堂消费记录处理 2.3 三餐分布饼图 2.4 工作日和非工作日分析 3.1 各专业不同性别消费水平分析 3.2 聚类分析及小结	1.1 背景与分析目标 2.1 会员信息表处理 2.2 销售流水表处理 3.1 会员年龄分析 3.2 不同年龄的消费能力 3.3 不同性别的消费情况 3.4 会员和非会员消费情况 3.5 商场会员年消费趋势 3.6 不同月份的消费趋势 3.7 每年每月的消费金额趋势 3.8 不同时刻的消费情况 4.1 用户画像介绍 4.2 会员基本信息标签 4.3 会员消费特征标签 4.4 会员商品偏好标签 4.5 生成用户画像 5.1 会员细分介绍 5.2 KMeans 算法实现会员聚类 5.3 结果分析	1 背景与目标 2 目标分析 3.1 YOLOv5 模型概况 3.2 YOLOv8 网络结构 3.3 数据增强 3.4 Backbone 和 Neck 结构 3.5 Head 结构和非极大值抑制 4.1 环境准备 4.2 模型和预训练权重下载 4.3 lableme 安装 4.4 数据标注 4.5 认识不同的数据格式 4.6 数据格式转化 4.7 数据可视化 时间：7.29 (9:00-18:00) 4.8 数据划分 4.9 数据配置文件修改 4.10 模型训练方式说明 4.11 模型训练 4.12 模型验证 4.13 模型预测 4.14 含油面积计算 4.15 图像分割过程总结 5.1 目标分析 5.2 工程准备 5.3 数据格式整理 5.4 图像分类模型训练 5.5 图像分类模型预测 5.6 图像分类过程总结 6 小结 时间：7.30 (16:00-18:00) 职业技术考试 跳转至线上上班课程安排

专题六 大数据技术应用实战(Hadoop+Spark)课程大纲

技能学习课程安排		
时间：报名成功后即可开始学习 Linux 操作系统基础 1. Linux 概述	时间：7.22 (9:00-18:00) Hive 大数据仓库 1. Hive 简介	时间：7.25 (9:00-18:00) 5. 配置 Spark IDEA 开发环境 5.1 搭建 Spark 开发环境

<ul style="list-style-type: none"> 2. Linux 系统安装 3. Linux 基本命令 4. Linux Vi 编辑器 5. Linux Shell 编程 <p>Java 编程基础</p> <ul style="list-style-type: none"> 1. 基础知识 2. 面向对象 3. 线程及异常处理 <p>Scala 编程基础</p> <ul style="list-style-type: none"> 1. Scala 简介 2. Scala 安装配置 3. Scala 基础语法 4. 函数 5. 面向对象编程 6. 文件读写 	<ul style="list-style-type: none"> 1.1 Hive 简介 2. Hive 安装配置 <ul style="list-style-type: none"> 2.1.1 Hive 安装配置之 MySQL 数据库安装 2.1.2 Hive 安装配置之 Hive 安装 2.2 Hive 实现单词计数 3. Hive 应用 <ul style="list-style-type: none"> 3.1.1 Hive 表定义 3.1.2 创建内部表与外部表 3.1.3 创建静态分区表和动态分区表 3.1.4 创建带数据的表和桶表 3.2 Hive 导入及导出 	<ul style="list-style-type: none"> 6. Spark SQL 应用 <ul style="list-style-type: none"> 6.1 Spark SQL 简介 6.2 Spark SQL 配置 6.3 从结构化数据文件创建 DataFrame 6.4 从外部数据库创建 DataFrame 6.5 从 RDD 创建 DataFrame 6.6 读取 Hive 表数据创建 DataFrame 6.7 读取学生成绩创建 DataFrame 6.8 常见 DataFrame API 操作 1 6.9 常见的 DataFrame 操作 2 6.10 常见的 DataFrame 操作 3 6.11 通过 DataFrame API 计算学生总分和平均分 6.12 保存 DataFrame 数据 6.13 保存学生成绩分析结果到 Hive 6.14 DataSet 基础操作 6.15 统计商品销量
<p style="color: red;">时间：7.20 (9:00-18:00)</p> <p>Hadoop 大数据基础</p>	<p style="color: red;">时间：7.23 (9:00-18:00)</p>	<p style="color: red;">时间：7.24 (9:00-18:00)</p>
<ul style="list-style-type: none"> 1. Hadoop 简介 <ul style="list-style-type: none"> 1.1 大数据介绍 1.2 Hadoop 核心组件 1.3 Hadoop 生态环境 1.4 Hadoop 应用场景 2. Hadoop 集群安装与部署 <ul style="list-style-type: none"> 2.1 安装虚拟机 2.2 安装 Java 2.3 Hadoop 集群部署模式 2.4 配置固定 IP 2.5 SSH 无密码登录 2.6 配置 Hadoop 集群 2.7 Hadoop 集群配置参数 2.8 Hadoop 集群启动与监控 	<ul style="list-style-type: none"> 3.3.1 Select 查询基本用法 1 3.3.2 Select 查询基本用法 2 3.3.3 内置函数应用 3.3.4 关联查询 <p>拓展</p> <ul style="list-style-type: none"> 4. Hive 开发 5. Hive 自定义函数 6. Hive 查询优化 <p>Spark 大数据技术应用</p> <ul style="list-style-type: none"> 1. Spark 入门 <ul style="list-style-type: none"> 1.1 Spark 入门 2. Spark 集群的安装配置 <ul style="list-style-type: none"> 2.1 Spark 安装部署 2.2 Spark 安装部署实战 3. Spark 架构及原理 <ul style="list-style-type: none"> 3.1 Spark 架构 3.2 Spark RDD 及 DAG 相关概念 4. Spark 编程基础 <ul style="list-style-type: none"> 4.1 创建 RDD 4.2 RDD 常用算子之 transformation 算子 (1) 4.3 RDD 常用算子之 transformation 算子 (2) 4.4 RDD 常用算子之键值对 RDD 算子 4.5 RDD 常用算子之 action 算子 	<p>拓展内容：</p> <ul style="list-style-type: none"> 1 Zookeeper 大数据分布式消息 2 HBase 非关系型数据库 3 Flume 数据采集 4 Kafka 消息系统 5 Flink 大数据实时处理 6 商品实时推荐系统 <p>特别内容：</p> <ul style="list-style-type: none"> ① ChatGPT 教学应用 ② 金牌助手 ChatGPT, 让应用开发更高效 ③ 大模型 (ChatGLM-6B) 产业应用漫步与畅想
<p style="color: red;">时间：7.21 (9:00-18:00)</p> <ul style="list-style-type: none"> 3. Hadoop 基础操作 <ul style="list-style-type: none"> 3.1 Hadoop 安全模式 3.2 Hadoop 集群基本信息介绍 3.3 HDFS 常用 Shell 操作 3.4 MapReduce 常用 Shell 操作 3.5 MapReduce 任务管理 3.6 Yarn 资源管理与任务调度 4. MapReduce 开发入门 <ul style="list-style-type: none"> 4.1 MapReduce 框架与设计构思 4.2 MapReduce 开发环境搭建 	<p style="color: red;">时间：7.24 (9:00-18:00)</p> <p>Spark 大数据技术应用</p> <ul style="list-style-type: none"> 1. Spark 入门 <ul style="list-style-type: none"> 1.1 Spark 入门 2. Spark 集群的安装配置 <ul style="list-style-type: none"> 2.1 Spark 安装部署 2.2 Spark 安装部署实战 3. Spark 架构及原理 <ul style="list-style-type: none"> 3.1 Spark 架构 3.2 Spark RDD 及 DAG 相关概念 4. Spark 编程基础 <ul style="list-style-type: none"> 4.1 创建 RDD 4.2 RDD 常用算子之 transformation 算子 (1) 4.3 RDD 常用算子之 transformation 算子 (2) 4.4 RDD 常用算子之键值对 RDD 算子 4.5 RDD 常用算子之 action 算子 	<p>特别内容：</p> <ul style="list-style-type: none"> ① ChatGPT 教学应用 ② 金牌助手 ChatGPT, 让应用开发更高效 ③ 大模型 (ChatGLM-6B) 产业应用漫步与畅想

4.3 MapReduce WordCount 源码分析 4.4 MapReduce API 分析 4.5 MapReduce 统计每天访问次数 4.6 MapReduce 按用户访问次数排序 拓展 5. MapReduce 编程进阶	子 4.6 文件读取与存储 4.7 统计用户停留时间最长的基站	
案例实战课程安排		
时间：7.26 (9:00-18:00) 航空客户价值分析 1. 案例背景与目标 2. 数据存储 3. 数据探索与预处理 4. 特征构造 5. 搭建 Spark 开发环境 6. KMeans 算法实现客户群体聚类划分 7. 客户群体特征分析及营销建议	时间：7.27 (9:00-18:00) 广电大数据用户画像 1. 项目背景与目标分析 1.1 背景介绍 1.2 目标分析 1.3 系统架构介绍 2. 数据存储 2.1 数据介绍 2.2 数据存储 2.3 开发环境搭建 3. 数据探索与清洗	时间：7.28 (9:00-18:00) 4. 用户画像标签计算 4.1 用户画像简介 4.2 消费内容 4.3 消费等级 4.4 产品名称 4.5 入网程度&业务名称 4.6 地区&语言偏好 5. SVM 预测用户是否挽留 5.1 SVM 介绍 5.2 特征构建 5.3 标签构建 5.4 模型构建与评估 时间：7.29 (16:00-18:00) 职业技术考试 跳转至线上班课程安排

专题七 数据采集与处理实战 (Python&八爪鱼) 课程大纲

技能学习课程安排		
时间：报名成功后即可开始学习 Python 编程基础 1 认识 Python 2 编写 Python 程序 3 认识 Python 数据结构 4 条件判断及分支语句 5 使用 def 定义函数 6 认识面向对象 7 读取文件数据 8 模块和第三方库 Python 数据分析基础 1 Python 数据分析概述 2 NumPy 数值计算基础	时间：7.22 (9:00-18:00) 3 使用 Pandas 进行数据预处理 3.1 合并数据 3.2 清洗数据 3.3 标准化数据 3.4 转换数据 时间：7.23 (9:00-18:00) Python 网络爬虫实战 1 Python 爬虫环境与爬虫简介 1.1 认识爬虫 1.2 认识反爬虫 1.3 配置 Python 爬虫环境	时间：7.25 (9:00-18:00) 5 模拟登录 5.1 表单登录方法实现模拟登录 5.2 Cookie 登录方法实现模拟登录 5.3 使用 Selenium 模拟登录 拓展 6 Scrapy 爬虫 6.1 认识 Scrapy 6.2 通过 Scrapy 爬取基本页面信息 6.3 通过 Scrapy 抓取跳转页面数据 7 拓展：终端协议及爬取工具介绍 拓展内容：

<p>时间：7.20 (9:00-18:00)</p> <p>Python 数据分析与应用</p> <p>1.1 掌握绘图基础语法与常用参数</p> <p>1.2 分析特征间关系</p> <p>1.3 分析特征内部数据分布与分散情况</p> <p>时间：7.21 (9:00-18:00)</p> <p>2 Pandas 统计分析基础</p> <p>2.1 Pandas 简介</p> <p>2.2 读写不同数据源的数据</p> <p>2.3 数据框与数据框元素</p> <p>2.4 转换与处理时间序列数据</p> <p>2.5 使用分组聚合进行组内计算</p> <p>2.6 创建透视表与交叉表</p>	<p>2 网页前端基础</p> <p>2.1 认识网络信息传输过程</p> <p>2.2 认识 HTTP</p> <p>3 简单静态网页爬取</p> <p>3.1 认识静态网页</p> <p>3.2 实现 HTTP 请求</p> <p>3.3 解析网页</p> <p>3.4 存储数据</p> <p>时间：7.24 (9:00-18:00)</p> <p>4 认识动态网页</p> <p>4.1 认识动态网页</p> <p>4.2 逆向分析爬取动态网页</p> <p>4.3 使用 Selenium 爬取动态网页</p>	<p>网站图像素材采集实战</p> <p>1 思路介绍</p> <p>2 单个图片文件爬取</p> <p>3 获取一个页面所有图片网址</p> <p>4 保存所有图片</p> <p>5 翻页爬取更多数据</p> <p>6 PDF 文件规律及问题</p> <p>7 PDF 翻页刷新的网址规律</p> <p>8 获取当前页所有图片网址</p> <p>9 翻页刷新爬取所有图片</p> <p>10 图片拼接成 PDF 文件</p> <p>特别内容：</p> <p>① ChatGPT 教学应用</p> <p>② 金牌助手 ChatGPT, 让应用开发更高效</p> <p>③ 大模型 (ChatGLM) 产业应用漫步与畅想</p>
<p>案例实战课程安排</p>		
<p>时间：7.26 (9:00-18:00)</p> <p>泰迪内推平台招聘信息采集与分析 (Python)</p> <p>1 背景与目标</p> <p>2 数据采集</p> <p>2.1 网页结构探索</p> <p>2.2 定位一级页面数据地址</p> <p>2.3 爬取及解析一级页面数据</p> <p>2.4 提取一级页面字段</p> <p>2.5 定位二级页面数据地址</p> <p>2.6 爬取及解析二级页面数据</p> <p>2.7 翻页爬取及数据保存</p> <p>3 数据处理</p> <p>3.1 读取已爬取完成的数据</p> <p>3.2 数据预处理操作</p> <p>4 分析与可视化</p> <p>4.1 招聘岗位对学历要求分析</p> <p>4.2 各行业的大数据招聘需求数量分析</p> <p>4.3 不同类型公司的薪资待遇分析</p> <p>4.4 小结</p>	<p>时间：7.27 (9:00-18:00)</p> <p>新浪财经数据采集 (八爪鱼工具)</p> <p>1 背景与目标</p> <p>2 目标分析</p> <p>3.1 八爪鱼采集器介绍与安装</p> <p>3.2 菜单栏和模板任务</p> <p>4.1 自定义任务主界面介绍</p> <p>4.2 新建自定义任务</p> <p>4.3 自动识别爬虫字段</p> <p>4.4 手动配置爬虫字段</p> <p>4.5 翻页设置</p> <p>4.6 二级页面爬取</p> <p>4.7 进入二级页面的注意事项</p> <p>4.8 数据预览操作</p> <p>4.9 流程图操作</p> <p>4.10 自定义任务保存</p> <p>5.1 数据采集与导出</p> <p>5.2 任务设置和保存</p> <p>6 小结</p>	<p>时间：7.28 (9:00-18:00)</p> <p>汽车质量投诉平台数据采集 (八爪鱼工具)</p> <p>1 背景与分析目标</p> <p>2 数据采集</p> <p>2.1 投诉网页分析和爬虫思路</p> <p>2.2 分析翻页网址构造</p> <p>2.3 自定义翻页网址列表</p> <p>2.4 配置二级页面跳转位置</p> <p>2.5 循环点击进入二级页面</p> <p>2.6 配置二级页面的爬虫字段</p> <p>2.7 修改字段名称</p> <p>2.8 任务保存和开始采集</p> <p>2.9 回顾：整体流程分析</p> <p>3 数据和任务导出</p> <p>3.1 查看爬虫任务明细</p> <p>3.2 数据导出保存</p> <p>4 总结</p> <p>时间：7.29 (9:00-18:00)</p> <p>微博疫情评论数据采集 (Python)</p> <p>1 背景与目标</p> <p>2.1 评论结构分析</p> <p>2.2 数据接口分析</p> <p>3.1 微博页面接口分析</p> <p>3.2 评论数据接口分析</p>

		3.3 评论回复数据爬取 3.4 单页微博及评论数据爬取 3.5 多线程爬虫 4 小结 时间：7.30（16:00-18:00） 职业技术考试 跳转至线上上班课程安排
--	--	---

专题八 商务数据分析实战（Python）课程大纲

时间：报名成功后即可开始学习 Python 编程基础 1 认识 Python 2 编写 Python 程序 3 认识 Python 数据结构 4 条件判断及分支语句 5 使用 def 定义函数 6 认识面向对象 7 读取文件数据 8 模块和第三方库 Python 数据分析基础 1 Python 数据分析概述 2 NumPy 数值计算基础 时间：7.20（9:00-18:00） Python 数据分析与应用 1 Matplotlib 数据可视化基础 1.1 掌握绘图基础语法与常用 1.2 分析特征间关系 1.3 分析特征内部数据分布与分散情况	时间：7.21（9:00-18:00） 2 Pandas 统计分析基础 2.1 Pandas 简介 2.2 读写不同数据源的数据 2.3 数据框与数据框元素 2.4 转换与处理时间序列数据 2.5 使用分组聚合进行组内计算 2.6 创建透视表与交叉表 时间：7.22（9:00-18:00） 3 使用 Pandas 进行数据预处理 3.1 合并数据 3.2 清洗数据 3.3 标准化数据 3.4 转换数据	时间：7.23（9:00-18:00） Python 数据分析实训 1 iris 数据处理实训 1.1 拓展学习资料&Python 环境介绍 1.2 读取数据&修改列名称 1.3 以 PythonConsole 方式执行代码 1.4 缺失值处理 1.5 重置索引 2 探索 chipotle 数据 2.1 数据读取及介绍 2.2 分组聚合 2.3 数据类型转化 2.4 求客单价 3 探索 Apple 公司股价数据 3.1 数据读取及介绍 3.2 找到最后一个交易日 3.3 日期探索及可视化 特别内容： ① ChatGPT 教学应用 ② 金牌助手 ChatGPT，让应用开发更高效 ③ 大模型 (ChatGLM) 产业应用漫步与畅想
--	---	--

案例实战 课程安排

时间：7.24（9:00-18:00） 商务数据分析概述 1. 商务数据分析基本概念 2. 商务数据分析应用场景 3. 商务数据分析主要工作及流程 4. 商务数据分析方法与工具 5. 商务数据分析岗位与行业介绍	时间：7.26（9:00-18:00） 聚类分析-对某航空公司客户分群 1 背景与目标 2 数据预处理 2.1 数据读取 2.2 剔除票价为空的记录	时间：7.28（9:00-18:00） 关联规则-购物篮商品推荐 1 背景与目标 2 目标分析 3 数据探索与处理 3.1 数据读取及筛选 3.2 将数据处理成购物篮形式并
---	---	--

<p>牛刀小试-T公司运营数据分析与业务决策</p> <p>1 背景与目标</p> <p>2 目标分析</p> <p>3 数据探索与处理</p> <p>3.1 数据探索与处理思路</p> <p>3.2 读取项目数据&处理业务标签</p> <p>3.3 日报表读取及上下半年数据拆分</p> <p>3.4 读取人员成本数据并与日报表拼接</p> <p>3.5 三表合并及异常值处理</p> <p>4 数据分析</p> <p>4.1 计算各类业务的核心指标</p> <p>4.2 各类业务指标分析及可视化</p> <p>4.3 各项目核心指标统计及分析</p> <p>4.4 各部门的核心指标计算分析</p> <p>5 分析结果与建议</p> <p>时间：7.25 (9:00-18:00)</p> <p>回归分析-某市财政收入预测</p> <p>1 背景与目标</p> <p>1.1 案例背景</p> <p>1.2 数据介绍</p> <p>2 目标分析</p> <p>3 线性回归分析</p> <p>3.1 线性回归分析介绍</p> <p>3.2 构建线性回归模型</p> <p>3.3 利用构建好的模型进行预测</p> <p>3.4 数据读取与拆分</p> <p>3.5 模型训练与预测</p> <p>4 分析结果与评估</p> <p>4.1 模型性能评估</p> <p>4.2 将预测结果可视化</p> <p>4.3 G市未来两年财政收入预测</p>	<p>2.3 剔除异常记录</p> <p>3 特征构造</p> <p>3.1 RFM模型介绍</p> <p>3.2 LRFMC模型</p> <p>3.3 构造入会时长特征</p> <p>3.4 剩余特征构造</p> <p>4 K-Means 客户分群</p> <p>4.1 使用 K-Means 算法进行客户分群</p> <p>4.2 获取 K-Means 聚类结果</p> <p>4.3 聚类结果可视化</p> <p>5 小结</p> <p>时间：7.27 (9:00-18:00)</p> <p>分类分析-运营用户流失判别</p> <p>1.1 背景与目标</p> <p>1.2 案例思路分析</p> <p>2.1 数据探索</p> <p>2.2 数据去重及删除无关属性</p> <p>2.3 用户分组及标签构建</p> <p>2.4 提取用户基本信息和在网时长</p> <p>2.5 处理合约是否有效</p> <p>2.6 处理合约计划到期时间</p> <p>2.7 其余变量处理</p> <p>2.8 特征拼接及缺失值处理</p> <p>2.9 数据保存</p> <p>3.1 特征选择介绍</p> <p>3.2 皮尔逊特征选择</p> <p>3.3 处理样本类别不均衡问题</p> <p>4.1 模型性能评估介绍</p> <p>4.2 模型构建及性能评估</p>	<p>保存</p> <p>4 关联规则分析</p> <p>4.1 关联规则介绍</p> <p>4.2 Apriori 算法流程</p> <p>4.3 提升度的概念</p> <p>5 关联规则模型构建</p> <p>5.1 数据加载及拆分</p> <p>5.2 挖掘强关联规则 (Apriori 算法实现)</p> <p>6 模型性能评估</p> <p>时间：7.29 (9:00-18:00)</p> <p>协同过滤-泰迪内推平台信息精准推荐应用</p> <p>1 背景与目标</p> <p>2 目标分析</p> <p>2.1 推荐思路分析</p> <p>2.2 基于物品的协同过滤推荐介绍</p> <p>3 工程实现</p> <p>3.1 EB 工具登录及简介</p> <p>3.2 创建空白工程</p> <p>3.3 导入数据</p> <p>3.4 筛选正文数据</p> <p>3.5 字符替换及记录去重</p> <p>3.6 划分训练集用户和测试集用户</p> <p>3.7 构造训练集和测试集数据</p> <p>3.8 构建模型</p> <p>3.9 推荐及性能评估</p> <p>4 小结</p> <p>时间：7.30 (16:00-18:00)</p> <p>职业技术考试</p> <p>跳转至线上课程安排</p>
---	--	---

专题九 PyTorch 深度学习实战课程大纲

技能学习课程安排		
<p>时间：报名成功后即可开始学习</p> <p>Python 编程基础</p>	<p>时间：7.21 (9:00-18:00)</p> <p>5 神经网络</p>	<p>时间：7.24 (9:00-18:00)</p> <p>PyTorch 深度学习原理与实现</p>

<p>1 认识 Python</p> <p>2 编写 Python 程序</p> <p>3 认识 Python 数据结构</p> <p>4 条件判断及分支语句</p> <p>5 使用 def 定义函数</p> <p>6 认识面向对象</p> <p>7 读取文件数据</p> <p>8 模块和第三方库</p> <p>Python 数据分析与应用</p> <p>1 Python 数据分析概述</p> <p>2 NumPy 数值计算基础</p> <p>3 Matplotlib 数据可视化基础</p> <p>4 Pandas 统计分析基础</p> <p>5 使用 Pandas 进行数据预处理</p> <p>时间：7.20 (9:00-18:00)</p> <p>Python 机器学习实战</p> <p>1 机器学习绪论</p> <p>1.1 引言</p> <p>1.2 基本术语</p> <p>1.3 假设空间&归纳偏好</p> <p>2 模型评估与选择</p> <p>2.1 经验误差与过拟合</p> <p>2.2 评估方法</p> <p>2.3 性能度量</p> <p>2.4 性能度量 Python 实现</p> <p>3 回归分析</p> <p>3.1 线性回归基本形式</p> <p>3.2 线性回归模型的 Python 实现</p> <p>3.3 波士顿房价预测的 Python 实现</p> <p>3.4 逻辑回归介绍</p> <p>3.5 研究生入学录取预测的 Python 实现</p> <p>4 决策树</p> <p>4.1 从女生相亲到决策树</p> <p>4.2 明天适合打球吗</p> <p>4.3 决策树拆分属性选择</p> <p>4.4 决策树算法家族</p> <p>4.5 泰坦尼克号生还者预测—数据预处理</p> <p>4.6 泰坦尼克号生还者预测—模型构建与预测</p>	<p>5.1 单个神经元介绍</p> <p>5.2 经典网络结构介绍</p> <p>5.3 神经网络工作流程演示</p> <p>5.4 如何修正网络参数-梯度下降法</p> <p>5.5 网络工作原理推导</p> <p>5.6 网络搭建准备</p> <p>5.7 样本从输入层到隐层传输的 Python 实现</p> <p>5.8 网络输出的 Python 实现</p> <p>5.9 单样本网络训练的 Python 实现</p> <p>5.10 全样本网络训练的 Python 实现</p> <p>5.11 网络性能评价</p> <p>5.12 调用 sklearn 实现神经网络</p> <p>时间：7.22 (9:00-18:00)</p> <p>PyTorch 框架基础实践</p> <p>1 PyTorch 简介</p> <p>2 张量操作</p> <p>2.1 创建张量</p> <p>2.2 张量与数组相互转化</p> <p>3 构建一个线性模型</p> <p>3.1 任务描述构建一个线性模型</p> <p>3.2 读取数据</p> <p>3.3 构建初始模型及损失函数</p> <p>3.4 test-构建优化器</p> <p>3.4 构建优化器</p> <p>3.5 最小化方差 (训练)</p> <p>3.6 执行多轮训练</p> <p>3.7 训练过程可视化</p> <p>时间：7.23 (9:00-18:00)</p> <p>4 识别手写数字</p> <p>4.1 案例目标与流程</p> <p>4.2 加载数据</p> <p>4.3 加工数据</p> <p>4.4 模型结构介绍</p> <p>4.5 构建模型</p> <p>4.6 模型配置</p> <p>4.7 模型训练</p> <p>4.8 执行多轮训练</p> <p>4.9 模型性能评估</p>	<p>1 引言</p> <p>2 卷积神经网络 CNN</p> <p>2.1 浅层神经网络的局限</p> <p>2.2 卷积操作</p> <p>2.3 卷积操作的优势</p> <p>2.4 池化及全连接</p> <p>2.5 高维输入及多 filter 卷积</p> <p>2.6 实现卷积操作</p> <p>2.7 将卷积结果可视化</p> <p>2.8 实现池化操作</p> <p>时间：7.25 (9:00-18:00)</p> <p>3 循环神经网络 RNN</p> <p>3.1 循环神经网络简介</p> <p>3.2 循环神经网络的常见结构</p> <p>4 长短时记忆网络 LSTM</p> <p>4.1 LSTM 的三个门</p> <p>4.2 LSTM 三个门的计算示例</p> <p>4.3 实现 LSTM 操作</p> <p>4.4 LSTM 返回值解读</p> <p>5 利用 LSTM 实现手写数字识别</p> <p>5.1 加载数据</p> <p>5.2 数据加工</p> <p>5.3 搭建循环神经网络</p> <p>5.4 模型配置</p> <p>5.5 模型训练</p> <p>5.6 模型性能验证</p> <p>特别内容：</p> <p>① ChatGPT 教学应用</p> <p>② 金牌助手 ChatGPT，让应用开发更高效</p> <p>③ 大模型 (ChatGLM) 产业应用漫步与畅想</p>
--	---	---

	4.10 保存模型 4.11 加载模型 4.12 模型应用	
案例实战课程安排		
<p>时间：7.26 (9:00-18:00)</p> <p>脑 PET 图像分析与疾病预测</p> <p>1 背景与目标</p> <p>2 数据处理</p> <p>2.1 图像读取及尺寸调整</p> <p>2.2 图像增强之翻转操作</p> <p>2.3 图像增强之旋转缩放</p> <p>2.4 获取所有照片路径</p> <p>2.5 批量获取照片数据</p> <p>2.6 将数据处理过程封装成函数</p> <p>3 模型构建</p> <p>3.1 定义卷积&池化层</p> <p>3.2 定义全连接层</p> <p>3.3 定义网络计算过程</p> <p>3.4 数据集维度调整及类型转化</p> <p>3.5 数据集分批及打乱操作</p> <p>3.6 模型构建及配置</p> <p>3.7 模型训练</p> <p>4 模型性能评估及应用</p> <p>4.1 模型性能评估及保存</p> <p>4.2 模型应用</p> <p>时间：7.27 (9:00-18:00)</p> <p>基于 textCNN 的公众健康问句分类</p> <p>1 背景与目标</p> <p>1.1 背景与目标</p> <p>1.2 目标分析</p> <p>2 数据探索与处理</p> <p>2.1 数据探索</p> <p>2.2 剔除无效字符及分词操作</p> <p>2.3 读取停用词表</p> <p>2.4 去除停用词</p> <p>2.5 为词语编号</p> <p>3 词嵌入 (Word2Vec)</p> <p>3.1 词嵌入 (Word2Vec) 介绍</p> <p>3.2 获取目标词向量矩阵</p> <p>3.3 保存处理好的数据</p> <p>4 模型构建 (textCNN)</p> <p>4.1 统一各样本的词语数量</p>	<p>时间：7.28 (9:00-18:00)</p> <p>基于大模型微调的命名实体识别</p> <p>1 背景与目标</p> <p>2 目标分析</p> <p>3 模型准备</p> <p>3.1 环境准备</p> <p>3.2 权重下载</p> <p>4 数据准备</p> <p>4.1 数据格式说明</p> <p>4.2 数据格式转化</p> <p>4.3 构建 Dataset</p> <p>5 模型微调训练</p> <p>5.1 构建模型训练函数</p> <p>5.2 构建模型测试函数</p> <p>5.3 模型训练</p> <p>5.4 模型测试</p> <p>6 小结</p>	<p>时间：7.29 (9:00-18:00)</p> <p>基于 YOLOv8 的岩石样本图像分割与识别</p> <p>1 背景与目标</p> <p>2 目标分析</p> <p>3.1 YOLOv5 模型概况</p> <p>3.2 YOLOv8 网络结构</p> <p>3.3 数据增强</p> <p>3.4 Backbone 和 Neck 结构</p> <p>3.5 Head 结构和非极大值抑制</p> <p>4.1 环境准备</p> <p>4.10 模型训练方式说明</p> <p>4.11 模型训练</p> <p>4.12 模型验证</p> <p>4.13 模型预测</p> <p>4.14 含油面积计算</p> <p>4.15 图像分割过程总结</p> <p>4.2 模型和预训练权重下载</p> <p>4.3 lableme 安装</p> <p>4.4 数据标注</p> <p>4.5 认识不同的数据格式</p> <p>4.6 数据格式转化</p> <p>4.7 数据可视化</p> <p>4.8 数据划分</p> <p>4.9 数据配置文件修改</p> <p>5.1 目标分析</p> <p>5.2 工程准备</p> <p>5.3 数据格式整理</p> <p>5.4 图像分类模型训练</p> <p>5.5 图像分类模型预测</p> <p>5.6 图像分类过程总结</p> <p>6 小结</p> <p>时间：7.30 (16:00-18:00)</p> <p>职业技术考试</p> <p>跳转至线上课程安排</p>

4.2 textCNN 介绍		
4.3 构建 textCNN		
4.4 建模前数据准备		
4.5 执行模型训练		
5 模型性能评估		
5.1 模型性能评估		
5.2 小结		

附件二 师资介绍

姜春茂，博士（后），福建理工大学计算机学院教授，硕士生导师。主要研究方向：智能信息处理与决策，多粒度云计算，汉俄自然语言处理等。近年来，主持和参与国家自然科学基金、黑龙江省自然科学基金、国家社科基金等 10 余项。在 Information sciences、Cognitive computation、Knowledge-based Systems、Applied Intelligence、电子学报等重要期刊发表论文 80 余篇，其中 SCI 检索 30 余篇。主持教学改革项目 8 项，包括省级重点，省级一般等。获得黑龙江省自然科学技术学术成果二等奖，黑龙江省高校科技三等奖。

郭森，深圳信息职业技术学院财经学院教授，毕业于华南理工大学计算机科学与工程学院，工学博士。郭森教授长期从事人工智能、金融大数据分析等领域的教学和研究工作，曾 2 次获广东省科学技术奖以及深圳市科技创新奖、深圳市科技进步奖各 1 次；主持并完成省、市级教科研项目 10 余项；主持过多项金融科技方面的国培、省培项目。获得国内外专利授权 10 余项。指导学生在全国大学生数学建模竞赛、全国大数据金融分析与应用竞赛、广东省金融科技技能大赛、广东省软件设计大赛、挑战杯创业计划大赛以及中国机器人大赛等比赛中取得过好成绩。

欧阳通达，中国科学院深圳先进技术研究院合成生物学研究所云实验室研发组长、高级系统架构设计师、后端工程师、华南师范大学工程管理硕士。具有丰富的信息系统开发经验、熟悉 Hadoop、Kafka、Redis、Hbase 和 Elasticsearch 等大数据框架，擅长智能机器人调度、机器学习和计算机视觉，熟练运用 ChatGPT 等大模型进行科研及开发工作。开发项目涉及工业互联网、物联网、自动化和生物多个领域。近两年取得三项发明专利和数十项软件著作权。现参与合成生物研究重大科技基础设施（总投资近 10 亿元）的建设工作。

樊老师，大数据开发工程师，Hortonworks 授权 Apache Hadoop 开发者认证培训讲师，Hadoop、Mahout 技术实践者和研究者；对 Hadoop 的 MapReduce 编程模型有深刻理解，同时

对 Mahout 技术有较深认识，对 Mahout 源码有深入研究，擅于 Mahout 中数据挖掘的 K 均值聚类算法、贝叶斯分类算法、FP 树关联规则算法的应用；主编《Mahout 算法解析与案例实战》、《Hadoop 数据分析与挖掘实战》、《Hadoop 与大数据挖掘》等图书专著；具有电信行业和银行业的项目经验和行业知识，主持中国电科院电力大数据平台、电能量数据挖掘与智能分析、客户服务智能分析系统等项目。

易云恒，西南交通大学希望学院专职教师，广西民族大学工学硕士，广西心境科技有限公司技术总监。具备人工智能开发、数据分析、软件设计及服务器运维经验。主要研究方向为自然语言处理，精通 Python、C# 等编程语言，熟悉 PaddlePaddle 与 PyTorch 等深度学习框架。掌握图像处理基础，发表中文核心论文一篇；熟悉强化学习，同样发表中文核心论文一篇。参与完成一个国家级项目、一个广西省级项目、两个四川省级项目，技术能力显著。指导学生获得全国数字媒体大赛二等奖、三等奖，以及多项省级比赛奖项。

胡博，西南交通大学希望学院专职教师，华东交通大学软件工程硕士，华为开发者学堂认证讲师。具备大数据与统计分析、移动应用开发、服务器集群作业处理等相关技术。主要研究方向为生物学大数据建模与分析，掌握 Python、Java、R 等编程语言，熟悉 Pytorch、TensorFlow、华为 MindSpore 等深度学习框架。掌握医学图像处理与数据分析，发表 EI 论文 1 篇，参与完成国家自然科学基金 1 项、江西省科技厅项目 1 项。指导学生获得全国数字媒体国赛、省赛一等奖各 1 项，全国大学生信息素养国赛二等奖 1 项、省赛共获奖 6 项，蓝桥杯四川省赛共获奖 9 项。

刘名军，广东泰迪智能科技股份有限公司技术副总经理，中山大学计算机科学与技术专业，高级信息系统项目管理师。专注于 hadoop 及相关子系统 (hbase/pig/hive), Cassandra, memcached 等云计算相关产品的研究；具有丰富的计算机软件、数据挖掘算法及应用的经验。曾主持电力行业统计分析决策系统、电力大数据平台、客户停电数据综合分析系统、异常用电智能诊断系统、贵阳民政部数据铁笼项目等多个百万级的项目管理和研发工作，具有丰富的项目管理和研发工作经验，给电力、民政、政府等行业进行数据挖掘应用培训十余场。

张敏，广东泰迪智能科技股份有限公司培训总监，高级信息系统项目管理师。具有丰富的大数据挖掘、人工智能教学和开发经验，曾为南方电网、国家电网、格力电器、珠江数码等多个大型企业提供项目研发与维护服务。参编数据挖掘与人工智能类教材 10 余本，作为主

讲老师参与国内高校和企业关于数据挖掘、人工智能相关培训班百余场。

周东平，广东泰迪智能科技股份有限公司产品总监，高级数据分析师，项目经理。从事数据行业多年，熟悉大数据、人工智能相关项目开发流程；具有丰富的大数据、人工智能产品与应用设计经验，对于大数据、人工智能教学具有深入研究。精通 Python 语言，策划主持编写 Python 方向大数据与人工智能图书 10 余本，包括《Python 数据分析与应用》、《Python3 智能数据分析快速入门》、《大数据数学基础（Python 语言描述）》等。在职期间参与项目涵盖电力、税务、金融、新闻、建筑等方向，具有丰富的行业经历。

陈四德，广东泰迪智能科技股份有限公司高级数据分析师，统计学专业，对数据统计分析 and 数据挖掘领域均有较强的理解和理论基础；有造价行业、游戏行业背景和丰富的项目经验，精通行业内的各种指标分析，擅于从多维度分析数据，逻辑性强；擅长 Python、R 语言、MySQL 数据库等工具，能熟练对数据进行数据处理和分析，掌握常用的数据挖掘算法如分类、聚类等，以及深度学习 TensorFlow 的使用。负责“网站会员流失预测”项目，完成数据处理，模型构建；负责“平台 BI 埋点数据入库及数据分析”项目，完成数据盘点、数据指标整理和把控；负责“游戏数据分析”项目，完成产出游戏生态日报、客户价值分群结果、用户流失的预警、用户画像指标的完善和维护，项目经验丰富。负责过西安交大城市学院、福建农林大学、国培师资培训、韩山师范学院数据分析就业班、湖南科技职业技术学院、武汉科技大学、广东机电职业技术学院国培、柳州城市职业技术学院第一届大数据职业技能竞赛指导、吉林大学珠海学院等培训项目，授课经验丰富。负责过“泰迪杯”数据挖掘挑战赛出题及赛题指导。大数据专业系列图书编写委员会成员，负责《Keras 与深度学习实战》、《Python 中文自然语言处理基础与实战》、《深度学习与计算机视觉实战》等书籍编写工作。

焦正升，广东泰迪智能科技股份有限公司资深项目研发工程师、高级信息系统项目经理、高级软件开发工程师，拥有 7 年相关从业经验；致力于信息技术的应用与传播，信息系统产业的发展。精通 JAVA 编程语言，熟悉 Spring Boot、Spring Cloud 等主流开发框架、MySQL 数据库、VUE 数据驱动渐进式框架等主流技术。参与《Hadoop 与大数据挖掘》、《Hadoop 大数据分析 with 挖掘实战》等图书的编写。拥有电力、电子政务、轻工环保、交通运输等多项领域的项目管理研发经验，项目团队为北京市信访办研发的“大数据助力智慧信访”系统获得第七届金铃奖-公共服务类“智能决策奖”。

郑素铃,从事大数据项目研发工作,对Hadoop大数据技术有较深的研究,熟练掌握Hadoop环境部署和Hadoop核心计算框架MapReduce的原理和应用。掌握Spark原理及编程,熟练使用Spark的图计算Graphx和算法库MLlib。对非结构化数据库HBase以及结构化数据库Hive有深刻的了解。掌握数据挖掘和机器学习的常用算法,熟悉数据挖掘流程,具备项目开发经验,如“数睿思网站用户画像研究”和“法律服务智能推荐系统”项目,在推荐系统方面比较有研究。先后参与了《Hadoop大数据开发基础》、《Spark大数据技术与应用》等图书编写工作。

吴嘉泳,一线大数据研发工程师。对Hadoop + Spark生态体系有深入研究,熟练掌握Hadoop环境部署和Hadoop核心计算框架MapReduce的原理和应用。掌握Spark原理及编程,熟练使用Spark的算法库MLlib。对非结构化数据库HBase以及结构化数据库Hive有深刻的了解。深度参与某电网公司内部客服优化系统开发,主要负责对系统中的海量文本数据进行处理和挖掘,利用Spark + Hive和相应组件实现潜在规律地挖掘。参与编写1+X相关图书编写工作,参与《Hadoop大数据开发基础》、《Spark大数据技术与应用》等图书编写。先后跟进负责第三期全国高校大数据与人工智能双师型骨干教师研修班和部分项目案例资源开发工作。

叶丽凡,广东泰迪智能科技股份有限公司高级大数据开发工程师,对Hadoop大数据技术有深入理解,熟悉HDFS分布式文件系统存储结构,熟练掌握Hadoop环境部署和MapReduce计算框架编程。对HBase、Hive数据库有深刻了解。掌握Spark原理及编程,熟悉Spark底层运行机制,并熟练使用Spark SQL即时查询框架和Spark MLlib算法库。深度参与过华南某广电公司大数据营销推荐系统开发,利用Hadoop + Spark + Hive为其中的400多万用户生成用户画像。参与编写《Hadoop与大数据挖掘》、《Spark大数据技术与应用》等图书。跟进负责全国高校大数据与人工智能双师型骨干教师研修班、韩山师范学院等高校的大数据培训课程。

温鼎,从事大数据研发工作,对Hadoop生态圈相关组件技术有深刻的认识,掌握Hadoop核心框架原理。熟练掌握Hadoop、Hive、Flume等大数据核心组件的环境搭建和使用。同时对Spark核心API core和SQL & DataFrame的使用有深刻的认识,掌握MLlib算法库的应用。掌握数据挖掘和机器学习十大算法的应用。对数据有较强的敏感度,掌握常用机器学习算法

原理。先后参与湖南商务职业技术学院 Hadoop 模块考题开发，深圳职业计算学院 PySpark 大数据课程资源开发，广东技术师范大学 Hadoop 大数据培训。