

关于举办全国数字人才技能提升师资培训班的通知

新一代信息技术正深刻改变着人们的生产、生活、学习和思维方式，其朝气蓬勃的产业发展态势和大力度的国家政策鼓励支持，无不揭示了数据智能技术的重要性，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，加快形成新质生产力，为高质量发展赋能蓄力，培养现代化数字技能人才，广东泰迪智能科技股份有限公司联合华南农业大学珠江学院人工智能学院共同开展全国数字人才技能提升师资研修班，旨在推动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。

该系列培训期限自2024年09月04日至2025年09月01日止。



一、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）

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

广州专题一 [大模型与 AIGC 技术应用实战](#)

基本信息	现场实战，1.15-1.19	拓展课程学习(线上云课堂)
时间： 1月14日-19日	1. 大模型与 AIGC 概述	1. AIGC 生产力提升
线下报到时间：1月14日	2. 提示工程	2. 专利撰写助手 Agent
地点：广州	3. AIGC 的教学应用	3. 个人学习助手 Agent
学时：共计 72 学时	4. AI Agent 基础与实践	4. 多模态 Agent：海报设计大师
证书：高级 AIGC 应用工程师 职业技术证书	6. 企业新人导师 Agent	5. 基于大模型的图像生成艺术
费用：3980 元/人	7. 视频课程讲义整理助手 Agent	
学习环境要求：Windows10 或以上操作系统（64 位），8G+内存	8. 实战演练-搭建一个 AI Agent	
【专题特色：紧贴大模型前沿应用、切实提升工作效率、全程以大模型应用为主、零门槛学习最新技术】		

广州专题二 [数据分析与挖掘实战（泰迪杯竞赛方向）](#)

基本信息	技能学习(线上云课堂)，1.13-1.15	现场案例实战，1.17-1.21
时间： 1月13日-21日	1. Python 编程基础	1. 机器学习核心技能串讲与巩固
线下报到时间：1月16日	2. Python 数据分析与应用	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）
地点：广州	3. Python 网络爬虫实战	3. 深度学习核心技能串讲与巩固
学时：共计 72 学时	4. Python 机器学习实战	4. 基于 YOLOv8 的岩石样本图像分割（2021 年挑战赛 B 题相关）
证书：高级大数据分析师职业 技术证书	5. PyTorch 框架基础实践	5. 2024 年挑战赛真题解读与优秀作品剖析
费用：3980 元/人	6. PyTorch 深度学习原理与实现(拓展)	
学习环境要求：Windows10 或以上操作系统（64 位），8G+内存	7. 特别内容：	
【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 ChatGPT 教学应用	
	② 金牌助手 ChatGPT，让应用开发更高效	
	③ 大模型(ChatGLM2-6B)产业应用漫步与畅想	

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。

线上专题一 [大模型与 AIGC 技术应用实战](#)

基本信息	技能学习	案例实战
时间：1 月 14 日-24 日 学时：共计 88 学时 证书：高级 AIGC 应用工程师 职业技术证书 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、全程以大模型应用为主、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. AIGC 生产力提升 5. AI Agent 基础与实践-认识 Agent 6. AI Agent 基础与实践-扣子智能体平台应用 7. 基于大模型的图像生成艺术（拓展）	1. 视频课程讲义整理助手 Agent 2. 专利撰写助手 Agent 3. 企业新人导师 Agent 4. 海报设计大师 Agent

线上专题二 [大模型应用开发实战（LangChain）](#)

基本信息	技能学习	案例实战
时间：1 月 14 日-24 日 学时：共计 88 学时 证书：高级大模型开发工程师 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：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。】	1. Python 编程基础 2. 大模型与 AIGC 概述 3. 提示工程 4. AIGC 的教学应用 5. LangChain 实战-基础使用 6. LangChain 实战-RAG 应用 7. LangChain 实战-Agent 应用 8. 基于 LangChain 的本地大模型调用	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 4. 基于本地大模型的面试助手开发

线上专题三 [大模型微调应用实战](#)

基本信息	技能学习	案例实战
时间：1 月 14 日-24 日 学时：共计 88 学时 证书：高级大模型应用工程师 职业技术证书 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	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. PyTorch 深度学习原理与实现（拓展） 7. 大模型原理与实现 8. 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）

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】

线上专题四 [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：1月14日-24日 学时：共计88学时 证书：高级大模型开发工程师 职业技术证书 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】	1. Python 编程基础 2. Python 数据分析与应用 3. 大模型与 AIGC 概述 4. 大模型开源平台介绍及使用 5. 部署自己的 Qwen2 大模型 6. 基于 LangChain 的本地大模型部署和应用 7. 基于 Ollama 和 Dify 的本地大模型的部署和应用开发	1. Qwen2 大模型部署与电商客服模型微调 2. 基于 Qwen1 的创意广告生成 3. 基于 LangChain 的人事面试助手开发

线上专题五 [AIGC 多模态技术应用实战](#)

基本信息	技能学习	案例实战
时间：1月14日-24日 学时：共计88学时 证书：高级 AIGC 应用工程师 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：从文本到视觉听觉的创作革命，AI 驱动的视觉与语音艺术】	1. 大模型与 AIGC 概述 2. 提示工程 3. 从文本到视觉的创作革命 4. AI 绘画应用实战-基础实践 5. AI 绘画应用实战-技巧进阶 6. AI 绘画应用实战-应用实战	1. AI 视频生成应用实战 2. AI 语音技术应用实战 3. 虚拟数字人技术应用实战 4. 多模态 Agent

线上专题六 [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：1月14日-23日 学时：共计80学时 证书：高级大数据技术应用工程师 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、	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）	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）

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。】	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
---	---

线上专题七 [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：1月14日-24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）

线上专题八 [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
时间：1月14日-24日 学时：共计88学时 证书：高级 Python 技术应用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。】	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： 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. 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. 微博评论数据采集（Python）

线上专题九 [商务数据分析实战（Python）](#)

基本信息	技能学习	案例实战
时间：1月14日-24日 学时：共计88学时 证书：高级大数据分析师职业技术证书 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。】	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： 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. 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

线上专题十 [计算机视觉应用实战（PyTorch）](#)

基本信息	技能学习	案例实战
时间：1月14日-24日 学时：共计88学时 证书：高级人工智能应用工程师职业技术证书 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：重点专注于计算机视觉领域。理论与实践相结合，深入解析深度学习原理，实战项目教学】	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. 计算机视觉实战 7. 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(ChatGLM2-6B)产业应用漫步与畅想	1. 脑 PET 图像分析和疾病预测 2. 基于 FaceNet 的人脸智能识别 3. 基于 YOLOX 的农田害虫图像检测与识别

线上专题十一 [文本分析与挖掘实战（PyTorch）](#)

基本信息	技能学习	案例实战
时间：1月14日-24日 学时：共计88学时 证书：高级人工智能应用工程师职业技术证书 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：重点专注于自然语言领域。理论与实践相结合，深	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. 自然语言处理实战 7. 特别内容： ① ChatGPT 教学应用 ② 金牌助手 ChatGPT，让应用开发更高效	1. 基于 textCNN 的公众健康问句分类 2. 网络舆情与网民情绪波动分析 3. 推荐系统受众性别智能识别

二、课程特色

1、紧贴前沿技术应用，畅想数据智能新未来

大模型赋能下的 AI Agent 极有可能是 AGI（通用人工智能）的重要实现形式，是 AGI 时代每个人都要掌握的硬核技能，本活动设置相应专题可让零基础学员快速、有效上手这一全新技能，并充分感受和理解新技术新趋势。

2、数字人才赋能产业+落地实操授课

课程全程强调动手实操，以实际操作/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。

3、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。

4、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。

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。

三、证书颁发

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。



四、报名及联系方式

- 1、报名材料：报名申请表、身份证正面人像照、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：***
电 话：*****

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： 联系人：								

附件一 线下班课程大纲

<广州专题一> 大模型与 AIGC 技术应用实战课程大纲

现场案例实战课程安排		
线下报到时间： 1 月 14 日 时间：1.15（8:30-11:30） 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（Prompt Engineering）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（Prompt Engineering）是什么 1.4 提示词编写原则与策略介绍 2 提示编写原则与策略 2.1 原则与策略 1：编写清晰的提示 2.2 原则与策略 2：提供参考示例 2.3 原则与策略 3：让模型一步步思考 2.4 原则与策略 4：调用外部工具 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	时间：1.16（8:30-17:00） AI Agent 基础与实践 1 Agent 是什么 1.1 AI Agent 的概念 1.2 走向 AGI 的 Agent 2 基于大模型的 Agent 3 智谱清言智能体平台 3.1 智谱清言智能体平台介绍 3.2 创建文献阅读助手 Agent 3.3 智能体的能力配置 3.4 智能体的知识库配置 4 创建一个 Prompt 优化工程师 5 扣子智能体平台搭建 6 Agent 实践：知冷知热的桃子姐姐 6.1 目标介绍-创建一个天气播报机器人 6.2 人设与回复逻辑编写 6.3 Bot 发布与调用 7 创建 workflow 7.1 workflow 介绍 7.2 创建 workflow 7.3 调用 workflow 8 Agent 实践：AI 资讯早知道 8.1 创建 workflow 8.2 workflow 配置与调试 8.3 优化及发布 workflow 8.4 创建并发布 Bot 9 图像流实践 9.1 图像流介绍 9.2 创建一个图像流 9.3 创建一个抠图大师 Agent 9.4 创建一个证件照大师 Agent 9.5 个性头像生成大师 Agent 9.5.1 推理原图的提示词 9.5.3 创建并发布 Bot 9.5.2 创建图像流	时间：1.17（8:30-11:30） 企业新人导师 Agent 1 背景与目标 2.1 RAG 的概念 2.2 RAG 数据准备阶段 2.3 RAG 应用阶段 3.1 Coze 平台中的 RAG 实现 3.2 知识库上传与自动分段 3.3 知识库自定义分段 3.4 Coze 知识库维护 3.5 知识库自动调用 3.6 知识库按需调用 3.7 Bot 发布到商店 3.8 （拓展）Bot 以 Web SDK 发布 4.1 第一个版本存在的问题 4.2 RAG 的优化 workflow 4.3 问题重写 4.4 知识库检索 4.5 重排和优化 4.6 答案生成 4.7 答案优化 4.8 workflow 运行与发布 4.9 Bot 优化与发布 5 小结 时间：1.17（14:00-17:00） 视频课程讲义整理助手 Agent 1.1 背景与目标 1.2 目标实现思路 2.1 通义听悟 AI 助手介绍 2.2 通义听悟提取视频文字稿 3.1 创建 video2doc workflow 3.2 Word 文字稿文档读取 3.3 提取标题 3.4 合并文字稿并去口语化 3.5 文字稿分段总结 3.6 视频课程讲义发布至飞书文档

3 总结 时间：1.15（14:00-17:00） AIGC 的教学应用 1 学习准备与声明 2 大模型辅助教案撰写 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		4.1 创建 video2text 工作流 4.2 获取视频链接和标题 4.3 提取在线视频文字稿 4.4 在线视频文字稿去口语化 4.5 文档生成与工作流发布 5.1 视频课程讲义整理助手 Bot 5.2 Bot 搭建与发布 6 小结 时间：1.18（8:30-17:00） 实战演练-搭建一个 AI Agent 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 时间：1.19（9:00-17:00） 1. 职业技术考试 2. 企业参观 跳转至线下课程安排
--	--	---

拓展课程安排		
AIGC 生产力提升 1 学习准备与声明 2 大模型辅助生成 PPT 2.1.1 一步到位生成自我介绍 PPT 2.1.2 基于文档制作 PPT 2.2.1 提炼并梳理文档中的核心内容 2.2.2 将 markdown 内容转为 PPT 3 大模型辅助生成思维导图 3.1 利用大模型生成思维导图 3.2 将已有内容提炼成思维导图 4 大模型辅助绘制流程图 专利撰写助手 Agent 1 发明专利技术交底书介绍与目标路径 2 上课环境和实际场景 3 构建工作流-找到问题的核心 4 构建工作流-找到解决方案 5 构建工作流-审核解决方案得到创新点	多模态 Agent：海报设计大师 1 背景与目标 2.1 任务分析 2.2 创建 holiday 图像流 2.3 图像生成 2.4 画板设计与调整 3.1 创建 create_holiday_poster 工作流 3.2 主体图提示词优化 3.3 节日祝福生成 3.4 循环生成多张海报 3.5 create_holiday 工作流测试 3.6 节日海报制作 Bot 搭建与发布 4.1 直播课宣传海报工作流分析 4.2 创建 poster 图像流 4.3 选择器设置 4.4 智能人物抠图 4.5 单人海报主体设计 4.6 单人海报添加其他元素 4.7 单人海报微调 4.8 双人海报主体设计	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 大模型绘制柱状图、折线图 4.2 大模型绘制饼图、雷达图 4.3 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

6 构建工作流-详细阐述 7 构建工作流-文本处理 8 上传云文档和完成 bot 机器人 个人学习助手 Agent 1.1 背景与目标 1.2 实现思路 2.1 网页搜索工作流实现步骤 2.2 问题重述 2.3 Coze 当中的提示词小技巧 2.4 网页搜索 2.5 内容总结 3.1 网页总结工作流实现步骤 3.2 网址提取 3.3 循环爬取网页内容并清洗 3.4 意图识别—总结或对比 3.5 内容生成与工作流发布 4.1 意图识别与内容生成 4.2 思维导图结构生成 4.3 生成思维导图 4.4 文档内容中增加思维导图 4.5 文档创建与发布 5.1 个人学习助手搭建 5.2 Bot 优化与发布 6 小结	4.9 双人海报添加其他元素 4.10 双人海报微调与工作流发布 5 直播课宣传海报 Bot 搭建 6 小结	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-diffusionwebui 7.3 魔搭平台的视频生成模型
---	---	--

〈广州专题二〉 数据分析与挖掘实战（泰迪杯竞赛方向） 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 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	时间：1.14（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 3 回归分析 3.1 线性回归基本形式 3.2 线性回归模型实现	时间：1.15（9:00-18:00） 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

4 Pandas 统计分析基础 5 使用 Pandas 进行数据预处理 时间：1.13（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 2 网页前端基础 2.1 认识网络信息传输过程 2.2 认识 HTTP 3 简单静态网页爬取 3.1 认识静态网页 3.2 实现 HTTP 请求 3.3 解析网页 3.4 存储数据 4 认识动态网页 4.1 认识动态网页 4.2 逆向分析爬取动态网页 4.3 使用 Selenium 爬取动态网页	4 神经网络 4.1 单个神经元介绍 4.2 经典网络结构介绍 4.3 神经网络工作流程演示 4.4 如何修正网络参数-梯度下降法 4.5 网络工作原理推导 4.6 网络搭建准备 4.7 样本从输入层到隐层传输的 Python 实现 4.8 网络输出的 Python 实现 4.9 单样本网络训练的 Python 实现 4.10 全样本网络训练的 Python 实现 4.11 网络性能评价 4.12 调用 sklearn 实现神经网络算法	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 4.12 模型应用 拓展内容： PyTorch 深度学习原理与实现 1 引言 2 卷积神经网络 CNN 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		
线下报到时间：1月16日 时间：1.17（8:30-17:00） 机器学习核心技能串讲与巩固 1 Python 爬虫环境与爬虫简介 2 网页前端基础 3 简单静态网页爬取 4 认识动态网页 5 机器学习绪论 6 模型评估与选择 7 回归分析 8 神经网络 时间：1.18（8:30-11:30） 泰迪内推平台招聘信息采集与分析（2023 年挑战赛 C 题） 1 背景与目标	时间：1.18（14:00-17:00） 深度学习基础串讲与巩固 1 PyTorch 简介 2 张量操作 3 构建一个线性模型 4 识别手写数字 时间：1.19（8:30-11:30） 深度学习原理串讲与巩固 1 深度学习引言 2 卷积神经网络 CNN 时间：1.19（14:00-17:00） 基于 YOLOv8 的岩石样本图像分割（2021 年挑战赛 B 题） 1 背景与目标 2 目标分析	时间：1.20（8:30-11:30） 基于 YOLOv8 的岩石样本图像分割（2021 年挑战赛 B 题） 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 时间：1.20（14:00-17:00） 专题讲座（线上直播） 2024 年挑战赛真题解读与优秀作品剖析

2 数据采集 2.1 网页结构探索 2.2 定位一级页面数据地址 2.3 爬取及解析一级页面数据 2.4 提取一级页面字段 2.5 定位二级页面数据地址 2.6 爬取及解析二级页面数据 2.7 翻页爬取及数据保存 3 数据处理 3.1 读取已爬取完成的数据 3.2 数据预处理操作 4 分析与可视化 4.1 招聘岗位对学历要求分析 4.2 各行业的大数据招聘需求数量分析 4.3 不同类型公司薪资待遇分析 4.4 小结	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 labelme 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：1.21（9:00-17:00） 1. 职业技术考试 2. 企业参观 跳转至线下班课程安排
--	--	--

附件二 线上班课程大纲

线上专题一 大模型与 AIGC 技术应用实战课程大纲

技能学习课程安排		
时间：1.14（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：1.17（9:00-18:00） AI Agent 基础与实践-认识 Agent 1 Agent 是什么 1.1 AI Agent 的概念 1.2 走向 AGI 的 Agent 2 基于大模型的 Agent 3 智谱清言智能体平台 3.1 智谱清言智能体平台介绍 3.2 创建文献阅读助手 Agent 3.3 智能体的能力配置 3.4 智能体的知识库配置 4 创建一个 Prompt 优化工程师 时间：1.18（9:00-18:00） AI Agent 基础与实践-扣子智能体平台应用 5 扣子智能体平台搭建 6 Agent 实践：知冷知热桃子姐姐 6.1 目标介绍-创建一个天气播	拓展内容 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 大模型绘制柱状图、折线图 4.2 大模型绘制饼图、雷达图 4.3 大模型绘制散点图、漏斗图、思维导图

2 提示编写原则与策略 2.1 原则与策略 1: 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 时间: 1.15 (9:00-18:00) AIGC 的教学应用 1 学习准备与声明 2 大模型辅助教案撰写 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 时间: 1.16 (9:00-18:00) AIGC 生产力提升 1 学习准备与声明 2 大模型辅助生成 PPT 2.1.1 一步到位生成自我介绍 PPT 2.1.2 基于文档制作 PPT 2.2.1 提炼并梳理文档中的核心内容 2.2.2 将 markdown 内容转为 PPT 3 大模型辅助生成思维导图 3.1 利用大模型生成思维导图 3.2 将已有内容提炼成思维导图 4 大模型辅助绘制流程图	报机器人 6.2 人设与回复逻辑编写 6.3 Bot 发布与调用 7 创建 workflow 7.1 workflow 介绍 7.2 创建 workflow 7.3 调用 workflow 8 Agent 实践: AI 资讯早知道 8.1 创建 workflow 8.2 workflow 配置与调试 8.3 优化及发布 workflow 8.4 创建并发布 Bot 时间: 1.19 (9:00-18:00) 9 图像流实践 9.1 图像流介绍 9.2 创建一个图像流 9.3 创建一个抠图大师 Agent 9.4 创建一个证件照大师 Agent 9.5 个性头像生成大师 Agent 9.5.1 推理原图的提示词 9.5.3 创建并发布 Bot 9.5.2 创建图像流 10 多 Agent 实践 10.1 多 Agent 模式介绍 10.2 创建一个多语种翻译 Bot 10.3 创建一个多任务图片处理专家	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-diffusionwebui 7.3 魔搭平台的视频生成模型
案例实战课程安排		
时间: 1.20 (9:00-18:00) 视频课程讲义整理助手 Agent	时间: 1.22 (9:00-18:00) 企业新人导师 Agent	时间: 1.23 (9:00-18:00) 多模态 Agent: 海报设计大师

1.1 背景与目标 1.2 目标实现思路 2.1 通义听悟 AI 助手介绍 2.2 通义听悟提取视频文字稿 3.1 创建 video2doc 工作流 3.2 Word 文字稿文档读取 3.3 提取标题 3.4 合并文字稿并去口语化 3.5 文字稿分段总结 3.6 视频课程讲义发布至飞书文档 4.1 创建 video2text 工作流 4.2 获取视频链接和标题 4.3 提取在线视频文字稿 4.4 在线视频文字稿去口语化和总结 4.5 文档生成与工作流发布 5.1 视频课程讲义整理助手 Bot 介绍 5.2 Bot 搭建与发布 6 小结 时间：1.21（9:00-18:00） 专利撰写助手 Agent 1 发明专利技术交底书介绍与目标路径 2 上课环境和实际场景 3 构建工作流-找到问题的核心 4 构建工作流-找到解决方案 5 构建工作流-审核解决方案得到创新点 6 构建工作流-详细阐述 7 构建工作流-文本处理 8 上传云文档和完成 bot 机器人	1 背景与目标 2.1 RAG 的概念 2.2 RAG 数据准备阶段 2.3 RAG 应用阶段 3.1 Coze 平台中的 RAG 实现 3.2 知识库上传与自动分段 3.3 知识库自定义分段 3.4 Coze 知识库维护 3.5 知识库自动调用 3.6 知识库按需调用 3.7 Bot 发布到商店 3.8 Bot 以 Web SDK 发布 4.1 第一个版本存在的问题 4.2 RAG 的优化工作流 4.3 问题重写 4.4 知识库检索 4.5 重排和优化 4.6 答案生成 4.7 答案优化 4.8 工作流运行与发布 4.9 Bot 优化与发布 5 小结	1 背景与目标 2.1 任务分析 2.2 创建 holiday 图像流 2.3 图像生成 2.4 画板设计与调整 3.1 创建 create_holiday_poster 工作流 3.2 主体图提示词优化 3.3 节日祝福生成 3.4 循环生成多张海报 3.5 create_holiday 工作流测试 3.6 节日海报制作 Bot 搭建与发布 4.1 直播课宣传海报工作流分析 4.2 创建 poster 图像流 4.3 选择器设置 4.4 智能人物抠图 4.5 单人海报主体设计 4.6 单人海报添加其他元素 4.7 单人海报微调 4.8 双人海报主体设计 4.9 双人海报添加其他元素 4.10 双人海报微调与工作流发布 5.1 create_poster 工作流分析 5.2 人物和二维码知识库 5.3 开始节点设置 5.4 知识库检索 5.5 知识库信息处理 5.6 海报文案提取 5.7 海报文案优化处理 5.8 打印提取到的文案消息 5.9 海报生成测试 5.10 选择器判断是否有文案 5.11 添加网络搜索 5.12 海报标题与正文生成 5.13 海报技术要点生成 5.14 使用生成的文案制作海报 5.15 直播课宣传海报 Bot 搭建 6 小结 时间：1.24（16:00-18:00） 职业技术考试 跳转至线上上班课程安排
---	--	--

线上专题二 大模型应用开发实战（LangChain）课程大纲

技能学习课程安排		
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 时间：1.14（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：1.15（9:00-18:00） 2 提示编写原则与策略 2.1 原则与策略 1：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 时间：1.16（9:00-18:00） AIGC 的教学应用 1 学习准备与声明 2 大模型辅助教案撰写 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 时间：1.17（9:00-18:00） LangChain 实战-基础使用 1.1 LangChain 简介 1.2 LangChain 环境安装 2.1 申请阿里云的 API KEY 2.2 LangChain 示例 2.3 链的使用 3.1 提示模板 3.2 模型调用 3.3 输出解析 3.4 开发情感极性分析应用 3.5 课堂练习：基于 LangChian 开发命名实体识别应用	时间：1.18（9:00-18:00） LangChain 实战-RAG 案例应用 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 时间：1.19（9:00-18:00） LangChain 实战-Agent 应用 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 时间：1.20（9:00-18:00） 基于 LangChain 的本地大模型调用 1.1 LangChain 本地大模型调用思路 1.2 环境准备 2.1 HuggingFacePipeline 使用 2.2 ChatHuggingFace 使用 2.3 自定义 LLM 2.4 HuggingFaceEmbeddings 使用 2.5 基于 Embedding 的文本相似度计算

案例实战课程安排		
时间：1.21（9:00-18:00） RAG 实战-论文阅读助手 1 背景与目标 2.1 模型准备 2.2 文档读取和预处理 2.3 文档拆分 2.4 向量存储 2.5 构建阅读问答链 2.6 阅读应用 2.7 RAG 增强技术 3 阅读机器人界面开发 4 小结	时间：1.22（9:00-18:00） Agent 实战-在线教育课程订单智能助手 1 背景与目标 2 目标分析 3.1 模型准备 3.2 创建表格智能体 3.3 工具准备 3.4 Agent 创建和使用 4 思维拓展：RAG 结合 Agent 5 小结	时间：1.23（9:00-18:00） 基于本地大模型的面试助手开发 1.1 目标分析 1.2 RAG 技术实现 2.1 环境准备 2.2 模型下载 2.3 LangChain 加载本地大模型 2.4 文档读取和切分方法 2.5 文档读取和分割实现 2.6 向量化存储 2.7 检索器 2.8 构建 RetrievalQA 问答链 3.1 Gradio 框架简介 3.2 页面搭建 4 小结 时间：1.24（16:00-18:00） 职业技术考试 跳转至线上上班课程安排

线上专题三 大模型微调应用实战课程大纲

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：1.14（9:00-18:00） 大模型与 AIGC 概述 1 大模型是什么	时间：1.16（9:00-18:00） 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	时间：1.19（9:00-18:00） 大模型微调技术应用 1. 大模型微调的概念 2.1 大模型微调方法 2.2 硬件配置要求说明 2.3 环境准备 2.4 模型准备 2.5 数据准备 3.1 全量微调 FFT-原理介绍 3.2 全量微调 FFT-配置训练参数 3.3 全量微调 FFT-构建训练器 3.4 全量微调 FFT-模型训练 3.5 全量微调 FFT-模型效果验证 4.1 参数高效微调 PEFT 4.2 Prompt Tuning 微调原理 4.3 Prompt Tuning 实现 4.4 Prompt Tuning 结果分析 4.5 Prompt Tuning 模型验证 4.6 P-Tuning 微调原理 4.7 P-Tuning 实现

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 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 3 回归分析 3.1 线性回归基本形式 3.2 线性回归模型的 Python 实现 时间：1.15（9:00-18:00） 4 人工神经网络 4.1 单个神经元介绍 4.2 经典网络结构介绍 4.3 神经网络工作流程演示 4.4 如何修正网络参数-梯度下降法 4.5 网络工作原理推导 4.6 网络搭建准备 4.7 样本从输入层到隐层传输的 Python 实现 4.8 网络输出的 Python 实现 4.9 单样本网络训练实现 4.10 全样本网络训练实现 4.11 网络性能评价 4.12 实现神经网络算法	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 4.12 模型应用 时间：1.17（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 并行计算能力 时间：1.18（9:00-18:00） 3 生成式预训练模型 GPT 3.1 生成式预训练模型 GPT 简介 3.2 GPT 预训练流程 3.3 GPT 微调流程 4 双向编码模型 BERT 4.1 BERT 简介 4.2 BERT 预训练流程 4.3 BERT 微调流程 5 Transformer 应用 5.1 HuggingFace 简介 5.2 用 pipeline 进行情感分析 5.3 将预训练模型下载至本地 5.4 用 pipeline 完成其他任务 5.5 完成机器翻译任务	4.8 P-Tuning v2 微调原理 4.9 P-Tuning v2 实现 时间：1.20（9:00-18:00） 4.10 LoRA 微调原理 4.11 LoRA 实现 4.12 AdaLoRA 原理 4.13 AdaLoRA 实现 4.14 IA3 原理与实现 4.15 BitFit 原理与实现 5.1 如何降低大模型显存占用 5.2 半精度 F16 与 BF16 5.3 半精度微调实现 5.4 INT8 量化原理 5.5 INT8 量化实现 5.6 QLoRA 原理 5.7 QLoRA 实现 6 小结 7.1 拓展 1: modelscope 平台启动 7.2 拓展 1: modelscope 微调步骤实现 8.1 拓展 2: LLaMA-Factory 界面微调工具 8.2 拓展 2: LLaMA-Factory 环境配置 8.3 拓展 2: LLaMA-Factory 配置和训练 8.4 拓展 2: LLaMA-Factory 其他功能使用 拓展内容： PyTorch 深度学习原理与实现 1 引言 2 循环神经网络 RNN 3 长短时记忆网络 LSTM
---	---	--

案例实战课程安排		
时间：1.21（9:00-18:00） 基于通义千问（Qwen）的创意广告生成 1 背景与目标 2.1 千问大模型介绍 2.2 modelscope 免费实例 2.3 创建虚拟环境 2.4 模块安装 2.5 模型下载和使用 2.6 模型文件路径说明 3.1 数据格式要求 3.2 数据格式转换 4.1 LoRA 技术 4.2 finetune 代码解读 4.3 LoRA 微调和结果 4.4 模型验证 4.5 模型验证的注意事项 5 小结 6 拓展：modelscope 环境的使用	时间：1.22（9:00-18:00） 基于大模型微调的命名实体识别 1 背景与目标 2.1 命名实体识别的概念 2.2 命名实体识别的方法 3.1 数据格式 3.10 NERDataset 结果返回 3.11 NERDataset 代码调试 3.12 构建最终的训练数据 3.2 环境配置 3.3 数据格式转换 3.4 Label-Studio 数据标注工具使用 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 Pytorch Lightning 框架 4.4 模型训练 4.5 指标评价 4.6 模型验证 5 小结	时间：1.23（9:00-18:00） 基于大模型的金融问句语义相似度计算 1 案例背景及挖掘目标 2 语义文本相似度 3.1 预处理：文件加载 3.2 预处理：错别字纠正 3.3 预处理：数据类型转化与长度统计 3.4 预处理：数据准备 3.5 加载预训练模型 3.6 定义损失函数 3.7 模型训练 3.8 模型预测 3.9 模型评估 4 小结 拓展内容： 基于 Transformer 的疫情问诊系统自动翻译实现 1.1 机器翻译发展历程 1.2 任务目标及实现流程 2.1 数据预处理 3 预训练模型加载 4 模型训练 5 模型推理 时间：1.24（16:00-18:00） 职业技术考试 跳转至线上上班课程安排

线上专题四 大模型私有化部署和应用开发实战课程大纲

课程安排		
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	时间：1.16（9:00-18:00） 4 Qwen2 大模型量化 4.1 量化的概念 4.2 半精度模型部署 4.3 低精度 INT8 量化部署 4.4 低精度 INT4 量化部署 4.5 GPTQ 量化介绍 4.6 GPTQ 量化实现 4.7 AWQ 量化介绍	时间：1.20（9:00-18:00） 基于 LangChain 的人事面试助手开发 1.1 目标分析 1.2 RAG 技术实现途径 2 RAG 技术流程 3.1 环境说明 3.2 模型下载

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：1.14（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 大模型开源平台介绍及使用 1 LLM 大模型基础认知 1.1 开篇 1.2 LLM 大模型基础认知 2 HuggingFace 社区介绍（拓展） 2.1 环境准备 2.2 HuggingFace 介绍 2.3 模型下载 2.4 模型推理的流程 2.5 模型推理实现 2.6 tokenizer 使用 2.7 模型生成参数 2.8 datasets 数据导入导出 2.9 datasets 其他用法 2.10 Trainer 2.11 拓展训练：构建一个情感分类模型 3 ModelScope 社区介绍 3.1 魔塔社区介绍 3.2 模型和数据下载	4.8 AWQ 量化实现 4.9 GGUF 模型 4.10 llama.cpp 环境安装 4.11 GGUF 量化实现 4.12 llama.cpp 调用 GGUF 模型 4.13 GGUF 模型 API 部署 4.14 Ollama 快速部署 GGUF 模型 4.15 模型量化小结 5 其他大模型部署（拓展） 5.1 InternVL 大模型部署 5.2 pai-diffusion-artist 部署 5.3 ChatTTS 部署 6 小结 时间：1.17（9:00-18:00） Qwen2 大模型部署与电商客服模型微调 1 背景与目标 2 微调方法介绍 3.1 环境准备 3.2 分析 Qwen2 大模型对电商问题的回复效果 3.3 数据格式转化 3.4 微调脚本 3.5 执行微调脚本 3.6 模型微调日志分析 3.7 模型合并 3.8 微调后的模型使用效果 3.9 API 部署和请求 3.10 web 界面部署 3.11 Qwen2 微调小结 4.1 swift 模型推理 4.2 swift 模型微调 4.3 swift 模型微调后使用 4.4 swift 界面化微调训练 4.5 swift 界面化的其他操作 4.6 swift 使用小结 时间：1.18（9:00-18:00） 基于 Qwen1 的创意广告生成 1 背景与目标 2.1 千问大模型介绍 2.2 modelscope 免费实例 2.3 创建虚拟环境	3.3 LangChain 加载本地大模型 3.4 文档读取和切分方法 3.5 文档读取和分割实现 3.6 向量化存储 3.7 检索器 3.8 构建 RetrievalQA 问答链 4.1 Gradio 框架简介 4.2 人事面试百事通页面搭建 5 小结 时间：1.21（9:00-18:00） 基于 Ollama 和 Dify 的本地大模型的部署和应用开发 1 背景与目标分析 2.1 Ollama 介绍与安装 2.2 Ollama 常用操作 2.3 通过 openai 调用 ollama 本地大模型 2.4 通过 ollama 模块调用 ollama 本地大模型 2.5 通过 LangChain 调用 ollama 本地大模型 2.6 Ollama 部署和使用 Embedding 模型 2.7 Modelfile 文件配置介绍 2.8 通过 ModelFile 部署本地大模型 2.9 练习：使用 LangChain 接入 Ollama 大模型实现 RAG 时间：1.22（9:00-18:00） 3.1 LM Studio 介绍与安装 3.2 LM Studio 使用 3.3 LM Studio 和 Ollama 使用小结 4.1 Dify 平台介绍 4.2 Docker desktop 安装 4.3 Docker desktop 环境配置 4.4 Dify 安装 4.5 Dify 界面 5.1 基于 Dify 的 RAG 实现流程
--	---	---

3.3 基于 modelscope 的模型推理 3.4 swift 模型推理 3.5 modelscope 算力平台介绍 4 小结 时间：1.15（9:00-18:00） 部署自己的 Qwen2 大模型 1 开篇 2 Github 源码下载 3.1 Github 源文件介绍 3.2 Qwen2 模型推理部署 3.3 CLI 部署 3.4 Gradio 界面部署 3.5 FaskAPI 部署思路 3.6 FaskAPI 部署 Qwen2 3.7 API 请求 1: curl 命令 3.8 API 请求 2: requests 请求 3.9 Flask 部署 Qwen2 3.10 xinference 介绍和安装 3.11 xinference 模型部署 3.12 xinference 模型 API 调用 3.13 xinference 界面化部署 3.14 vLLM 框架介绍 3.15 模型部署小结	2.4 模块安装 2.5 模型下载和使用 2.6 模型文件路径说明 3.1 数据格式要求 3.2 数据格式转换 4.1 LoRA 技术 4.2 finetune 代码解读 4.3 LoRA 微调和结果 4.4 模型验证 4.5 模型验证的注意事项 5 小结 时间：1.19（9:00-18:00） 基于 Langchain 的本地大模型部署和应用 1.1 LangChain 介绍 1.2 环境准备 2.1 HuggingFacePipeline 使用 2.2 ChatHuggingFace 使用 2.3 自定义 LLM 2.4 链的构建和使用 2.5 链的使用示例 2.6 函数式链的使用 2.7HuggingFaceEmbeddings 使用 2.8 基于 Embedding 的文本相似度计算 3.1 在线大模型使用说明 3.2 在线大模型使用示例 4.1 情感分析应用思路 4.2 不同模型的效果比较 4.3 分析评论的情感倾向 4.4 命名实体提取介绍 4.5 命名实体识别实现 5 小结	5.2 模型配置须知 5.3 模型配置实现 5.4 创建通用聊天机器人 时间：1.23（9:00-18:00） 5.5 知识库创建 5.6 人事面试助手应用搭建 5.7 Dify 应用部署（网页部署） 5.8 Dify 应用部署（微信部署） 5.9 Dify 部署小结 6.1 工作流 6.2 网页提取工作流 6.3 Agent 使用 7 小结 时间：1.24（16:00-18:00） 职业技术考试 跳转至线上班课程安排
--	--	---

线上专题五 AIGC 多模态技术应用实战课程大纲

技能学习课程安排		
时间：1.14（9:00-18:00） 大模型与 AIGC 概述 1 大模型是什么	时间：1.15（9:00-18:00） 从文本到视觉的创作革命 1 视觉创作的变革	时间：1.17（9:00-18:00） AI 绘画应用实战-技巧进阶 3 技巧进阶

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、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 2 提示编写原则与策略 2.1 原则与策略 1: 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	1.1 文本创作 1.2 视频创作 2 AI 文生图技术 2.1 文生图技术发展路线 2.2 扩散模型 2.3 CLIP 模型 2.4 扩散模型的典型代表—DALL-E 2 模型 2.5 可控条件 2.6 文生图模型评测 3 视频生成技术 3.1 文生图到文生视频的思考 3.2 视频生成的里程碑: Sora 模型 3.3 Sora 模型架构 3.4 文生视频的模型发展方向 3.5 文生视频模型评测 4 小结 时间: 1.16 (9:00-18:00) AI 绘画应用实战-基础实践 1 AI 绘画平台与提示词 1.1 即梦 AI 平台 1.2 提示词技巧 2 通用 AI 绘画一场景实践 2.1 古诗词变美图 2.2 电影海报制作 2.3 海报素材制作 2.4 微缩景观绘画 2.5 人像写真绘画 2.6 产品图设计	3.1 图像控制 3.2 使用图生图制作产品宣传 3.3 一键做同款 3.4 图片反推提示词 3.5 主体一致性 4 图像生成编辑 4.1 局部重绘 4.2 涂抹消除 4.3 画质提升 4.4 智能扩图、智能抠图 4.5 智能画布 时间: 1.18 (9:00-18:00) AI 绘画应用实战-应用实战 5 画质增强与修复 5.1 即梦 AI 图像修复问题 5.2 画质增强修复技术体系 5.3 追影-放大镜图像增强工具 5.4 AI 老照片修复 6 艺术字及艺术字海报设计 6.1 现阶段 AI 绘画的文字识别问题 6.2 锦书 WordArt 模型介绍 6.3 文字变形模型框架 6.4 文字仿写模型框架 6.5 文字纹理模型框架 6.6 WordArt 模型实战 7 AI 虚拟试衣 7.1 传统虚拟试衣模型 7.2 最新虚拟试衣模型技术进展 7.3 可图 AI 虚拟试衣使用 7.4 开源虚拟试衣模型
案例实战课程安排		
时间: 1.19 (9:00-18:00) AI 视频生成应用实战 1 文生视频 1.1 文生视频平台 1.2 文生视频提示词 1.3 文生视频实战 2 图生视频 2.1 图生视频提示词 2.2 图生视频技巧 2.3 图生视频实战	时间: 1.21 (9:00-18:00) 5 音频降噪 5.1 音频降噪原理 5.2 音频降噪模型 5.3 音频降噪应用 6 说话人识别技术应用 6.1 说话人识别技术应用场景 6.2 说话人识别模型 6.3 说话人识别实现	时间: 1.23 (9:00-18:00) 多模态 Agent 1 背景与目标 2.1 任务分析 2.2 创建 holiday 图像流 2.3 图像生成 2.4 画板设计与调整 3.1 创建 create_holiday_poster 工作流 3.2 主体图提示词优化 3.3 节日祝福生成

时间：1.20（9:00-18:00） AI 语音技术应用实战 1 语音识别技术介绍 1.1 语音识别技术概念 1.2 语音识别模型 1.3 语音识别实战 2 文本生成语音技术介绍 2.1 语音生成概念 2.2 语音生成模型 2.3 语音生成实战 3 语音克隆技术介绍与实践 3.1 语音克隆技术概念 3.2 语音克隆模型 3.3 语音克隆实战 3.4 AI 语音克隆的伦理与法律 4 语音克隆应用软件 4.1 商业软件介绍 4.2 语音克隆工具实现	时间：1.22（9:00-18:00） 虚拟数字人技术应用实战 1 数字人应用场景 2 数字人技术概述 3 数字人技术进展 4 数字人实战（张奥多） 5 虚拟数字人应用软件介绍	3.4 循环生成多张海报 3.5 create_holiday 工作流测试 3.6 节日海报制作 Bot 搭建与发布 4.1 直播课宣传海报工作流分析 4.2 创建 poster 图像流 4.3 选择器设置 4.4 智能人物抠图 4.5 单人海报主体设计 4.6 单人海报添加其他元素 4.7 单人海报微调 4.8 双人海报主体设计 4.9 双人海报添加其他元素 4.10 双人海报微调与工作流发布 5.1 create_poster 工作流分析 5.2 人物和二维码知识库 5.3 海报标题与正文生成 5.4 海报技术要点生成 5.5 使用生成的文案制作海报 5.6 直播课宣传海报 Bot 搭建 6 小结 时间：1.24（16:00-18:00） 职业技术考试 跳转至线上课程安排
--	--	--

线上专题六 Hadoop+Spark 大数据技术应用实战课程大纲

技能学习课程安排		
时间：报名成功后即可开始学习 Linux 操作系统基础 1. Linux 概述 2. Linux 系统安装 3. Linux 基本命令 4. Linux Vi 编辑器 5. Linux Shell 编程 Java 编程基础 1. 基础知识 2. 面向对象 3. 线程及异常处理 Scala 编程基础 1. Scala 简介 2. Scala 安装配置 3. Scala 基础语法 4. 函数	时间：1.16（9:00-18:00） Hive 大数据仓库 1. Hive 简介 1.1 Hive 简介 2. Hive 安装配置 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 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	时间：1.19（9:00-18:00） 5. 配置 Spark IDEA 开发环境 5.1 搭建 Spark 开发环境 6. Spark SQL 应用 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

5. 面向对象编程 6. 文件读写 时间：1.14（9:00-18:00） 专题讲座 计算机视觉技术及其应用 Hadoop 大数据基础 1. Hadoop 简介 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 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 时间：1.15（9:00-18:00） 3. Hadoop 基础操作 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 4.1 MapReduce 框架与设计构思 4.2 MapReduce 开发环境搭建 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	3.2 Hive 导入及导出 时间：1.17（9:00-18:00） 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 拓展 4. Hive 开发 5. Hive 自定义函数 6. Hive 查询优化 时间：1.18（9:00-18:00） Spark 大数据技术应用 1. Spark 入门 1.1 Spark 入门 2. Spark 集群的安装配置 2.1 Spark 安装部署 2.2 Spark 安装部署实战 3. Spark 架构及原理 3.1 Spark 架构 3.2 Spark RDD 及 DAG 相关概念 4. Spark 编程基础 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 4.6 文件读取与存储 4.7 统计用户停留时间最长的基站	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 拓展内容： 1 HBase 非关系型数据库 2 Zookeeper 大数据分布式消息 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 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 (ChatGLM2-6B) 产业应用漫步与畅想
案例实战课程安排		

时间：1.20（9:00-18:00） 航空客户价值分析 1. 案例背景与挖掘目标 2. 数据存储 3. 数据探索与数据清洗 4. 构建航空客户价值分析关键特征 5. 航空客户分群 6. 会员营销策略与建议	时间：1.21（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：1.22（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：1.23（16:00-18:00） 职业技术考试 跳转至线上班课程安排
---	--	---

线上专题七 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		
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 3 Matplotlib 数据可视化基础 4 Pandas 统计分析基础 时间：1.14（9:00-18:00） Python 数据分析与应用 1 使用 Pandas 进行数据预处理 1.1 合并数据 1.2 清洗数据 1.3 标准化数据 1.4 转换数据	时间：1.16（9:00-18:00） 5 人工神经网络 5.1 单个神经元介绍 5.2 经典网络结构介绍 5.3 神经网络工作流程演示 5.4 修正网络参数-梯度下降法 5.5 网络工作原理推导 5.6 网络搭建准备 5.7 样本从输入层到隐层传输的 Python 实现 5.8 网络输出的 Python 实现 5.9 单样本网络训练实现 5.10 全样本网络训练实现 5.11 网络性能评价 5.12 sklearn 实现神经网络 时间：1.17（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 2 网页前端基础	时间：1.19（9:00-18:00） 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

时间：1.15（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 3 回归分析 3.1 线性回归基本形式 3.2 线性回归模型实现 4 决策树 4.1 从女生相亲到决策树 4.2 明天适合打球吗 4.3 决策树拆分属性选择 4.4 决策树算法家族 4.5 泰坦尼克号生还者预测—数据预处理 4.6 泰坦尼克号生还者预测—模型构建与预测	2.1 认识网络信息传输过程 2.2 认识 HTTP 3 简单静态网页爬取 3.1 认识静态网页 3.2 实现 HTTP 请求 3.3 解析网页 3.4 存储数据 时间：1.18（9:00-18:00） 4 认识动态网页 4.1 认识动态网页 4.2 逆向分析爬取动态网页 4.3 使用 Selenium 爬取动态网页	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 4.12 模型应用 拓展内容： PyTorch 深度学习原理与实现 1 引言 2 循环神经网络 RNN 3 长短期记忆网络 LSTM 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：1.20（9:00-18:00） 学生校园卡消费行为分析（2020 年技能赛 A 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	时间：1.21（9:00-18:00） 泰迪内推平台招聘信息采集与分析（2023 年挑战赛 C 题） 1 背景与目标 2 数据采集 2.1 网页结构探索 2.2 定位一级页面数据地址 2.3 爬取及解析一级页面数据 2.4 提取一级页面字段 2.5 定位二级页面数据地址 2.6 爬取及解析二级页面数据 2.7 翻页爬取及数据保存 3 数据处理 3.1 读取已爬取完成的数据 3.2 数据预处理操作 4 分析与可视化 4.1 招聘岗位对学历要求分析 4.2 各行业的大数据招聘需求数量分析	时间：1.22（9:00-18:00） 基于 YOLOv8 的岩石样本图像分割（2021 年挑战赛 B 题） 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 lablme 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：1.23（9:00-18:00）

	4.3 不同类型公司的薪资待遇分析 4.4 小结	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 岩石样本分类（拓展） 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：1.24（16:00-18:00） 职业技术考试 跳转至线上课程安排
--	-----------------------------	---

线上专题八 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：1.14（9:00-18:00） Python 数据分析与应用 1 Matplotlib 数据可视化基础 1.1 绘图基础语法与常用参数 1.2 分析特征间关系 1.3 分析特征内部数据分布与分	时间：1.16（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：1.17（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 2 网页前端基础 2.1 认识网络信息传输过程 2.2 认识 HTTP 3 简单静态网页爬取 3.1 认识静态网页 3.2 实现 HTTP 请求 3.3 解析网页	时间：1.19（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 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

散情况 时间：1.15（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	3.4 存储数据 时间：1.18（9:00-18:00） 4 认识动态网页 4.1 认识动态网页 4.2 逆向分析爬取动态网页 4.3 使用Selenium爬取动态网页	
案例实战课程安排		
时间：1.20（9:00-18:00） 泰迪内推平台招聘信息采集与分析（Python） 1 背景与目标 2 数据采集 2.1 网页结构探索 2.2 定位一级页面数据地址 2.3 爬取及解析一级页面数据 2.4 提取一级页面字段 2.5 定位二级页面数据地址 2.6 爬取及解析二级页面数据 2.7 翻页爬取及数据保存 3 数据处理 3.1 读取已爬取完成的数据 3.2 数据预处理操作 4 分析与可视化 4.1 招聘岗位对学历要求分析 4.2 各行业的大数据招聘需求数量分析 4.3 不同类型公司的薪资待遇分析 4.4 小结	时间：1.21（9:00-18:00） 新浪财经数据采集（八爪鱼工具） 1 背景与目标 2 目标分析 3.1 八爪鱼采集器介绍与安装 3.2 菜单栏和模板任务 4.1 自定义任务主界面介绍 4.2 新建自定义任务 4.3 自动识别爬虫字段 4.4 手动配置爬虫字段 4.5 翻页设置 4.6 二级页面爬取 4.7 进入二级页面的注意事项 4.8 数据预览操作 4.9 流程图操作 4.10 自定义任务保存 5.1 数据采集与导出 5.2 任务设置和保存 6 小结	时间：1.22（9:00-18:00） 汽车质量投诉平台数据采集（八爪鱼工具） 1 背景与分析目标 2 数据采集 2.1 投诉网页分析和爬虫思路 2.2 分析翻页网址构造 2.3 自定义翻页网址列表 2.4 配置二级页面跳转位置 2.5 循环点击进入二级页面 2.6 配置二级页面的爬虫字段 2.7 修改字段名称 2.8 任务保存和开始采集 2.9 回顾：整体流程分析 3 数据和任务导出 3.1 查看爬虫任务明细 3.2 数据导出保存 时间：1.23（9:00-18:00） 微博评论数据采集（Python） 1 背景与目标 2.1 评论结构分析 2.2 数据接口分析 3.1 微博页面接口分析 3.2 评论数据接口分析 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：1.24（16:00-18:00） 职业技术考试 跳转至线上课程安排

线上专题九 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 3 Matplotlib 数据可视化基础 时间：1.14（9:00-18:00） Python 数据分析与应用 1 Pandas 统计分析基础 1.1 Pandas 简介 1.2 读写不同数据源的数据 1.3 数据框与数据框元素 1.4 转换与处理时间序列数据 1.5 使用分组聚合进行组内计算 1.6 创建透视表与交叉表	时间：1.15（9:00-18:00） 2 使用 Pandas 进行数据预处理 2.1 合并数据 2.2 清洗数据 2.3 标准化数据 2.4 转换数据 时间：1.16（9:00-18:00） 3 使用 scikit-learn 构建模型 3.1 使用 sklearn 转换器处理数据 3.1.1 Scikit-Learn 简介 3.1.2 获取及认识 datasets 中的数据 3.1.3 将数据集划分为训练集和测试集 3.1.4 利用转化器进行数据转化操作 3.2 构建并评价聚类模型 3.3 构建并评价分类模型 3.4 构建并评估回归模型 3.4.1 构建回归模型 3.4.2 评价回归模型	时间：1.17（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：1.18（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 牛刀小试-T 公司运营数据分析与业务决策 1 背景与目标 2 目标分析 3 数据探索与处理 3.1 数据探索与处理思路 3.2 读项目数据&处理业务标签 3.3 日报表读取及上下半年数据	时间：1.20（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 2.3 剔除异常记录 3 特征构造 3.1 RFM 模型介绍 3.2 LRFMC 模型 3.3 构造入会时长特征 3.4 剩余特征构造 4 K-Means 客户分群 4.1 使用 K-Means 算法进行客户	时间：1.22（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保存 4 关联规则分析 4.1 关联规则介绍 4.2 Apriori 算法流程 4.3 提升度的概念 5 关联规则模型构建 5.1 数据加载及拆分 5.2 挖掘强关联规则 (Apriori)

拆分 3.4 读取人员成本数据并与日报表拼接 3.5 三表合并及异常值处理 4 数据分析 4.1 计算各类业务的核心指标 4.2 各类业务指标分析及可视化 4.3 各项目核心指标统计及分析 4.4 各部门的核心指标计算分析 5 分析结果与建议 时间：1.19（9:00-18:00） 回归分析-某市财政收入预测 1 背景与目标 1.1 案例背景 1.2 数据介绍 2 目标分析 3 线性回归分析 3.1 线性回归分析介绍 3.2 构建线性回归模型 3.3 利用构建好的模型进行预测 3.4 数据读取与拆分 3.5 模型训练与预测 4 分析结果与评估 4.1 模型性能评估 4.2 将预测结果可视化 4.3 G 市未来两年财政收入预测	分群 4.2 获取 K-Means 聚类结果 4.3 聚类结果可视化 5 小结 时间：1.21（9:00-18:00） 分类分析-运营商用户流失判别 1.1 背景与目标 1.2 案例思路分析 2.1 数据探索 2.2 数据去重及删除无关属性 2.3 用户分组及标签构建 2.4 提取用户基本信息和在网时长 2.5 处理合约是否有效 2.6 处理合约计划到期时间 2.7 其余变量处理 2.8 特征拼接及缺失值处理 2.9 数据保存 3.1 特征选择介绍 3.2 皮尔逊特征选择 3.3 处理样本类别不均衡问题 4.1 模型性能评估介绍 4.2 模型构建及性能评估	算法实现) 6 模型性能评估 时间：1.23（9:00-18:00） 协同过滤-泰迪内推平台信息精准推荐应用 1 背景与目标 2 目标分析 2.1 推荐思路分析 2.2 基于物品的协同过滤推荐介绍 3 工程实现 3.1 EB 工具登录及简介 3.2 创建空白工程 3.3 导入数据 3.4 筛选正文数据 3.5 字符替换及记录去重 3.6 划分训练集用户和测试集用户 3.7 构造训练集和测试集数据 3.8 构建模型 3.9 推荐及性能评估 4 小结 时间：1.24（16:00-18:00） 职业技术考试 跳转至线上班课程安排
--	--	---

线上专题十 计算机视觉应用实战（PyTorch）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	时间：1.15（9:00-18:00） 5 人工神经网络 5.1 单个神经元介绍 5.2 经典网络结构介绍 5.3 神经网络工作流程演示 5.4 如何修正网络参数-梯度下降法 5.5 网络工作原理推导 5.6 网络搭建准备 5.7 样本从输入层到隐层传输的 Python 实现	时间：1.17（9:00-18:00） PyTorch 深度学习原理与实现 1 引言 2 卷积神经网络 CNN 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

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：1.14（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 3 回归分析 3.1 线性回归基本形式 3.2 线性回归模型的 Python 实现 4 决策树 4.1 从女生相亲到决策树 4.2 明天适合打球吗 4.3 决策树拆分属性选择 4.4 决策树算法家族 4.5 泰坦尼克号生还者预测—数据预处理 4.6 泰坦尼克号生还者预测—模型构建与预测	5.8 网络输出的 Python 实现 5.9 单样本网络训练的 Python 实现 5.10 全样本网络训练的 Python 实现 5.11 网络性能评价 5.12 调用 sklearn 实现神经网络 时间：1.16（9:00-18:00） 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 4.12 模型应用	2.8 实现池化操作 时间：1.18（9:00-18:00） 3 循环神经网络 RNN 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 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 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 5.5 模型训练 5.6 模型性能验证 时间：1.19（9:00-18:00） 计算机视觉实战 1 概述 2.1 图像基础 2.2 读写图像 3.1 几何变换-图像平移和旋转 3.2 几何变换-最近邻插值 3.3 几何变换-其他插值方法介绍 4.1 灰度处理-线性变换 4.2 灰度处理-非线性变换 4.3 灰度处理-直方图均衡化 4.4 图像二值化 5.1 图像平滑 5.2 图像锐化-Sobel 算子 5.3 图像锐化-其他算法 6. 练习 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 (ChatGLM2-6B) 产业应用漫步与畅想
案例实战课程安排		

时间：1.20（9:00-18:00） 脑 PET 图像分析与疾病预测 - 背景与目标 - 数据处理 - 图像读取及尺寸调整 - 图像增强之翻转操作 - 图像增强之旋转缩放 - 获取所有照片路径 - 批量获取照片数据 - 将数据处理过程封装成函数 - 模型构建 - 定义卷积&池化层 - 定义全连接层 - 定义网络计算过程 - 数据集维度顺序调整及类型转化 - 数据集分批及打乱操作 - 模型构建及配置 - 模型训练 - 模型性能评估及应用 - 模型性能评估及保存 - 模型应用	时间：1.21（9:00-18:00） 基于 FaceNet 的人脸智能识别 - 背景与目标 - 背景与目标 - 目标分析 - 开发环境和工程结构介绍 - 人脸检测 - MTCNN 人脸检测介绍 - 执行人脸检测操作 - 人脸对齐 - 执行人脸对齐操作 - 人脸检测及对齐代码整理 - 人脸特征提取 - FaceNet 溯源-计算机视觉领域的部分大事件 - FaceNet 介绍 - 执行人脸特征提取操作 - 人脸识别 - 获取后台数据库中的人脸数据 - 获取后台数据库人脸数据脚本解读 - 完成人脸识别操作 - 代码整理与结果可视化 - 小结	时间：1.22（9:00-18:00） 基于 YOLOX 的农田害虫图像检测与识别 - 背景与目标 - 目标分析 - 目标检测任务的难点与挑战 - 目标检测方法发展历程 - 经典二阶段&一阶段算法 - YOLO 目标检测 - 将目标检测转化为回归任务 - YOLOv1 中训练样本标签设计 - YOLOv1 网络结构及输出解读 - YOLOv1 损失函数介绍 - YOLOv1 的缺点 - YOLOv3 及 YOLOX 介绍 - 目标检测常用数据集与性能评价指标介绍 - 项目目标完成步骤介绍 - 数据探索与处理 - 配套资源说明 - 数据探索 - CopyAndPast 数据增强介绍 - CopyAndPast 数据增强实现 - 汇总照片数据 时间：1.23（9:00-18:00） - 数据加工 - 数据加工介绍 - 数据加工实现 - 环境搭建与模型训练 - 创建虚拟开发环境 - 启动虚拟环境并为其安装依赖库 - 安装 YOLOX - 模型训练 - 模型应用 - 模型应用 - 小结 - 拓展延伸 时间：1.24（16:00-18:00） 职业技术考试 跳转至线上课程安排
--	--	--

线上专题十一 文本分析与挖掘实战（PyTorch）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：1.14（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 3 回归分析 3.1 线性回归基本形式 3.2 线性回归模型的 Python 实现 4 人工神经网络 4.1 单个神经元介绍 4.2 经典网络结构介绍 4.3 神经网络工作流程演示 4.4 如何修正网络参数-梯度下降法 4.5 网络工作原理推导 4.6 网络搭建准备 4.7 样本从输入层到隐层传输的	时间：1.15（9:00-18:00） 5.1 聚类分析概述 5.2 相似性度量 5.3 K-Means 聚类分析算法介绍 5.4 利用 K-Means 算法对鸢尾花进行聚类 5.5 聚类结果的性能度量 5.6 调用 sklearn 实现聚类分析 6.1 间隔与支持向量 6.2 对偶问题 6.3 核函数 6.4 软间隔与正则化 6.5 支持向量机算法的 Python 实现 时间：1.16（9:00-18:00） 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 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	时间：1.17（9:00-18:00） PyTorch 深度学习原理与实现 1 引言 2 卷积神经网络 CNN 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 时间：1.18（9:00-18:00） 3 循环神经网络 RNN 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 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 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 5.5 模型训练 5.6 模型性能验证 时间：1.19（9:00-18:00） 自然语言处理实战 1 自然语言处理概述 1.1 自然语言处理概述 2 NLP 的基本流程 2.1 NLP 的基本流程 2.2 语料字符处理 2.3 分词和停用词处理 2.4 N 元语法模型 2.5 隐马尔可夫概述

Python 实现 4.8 网络输出的 Python 实现 4.9 单样本网络训练的 Python 实现 4.10 全样本网络训练的 Python 实现 4.11 网络性能评价 4.12 调用 sklearn 实现神经网络算法	4.9 模型性能评估 4.10 保存模型 4.11 加载模型 4.12 模型应用	时间：1.20（9:00-18:00） 2.6 jieba 分词 2.7 去停用词 3 文本向量化 3.2 词袋模型 3.3 TF-IDF 3.4 Word2Vec 模型 3.5 Doc2Vec 模型 特别内容 1. 金牌助手 ChatGPT, 让应用开发更高效
案例实战课程安排		
时间：1.21（9:00-18:00） 基于 textCNN 的公众健康问句分类 1 背景与目标 1.1 背景与目标 1.2 目标分析 2 数据探索与处理 2.1 数据探索 2.2 剔除无效字符及分词操作 2.3 读取停用词表 2.4 去除停用词 2.5 为词语编号 3 词嵌入（Word2Vec） 3.1 词嵌入（Word2Vec）介绍 3.2 获取目标词向量矩阵 3.3 保存处理好的数据 4 模型构建（textCNN） 4.1 统一各样本的词语数量 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	时间：1.22（9:00-18:00） 网络舆情与网民情绪波动分析 1 背景与目标 2 数据准备 2.1 数据基本介绍 2.2 词嵌入介绍 2.3 进行词向量训练 2.4 构建词向量矩阵 2.5 获取编码后的语料库 2.6 对各样本执行 padding 操作 2.7 拆分数据并将其转为模型所需格式 3 模型训练与性能验证 3.1 Embedding 层介绍 3.2 定义 Embedding 层算子 3.3 定义 LSTM 层算子 3.4 定义全连接层算子 3.5 定义网络计算流程 3.6 模型构建及配置 3.7 模型训练 3.8 模型性能评估 4 小结	时间：1.23（9:00-18:00） 推荐系统受众性别智能识别 1 背景与目标 1.1 项目目标 1.2 数据介绍 1.3 目标分析 2 数据探索与处理 2.1 数据读取与探索 2.2 获取用户点击流介绍 2.3 统计用户的 id 点击次数 2.4 获取用户点击流 2.5 提取并保存样本标签 2.6 封装数据处理函数 3 词嵌入 3.1 词嵌入准备 3.2 词向量训练 3.3 构建词向量矩阵 3.4 对点击流进行编码 3.5 统一点击流的长度 3.6 保存词向量及点击流数据 4 模型构建与训练 4.1 建模前数据准备 4.2 模型训练 4 模型性能评估与优化 5.1 模型性能评估 5.2 模型优化 时间：1.24（16:00-18:00） 职业技术考试 跳转至线上课程安排

附件三 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。

杨坦，中山大学博士，华南师范大学教师，泰迪杯数据挖掘挑战赛专家组成员。精通人工智能，模式识别，图像处理等专业技术；兼有大型高科技企业和高校的工作经历，主要从事推荐系统研究和开发、数据挖掘研究等方面的工作，并将研究成果应用于精准营销、交叉营销等场景取得了显著的成效；图书《MATLAB 数据分析与挖掘实战》、《R 语言与数据挖掘》的主要作者之一；多年大数据挖掘研究、应用经验和电信、电子商务、互联网等行业背景。

欧阳通达，中国科学院深圳先进技术研究院合成生物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。

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，广州市科技专家库入库专家、泰迪科技培训总监、高级信息系统项目管理师、华为昇思 MindSpore 布道师、华南农业大学校外硕士生导师。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，有丰富的人工智能 / 大数据授课经验，100 次以上担任全国高校大数据与人工智能师资培训主讲老师，500+场学生培训讲师经验，学员总数 50000+，主讲课程受到学员一致好评。

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，对数据分析和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“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。

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 大数据分析与应用》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师资研修班、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。