

全国高校大数据与人工智能师资研修班 邀请函

(2022 年第五期)

九大专题：商务数据分析实战、数据采集与处理实战 (Python)、大数据分析
与机器学习实战 (Python)、深度学习推荐系统实战、健康医疗大数据应用、大数
据技术应用实战 (Hadoop+Spark)、人工智能实战 (计算机视觉方向)、人工智能
实战 (自然语言处理方向)、智能语音实战

主办单位：泰迪杯数据挖掘挑战赛组委会

承办单位：广东泰迪智能科技股份有限公司

协办单位：人民邮电出版社有限公司

北京泰迪云智信息技术研究院

互联网、大数据、云计算、人工智能等现代信息技术深刻改变着人类的生产、生活、学习乃至思维方式，深刻展示了世界发展的前景。目前各院校的大数据和人工智能专业教师匮乏、相关落地动手实战应用能力欠缺、授课过程中相关行业实战案例项目缺失等，为加快建设大数据、人工智能相关专业教师队伍，推动各院校建立人才培训和评价体系，特推出全国高校大数据与人工智能师资研修班，每年在全国范围内滚动开展，截止目前已在全国巡回举办 50 余场，参训教师近 6000 人次。2022 年第五期全国高校大数据与人工智能师资研修班将开设九大专题方向，本期研修班以线上云课堂形式举办，现将有关安排通知如下。

一、课程特色

1、本研修班课程全程强调动手实操，内容以代码落地为主，以理论讲解为根，以公式推导为辅，通过讲解企业级案例，真正的让学员把所学内容和工作实际有效结合、更好地进行教育教学工作。

2、核心课程部分由讲师手把手一起进行实操演练，在具体应用场景中全面掌握相关技能，助力实训教学工作、实际动手的能力。视频制作精良，讲师真人出镜，全面解析专业必备技能，为相关课程开设和备课、应对科研和项目开发打下坚实基础。

3、课程设有答疑交流讨论群，培训期间助教全程辅助教学，每天提供 10 小时的实时在线答疑辅导，并进行答疑文档汇总，帮助学员更好地总结学习。

4、本课程配套有基础知识内容，即使零基础学员快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。

5、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！视频内容支持六个月内免费回看，以便复习和参考。

6、全面实践大数据/人工智能项目流程，包括数据采集、数据存储管理、数据探索、数据处理、特征工程、数据建模等课程，提供知识讲解，助力夯实理论基础，掌握核心技术。

7、参加线上课程学习的学员，如后续本人参加线下课程继续深造，持线上缴费凭证可享受一次免费学习机会（仅限参加同一培训专题方向的线下课程）。

二、课程安排

专题一 商务数据分析实战	学习时间：07月20日-07月29日，共计80学时
证书颁发：高级数据分析师职业技术证书	费用：1980元（报名费、学习费、资料费、证书费等）
学习环境要求：Windows10或以上操作系统（64位），不可使用Mac和linux系统	
课程模块：Excel 数据分析基础与实战、Power BI 数据分析与可视化、实战案例：财务分析在纳税评估中的应用（Excel）、实战案例：汽车大数据综合案例分析（Power BI）、实战案例：新零售智能售货机可视化项目（Power BI）、实战案例：学生校园卡消费行为分析（Excel）。 详见附件一 商务数据分析实战课程大纲	
专题二 数据采集与处理实战（Python）	学习时间：07月20日-07月29日，共计80学时
证书颁发：高级Python技术应用工程师职业技术证书	费用：1980元（报名费、学习费、资料费、证书费等）
学习环境要求：Windows7或以上操作系统（64位），4G+内存	
课程模块：Python 编程基础、Python 数据分析与应用、Pandas 数据分析基础、Python 数据可视化、Python 网络爬虫实战、实战案例：Python 爬虫助力疫情数据追踪、实战案例：某品牌手机的京东评论数据采集与分析、实战案例：泰迪内推平台招聘信息采集与分析、拓展自学篇：网站图像素材采集实战。 详见附件二 数据采集与处理实战（Python）课程大纲	

专题三 大数据分析机器学习实战 (Python)	学习时间：07月20日-07月29日，共计80学时
证书颁发：高级机器学习工程师职业技术证书	费用：1980元（报名费、学习费、资料费、证书费等）
学习环境要求：Windows7或以上操作系统（64位），4G+内存	
课程模块：Python编程基础、Python数据分析与应用、Pandas数据分析基础、Python数据可视化、Python机器学习实战、实战案例：学生校园消费行为分析、实战案例：运营商用户流失预测、实战案例：百货商场用户画像描绘与价值分析、实战案例：泰迪内推平台信息精准推荐应用（基于泰迪建模平台实现）。 详见附件三 大数据分析机器学习实战 (Python) 课程大纲	
专题四 深度学习推荐系统实战	学习时间：07月20日-07月30日，共计88学时
证书颁发：高级大数据分析师职业技术证书	费用：1980元（报名费、学习费、资料费、证书费等）
学习环境要求：Windows10或以上操作系统（64位），4G+内存	
课程模块：Python编程基础、Python数据分析与应用、Python机器学习实战、智能推荐算法、BP神经网络与TensorFlow2实战、TensorFlow2深度学习原理与实战、实战案例：020优惠券使用预测、实战案例：教育平台的线上课程智能推荐策略、实战案例：基于深度学习的推荐系统受众性别预测、实战案例：泰迪内推平台信息精准推荐应用（基于泰迪建模平台实现）。 详见附件四 深度学习推荐系统实战课程大纲	
专题五 健康医疗大数据应用	学习时间：07月20日-08月01日，共计104学时
证书颁发：高级大数据应用工程师职业技术证书	费用：1980元（报名费、学习费、资料费、证书费等）
学习环境要求：Windows10或以上操作系统（64位），4G+内存	
课程模块：Python编程基础、AI医疗概论、Python数据分析与应用、Python机器学习实战、BP神经网络与TensorFlow2实战、TensorFlow2深度学习原理与实战、实战案例：公众健康问句分类、实战案例：脑PET图像分析和疾病预测、实战案例：基于深度学习的肝脏肿瘤分割。 详见附件五 健康医疗大数据应用课程大纲	
专题六 大数据技术应用实战 (Hadoop+Spark)	学习时间：08月01日-08月15日，共计112学时
证书颁发：高级大数据技术应用职业技术证书	费用：2980元（报名费、学习费、资料费、证书费等）
学习环境要求：Windows10或以上操作系统（64位），8G+内存，100G+可用磁盘	
课程模块：Linux操作系统基础、Java编程基础、Hadoop大数据基础、Hive大数据仓库、综合实战：航空客户乘机数据预处理、HBase非关系型数据库、Scala编程基础、Spark大数据技术与应用、项目实战：广电大数据用户画像（Hadoop+Spark+Hive），拓展课程：Zookeeper分布式服务框架、Flume大数据采集与传输、Kafka大数据流处理、Flink大数据实时处理、商品实时推荐系统（Flume+Kafka+Flink）。 详见附件六 Hadoop+Spark 大数据技术应用实战课程大纲	
专题七 人工智能实战（计算机视觉方向）	学习时间：08月01日-08月12日，共计88学时
证书颁发：高级人工智能应用工程师职业技术证书	费用：2480元（报名费、学习费、资料费、证书费等）
学习环境要求：Windows10或以上操作系统（64位），4G+内存	
课程模块：Python编程基础、Python数据分析与应用、Python机器学习算法原理与实现、TensorFlow2实战、TensorFlow2深度学习原理与实现、计算机视觉实战、图像处理实战：脑PET图像分析和疾病预测、计算机视觉实战：动态人脸智能识别、实战案例：基于GAN的图像风格转换 详见附件七 人工智能实战（计算机视觉方向）课程大纲	
专题八 人工智能实战（自然语言处理方向）	学习时间：08月01日-08月12日，共计92学时
证书颁发：高级人工智能应用工程师职业技术证书	费用：2480元（报名费、学习费、资料费、证书费等）

学习环境要求： Windows10 或以上操作系统（64 位），4G+内存

课程模块：Python 编程基础、Python 数据分析与应用、Python 机器学习算法原理与实现、TensorFlow2 实战、TensorFlow2 深度学习原理与实现、自然语言处理实战、文本处理实战：公众健康问句分类、实战案例：搭建一个属于自己的聊天机器人、综合实战：基于深度学习的推荐系统受众性别预测。

详见附件八 人工智能实战（自然语言处理方向）课程大纲

专题九 智能语音实战

学习时间：08 月 01 日-08 月 11 日，共计 88 学时

证书颁发：高级大数据分析师职业技术证书

费用：2480 元（报名费、学习费、资料费、证书费等）

学习环境要求： Windows10 或以上操作系统（64 位），4G+内存

课程模块：Python 编程基础、Python 数据分析与应用、Python 机器学习算法原理与实现、TensorFlow2 实战、TensorFlow2 深度学习原理及实战、语音处理实战、实战案例：语音识别中的 HMM 声学模型、实战案例：英文数字语音识别、实战案例：中文语音识别。

详见附件九 智能语音实战课程大纲

三、师资介绍

郝志峰，汕头大学校长、教授、博士生导师。泰迪杯数据挖掘挑战赛组委会主任。教育部 2018-2022 高等学校大学数学课程教学指导委员会副主任委员。入选教育部“新世纪人才支持计划”、广东省“千百十”工程省级人选。郝志峰教授主要从事代数学及其应用、数学建模、教育信息化等领域的研究，先后主持 NSFC 联合基金、国家“新世纪人才支持计划”、国家自然科学基金、教育部优秀青年教师基金、教育部霍英东基金、广东省科技攻关重大项目、广东省自然科学基金、国家教育科学“十五”规划项目等省部级以上项目 20 余项，先后赴美国、英国、德国、日本、泰国和香港等地区访问讲学。曾获教育部自然科学奖二等奖、教育部（原国家教委）霍英东青年教师奖、广东省科技进步奖一等奖等奖项。2009 年主持的“大学数学立体化教育资源与集成系统的研究和实践”获国家优秀教学成果奖二等奖。

刘政永，吉林大学经济学学士，厦门大学经济学硕士，河北金融学院大数据科学学院副教授，科研处战略管理岗，长期从事数据挖掘分析工作，致力于运用 Python、R 进行金融数据挖掘，有较丰富的数据挖掘等理论知识与实践经验。现为河北金融学院大数据科学学院副教授，曾担任河北金融学院经济贸易系经济统计学专业任教教研室主任三年。近年来致力于金融数据挖掘。近三年来给金融硕士讲授《金融数据挖掘》课程，运用 Python 进行金融数据挖掘分析；给本科生讲授《Python 经济数据分析》、《国内外经济预测分析实训》等课程，运用 Python 进行经济数据分析、挖掘、预测等；参加天池 AI 实训平台师资培训、《python 金融大数据分析》培训；带领学生参加“泰迪杯”数据分析与挖掘比赛并获三等奖。从 2020 年

开始担任财经数据科学实战训练营讲师，讲授金融风控方面的实战案例。

樊老师，广东泰迪智能科技股份有限公司特聘讲师，某国际银行大数据开发工程师，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。

张敏，广东泰迪智能科技股份有限公司、培训总监，从事用户数据分析和数据挖掘工作六年，具有丰富的大数据挖掘理论及实践培训经验，对数据具有较高的敏感度，根据数据对其进行全面的统计分析。精通 Python、R 语言、MATLAB 等多种数据挖掘工具。擅长市场发展情况监控、精确营销方面的数据挖掘工作。有为南方电网、珠江数码等大型企业长期提供实施服务的经验，主导了电子商务网站用户行为分析及网页智能推荐服务、中医证型关联规则挖掘、电信业务话单量预测、航空公司客户价值分析等多个项目。2017 年“泰迪杯数据挖掘挑战赛教练员培训”主讲讲师，2018 年广东省 Python 与深度学习技术师资培训班主讲讲师，2019 年第一/二/三期全国高校大数据核心技术与应用师资研修班主讲讲师、2020 年第一/二/三/四/五/六期全国高校大数据与人工智能师资研修班主讲讲师，2020 年国培班主讲讲师，先后负责过西安理工大学、广东工业大学、广西师范学院、广西科技大学、闽江学院、广东石油化工学院、上海健康医学院等高校实训课程及德生科技等企业内训和数据挖掘就业班的课程。组织、参与编写图书《Python 编程基础》、《Python 数据分析与应用》、《R 语言编程基础》等。

律波，广东泰迪智能科技有限公司高级数据分析工程师，应用统计学硕士，有较强的统计学、数学、数据挖掘理论功底；精通 R、Python、Power BI、Excel 等数据挖掘分析工具，具有丰富的培训和项目经验，擅长从数据中发掘规律，对数据具有较高的敏感度，逻辑思维能力强，擅长数据可视化，机器学习、深度学习等算法原理的实现，如神经网络、SVM、决策树、贝叶斯等；负责“珠江数码大数据营销推荐应用”项目，完成标签库的构建及产品推荐模型；负责“京东电商产品评论情感分析”项目，完成了评论数据情感评价模型、LDA 主题模型的构建；通过项目案例的转换；负责多个本科类院校数据分析软件培训和毕业生数据分

析培训，先后负责广西科技大学、闽江学院、广东石油化工、韩山师范学院、广西师范大学等数据分析软件培训及实训等。多次负责“泰迪杯”数据挖掘大赛题目的构思和实现、赛前培训。大数据专业系列图书编写委员会成员，负责《R语言与数据挖掘》、《Python实训案例》、《Excel可视化案例》等书籍编写工作。

杨惠，广东泰迪智能科技股份有限公司高级数据分析师，从事人工智能工作多年，擅长计算机视觉和自然语言处理，熟悉常用深度学习算法原理及应用，如神经网络、SVM、强化学习等算法；精通TensorFlow、Python、MATLAB等常用数据挖掘处理工具。具有丰富的实践项目经验。如“智能聊天客服”项目，“车牌智能识别”项目，“京东商城产品评论情感分析”项目，“珠江数码大数据营销推荐应用”项目；“电子商务网站智能推荐服务”项目；“基于Seq2Seq注意力模型实现聊天机器人”项目。具备丰富的培训经验，曾为多家企业、院校服务过专业培训工作。如PPV商业培训、泰迪大数据师资培训、珠海城职院数据分析培训；2018年第一、三、五期全国高校大数据核心技术与应用师资研修班主讲讲师、2019年第一、三、五期全国高校大数据与人工智能师资研修班主讲讲师，2019年国家电网大数据竞赛河北省电力系统培训班主讲讲师。大数据专业系列图书编写委员会成员，负责《Python数据分析与应用》、《R语言编程基础》、《TensorFlow2深度学习实战》、《深度学习与计算机视觉实战》等书籍编写工作。

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“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。

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，广东泰迪智能科技股份有限公司特聘讲师。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。

周津，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 生态圈技术有深入理解，熟练掌握 Hadoop 环境部署和 Hadoop 核心计算框架 MapReduce 的原理和 API 应用，以及 HDFS 分布式文件系统存储结构。对 HBase、Hive 数据库有深刻理解。掌握 SparkCore、SparkSQL、SparkStreaming 原理及编程，熟悉 Spark 底层运行机制。曾深度参与江苏省纪委大数据平台与广东省公安厅电子数据取证大数据平台的开发和建设。

吴嘉泳，广东泰迪智能科技股份有限公司大数据研发工程师，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师资研修班和部分项目案例资源开发工作。

四、证书颁发

学员经在线培训并考试合格后，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、本期研修班两专题及以上联报者可享受九折优惠。
- 4、联系方式

联系人：曾爱枝

微 信：antonia602501

电 话：13246821827

邮 箱：zengaizhi@tipdm.com



全国高校大数据与人工智能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					发票内容		
纳税号					电子发票接收邮箱		
联系人			电 话			邮 箱	
以下表格中要求提供的信息为申报职业资格证书使用，请真实完整填写。							
姓 名	性 别	职 务	毕业院校	最高学历	手机号	电子邮箱	专题选择
费用支付方式	1、电汇到指定账号。 2、扫码支付（报名后联系工作人员索要支付码）。 3、付款时请注明“大数据研修班+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						

附件一 商务数据分析实战课程大纲

核心课程篇

时间	课程内容	学习平台
专题讲座		
07月20日 18:30-22:00	主讲：郝志峰 主题内容：大数据视角下数字孪生与元宇宙的思考	泰迪云课堂
第一课 Excel 数据分析基础与实战		
07月20日 18:30-22:00	1.1 认识数据分析 1.2 认识 Excel 2016 2.1 获取文本数据 2.2 从数据库获取数据 3.1 排序 3.2 筛选 3.3 分类汇总 4.1 认识公式和函数 4.2 数组公式 4.3 日期和时间函数 4.4 数学函数 4.5 统计函数 4.6 文本函数 4.7 逻辑函数 5.1 透视表的创建和修改 5.2 透视表的操作 5.3 透视图的操作	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
07月21日 18:30-22:00	6.1 对比分析 6.2 趋势分析 6.3 饼图 6.4 散点图 6.5 雷达图 7.1 案例背景 7.2 数据预处理 8 商品销售分析 9 库存分析 10 用户分析	泰迪云课堂

	11 分析报告	
	操作演练	个人 PC
	在线答疑	微信群
第二课 Power BI 数据分析与可视化		
07 月 22 日 18:30-22:00	1 数据分析与可视化概述 1.1 认识数据分析 (Power BI) 1.2 常用数据可视化软件 1.3 认识 Power BI 2 数据获取 2.1 获取数据的方式 2.2 数据获取 3 M 语言数据建模与处理 3.1 编辑器和 M 语言 3.2 获取网络分页数据 3.3 清洗数据 3.3 数据集成 3.4 数据转换 3.5 规约数据	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
07 月 23 日 18:30-22:00	4 DAX 语言数据处理 4.1 DAX 语言 4.2 DAX 语言处理表间关系 4.3 DAX 函数使用实例 4.4. 数据查询操作 5 数据分析可视化 5.1 可视化 5.2 认识对比分析 5.3 对比分析图表绘制操作	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
07 月 24 日 18:30-22:00	5.4 认识结构分析 5.5 结构分析可视化操作 5.6 相关分析 5.7 相关分析可视化操作 5.8 描述性分析 5.9 描述性分析可视化操作 5.10 KPI 图表分析	泰迪云课堂

	<ul style="list-style-type: none"> 5.11KPI 分析可视化操作 6 数据分析报表 <ul style="list-style-type: none"> 6.1 认识 Power BI 报表 6.2 完整的分析报表：会员数据分析 6.3 完整分析报表的操作 7 Power BI 移动版数据部署 <ul style="list-style-type: none"> 7.1 移动版发布一份分析报表 7.2 仪表板的使用 	
	操作演练	个人 PC
	在线答疑	微信群

第三课 财务分析在纳税评估中的应用 (Excel)

07月25日 18:30-22:00	<ul style="list-style-type: none"> 1 案例背景与目标 <ul style="list-style-type: none"> 1.1 案例背景 2 获取和整理财务数据 <ul style="list-style-type: none"> 2.1 创建财务分析和纳税评估文件 2.2 获取财务数据分析的指标数据 2.3 整理财务分析的指标数据 2.4 采购渠道提取 3 财务指标分析和税务评估 <ul style="list-style-type: none"> 3.1 数据探索分析 3.2 共同比分析 3.3 趋势分析 3.4 财务比率分析 3.5 确定重点评估领域 3.6 重点评估领域审计 3.7 问题的发现及财务报表的调整 4 小结 	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群

第四课 实战案例：汽车大数据综合案例分析 (Power BI)

07月26日 18:30-22:00	<ul style="list-style-type: none"> 1 了解汽车销售行业现状 2 数据预处理 <ul style="list-style-type: none"> 2.1 数据清洗 2.2 数据归约 2.3 数据建模 3 数据分析可视化 <ul style="list-style-type: none"> 3.1 区域汽车销售分析及可视化 3.2 各种类别的汽车销售分析及可视化 	泰迪云课堂
-----------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------

	4 发布到 web 端	
	操作演练	个人 PC
	在线答疑	微信群
第五课 实战案例：新零售智能售货机可视化项目（Power BI）		
07 月 27 日 18:30-22:00	1 案例背景 2 数据预处理 2.1 数据读取与数据清洗 2.2 数据规约 2.3 数据建模 3 销售分析及可视化 4 库存分析和可视化 5 用户分析和可视化 6 数据部署	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
第六课 实战案例：学生校园卡消费行为分析（Excel）		
07 月 28 日 18:30-22:00	1 案例背景与目标 1.1 案例背景解读 1.2 挖掘目标分析 2 数据预处理 2.1 预处理：读取数据和异常值 2.2 预处理：缺失值 2.3 预处理：重复值与合并数据 3 数据分析 3.1 食堂消费数据分析 3.2 学生消费行为分析 4 小结	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
第七课 在线考试		
07 月 29 日 19:00-21:00	高级数据分析师职业技术证书在线考试	泰迪云课堂

附件二 数据采集与处理实战（Python）课程大纲

基础篇（报名成功后即可开始学习）

时间	课程内容	学习平台
正式培训前	<p>Python 编程基础</p> <p>1 准备工作</p> <p>2 列表操作</p> <p>3 程序流程控制语句</p> <p>4 字符串操作</p> <p>4.1 字符串及其索引&切片</p> <p>4.2 字符串的常见方法</p> <p>4.3 字典的创建及索引</p> <p>4.4 字典常用操作</p> <p>4.5 字典推导式</p> <p>5 Python 文件读取操作</p> <p>5.1 Python 读取文件</p> <p>5.2 练习 3: 统计小说中的单词频次</p> <p>6 函数</p> <p>6.1 Python 函数自定义</p> <p>6.2 练习 4: 自定义求序列偶数个数的函数</p> <p>7 面向对象与模块</p> <p>7.1 Python 方法与函数对比介绍</p> <p>7.2 Python 面向对象示例</p> <p>7.3 Python 模块使用</p> <p>7.4 第三方库的安装与调用</p> <p>8 注意事项</p> <p>8.1 Python 工作路径说明</p> <p>8.2 模块命名及存放路径的注意事项</p> <p>8.3 结语</p>	泰迪云课堂
正式培训前	<p>Python 数据分析与应用</p> <p>1 Python 数据分析概述</p> <p>1.1 认识数据分析</p> <p>1.2 熟悉 Python 数据分析的工具</p> <p>1.3 安装 anaconda 与启动 Jupyter Notebook</p> <p>1.4 掌握 Jupyter Notebook 常用功能</p> <p>2 NumPy 数值计算基础</p> <p>2.1 掌握 NumPy 数组对象</p> <p>2.1.1 NumPy 简介</p> <p>2.1.2 数组创建及基础属性</p> <p>2.1.3 初识数组的特点</p> <p>2.1.4 创建常用数组</p>	泰迪云课堂

	2.1.5 数组数据类型 2.1.6 生成随机数 2.1.7 一维数组的索引 2.1.8 逻辑型索引 2.1.9 多维数组的索引 2.1.10 求解距离矩阵 2.1.11 变化数组 shape 2.2 掌握 NumPy 矩阵与通用函数 2.2.1 NumPy 矩阵介绍 2.2.2 NumPy 通用函数介绍 2.2.3 通用函数的广播机制 2.3 利用 NumPy 进行统计分析 2.3.1 NumPy 读写二进制文件 2.3.2 NumPy 读写 txt 文件 2.3.3 利用 NumPy 对数据进行简单统计分析	
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

核心课程篇

时间	课程内容	学习平台
专题讲座		
07月20日 18:30-22:00	主讲：郝志峰 讲座主题：大数据视角下数字孪生与元宇宙的思考	泰迪云课堂
第一课 Pandas 数据分析基础		
07月20日 18:30-22:00	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 1.2.1 Pandas 读取文本数据 1.2.2 存储数据框 1.2.3 Pandas 读取 excel 文件 1.2.4 将数据框存储为 excel 文件 1.3 数据框与数据框元素 1.3.1 构建数据框 1.3.2 查看数据框的常用属性 1.3.3 按行列顺序访问数据框中的元素 1.3.4 按行列名称访问数据框中的元素 1.3.5 修改数据框中的元素 1.3.6 删除数据框中的元素 1.3.7 描述分析数据框中的元素 1.4 转换与处理时间序列数据	泰迪云课堂

	<ul style="list-style-type: none"> 1.4.1 转换成时间类型数据 1.4.2 时间类型数据的常用操作 1.5 使用分组聚合进行组内计算 <ul style="list-style-type: none"> 1.5.1 groupby 分组操作 1.5.2 agg 聚合操作 1.6 创建透视表与交叉表 <ul style="list-style-type: none"> 1.6.1 生成透视表 1.6.2 生成交叉表 	
	操作演练	个人 PC
	在线答疑	微信群
07月21日 18:30-22:00	<ul style="list-style-type: none"> 2 使用 Pandas 进行数据预处理 <ul style="list-style-type: none"> 2.1 合并数据 <ul style="list-style-type: none"> 2.1.1 表堆叠 2.1.2 主键合并 2.1.3 重叠合并 2.2 清洗数据 <ul style="list-style-type: none"> 2.2.1 检测与处理重复值 2.2.2 检测与处理缺失值 2.2.3 检测与处理异常值 2.3 标准化数据 2.4 转换数据 <ul style="list-style-type: none"> 2.4.1 哑变量处理 2.4.2 离散化连续型数据 	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
第二课 Python 数据可视化		
07月22日 18:30-22:00	<ul style="list-style-type: none"> 1.1 Matplotlib 绘制流程说明 1.2 添加文本和修改绘图风格 1.3 rc 参数 1.4 散点图 1.5 折线图 1.6 直方图和条形图 1.7 饼图 1.8 箱线图 1.9 人口特征间分布 1.10 人口各个特征分布 	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群

拓展自学篇		
自行安排	2.1 seaborn 基础介绍 2.2 seaborn 简单绘图 2.3 seaborn 绘图风格 2.4 调色板 1 2.5 调色板 2 2.6 关系图 2.7 分类图 2.8 分布图 2.9 回归图 2.10 矩阵图 2.11 网格图 3.1Pyecharts 基础介绍 3.2Pyecharts 绘制日历图 3.3Pyecharts 绘制漏斗图 3.4Pyecharts 绘制仪表盘 3.5Pyecharts 绘制水球图 3.6Pyecharts 绘制关系图 3.7Pyecharts 绘制直角坐标系图表 3.8Pyecharts 绘制饼图 3.9Pyecharts 绘制雷达图 3.10Pyecharts 绘制词云图 3.11Pyecharts 绘制柱状图 3.12Pyecharts 绘制树形图 3.13Pyecharts 绘制地理图表	泰迪云课堂
第三课 Python 网络爬虫实战		
07 月 23 日 18:30-22:00	1Python 爬虫环境与爬虫简介 1.1 认识爬虫 1.1.1 常见网上冲浪过程 1.1.2 爬虫的概念 1.1.3 爬虫合法性 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 2.1.1 网络传输模型 2.1.2 网络信息传输过程 2.2 认识 HTTP	泰迪云课堂

	<ul style="list-style-type: none"> 2.2.1 认识 HTTP 2.2.2 熟悉 Cookie 3 简单静态网页爬取 3.1 认识静态网页 3.2 实现 HTTP 请求 3.2.1 创建工程 3.2.2 生成 HTTP 请求 3.2.3 完善 HTTP 请求 3.3 解析网页 3.3.1 为什么要解析网页 3.3.2 初识 Xpath 3.3.3 Xpath 相对路径及属性查找 3.3.4 使用 BeautifulSoup 解析网页 3.3.5 网页解析小结 3.4 存储数据 3.4.1 认识 chrome 开发者工具 3.4.2 任务演练：爬取并存储泰迪科技官网首页数据 	
	操作演练	个人 PC
	在线答疑	微信群
<p>07 月 24 日 18:30-22:00</p>	<ul style="list-style-type: none"> 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 4.2.1 通过网页源码追踪目标数据文件地址 4.2.2 通过开发者工具追踪目标数据文件地址 4.2.3 爬取数据并进行保存 4.3 使用 Selenium 库爬取动态网页 4.3.1 搭建 Selenium 环境 4.3.2 利用 Selenium 获取网页数据 4.3.3 利用 Selenium 控制点击操作 5 模拟登录 5.1 使用表单登录方法实现模拟登录 5.1.1 模拟登录的过程 5.1.2 查找提交入口和表单数据 5.1.3 提交表单完成模拟登录 5.1.4 使用表单登录的注意事项 5.2 使用 Cookie 登录方法实现模拟登录 5.3 使用 Selenium 模拟登录 	<p>泰迪云课堂</p>
	操作演练	个人 PC

	在线答疑	微信群
07月25日 18:30-22:00	6 Scrapy 爬虫 6.1 认识 Scrapy 6.2 通过 Scrapy 爬取基本页面信息 6.2.1 创建项目 6.2.2 指定字段及创建 spiders 6.2.3 完成 spiders 编写 6.2.4 运行程序保存数据 6.3 通过 Scrapy 抓取跳转页面数据 6.3.1 任务介绍及项目创建 6.3.2 获取所有页面的 url 6.3.3 获取每个页面的新闻二次页面 url 6.3.4 提取各新闻二次页面中的目标数据 6.3.5 运行程序保存数据 7. 拓展：终端协议及爬取工具介绍	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
	第四课 数据采集与处理实战：Python 爬虫助力疫情数据追踪	
07月26日 18:30-22:00	1.1 案例背景与挖掘目标 2 获取疫情数据 2.1 了解 Python 获取网络数据方法 2.2 了解本案例数据获取的方法 2.3 获取实时疫情页面网页源码 2.4 解析 JSON 数据 2.5 爬取国内疫情数据 2.6 爬取国外疫情数据 2.7 单页数据整理 3 疫情数据分析 3.1 国内数据各地区对比 3.2 国内数据随时间对比 3.3 全国数据各地区对比 3.4 全国数据随时间对比 3.5 任务实现 3.5.1 特征处理 3.5.2 各国每日死亡率 3.5.3 各国每日治愈率 3.5.4 各国每日新增确诊案例 3.5.5 各国每日新增死亡案例	泰迪云课堂

	3.5.6 各国每日新增治愈案例 4 疫情数据可视化及前端部署 4.1 疫情数据接口整合 4.2 疫情可视化前端实现 4.3 疫情可视化后端实现及本地部署 4.4 不同评分数量与时间的关系 5 小结	
	操作演练	个人 PC
	在线答疑	微信群
第五课 数据采集与处理实战：某品牌手机的京东评论数据采集与分析		
07月27日 18:30-22:00	1 项目背景与挖掘目标 2.1 通过谷歌开发者工具获取 URL 2.2 发送 HTTP 请求与转换成 JSON 数据 2.3 提取信息并保存到本地 2.4 循环采集数据：保存数据 2.4 循环采集数据：自定义函数 3.1 数据预处理 3.2 数据预处理实现 4.1 评论数据词云绘制 4.2 好评、差评词云绘制与分析 4.3 不同颜色商品购买比例分析 4.4 不同配置商品购买比例分析 4.5 销售数量和评论数量和日期的关系 4.6 销售数量和评论数量和时间的关系 4.7 购买与评论时间间隔统计分析 4.8 不同渠道的销售比例	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
第六课 数据采集与处理实战：泰迪内推平台招聘信息采集与分析		
07月28日 18:30-22:00	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	泰迪云课堂

	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	
	操作演练	个人 PC
	在线答疑	微信群
第七课 在线考试		
07月29日 19:00-21:00	高级 Python 技术应用职业技术证书在线考试	泰迪云课堂

附件三 大数据分析 with 机器学习实战 (Python) 课程大纲

基础篇 (报名成功后即可开始学习)

时间	课程内容	学习平台
正式培训前	Python 编程基础 1 准备工作 2 列表操作 3 程序流程控制语句 4 字符串操作 4.1 字符串及其索引&切片 4.2 字符串的常见方法 4.3 字典的创建及索引 4.4 字典常用操作 4.5 字典推导式 5 Python 文件读取操作 5.1 Python 读取文件 5.2 练习 3: 统计小说中的单词频次 6 函数 6.1 Python 函数自定义 6.2 练习 4: 自定义求序列偶数个数的函数 7 面向对象与模块 7.1 Python 方法与函数对比介绍	泰迪云课堂

	<p>7.2 Python 面向对象示例</p> <p>7.3 Python 模块使用</p> <p>7.4 第三方库的安装与调用</p> <p>8 注意事项</p> <p>8.1 Python 工作路径说明</p> <p>8.2 模块命名及存放路径的注意事项</p> <p>8.3 结语</p>	
正式培训前	<p>Python 数据分析与应用</p> <p>1 Python 数据分析概述</p> <p>1.1 认识数据分析</p> <p>1.2 熟悉 Python 数据分析的工具</p> <p>1.3 安装 anaconda 与启动 Jupyter Notebook</p> <p>1.4 掌握 Jupyter Notebook 常用功能</p> <p>2 NumPy 数值计算基础</p> <p>2.1 掌握 NumPy 数组对象</p> <p>2.1.1 NumPy 简介</p> <p>2.1.2 数组创建及基础属性</p> <p>2.1.3 初识数组的特点</p> <p>2.1.4 创建常用数组</p> <p>2.1.5 数组数据类型</p> <p>2.1.6 生成随机数</p> <p>2.1.7 一维数组的索引</p> <p>2.1.8 逻辑型索引</p> <p>2.1.9 多维数组的索引</p> <p>2.1.10 求解距离矩阵</p> <p>2.1.11 变化数组 shape</p> <p>2.2 掌握 NumPy 矩阵与通用函数</p> <p>2.2.1 NumPy 矩阵介绍</p> <p>2.2.2 NumPy 通用函数介绍</p> <p>2.2.3 通用函数的广播机制</p> <p>2.3 利用 NumPy 进行统计分析</p> <p>2.3.1 NumPy 读写二进制文件</p> <p>2.3.2 NumPy 读写 txt 文件</p> <p>2.3.3 利用 NumPy 对数据进行简单统计分析</p>	泰迪云课堂

核心课程篇

时间	课程内容	学习平台
专题讲座		

07月20日 18:30-22:00	主讲：郝志峰 讲座主题：大数据视角下数字孪生与元宇宙的思考	泰迪云课堂
第一课 Pandas 数据分析基础		
07月20日 18:30-22:00	<ul style="list-style-type: none"> 1 Pandas 统计分析基础 <ul style="list-style-type: none"> 1.1 Pandas 简介 1.2 读写不同数据源的数据 <ul style="list-style-type: none"> 1.2.1 Pandas 读取文本数据 1.2.2 存储数据框 1.2.3 Pandas 读取 excel 文件 1.2.4 将数据框存储为 excel 文件 1.3 数据框与数据框元素 <ul style="list-style-type: none"> 1.3.1 构建数据框 1.3.2 查看数据框的常用属性 1.3.3 按行列顺序访问数据框中的元素 1.3.4 按行列名称访问数据框中的元素 1.3.5 修改数据框中的元素 1.3.6 删除数据框中的元素 1.3.7 描述分析数据框中的元素 1.4 转换与处理时间序列数据 <ul style="list-style-type: none"> 1.4.1 转换成时间类型数据 1.4.2 时间类型数据的常用操作 1.5 使用分组聚合进行组内计算 <ul style="list-style-type: none"> 1.5.1 groupby 分组操作 1.5.2 agg 聚合操作 1.6 创建透视表与交叉表 <ul style="list-style-type: none"> 1.6.1 生成透视表 1.6.2 生成交叉表 2 使用 Pandas 进行数据预处理 <ul style="list-style-type: none"> 2.1 合并数据 <ul style="list-style-type: none"> 2.1.1 表堆叠 2.1.2 主键合并 2.1.3 重叠合并 2.2 清洗数据 <ul style="list-style-type: none"> 2.2.1 检测与处理重复值 2.2.2 检测与处理缺失值 2.2.3 检测与处理异常值 2.3 标准化数据 2.4 转换数据 	泰迪云课堂

	2.4.1 哑变量处理	
	2.4.2 离散化连续型数据	
	操作演练	个人 PC
	在线答疑	微信群
第二课 Python 数据可视化		
07月21日 18:30-22:00	1.1 Matplotlib 绘制流程说明	泰迪云课堂
	1.2 添加文本和修改绘图风格	
	1.3 rc 参数	
	1.4 散点图	
1.5 折线图		
1.6 直方图和条形图		
1.7 饼图		
1.8 箱线图		
1.9 人口特征间分布		
1.10 人口各个特征分布		
	操作演练/作业	个人 PC
	在线答疑	微信群
拓展自学篇		
自行安排	2.1 seaborn 基础介绍	泰迪云课堂
	2.2 seaborn 简单绘图	
	2.3 seaborn 绘图风格	
	2.4 调色板 1	
	2.5 调色板 2	
	2.6 关系图	
	2.7 分类图	
	2.8 分布图	
	2.9 回归图	
	2.10 矩阵图	
	2.11 网格图	
	3.1 Pyecharts 基础介绍	
	3.2 Pyecharts 绘制日历图	
	3.3 Pyecharts 绘制漏斗图	
	3.4 Pyecharts 绘制仪表盘	
	3.5 Pyecharts 绘制水球图	
	3.6 Pyecharts 绘制关系图	
	3.7 Pyecharts 绘制直角坐标系图表	
	3.8 Pyecharts 绘制饼图	
3.9 Pyecharts 绘制雷达图		

	3.10Pyecharts 绘制词云图 3.11Pyecharts 绘制柱状图 3.12Pyecharts 绘制树形图 3.13Pyecharts 绘制地理图表	
第三课 Python 机器学习实战		
07月22日 18:30-22:00	1.1 引言 1.2 基本术语 1.3 假设空间&归纳偏好 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 3.1 线性回归基本形式 3.2 线性回归模型的 Python 实现 3.3 波士顿房价预测的 Python 实现 3.4 逻辑回归介绍 3.5 研究生入学录取预测的 Python 实现	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
07月23日 18:30-22:00	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 5.1 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 scikit-learn 实现聚类分析	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
07月24日 18:30-22:00	6.1 单个神经元介绍 6.2 经典网络结构介绍 6.3 神经网络工作流程演示 6.4 如何修正网络参数-梯度下降法 6.5 网络工作原理推导	泰迪云课堂

6.6 网络搭建准备	
6.7 样本从输入层到隐层传输的 Python 实现	
6.8 网络输出的 Python 实现	
6.9 单样本网络训练的 Python 实现	
6.10 全样本网络训练的 Python 实现	
6.11 网络性能评价	
6.12 调用 scikit-learn 实现神经网络算法	
操作演练/作业	个人 PC
在线答疑	微信群

拓展自学篇

自行安排	<p>7 支持向量机</p> <p>7.1 间隔与支持向量</p> <p>7.2 对偶问题</p> <p>7.3 核函数</p> <p>7.4 软间隔与正则化</p> <p>7.5 支持向量机算法的 Python 实现</p> <p>8 集成学习</p> <p>8.1 集成学习基本概念</p> <p>8.2 并行集成算法-Bagging&RandomForest</p> <p>8.3 串行集成算法-Boosting 算法流程</p> <p>8.4 串行集成算法-Boosting 代码实现</p> <p>8.5 Stacking 算法流程</p> <p>8.6 Stacking 代码实现</p>	泰迪云课堂
------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------

第四课 实战案例：学生校园消费行为分析

07月25日 18:30-22:00	<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 缺失值处理</p> <p>2.4 异常值处理</p> <p>2.5 数据合并</p> <p>3 数据可视化分析</p> <p>3.1 食堂就餐行为分析</p> <p>3.2 学生早中晚餐就餐地点占比饼图</p> <p>3.3 工作日和非工作日食堂就餐时间曲线图</p> <p>3.4 不同专业、不同性别的学生消费分析</p> <p>4 模型构建</p> <p>4.1 构建聚类模型</p>	泰迪云课堂
-----------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------

	4.2 模型性能评估 5 小结	
	操作演练/作业	个人 PC
	在线答疑	微信群
第五课 实战案例：运营商流失用户分析与预测		
07月26日 18:30-22:00	1.1 背景与目标	泰迪云课堂
	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	
	操作演练/作业	个人 PC
	在线答疑	微信群
第六课 实战案例：百货商场用户画像描绘与价值分析		
07月27日 18:30-22:00	1.1 背景与分析目标	泰迪云课堂
	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 K-Means 算法实现会员聚类 5.3 结果分析	
	操作演练/作业	个人 PC
	在线答疑	微信群
第七课 实战案例：泰迪内推平台信息精准推荐应用（基于泰迪建模平台实现）		
07月28日 18:30-22:00	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	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第八课 在线考试		
07月29日 19:00-21:00	高级机器学习工程师职业技术证书在线考试	泰迪云课堂

附件四 深度学习推荐系统实战课程大纲

基础篇（报名成功后即可开始学习）

时间	课程内容	学习平台
正式培训前	Python 编程基础 1 准备工作 2 列表操作	泰迪云课堂

	<ul style="list-style-type: none"> 3 程序流程控制语句 4 字符串操作 <ul style="list-style-type: none"> 4.1 字符串及其索引&切片 4.2 字符串的常见方法 4.3 字典的创建及索引 4.4 字典常用操作 4.5 字典推导式 5 Python 文件读取操作 <ul style="list-style-type: none"> 5.1 Python 读取文件 5.2 练习 3: 统计小说中的单词频次 6 函数 <ul style="list-style-type: none"> 6.1 Python 函数自定义 6.2 练习 4: 自定义求序列偶数个数的函数 7 面向对象与模块 <ul style="list-style-type: none"> 7.1 Python 方法与函数对比介绍 7.2 Python 面向对象示例 7.3 Python 模块使用 7.4 第三方库的安装与调用 8 注意事项 <ul style="list-style-type: none"> 8.1 Python 工作路径说明 8.2 模块命名及存放路径的注意事项 8.3 结语 	
正式培训前	<p>Python 数据分析与应用</p> <ul style="list-style-type: none"> 1 Python 数据分析概述 <ul style="list-style-type: none"> 1.1 认识数据分析 1.2 熟悉 Python 数据分析的工具 1.3 安装 anaconda 与启动 Jupyter Notebook 1.4 掌握 Jupyter Notebook 常用功能 2 NumPy 数值计算基础 <ul style="list-style-type: none"> 2.1 掌握 NumPy 数组对象 <ul style="list-style-type: none"> 2.1.1 NumPy 简介 2.1.2 数组创建及基础属性 2.1.3 初识数组的特点 2.1.4 创建常用数组 2.1.5 数组数据类型 2.1.6 生成随机数 2.1.7 一维数组的索引 2.1.8 逻辑型索引 	泰迪云课堂

- 2.1.9 多维数组的索引
- 2.1.10 求解距离矩阵
- 2.1.11 变化数组 shape
- 2.2 掌握 NumPy 矩阵与通用函数
 - 2.2.1 NumPy 矩阵介绍
 - 2.2.2 NumPy 通用函数介绍
 - 2.2.3 通用函数的广播机制
- 2.3 利用 NumPy 进行统计分析
 - 2.3.1 NumPy 读写二进制文件
 - 2.3.2 NumPy 读写 txt 文件
 - 2.3.3 利用 NumPy 对数据进行简单统计分析
- 3 Matplotlib 数据可视化基础
 - 3.1 掌握绘图基础语法与常用参数
 - 3.1.1 Matplotlib 介绍
 - 3.1.2 基础图形绘制
 - 3.1.3 常用参数设置
 - 3.2 分析特征间关系
 - 3.2.1 绘制散点图
 - 3.2.2 散点图参数设置
 - 3.2.3 绘制折线图
 - 3.3 分析特征内部数据分布与分散情况
 - 3.3.1 绘制直方图
 - 3.3.2 绘制饼图
 - 3.3.3 绘制箱线图
- 4 Pandas 统计分析基础
 - 4.1 Pandas 简介
 - 4.2 读写不同数据源的数据
 - 4.2.1 Pandas 读取文本数据
 - 4.2.2 存储数据框
 - 4.2.3 Pandas 读取 excel 文件
 - 4.2.4 将数据框存储为 excel 文件
 - 4.3 数据框与数据框元素
 - 4.3.1 构建数据框
 - 4.3.2 查看数据框的常用属性
 - 4.3.3 按行列顺序访问数据框中的元素
 - 4.3.4 按行列名称访问数据框中的元素
 - 4.3.5 修改数据框中的元素
 - 4.3.6 删除数据框中的元素

<ul style="list-style-type: none"> 4.3.7 描述分析数据框中的元素 4.4 转换与处理时间序列数据 <ul style="list-style-type: none"> 4.4.1 转换成时间类型数据 4.4.2 时间类型数据的常用操作 4.5 使用分组聚合进行组内计算 <ul style="list-style-type: none"> 4.5.1 groupby 分组操作 4.5.2 agg 聚合操作 4.6 创建透视表与交叉表 <ul style="list-style-type: none"> 4.6.1 生成透视表 4.6.2 生成交叉表 5 使用 Pandas 进行数据预处理 <ul style="list-style-type: none"> 5.1 合并数据 <ul style="list-style-type: none"> 5.1.1 表堆叠 5.1.2 主键合并 5.1.3 重叠合并 5.2 清洗数据 <ul style="list-style-type: none"> 5.2.1 检测与处理重复值 5.2.2 检测与处理缺失值 5.2.3 检测与处理异常值 5.3 标准化数据 5.4 转换数据 <ul style="list-style-type: none"> 5.4.1 哑变量处理 5.4.2 离散化连续型数据 	
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

核心课程篇

时间	课程内容	学习平台
专题讲座		
07月20日 18:30-22:00	主讲：郝志峰 讲座主题：大数据视角下数字孪生与元宇宙的思考	泰迪云课堂
第一课 Python 机器学习算法原理与实现		
07月20日 18:30-22:00	<ul style="list-style-type: none"> 1 机器学习绪论 <ul 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 <ul 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 	泰迪云课堂

	2.4 性能度量 Python 实现	
	操作演练	个人 PC
	在线答疑	微信群
07 月 21 日 18:30-22:00	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 3.3 波士顿房价预测的 Python 实现 3.4 逻辑回归介绍 3.5 研究生入学录取预测的 Python 实现 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	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
拓展自学篇		
自行安排	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 集成学习 6.1 集成学习基本概念 6.2 并行集成算法-Bagging&RandomForest 6.3 串行集成算法-Boosting 算法流程 6.4 串行集成算法-Boosting 代码实现 6.5 Stacking 算法流程 6.6 Stacking 代码实现	泰迪云课堂
第二课 智能推荐算法		
07 月 22 日 18:30-22:00	1 智能推荐系统简介 1.1 智能推荐简介 1.2 基于内容的推荐 1.3 协同过滤推荐 1.4 基于深度学习的推荐&复合推荐 1.5 智能推荐系统性能度量	泰迪云课堂

2 关联规则	
2.1 关联规则简介	
2.2 事务和项集	
2.3 支持度与置信度	
2.4 Apriori 算法流程	
2.5 Apriori 算法实现	
3 协同过滤	
3.1 协同过滤推荐算法流程	
3.2 新闻推荐数据介绍	
3.3 数据准备	
3.4 模型构建	
3.5 构建测试集用户网址浏览字典	
3.6 执行推荐操作	
3.7 推荐结果评估	
4 FM&Ranking 模型	
操作演练	个人 PC
在线答疑	微信群

第三课 BP 神经网络与 TensorFlow2 实战

07 月 23 日 18:30-22:00	1 BP 神经网络	
	1.1 单个神经元介绍	
	1.2 经典网络结构介绍	
	1.3 神经网络工作流程演示	
	1.4 如何修正网络参数. 梯度下降	
	1.5 网络工作原理推导	
	1.6 网络搭建准备	泰迪云课堂
	1.7 样本从输入层到隐层传输的 Python 实现	
	1.8 网络输出的 Python 实现	
	1.9 单样本网络训练的 Python 实现	
	1.10 全样本网络训练的 Python 实现	
	1.11 网络性能评价	
	1.12 调用 scikit-learn 实现神经网络算法	
	操作演练/作业	个人 PC
	在线答疑	微信群
07 月 24 日 18:30-22:00	2 任务 1: 构建一个线性模型	
	2.1 TensorFlow2 介绍	
	2.2 TensorFlow2 常用数据类型和操作	泰迪云课堂
	2.3 初始化模型	
	2.4 构建损失函数	

	2.5 模型训练及可视化 2.6 使用高阶 API-Keras 3 任务 2: MNIST 手写数字识别 3.1 数据读取及探索 3.2 交叉熵 3.3 模型构建及训练 3.4 调用保存好的模型对新样本进行预测	
	操作演练/作业	个人 PC
	在线答疑	微信群
第四课 TensorFlow 2 深度学习原理与实战		
07 月 25 日 18:30-22:00	1.1 深度神经网络-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 实现池化操作 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	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第五课 实战案例: O2O 优惠券使用预测		
07 月 26 日 18:30-22:00	1 背景与目标 2.1 线下训练集数据介绍 2.2 线上训练集数据介绍 2.3 测试数据介绍 2.4 项目流程介绍 3.1 构建正样本 3.2 构建负样本 3.3 构建样本标签 4.1 特征构建介绍 4.2 处理 Discount_rate 列 4.3 特征 1-折扣率 4.4 特征 2-商户与用户之间的距离 5.1 建模前数据准备	泰迪云课堂

	5.2 初级模型构建 5.3 ROC 曲线与 AUC 值 5.4 模型性能评估 5.5 训练函数封装 5.6 模型预测 5.7 预测函数封装 6.1 特征 3-优惠券流行度 6.2 特征 4-用户在商家中的消费次数 6.3 如何进行特征拼接 6.4 拼接训练集的特征 6.5 拼接测试集的特征 7.1 模型训练 7.2 预测 7.3 代码整理 7.4 结果提交	
	操作演练/作业	个人 PC
	在线答疑	微信群
第六课 实战案例：教育平台的线上课程智能推荐策略		
07 月 27 日 18:30-22:00	1 数据读取 2 数据探索与预处理 2.1 缺失值处理 2.2 异常值处理 2.3 重复值处理 3 统计分析与可视化 3.1 用户分布分析 3.2 用户活跃度分析 3.3 用户流失率分析 3.4 用户参与课程情况分析 3.5 用户选择课程情况分析 3.6 课程收费差异与用户学习进度相关性分析 4 智能推荐 4.1 线上课程智能推荐 4.2 线上管理决策建议	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第七课 实战案例：基于深度学习的推荐系统受众性别预测		
07 月 28 日 18:30-22:00	1 项目文件结构及编辑环境介绍 2 数据获取与探索分析答案讲解	泰迪云课堂

	<ul style="list-style-type: none"> 3.1 理解用户单击流相关概念 3.2 获取用户的各单击流数据 4 对各单击流数据进行探索 5.1 对用户单击流数据进行预处理 5.2 进行词向量训练 5.3 对用户的单击流进行编码及 padding 操作 5.4 将词向量矩阵做相应排序并储存 5.5 将单击流数据转化为二维样本数据并存储 6.1 建模前数据整理 6.2 模型训练及性能评估 7.1 模型优化-并行网络结构介绍 7.2 模型优化 	
	操作演练/作业	个人 PC
	在线答疑	微信群
第八课 实战案例：泰迪内推平台信息精准推荐应用（基于泰迪建模平台实现）。		
07月29日 18:30-22:00	<ul style="list-style-type: none"> 1 背景与目标 2 目标分析 <ul style="list-style-type: none"> 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 <ul style="list-style-type: none"> 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 	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第九课 在线考试		
07月30日 19:00-21:00	高级大数据分析师职业技术证书在线考试	泰迪云课堂

附件五 健康医疗大数据应用课程大纲

基础篇（报名成功后即可开始学习）

时间	课程内容	学习平台
正式培训前	Python 编程基础 1 准备工作 2 列表操作 3 程序流程控制语句 4 字符串操作 4.1 字符串及其索引&切片 4.2 字符串的常见方法 4.3 字典的创建及索引 4.4 字典常用操作 4.5 字典推导式 5 Python 文件读取操作 5.1 Python 读取文件 5.2 练习 3：统计小说中的单词频次 6 函数 6.1 Python 函数自定义 6.2 练习 4：自定义求序列偶数个数的函数 7 面向对象与模块 7.1 Python 方法与函数对比介绍 7.2 Python 面向对象示例 7.3 Python 模块使用 7.4 第三方库的安装与调用 8 注意事项 8.1 Python 工作路径说明 8.2 模块命名及存放路径的注意事项 8.3 结语	泰迪云课堂

核心课程篇

时间	课程内容	学习平台
专题讲座		
07月20日 18:30-22:00	主讲：郝志峰 讲座主题：大数据视角下数字孪生与元宇宙的思考	泰迪云课堂
第一课 AI 医疗概率		
07月20日 18:30-22:00	1 医疗人工智能行业相关概述 1.1 医疗人工智能基本介绍 1.1.1 医疗人工智能的诞生	泰迪云课堂

	<ul style="list-style-type: none"> 1.1.2 医疗人工智能发展阶段 1.1.3 医疗人工智能的构成要素 1.2 医疗人工智能的价值分析 <ul style="list-style-type: none"> 1.2.1 缓解漏诊误诊问题 1.2.2 弥补资源供需缺口 1.2.3 提供健康顾问服务 1.2.4 提升制药效率 1.2.5 提升外科手术精准度 1.3 医疗人工智能产业链分析 	
	操作演练	个人 PC
	在线答疑	微信群

第二课 Python 数据分析与应用

<p>07 月 21 日 18:30-22:00</p>	<ul style="list-style-type: none"> 1 Python 数据分析概述 <ul style="list-style-type: none"> 1.1 认识数据分析 1.2 熟悉 Python 数据分析的工具 1.3 安装 anaconda 与启动 jupyter notebook 1.4 掌握 jupyter notebook 常用功能 2 NumPy 数值计算基础 <ul style="list-style-type: none"> 2.1 掌握 NumPy 数组对象 <ul style="list-style-type: none"> 2.1.1 NumPy 简介 <ul style="list-style-type: none"> 2.1.2 数组创建及基础属性 2.1.3 初识数组的特点 2.1.4 创建常用数组 2.1.5 数组数据类型 2.1.6 生成随机数 2.1.7 一维数组的索引 2.1.8 逻辑型索引 2.1.9 多维数组的索引 2.1.10 求解距离矩阵 2.1.11 变化数组 shape 2.2 掌握 NumPy 矩阵与通用函数 <ul style="list-style-type: none"> 2.2.1 NumPy 矩阵介绍 2.2.2 NumPy 通用函数介绍 2.2.3 通用函数的广播机制 2.3 利用 NumPy 进行统计分析 <ul style="list-style-type: none"> 2.3.1 NumPy 读写二进制文件 2.3.2 NumPy 读写 txt 文件 2.3.3 利用 NumPy 对数据进行简单统计分析 	<p>泰迪云课堂</p>
----------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------

	操作演练	个人 PC
	在线答疑	微信群
07月22日 18:30-22:00	3 Matplotlib 数据可视化基础 3.1 掌握绘图基础语法与常用参数 3.1.1 Matplotlib 介绍 3.1.2 基础图形绘制 3.1.3 常用参数设置 3.2 分析特征间关系 3.2.1 绘制散点图 3.2.2 散点图参数设置 3.2.3 绘制折线图 3.3 分析特征内部数据分布与分散情况 3.3.1 绘制直方图 3.3.2 绘制饼图 3.3.3 绘制箱线图	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
07月23日 18:30-22:00	4 Pandas 统计分析基础 4.1 Pandas 简介 4.2 读写不同数据源的数据 4.2.1 Pandas 读取文本数据 4.2.2 存储数据框 4.2.3 Pandas 读取 excel 文件 4.2.4 将数据框存储为 excel 文件 4.3 数据框与数据框元素 4.3.1 构建数据框 4.3.2 查看数据框的常用属性 4.3.3 按行列顺序访问数据框中的元素 4.3.4 按行列名称访问数据框中的元素 4.3.5 修改数据框中的元素 4.3.6 删除数据框中的元素 4.3.7 描述分析数据框中的元素	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
第三课 Python 机器学习实战		
07月24日 18:30-22:00	1.1 引言 1.2 基本术语 1.3 假设空间&归纳偏好	泰迪云课堂

	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	
	操作演练	个人 PC
	在线答疑	微信群
07 月 25 日 18:30-22:00	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 3.3 波士顿房价预测的 Python 实现 3.4 逻辑回归介绍 3.5 研究生入学录取预测的 Python 实现 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	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
拓展自学篇		
自行安排	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 集成学习 6.1 集成学习基本概念 6.2 并行集成算法-Bagging&RandomForest 6.3 串行集成算法-Boosting 算法流程 6.4 串行集成算法-Boosting 代码实现 6.5 Stacking 算法流程 6.6 Stacking 代码实现	泰迪云课堂
第四课 BP 神经网络与 TensorFlow2 实战		
07 月 26 日 18:30-22:00	1 BP 神经网络 1.1 单个神经元介绍 1.2 经典网络结构介绍	泰迪云课堂

	<ul style="list-style-type: none"> 1.3 神经网络工作流程演示 1.4 如何修正网络参数. 梯度下降 1.5 网络工作原理推导 1.6 网络搭建准备 1.7 样本从输入层到隐层传输的 Python 实现 1.8 网络输出的 Python 实现 1.9 单样本网络训练的 Python 实现 1.10 全样本网络训练的 Python 实现 1.11 网络性能评价 1.12 调用 scikit-learn 实现神经网络算法 	
	操作演练/作业	个人 PC
	在线答疑	微信群
07月27日 18:30-22:00	<ul style="list-style-type: none"> 2 任务 1: 构建一个线性模型 2.1 TensorFlow2 介绍 2.2 TensorFlow2 常用数据类型和操作 2.3 初始化模型 2.4 构建损失函数 2.5 模型训练及可视化 2.6 使用高阶 API-Keras 3 任务 2: MNIST 手写数字识别 3.1 数据读取及探索 3.2 交叉熵 3.3 模型构建及训练 3.4 调用保存好的模型对新样本进行预测 	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第五课 TensorFlow 2 深度学习原理与实战		
07月28日 18:30-22:00	<ul style="list-style-type: none"> 1.1 深度神经网络-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 实现池化操作 3 循环神经网络 RNN 3.1 循环神经网络简介 	泰迪云课堂

	3.2 循环神经网络的常见结构	
	操作演练/作业	个人 PC
	在线答疑	微信群
第六课 实战案例：公众健康问句分类		
07月29日 18:30-22:00	1 背景与挖掘目标	泰迪云课堂
	2.1 数据探索	
	2.2 分词和去停用词	
	2.3 词序列化和长度对齐	
	3.1 预训练词向量	
	3.2 词向量矩阵	
	4.1 textCNN 模型搭建	
	4.2 模型训练	
	4.3 模型评估和总结	
	操作演练/作业	个人 PC
	在线答疑	微信群
第七课 实战案例：脑 PET 图像分析和疾病预测		
07月30日 18:30-22:00	1 背景与目标	泰迪云课堂
	2.1 图片统一规格	
	2.2 图片增强	
	3 图片批处理	
	4.1 AlexNet 网络介绍	
	4.2 搭建 AlexNet 网络层(上)	
	4.3 搭建 AlexNet 网络层(下)	
	4.4 训练及预测	
	5 模型验证及小结	
	操作演练/作业	个人 PC
	在线答疑	微信群
第八课 实战案例：基于深度学习的肝脏肿瘤分割		
07月31日 18:30-22:00	1.1 案例背景	泰迪云课堂
	1.2 数据说明与目标	
	2.1.1 数据预处理流程介绍	
	2.1.2 打开工程文件	
	2.1.3 DICOM 医学图像数据加载	
	2.2.1 windowing 方法介绍与参数选择	
	2.2.2 windowing 方法实现	
	2.3 CT 对比增强：直方图均衡化	
	2.4 掩膜图像加载	
2.5 自定义提取肿瘤掩膜图位置的函数		

	2.6 文件保存 2.7 循环操作 20 名病人 CT 图像 3.1 数据增强：定义图像生成器 3.2 数据增强：图像数据增强设置及可视化 3.3 数据增强：归一化与自定义增强函数 4.1.1 U-Net 网络结构介绍 4.1.2 U-Net 网络结构实现 4.2 模型编译 5.1 模型训练 5.2 模型训练结果展示 6 模型预测 7 拓展思考	
	操作演练/作业	个人 PC
	在线答疑	微信群
第九课 在线考试		
08 月 01 日 19:00-21:00	高级大数据应用工程师职业资格证书在线考试	泰迪云课堂

附件六 Hadoop+Spark 大数据技术应用实战课程大纲

基础篇（报名成功后即可开始学习）

时间	课程内容	学习平台
正式培训前	Linux 操作系统基础 1. Linux 简介： 1.1 Linux 简介与发展历程 1.2 Linux 常见版本 2. Linux 基本命令： 2.1 Linux 文件系统树 2.2 创建与删除目录或文件 2.3 查看文件内容 2.4 创建硬链接和符号链接 2.5 帮助命令 2.6 重定向相关命令 2.7 用户及用户组 2.8 权限相关命令 2.9 用户、用户组及权限实践操作进程	泰迪云课堂

	<p>3. Linux Vi 编辑器:</p> <p>3.1 模式介绍与常见快捷命令</p> <p>3.2 一般模式快捷键</p> <p>3.3 编辑模式与指令模式快捷键</p>	
正式培训前	<p>Java 编程基础</p> <p>1. 基础知识:</p> <p>1.1 Java 简介</p> <p>1.2 Java 安装配置</p> <p>1.3 Java 基础语法</p> <p>1.4.1 Java 基本数据类型&引用数据类型</p> <p>1.4.2 Java 变量</p> <p>1.5 Java String 类</p> <p>1.6.1 Java 运算符-赋值运算符&算术运算符</p> <p>1.6.2 Java 运算符-关系、逻辑、其他运算符</p> <p>1.7 循环、条件、循环中断</p> <p>1.8 Java 数组</p> <p>1.9 Java 函数</p> <p>1.10.1 Java 集合简介</p> <p>1.10.2 Java 集合之 List</p> <p>1.10.3 Java 集合之 Set</p> <p>1.10.4 Java 集合之 Map</p> <p>1.11 Java 泛型</p> <p>2. 面向对象:</p> <p>2.1 Java 类、对象</p> <p>2.2.1 Java 类封装</p> <p>2.2.2 Java 构造函数</p> <p>2.2.3 Java 类继承</p> <p>2.2.4 Java 类多态</p> <p>2.3.1 Java 抽象类</p> <p>2.3.2 Java 接口</p> <p>3. 线程及异常处理:</p> <p>3.1 Java 多线程</p> <p>3.2 Java 包、异常</p>	泰迪云课堂

核心课程篇

时间	课程内容	学习平台
第一课 Hadoop 大数据基础		
08月01日	1. Hadoop 简介	泰迪云课堂

18:30-22:00	<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 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 	
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
08月02日 18:30-22:00	<ul style="list-style-type: none"> 2.Hadoop 集群安装与部署 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 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 	泰迪云课堂
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
08月03日 18:30-22:00	<ul style="list-style-type: none"> 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 5.MapReduce 编程进阶 5.1 Mapreduce 输出文件格式及序列化文件生成 5.2 Mapreduce 输入文件格式及序列化文件读取 5.3 Mapreduce 使用 Combiner 	泰迪云课堂

	5.3 Mapreduce 使用 Partitioner 5.4 Mapreeuce 本地提交任务	
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
第二课 Hive 大数据仓库		
08月04日 18:30-22:00	1. Hive 简介 1. Hive 简介 2. Hive 安装配置 2.1 Hadoop 集群介绍与 MySQL 安装 2.2 Hive 安装配置 2.3 Hive 单词计数 3. Hive 应用 3.1 任务讲解与数据库定义 3.2 创建内表 3.3 创建外表 3.4 创建泰坦尼克乘客信息表 3.5 创建静态分区表 3.6 创建动态分区表 3.7 创建带数据的表和桶表 3.8 从文件系统导入数据到 Hive 3.9 从其他表查询数据导入 Hive	泰迪云课堂
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
	3. Hive 应用 3.10 导出表数据到文件系统 3.11 创建职工信息表 3.12 select 查询入门 3.13 泰坦尼克号人员分布统计 3.14 select 连接查询和子查询 3.15 分析影响生存率关系 3.16 Hive 语句执行顺序 4. Hive 开发 4.1 配置 Hive 的 Java 开发环境 4.2 创建订单表和物品类别表 4.3 统计顾客数和商品销量 4.4 统计各类型商品购买情况	泰迪云课堂

	<ul style="list-style-type: none"> 4.5 统计销量排名前 20 的商品 4.6 统计各类型商品排名前 10 的热销商品 4.7 保存客户订单数据到 HDFS 	
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
08月06日 18:30-22:00	<ul style="list-style-type: none"> 5. Hive 自定义函数 5.1 自定义函数简介 5.2 自定义 UDF 5.3 自定义 UDAF 5.4 自定义 UDTF 6. Hive 查询优化 6.1 视图 6.2 索引 6.3 存储格式 综合案例：航空客户乘机数据预处理 1. 案例背景及挖掘目标介绍 2. 数据探索分析 3. 数据清洗 4. 属性规约与数据变换 	泰迪云课堂
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
第三课 HBase 非关系型数据库		
08月07日 18:30-22:00	<ul style="list-style-type: none"> 1. HBase 简介 1.1 认识 HBase 1.2 HBase 适用场景 2. HBase 安装配置 2.1 HBase 集群拓扑简介 2.2 Zookeeper 集群安装配置 2.3 HBase 安装配置 2.4 HBase 监控界面讲解 3. HBase 架构原理 3.1 HBase 数据模型 3.2 HBase 架构 3.3 HBase 读写流程 4. HBase Shell 基本操作 4.1 命令空间操作 	泰迪云课堂

	<ul style="list-style-type: none"> 4.2 创建表 4.3 修改和删除表 4.4 插入数据到表中 4.5 查看表数据 4.6 删除表数据 	
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
08月08日 18:30-22:00	<ul style="list-style-type: none"> 5. HBase 表的模式设计 5. HBase 表设计 6. HBase Java API 应用 6.1 搭建 HBase 开发环境 6.2 建立与 HBase 数据库连接实例 6.3 创建或删除表应用实例 6.4 设置 Region 分割点应用实例 6.5 修改表结构应用实例 6.6 写入数据应用实例 6.7 查看一行数据应用实例 6.8 扫描表数据应用实例 6.9 删除数据应用实例 6.10 创建职工信息表 emp 6.11 写入数据到职工信息表 6.12 查找指定范围职工信息 	泰迪云课堂
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
第四课 Scala 编程基础		
08月09日 18:30-22:00	<ul style="list-style-type: none"> 1. Scala 简介 1.1. Scala 简介 2. Scala 安装配置 2.1 Scala 线上环境介绍 2.2 在 Windows 上搭建 Scala 2.3 在 Linux 上搭建 Scala 2.4 Scala 插件安装 2.5 创建第一个 Scala 工程 3. Scala 基础语法 3.1 Scala 类型、变量、运算符 3.2 数组 	泰迪云课堂

	<ul style="list-style-type: none"> 3.3 if 条件判断 3.4 循环语句 3.5 实现冒泡排序法 3.6 列表 List 3.7 列表 ListBuffer 3.8 集合 Set 3.9 计算最大公约数和最小公倍数 3.10 映射 Map 3.11 元组 Tuple 	
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
08月10日 18:30-22:00	<ul style="list-style-type: none"> 4. 函数 <ul style="list-style-type: none"> 4.1 函数定义 4.2 匿名函数 4.3 高阶函数 4.4 集合中的高阶函数 4.5 递归与尾递归 5. 面向对象编程 <ul style="list-style-type: none"> 5.1 面向对象编程简介 5.2 类的定义 5.3 定义类实现有理数的运算 5.4 单例对象 5.5 样例类和模式匹配 6. 文件读写 <ul style="list-style-type: none"> 6.1 文件读写 6.2 手机号码段查询 	泰迪云课堂
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
第五课 Spark 大数据技术与应用		
08月11日 18:30-22:00	<ul style="list-style-type: none"> 1. Spark 入门 <ul style="list-style-type: none"> 1. Spark 介绍 2. Spark 集群的安装配置 <ul style="list-style-type: none"> 2.1 Spark 集群拓扑 2.2 Spark 安装配置 2.3 运行第一个 Spark 程序 3. Spark 架构及原理 	泰迪云课堂

	<ul style="list-style-type: none"> 3.1 Spark 架构 3.2 Spark RDD 4. Spark Shell 应用 <ul style="list-style-type: none"> 4.1 启动 Spark Shell 4.2 创建 RDD 4.3 读取学生成绩创建 RDD 4.4 map 与 flatMap 算子应用 4.5 sortBy 与 filter 算子应用 4.6 取出成绩排名前 5 的学生信息 4.7 交集与并集计算的算子应用 4.8 获取成绩 100 分的学生信息 4.9 键值对 RDD 常用算子 4.10 计算学生成绩总分和平均分 4.11 存储 RDD 4.12 统计用户停留时间最长的基站 	
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
08 月 12 日 18:30-22:00	<ul style="list-style-type: none"> 5. Spark 开发环境配置 <ul style="list-style-type: none"> 5.1 配置 Spark 开发环境 5.2 Spark 实现单词计数 5.3 使用本地模式运行 Spark 程序 5.4 使用集群模式运行 Spark 程序 5.5 设置应用程序使用的集群资源 6. Spark SQL 应用 <ul style="list-style-type: none"> 6.1 Spark SQL 简介和环境配置 6.2 读取结构化文件创建 DataFrame 6.3 读取外部数据库数据创建 DataFrame 6.4 读取 RDD 创建 DataFrame 6.5 读取 Hive 表数据创建 RDD 6.6 读取学生成绩创建 DataFrame 6.7 常见 DataFrame API 应用 6.8 通过 DataFrame API 计算学生总分和平均分 6.9 保存 DataFrame 为结构化文件 6.10 保存 DataFrame 到外部数据库和 Hive 数据库 6.11 保存学生成绩分析结果到 Hive 6.12 创建 DataSet 6.13 通过商品销售数据创建 DataSet 	泰迪云课堂

	6.14 统计商品销量	
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
第六课 项目实战：广电大数据用户画像 (Hadoop + Spark + Hive)		
08月13日 18:30-22:00	1. 项目背景与需求分析 1.1 广电大数据用户画像案例背景 1.2 广电大数据用户画像项目目标 1.3 系统架构与项目选型 2. 数据存储 2.1 数据描述 2.2 数据存储 3. 数据探索分析 3.1 数据探索分析 3.2 数据探索分析实战1 3.3 数据探索分析实战2 3.4 数据探索分析实战3 4. 数据预处理 4.1 数据处理规则分析 4.2 数据预处理实战	泰迪云课堂
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群
08月14日 18:30-22:00	5. 用户画像标签计算 5.1 用户画像及规则简介 5.2 用户画像实战-消费内容计算 5.3 用户画像实战-消费水平计算 5.4 用户画像实战-销售品计算 5.5 用户画像实战-宽带产品带宽 5.6 用户画像实战-用户信息表标签计算 5.7 用户画像实战-地区偏好和语言偏好 6. SVM 预测用户是否挽留 6.1 SVM 模型简介 6.2 SVM 模型数据处理 6.3 SVM 模型构建与评估	泰迪云课堂
	操作演练/作业	泰迪大数据编程实训平台
	在线答疑	微信群

拓展自学篇		
自行安排	1 Zookeeper 分布式服务框架 2 Flume 大数据采集与传输 3 Kafka 大数据流处理 4 Flink 大数据实时处理 5 商品实时推荐系统 (Flume + Kafka + Flink)	泰迪云课堂
第七课 在线考试		
08月15日 19:00-21:00	高级大数据技术应用职业技术证书在线考试	泰迪云课堂

附件七 人工智能实战（计算机视觉方向）课程大纲

基础篇（报名成功后即可开始学习）

时间	课程内容	学习平台
正式培训前	Python 编程基础 1 准备工作 2 列表操作 3 程序流程控制语句 4 字符串操作 4.1 字符串及其索引&切片 4.2 字符串的常见方法 4.3 字典的创建及索引 4.4 字典常用操作 4.5 字典推导式 5 Python 文件读取操作 5.1 Python 读取文件 5.2 练习 3：统计小说中的单词频次 6 函数 6.1 Python 函数自定义 6.2 练习 4：自定义求序列偶数个数的函数 7 面向对象与模块 7.1 Python 方法与函数对比介绍 7.2 Python 面向对象示例 7.3 Python 模块使用 7.4 第三方库的安装与调用 8 注意事项	泰迪云课堂

	8.1 Python 工作路径说明 8.2 模块命名及存放路径的注意事项 8.3 结语	
正式培训前	Python 数据分析与应用 1 Python 数据分析概述 1.1 认识数据分析 1.2 熟悉 Python 数据分析的工具 1.3 安装 anaconda 与启动 Jupyter Notebook 1.4 掌握 Jupyter Notebook 常用功能 2 NumPy 数值计算基础 2.1 掌握 NumPy 数组对象 2.1.1 NumPy 简介 2.1.2 数组创建及基础属性 2.1.3 初识数组的特点 2.1.4 创建常用数组 2.1.5 数组数据类型 2.1.6 生成随机数 2.1.7 一维数组的索引 2.1.8 逻辑型索引 2.1.9 多维数组的索引 2.1.10 求解距离矩阵 2.1.11 变化数组 shape 2.2 掌握 NumPy 矩阵与通用函数 2.2.1 NumPy 矩阵介绍 2.2.2 NumPy 通用函数介绍 2.2.3 通用函数的广播机制 2.3 利用 NumPy 进行统计分析 2.3.1 NumPy 读写二进制文件 2.3.2 NumPy 读写 txt 文件 2.3.3 利用 NumPy 对数据进行简单统计分析 3 Matplotlib 数据可视化基础 3.1 掌握绘图基础语法与常用参数 3.1.1 Matplotlib 介绍 3.1.2 基础图形绘制 3.1.3 常用参数设置 3.2 分析特征间关系 3.2.1 绘制散点图 3.2.2 散点图参数设置	泰迪云课堂

- 3.2.3 绘制折线图
- 3.3 分析特征内部数据分布与分散情况
 - 3.3.1 绘制直方图
 - 3.3.2 绘制饼图
 - 3.3.3 绘制箱线图
- 4 Pandas 统计分析基础
 - 4.1 Pandas 简介
 - 4.2 读写不同数据源的数据
 - 4.2.1 Pandas 读取文本数据
 - 4.2.2 存储数据框
 - 4.2.3 Pandas 读取 excel 文件
 - 4.2.4 将数据框存储为 excel 文件
 - 4.3 数据框与数据框元素
 - 4.3.1 构建数据框
 - 4.3.2 查看数据框的常用属性
 - 4.3.3 按行列顺序访问数据框中的元素
 - 4.3.4 按行列名称访问数据框中的元素
 - 4.3.5 修改数据框中的元素
 - 4.3.6 删除数据框中的元素
 - 4.3.7 描述分析数据框中的元素
 - 4.4 转换与处理时间序列数据
 - 4.4.1 转换成时间类型数据
 - 4.4.2 时间类型数据的常用操作
 - 4.5 使用分组聚合进行组内计算
 - 4.5.1 groupby 分组操作
 - 4.5.2 agg 聚合操作
 - 4.6 创建透视表与交叉表
 - 4.6.1 生成透视表
 - 4.6.2 生成交叉表
- 5 使用 Pandas 进行数据预处理
 - 5.1 合并数据
 - 5.1.1 表堆叠
 - 5.1.2 主键合并
 - 5.1.3 重叠合并
 - 5.2 清洗数据
 - 5.2.1 检测与处理重复值
 - 5.2.2 检测与处理缺失值
 - 5.2.3 检测与处理异常值

	5.3 标准化数据 5.4 转换数据 5.4.1 哑变量处理 5.4.2 离散化连续型数据	
--	--------------------------------------------------------	--

核心课程篇

时间	课程内容	学习平台
专题讲座		
08月01日 18:30-22:00	主讲：郝志峰 讲座主题：大数据视角下数字孪生与元宇宙的思考	泰迪云课堂
第一课 Python 机器学习算法原理与实现		
08月01日 18:30-22:00	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	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
08月02日 18:30-22:00	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 3.3 波士顿房价预测的 Python 实现 3.4 逻辑回归介绍 3.5 研究生入学录取预测的 Python 实现 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	泰迪云课堂

4.11 网络性能评价	
4.12 调用 sklearn 实现神经网络算法	
操作演练	个人 PC
在线答疑	微信群

第二课 TensorFlow2 实战

08月03日 18:30-22:00	1 任务 1: 构建一个线性模型	泰迪云课堂
	1.1 tensorflow 介绍	
	1.2 tensorflow2 常用数据类型和操作	
	1.3 初始化模型	
	1.4 构建损失函数	
	1.5 模型训练及可视化	
	1.6 使用高阶 API-keras	
	2 任务 2: mnist 手写数字识别	
	2.1 数据读取及探索	
	2.2 交叉熵	
	2.3 模型构建及训练	
	2.4 调用保存好的模型对新样本进行预测	
	操作演练	个人 PC
	在线答疑	微信群

第三课 TensorFlow 2 深度学习原理与实现

08月04日 18:30-22:00	1.1 深度神经网络-引言	泰迪云课堂
	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 实现池化操作	
	2.8 作业: 使用 CNN 实现	
	3.1 生成对抗神经网络简介	
	3.2 生成对抗神经网络训练流程	
	3.3 生成手写数字照片-数据准备&定义生成器网络	
	3.4 生成手写数字照片-判别器网络结构	
	3.5 生成手写数字照片-网络传输及网络损失值	
	3.6 生成手写数字照片-模型训练及效果展示	
	操作演练/作业	
在线答疑	微信群	

拓展自学篇		
自行安排	4 循环神经网络 RNN 4.1 循环神经网络简介 4.2 循环神经网络的常见结构 5 长短时记忆网络 LSTM 5.1 LSTM 的三个门 5.2 LSTM 三个门的计算示例 5.3 利用 RNN&LSTM 实现 mnist 手写数字识别	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第四课 计算机视觉实战		
08月05日 18:30-22:00	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	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
08月06日 18:30-22:00	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	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
08月07日 18:30-22:00	7 图像批处理任务示例 8.1 OpenCV 循环读取-批量获取图像路径 8.2 OpenCV 循环读取-批量读取图片和标签 8.3 OpenCV 循环读取-数据应用示例 9.1 图像文件读取-参数说明 9.2 图像文件读取-读取训练和测试数据 9.3 图像文件读取-查看数据情况 9.4 图像文件读取-数据应用示例 10.1 ImageDataGenerator 图像增强-方法介绍	

	10.2 ImageDataGenerator 图像增强-单张图像增强 10.3 ImageDataGenerator 图像增强-多张图像增强 10.4 ImageDataGenerator 图像增强-从文件夹读取图片并增强 10.5 ImageDataGenerator 图像增强-数据应用示例	
	操作演练/作业	个人 PC
	在线答疑	微信群
第五课 图像处理实战：脑 PET 图像分析和疾病预测		
08月08日 18:30-22:00	1 背景与目标 2.1 图片统一规格 2.2 图片增强 3 图片批处理 4.1 AlexNet 网络介绍 4.2 搭建 AlexNet 网络层(上) 4.3 搭建 AlexNet 网络层(下) 4.4 训练及预测 5 模型验证及小结	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第六课 计算机视觉实战：动态人脸智能识别		
08月09日 18:30-22:00	1 背景与目标 2.1 调用电脑摄像头拍照 2.2 实现批量拍照 2.3 封装拍照操作的代码 3.1 人脸检测介绍 3.2 实现照片中人脸检测 3.3 过滤人脸不全的照片 3.4 人脸照片的灰度转换及存储 3.5 批量照片处理及存储 3.6 将拍照及数据处理代码封装成类 4.1 获取人名及对应文件夹路径 4.2 自定义独热编码函数 4.3 确认样本和标签对应关系 4.4 定义照片数据转数组数据的函数 4.5 将数据整理代码封装成类	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
08月10日	5.1 建模介绍	泰迪云课堂

18:30-22:00	5.10 模型训练代码实现 5.11 执行模型训练 5.12 模型性能调优 5.2 定义模型搭建所需的常用方法 5.3 搭建第一个卷积层结构 5.4 测试网络结构 5.5 搭建第二个和第三个卷积层 5.6 搭建第一个全连接层 5.7 搭建网络输出层 5.8 定义模型训练的损失函数-交叉熵 5.9 模型训练准备 6.1 调用保存好的模型 6.2 模型性能评价结果	
	操作演练/作业	个人 PC
	在线答疑	微信群
第七课 实战案例：基于 GAN 的图像风格转换		
08月11日 18:30-22:00	1 背景与目标 2 过程与步骤 2.1 数据读取 2.2 数据预处理 2.3 数据集整理 3 模型构建 3.1 GAN 网络搭建 3.2 定义 GAN 损失函数 3.3 训练网络 3.4 结果分析 4 小结	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第八课 在线考试		
08月12日 19:00-21:00	高级人工智能应用工程师职业技术证书在线考试	泰迪云课堂

附件八 人工智能实战（自然语言处理方向）课程大纲

基础篇（报名成功后即可开始学习）

时间	课程内容	学习平台
正式培训前	<p>Python 编程基础</p> <ul style="list-style-type: none"> 1 准备工作 2 列表操作 3 程序流程控制语句 4 字符串操作 <ul style="list-style-type: none"> 4.1 字符串及其索引&切片 4.2 字符串的常见方法 4.3 字典的创建及索引 4.4 字典常用操作 4.5 字典推导式 5 Python 文件读取操作 <ul style="list-style-type: none"> 5.1 Python 读取文件 5.2 练习 3：统计小说中的单词频次 6 函数 <ul style="list-style-type: none"> 6.1 Python 函数自定义 6.2 练习 4：自定义求序列偶数个数的函数 7 面向对象与模块 <ul style="list-style-type: none"> 7.1 Python 方法与函数对比介绍 7.2 Python 面向对象示例 7.3 Python 模块使用 7.4 第三方库的安装与调用 8 注意事项 <ul style="list-style-type: none"> 8.1 Python 工作路径说明 8.2 模块命名及存放路径的注意事项 8.3 结语 	泰迪云课堂
正式培训前	<p>Python 数据分析与应用</p> <ul style="list-style-type: none"> 1 Python 数据分析概述 <ul style="list-style-type: none"> 1.1 认识数据分析 1.2 熟悉 Python 数据分析的工具 1.3 安装 anaconda 与启动 Jupyter Notebook 1.4 掌握 Jupyter Notebook 常用功能 2 NumPy 数值计算基础 <ul style="list-style-type: none"> 2.1 掌握 NumPy 数组对象 <ul style="list-style-type: none"> 2.1.1 NumPy 简介 2.1.2 数组创建及基础属性 2.1.3 初识数组的特点 2.1.4 创建常用数组 	泰迪云课堂

- 2.1.5 数组数据类型
- 2.1.6 生成随机数
- 2.1.7 一维数组的索引
- 2.1.8 逻辑型索引
- 2.1.9 多维数组的索引
- 2.1.10 求解距离矩阵
- 2.1.11 变化数组 shape
- 2.2 掌握 NumPy 矩阵与通用函数
 - 2.2.1 NumPy 矩阵介绍
 - 2.2.2 NumPy 通用函数介绍
 - 2.2.3 通用函数的广播机制
- 2.3 利用 NumPy 进行统计分析
 - 2.3.1 NumPy 读写二进制文件
 - 2.3.2 NumPy 读写 txt 文件
 - 2.3.3 利用 NumPy 对数据进行简单统计分析
- 3 Matplotlib 数据可视化基础
 - 3.1 掌握绘图基础语法与常用参数
 - 3.1.1 Matplotlib 介绍
 - 3.1.2 基础图形绘制
 - 3.1.3 常用参数设置
 - 3.2 分析特征间关系
 - 3.2.1 绘制散点图
 - 3.2.2 散点图参数设置
 - 3.2.3 绘制折线图
 - 3.3 分析特征内部数据分布与分散情况
 - 3.3.1 绘制直方图
 - 3.3.2 绘制饼图
 - 3.3.3 绘制箱线图
- 4 Pandas 统计分析基础
 - 4.1 Pandas 简介
 - 4.2 读写不同数据源的数据
 - 4.2.1 Pandas 读取文本数据
 - 4.2.2 存储数据框
 - 4.2.3 Pandas 读取 excel 文件
 - 4.2.4 将数据框存储为 excel 文件
 - 4.3 数据框与数据框元素
 - 4.3.1 构建数据框
 - 4.3.2 查看数据框的常用属性

<ul style="list-style-type: none"> 4.3.3 按行列顺序访问数据框中的元素 4.3.4 按行列名称访问数据框中的元素 4.3.5 修改数据框中的元素 4.3.6 删除数据框中的元素 4.3.7 描述分析数据框中的元素 4.4 转换与处理时间序列数据 <ul style="list-style-type: none"> 4.4.1 转换成时间类型数据 4.4.2 时间类型数据的常用操作 4.5 使用分组聚合进行组内计算 <ul style="list-style-type: none"> 4.5.1 groupby 分组操作 4.5.2 agg 聚合操作 4.6 创建透视表与交叉表 <ul style="list-style-type: none"> 4.6.1 生成透视表 4.6.2 生成交叉表 5 使用 Pandas 进行数据预处理 <ul style="list-style-type: none"> 5.1 合并数据 <ul style="list-style-type: none"> 5.1.1 表堆叠 5.1.2 主键合并 5.1.3 重叠合并 5.2 清洗数据 <ul style="list-style-type: none"> 5.2.1 检测与处理重复值 5.2.2 检测与处理缺失值 5.2.3 检测与处理异常值 5.3 标准化数据 5.4 转换数据 <ul style="list-style-type: none"> 5.4.1 哑变量处理 5.4.2 离散化连续型数据 	
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

核心课程篇

时间	课程内容	学习平台
专题讲座		
08月01日 18:30-22:00	主讲：郝志峰 讲座主题：大数据视角下数字孪生与元宇宙的思考	泰迪云课堂
第一课 Python 机器学习算法原理与实现		
08月01日 18:30-22:00	1 机器学习绪论 <ul 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 	泰迪云课堂

	<ul style="list-style-type: none"> 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 	
	操作演练	个人 PC
	在线答疑	微信群
08 月 02 日 18:30-22:00	<ul style="list-style-type: none"> 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 3.3 波士顿房价预测的 Python 实现 3.4 逻辑回归介绍 3.5 研究生入学录取预测的 Python 实现 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 	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
第二课 TensorFlow2 实战		
08 月 03 日 18:30-22:00	<ul style="list-style-type: none"> 1 任务 1: 构建一个线性模型 1.1 tensorflow 介绍 1.2 tensorflow2 常用数据类型和操作 1.3 初始化模型 1.4 构建损失函数 1.5 模型训练及可视化 1.6 使用高阶 API-keras 2 任务 2: mnist 手写数字识别 2.1 数据读取及探索 2.2 交叉熵 	泰迪云课堂

	2.3 模型构建及训练	
	2.4 调用保存好的模型对新样本进行预测	
	操作演练	个人 PC
	在线答疑	微信群
第三课 TensorFlow 2 深度学习原理与实现		
08月04日 18:30-22:00	1.1 神经网络-引言	泰迪云课堂
	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 实现池化操作	
	操作演练/作业	个人 PC
	在线答疑	微信群
08月05日 18:30-22:00	3 循环神经网络 RNN	泰迪云课堂
	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 利用 RNN&LSTM 实现 mnist 手写数字识别		
	操作演练/作业	个人 PC
	在线答疑	微信群
拓展自学篇		
自行安排	5.1 生成对抗神经网络简介	泰迪云课堂
	5.2 生成对抗神经网络训练流程	
	5.3 生成手写数字照片-数据准备&定义生成器网络	
	5.4 生成手写数字照片-判别器网络结构	
5.5 生成手写数字照片-网络传输及网络损失值		
5.6 生成手写数字照片-模型训练及效果展示		
	操作演练/作业	个人 PC
	在线答疑	微信群
第四课 自然语言处理实战		
08月06日 18:30-22:00	1 自然语言处理概述	泰迪云课堂
	1.1 自然语言处理概述	
	2 NLP 的基本流程	

	<ul style="list-style-type: none"> 2.1 NLP 的基本流程 2.2 语料字符处理 <ul style="list-style-type: none"> 2.2.1 语料字符处理-字符串函数 2.2.2 正则表达式 2.2.3 正则表达式应用 2.3 分词和停用词处理 <ul style="list-style-type: none"> 2.3.1 分词：正向最大匹配法 2.3.2 分词：正向最大匹配法实现 2.3.3 分词：逆向最大匹配法 2.4.1N 元语法模型 2.4.2 N 元语法模型与分词 2.5.1 隐马尔可夫概述 2.5.2 隐马尔可夫实例 2.5.3 Viterbi 算法 2.5.4 Viterbi 算法应用 	
	操作演练/作业	个人 PC
	在线答疑	微信群
08 月 07 日 18:30-22:00	<ul style="list-style-type: none"> 2.6 jieba 分词 2.7 去停用词 3 文本向量化 <ul style="list-style-type: none"> 3.1.1 向量化与独热编码 3.1.2 基于字典的 one-hot 实现 3.1.3 基于 keras 的 one-hot 表示 3.2 词袋模型 3.3TF-IDF 3.4.1 Word2Vec 模型概述 3.4.2 Word2Vec 模型训练 3.4.3 Word2Vec 模型应用 3.5 Doc2Vec 模型 	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第五课 文本处理实战：公众健康问句分类		
08 月 08 日 18:30-22:00	<ul style="list-style-type: none"> 1 背景与挖掘目标 2.1 数据探索 2.2 分词和去停用词 2.3 词序列化和长度对齐 <ul style="list-style-type: none"> 3.1 预训练词向量 3.2 词向量矩阵 	泰迪云课堂

	4.1 textCNN 模型搭建 4.2 模型训练 4.3 模型评估和总结	
	操作演练/作业	个人 PC
	在线答疑	微信群
第六课 案例实战：搭建一个属于自己的聊天机器人		
08月09日 18:30-22:00	1 案例背景与目标 1.1 背景与目标 1.2 流程介绍 2 语料库预处理 2.1 语料库介绍 2.2 工程和数据准备 2.3 语料分词处理 2.4 预处理结果保存 2.5 语料预处理代码整理 3 模型构建 3.1 模型介绍 3.2 构建哈希表 3.3 问答数据向量化 3.4 问答向量长度填充	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
08月10日 18:30-22:00	3.5 模型准备 3.6 模型编译 3.7 训练步 3.8 迭代训练 3.9 模型预测的句子分词 3.10 模型预测输入输出 3.11 模型整理、类化 4 网页前端展示 5 拓展思考	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第七课 实战案例：基于深度学习的推荐系统受众性别预测		
08月11日 18:30-22:00	1 任务 1-了解背景与目标 2 数据探索 2.1 任务 2 说明-数据探索 2.2 项目文件结构及编辑环境介绍	泰迪云课堂

	2.3 数据获取与探索分析 3 获取用户单击流数据 3.1 任务 3 说明-获取用户单击流数据 3.2 理解用户单击流相关概念 3.3 获取用户的各单击流数据 4 实现词嵌入操作-1 4.1 任务 4 说明-实现词嵌入操作-1 4.2 对各单击流数据进行探索 5 实现词嵌入操作-2 5.1 任务 5 说明-实现词嵌入操作-2 5.2 对用户单击流数据进行预处理 5.3 进行词向量训练 5.4 对用户的单击流进行编码及 padding 操作 5.5 将词向量矩阵做相应排序并储存 5.6 将单击流数据转化为二维样本数据并存储 6 模型构建 6.1 任务 6 说明-模型构建 6.2 建模前数据整理 6.3 模型训练及性能评估 7 模型优化 7.1 任务 7 说明-模型优化 7.2 模型优化-并行网络结构介绍 7.3 模型优化	
	操作演练/作业	个人 PC
	在线答疑	微信群
第八课 在线考试		
08月12日 19:00-21:00	高级人工智能应用工程师职业技术证书在线考试	泰迪云课堂

附件九 智能语音实战课程大纲

基础篇（报名成功后即可开始学习）

时间	课程内容	学习平台
正式培训前	Python 编程基础 1 准备工作 2 列表操作	泰迪云课堂

	<ul style="list-style-type: none"> 3 程序流程控制语句 4 字符串操作 <ul style="list-style-type: none"> 4.1 字符串及其索引&切片 4.2 字符串的常见方法 4.3 字典的创建及索引 4.4 字典常用操作 4.5 字典推导式 5 Python 文件读取操作 <ul style="list-style-type: none"> 5.1 Python 读取文件 5.2 练习 3: 统计小说中的单词频次 6 函数 <ul style="list-style-type: none"> 6.1 Python 函数自定义 6.2 练习 4: 自定义求序列偶数个数的函数 7 面向对象与模块 <ul style="list-style-type: none"> 7.1 Python 方法与函数对比介绍 7.2 Python 面向对象示例 7.3 Python 模块使用 7.4 第三方库的安装与调用 8 注意事项 <ul style="list-style-type: none"> 8.1 Python 工作路径说明 8.2 模块命名及存放路径的注意事项 8.3 结语 	
正式培训前	<p>Python 数据分析与应用</p> <ul style="list-style-type: none"> 1 Python 数据分析概述 <ul style="list-style-type: none"> 1.1 认识数据分析 1.2 熟悉 Python 数据分析的工具 1.3 安装 anaconda 与启动 Jupyter Notebook 1.4 掌握 Jupyter Notebook 常用功能 2 NumPy 数值计算基础 <ul style="list-style-type: none"> 2.1 掌握 NumPy 数组对象 <ul style="list-style-type: none"> 2.1.1 NumPy 简介 <ul style="list-style-type: none"> 2.1.2 数组创建及基础属性 2.1.3 初识数组的特点 2.1.4 创建常用数组 2.1.5 数组数据类型 2.1.6 生成随机数 2.1.7 一维数组的索引 2.1.8 逻辑型索引 	泰迪云课堂

<ul style="list-style-type: none"> 2.1.9 多维数组的索引 2.1.10 求解距离矩阵 2.1.11 变化数组 shape 2.2 掌握 NumPy 矩阵与通用函数 <ul style="list-style-type: none"> 2.2.1 NumPy 矩阵介绍 2.2.2 NumPy 通用函数介绍 2.2.3 通用函数的广播机制 2.3 利用 NumPy 进行统计分析 <ul style="list-style-type: none"> 2.3.1 NumPy 读写二进制文件 2.3.2 NumPy 读写 txt 文件 2.3.3 利用 NumPy 对数据进行简单统计分析 3 Matplotlib 数据可视化基础 <ul style="list-style-type: none"> 3.1 掌握绘图基础语法与常用参数 <ul style="list-style-type: none"> 3.1.1 Matplotlib 介绍 3.1.2 基础图形绘制 3.1.3 常用参数设置 3.2 分析特征间关系 <ul style="list-style-type: none"> 3.2.1 绘制散点图 3.2.2 散点图参数设置 3.2.3 绘制折线图 3.3 分析特征内部数据分布与分散情况 <ul style="list-style-type: none"> 3.3.1 绘制柱状图 3.3.2 绘制饼图 3.3.3 绘制箱线图 	
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

核心课程篇

时间	课程内容	学习平台
专题讲座		
08月01日 18:30-22:00	主讲：郝志峰 讲座主题：大数据视角下数字孪生与元宇宙的思考	泰迪云课堂
第一课 Python 机器学习算法原理与实现		
08月01日 18:30-22:00	<ul style="list-style-type: none"> 1 机器学习绪论 <ul 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 <ul style="list-style-type: none"> 2.1 经验误差与过拟合 2.2 评估方法 	泰迪云课堂

	2.3 性能度量 2.4 性能度量 Python 实现	
	操作演练	个人 PC
	在线答疑	微信群
08 月 02 日 18:30-22:00	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 3.3 波士顿房价预测的 Python 实现 3.4 逻辑回归介绍 3.5 研究生入学录取预测的 Python 实现 4 聚类分析 4.1 聚类分析概述 4.2 相似性度量 4.3 K-Means 聚类分析算法介绍 4.4 利用 K-Means 算法对鸢尾花进行聚类 4.5 聚类结果的性能度量 4.6 调用 scikit-learn 实现聚类分析	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
第二课 TensorFlow2 实战		
08 月 03 日 18:30-22:00	1 任务 1: 构建一个线性模型 1.1 TensorFlow 介绍 1.2 TensorFlow2 常用数据类型和操作 1.3 初始化模型 1.4 构建损失函数 1.5 模型训练及可视化 1.6 使用高阶 API-Keras 2 任务 2: MNIST 手写数字识别 2.1 数据读取及探索 2.2 交叉熵 2.3 模型构建及训练 2.4 调用保存好的模型对新样本进行预测	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
第三课 TensorFlow 2 深度学习原理与实战		
08 月 04 日 18:30-22:00	1.1 深度神经网络-引言 2 卷积神经网络 CNN 2.1 浅层神经网络的局限	泰迪云课堂

	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 实现池化操作 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 利用 RNN&LSTM 实现 mnist 手写数字识别	
	操作演练/作业	个人 PC
	在线答疑	微信群
第四课 语音处理实战		
08 月 05 日 18:30-22:00	1.1 语音识别的发展 1.2 语音识别的通用流程 2.1 语音信号参数 2.2 wave 语音读取 2.3 wave 语音保存 2.4 librosa 读写音频 2.5 音频播放（拓展） 2.6 录音（拓展） 2.7 时域&频域 2.8 Mel 谱图 3.1 噪声 3.2 降噪处理 3.3 静音处理 split	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
	08 月 06 日 18:30-22:00	3.4 静音处理 trim 3.5 预加重 3.6 分帧 3.7 加窗 3.8 傅里叶变换 3.9 Mel 滤波器组特征 3.10 Log-Mel 滤波器组特征

	3.11 MFCC 特征 4.1 语音识别的一般过程 4.2 语音标注 4.3 声学模型 4.4 语言模型 4.5 端到端的语音识别	
	操作演练/作业	个人 PC
	在线答疑	微信群
第五课 实战案例：语音识别中的 HMM 声学模型		
08 月 07 日 18:30-22:00	1 背景和思路 2 语音文件处理 3.1 批量提取 mfcc 特征 3.2 特征对齐 4.1 HMM 模型 4.2 HMM 模型训练 4.3 HMM 模型预测	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第六课 实战案例：英文数字语音识别		
08 月 08 日 18:30-22:00	1 案例背景和挖掘目标 2.1 语音数据获取 2.2 数据划分 2.3 MFCC 特征提取 2.4 MFCC 特征处理 2.5 特征维度分析 2.6 标准化和数据保存 2.7 数据预处理总结 3.1 构建网络 3.2 模型编译 3.3 模型训练 3.4 模型训练过程 3.5 模型测试 3.6 课后拓展：模型应用 4 小结	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第七课 实战案例：中文语音识别		
08 月 09 日	1 背景与任务目标	泰迪云课堂

18:30-22:00	2.1 音频读取 2.2 均方根去除静音 2.3 MFCC 特征提取 2.4 特征归一化处理 2.5 特征对齐 2.6 语音标签和字典获取 2.7 特征对齐方法 3.1 CTC 模型介绍	
	操作演练/作业	个人 PC
	在线答疑	微信群
08 月 10 日 18:30-22:00	3.2 批次数据构造 3.3 CTC 模型输入数据构造 3.4 数据集划分 4.1 双向 GRU 模型 4.2 GRU 模型训练 4.3 wavenet 模型 4.4 wavenet 模型训练 5.1 模型预测 5.2 拓展和小结	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第八课 在线考试		
08 月 11 日 19:00-21:00	高级大数据分析师职业技术证书在线考试	泰迪云课堂