

全国高校大数据与人工智能师资研修班 邀请函

(2022 年第四期)

四大专题：数据采集与处理实战 (Python)、大数据分析 with 机器学习实战、
自然语言处理实战 (Python)、计算机视觉处理实战 (Python)

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 年第四期全国高校大数据与人工智能师资研修班将开设四大专题方向，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，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,讲师真人出境,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 数据采集与处理实战 (Python)	学习时间: 05月20日-05月30日, 共计88学时
证书颁发: 高级 Python 技术应用工程师	费用: 1980元 (报名费、学习费、资料费、证书费等)
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、数据采集与处理实战:《流浪地球》豆瓣影评采集、数据采集与处理实战:农产品信息采集与分析、数据采集与处理实战: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。 详见附件一 数据采集与处理实战 (Python) 课程大纲	
专题二 大数据分析 with 机器学习实战 (Python)	学习时间: 05月21日-05月31日, 共计88学时
证书颁发: 高级机器学习工程师	费用: 1980元 (报名费、学习费、资料费、证书费等)
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:百货商场用户画像描绘与价值分析、实战案例:020 优惠券使用预测、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)。 详见附件二 大数据分析 with 机器学习实战 (Python) 课程大纲	
专题三 自然语言处理实战 (Python)	学习时间: 05月21日-05月31日, 共计88学时
证书颁发: 高级大数据技术应用	费用: 2480元 (报名费、学习费、资料费、证书费等)
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、实战案例: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。 详见附件三 自然语言处理实战 (Python) 课程大纲	
专题四 计算机视觉处理实战 (Python)	学习时间: 05月20日-05月31日, 共计96学时

证书颁发：**高级大数据分析师**

费用：2480 元（报名费、学习费、资料费、证书费等）

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、TensorFlow 2 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：基于深度学习的肝脏肿瘤分割、实战案例：基于 GAN 的图像风格转换
详见附件四 人工智能实战（计算机视觉方向） 课程大纲

三、师资介绍

方海涛，中国科学院数学与系统科学研究院研究员，博士生导师，《控制理论与应用》杂志副主编，泰迪杯数据挖掘挑战赛专家组成员。主要研究兴趣包括：系统估计、优化与控制等。

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，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 年第一/二/三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、湖南省、甘肃省电力系统培训班主讲讲师，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。

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。大数据专业系列图书编写委员会成员，负责《Keras 与深度学习实战》、《Python 中文自然语言处理基础与实战》、《深度学习与计算机视觉实战》等书籍编写工作。

四、证书颁发

学员经在线培训并考试合格后，可以获得本期专题对应职业技术证书。



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

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技术证书使用，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 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	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
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 与掌握 Jupyter Notebook 常用功能 2 NumPy 数值计算基础	泰迪云课堂

<ul style="list-style-type: none"> 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 	
---	--

核心课程篇

时间	课程内容	学习平台
专题讲座		
05月20日 18:30-22:00	主讲：方海涛 主题内容：随机最优控制与人工智能	泰迪云课堂
第一课 Pandas 数据分析基础		
05月20日 18:30-22:00	<ul style="list-style-type: none"> 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 	泰迪云课堂

	<ul style="list-style-type: none"> 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 	
	操作演练	个人 PC
	在线答疑	微信群
05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		
05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 	泰迪云课堂

	1.8 箱线图 1.9 人口特征间分布 1.10 人口各个特征分布 2.1Pyecharts 基础介绍 2.2Pyecharts 绘制日历图 2.3Pyecharts 绘制漏斗图 2.4Pyecharts 绘制仪表盘 2.5Pyecharts 绘制水球图 2.6Pyecharts 绘制关系图 2.7Pyecharts 绘制直角坐标系图表 2.8Pyecharts 绘制饼图 2.9Pyecharts 绘制雷达图 2.10Pyecharts 绘制词云图 2.11Pyecharts 绘制柱状图 2.12Pyecharts 绘制树形图 2.13Pyecharts 绘制地理图表	
	操作演练	个人 PC
	在线答疑	微信群
拓展自学篇		
自行安排	3.1 seaborn 基础介绍 3.2 seaborn 简单绘图 3.3 seaborn 绘图风格 3.4 调色板 1 3.5 调色板 2 3.6 关系图 3.7 分类图 3.8 分布图 3.9 回归图 3.10 矩阵图 3.11 网格图	泰迪云课堂
第三课 Python 网络爬虫实战		
05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	泰迪云课堂

	<p>2 网页前端基础</p> <p>2.1 认识网络信息传输过程</p> <p>2.1.1 网络传输模型</p> <p>2.1.2 网络信息传输过程</p> <p>2.2 认识 HTTP</p> <p>2.2.1 认识 HTTP</p> <p>2.2.2 熟悉 Cookie</p> <p>3 简单静态网页爬取</p> <p>3.1 认识静态网页</p> <p>3.2 实现 HTTP 请求</p> <p>3.2.1 创建工程</p> <p>3.2.2 生成 HTTP 请求</p> <p>3.2.3 完善 HTTP 请求</p> <p>3.3 解析网页</p> <p>3.3.1 为什么要解析网页</p> <p>3.3.2 初识 Xpath</p> <p>3.3.3 Xpath 相对路径及属性查找</p> <p>3.3.4 使用 BeautifulSoup 解析网页</p> <p>3.3.5 网页解析小结</p> <p>3.4 存储数据</p> <p>3.4.1 认识 chrome 开发者工具</p> <p>3.4.2 任务演练：爬取并存储泰迪科技官网首页数据</p>	
	操作演练	个人 PC
	在线答疑	微信群
<p>05 月 24 日</p> <p>18:30-22:00</p>	<p>4 认识动态网页</p> <p>4.1 认识动态网页</p> <p>4.2 逆向分析爬取动态网页</p> <p>4.2.1 通过网页源码追踪目标数据文件地址</p> <p>4.2.2 通过开发者工具追踪目标数据文件地址</p> <p>4.2.3 爬取数据并进行保存</p> <p>4.3 使用 Selenium 库爬取动态网页</p> <p>4.3.1 搭建 Selenium 环境</p> <p>4.3.2 利用 Selenium 获取网页数据</p> <p>4.3.3 利用 Selenium 控制点击操作</p> <p>5 模拟登录</p> <p>5.1 使用表单登录方法实现模拟登录</p> <p>5.1.1 模拟登录的过程</p> <p>5.1.2 查找提交入口和表单数据</p>	<p>泰迪云课堂</p>

	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	
	操作演练	个人 PC
	在线答疑	微信群
05 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：《流浪地球》豆瓣影评采集		
05 月 26 日 18:30-22:00	1.1 案例背景与挖掘目标 2.1 短评数据爬取介绍 2.2 安装 selenium 及配置 chrome driver 2.3 获取用户名 2.4 获取短评正文 2.5 设置 cookies 2.6 获取用户居住地和入会时间信息 2.7 单页数据整理 2.8 自定义获取单页数据的函数 2.9 判定网页是否已被加载 2.10 翻页爬取 2.11 代码整理及小结	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
05 月 27 日	3.1 短评正文数据预处理	泰迪云课堂

	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	
	操作演练	个人 PC
	在线答疑	微信群
拓展自学篇：网站图像素材采集实战		
自行安排	1 思路介绍 2 单个图片文件爬取 3 获取一个页面所有图片网址 4 保存所有图片 5 翻页爬取更多数据 6 PDF 文件规律及问题 7 PDF 翻页刷新的网址规律 8 获取当前页所有图片网址 9 翻页刷新爬取所有图片 10 图片拼接成 PDF 文件	泰迪云课堂
第七课 在线考试		
05月30日 19:00-21:00	高级 Python 技术应用工程师技术证书在线考试	泰迪云课堂

附件二 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	泰迪云课堂

	<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>5Python 文件读取操作</p> <p>5.1Python 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>6.1Python 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>7.1Python 方法与函数对比介绍</p> <p>7.2Python 面向对象示例</p> <p>7.3Python 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>8.1Python 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>	泰迪云课堂

	<ul style="list-style-type: none"> 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 	
--	---	--

核心课程篇

时间	课程内容	学习平台
专题讲座		
05月21日 18:30-22:00	主讲：方海涛 主题内容：随机最优控制与人工智能	泰迪云课堂
第一课 Pandas 数据分析基础		
05月21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 	泰迪云课堂

	<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 2.1 合并数据 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 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 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

05月22日
18:30-22:00

- 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
- 2.1 Pyecharts 基础介绍
- 2.2 Pyecharts 绘制日历图
- 2.3 Pyecharts 绘制漏斗图
- 2.4 Pyecharts 绘制仪表盘
- 2.5 Pyecharts 绘制水球图
- 2.6 Pyecharts 绘制关系图
- 2.7 Pyecharts 绘制直角坐标系图表
- 2.8 Pyecharts 绘制饼图
- 2.9 Pyecharts 绘制雷达图
- 2.10 Pyecharts 绘制词云图

泰迪云课堂

	<ul style="list-style-type: none"> 2.11 Pyecharts 绘制柱状图 2.12 Pyecharts 绘制树形图 2.13 Pyecharts 绘制地理图表 	
	操作演练	个人 PC
	在线答疑	微信群
拓展自学篇		
自行安排	<ul style="list-style-type: none"> 3.1 seaborn 基础介绍 3.2 seaborn 简单绘图 3.3 seaborn 绘图风格 3.4 调色板 1 3.5 调色板 2 3.6 关系图 3.7 分类图 3.8 分布图 3.9 回归图 3.10 矩阵图 3.11 网格图 	泰迪云课堂
第三课 Python 机器学习算法原理与实现		
05月23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 2.4 性能度量 Python 实现 	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
05月24日 18:30-22:00	<ul style="list-style-type: none"> 3 回归分析 <ul style="list-style-type: none"> 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 决策树 <ul style="list-style-type: none"> 4.1 从女生相亲到决策树 4.2 明天适合打球吗 	泰迪云课堂

	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	
	操作演练/作业	个人 PC
	在线答疑	微信群
05月25日 18:30-22:00	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 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 实现神经网络算法	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
05月26日 18:30-22:00	6 最近邻算法 (KNN) 6.1 KNN 算法介绍 6.2 KNN 算法解决鸢尾花分类问题 7 朴素贝叶斯 7.1 非洲人还是北美人 7.2 为什么有“朴素”二字 7.3 拉普拉斯修正 7.4 用高斯朴素贝叶斯算法解决鸢尾花分类问题 8 聚类分析 8.1 聚类分析概述 8.2 相似性度量 8.3 K-Means 聚类分析算法介绍 8.4 利用 K-Means 算法对鸢尾花进行聚类 8.5 聚类结果的性能度量 8.6 调用 sklearn 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

拓展自学篇		
自行安排	9 集成学习 9.1 集成学习基本概念 9.2 并行集成算法-Bagging&RandomForest 9.3 串行集成算法-Boosting 算法流程 9.4 串行集成算法-Boosting 代码实现 9.5 Stacking 算法流程 9.6 Stacking 代码实现	泰迪云课堂
第四课 实战案例：学生校园消费行为分析		
05月27日 18:30-22:00	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 缺失值处理 2.4 异常值处理 2.5 数据合并 3 数据可视化分析 3.1 食堂就餐行为分析 3.2 学生早中晚餐就餐地点占比饼图 3.3 工作日和非工作日食堂就餐时间曲线图 3.4 不同专业、不同性别的学生消费分析 4 模型构建 4.1 构建聚类模型 4.2 模型性能评估 5 小结	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：百货商场用户画像描绘与价值分析		
05月28日 18:30-22:00	1.1 背景与分析目标 2.1 会员信息表处理 2.2 销售流水表处理 3.1 会员年龄分析 3.2 不同年龄的消费能力 3.3 不同性别的消费情况 3.4 会员和非会员消费情况 3.5 商场会员年消费趋势 3.6 不同月份的消费趋势 3.7 每年每月的消费金额趋势	泰迪云课堂

	3.8 不同时刻的消费情况	
	操作演练/作业	个人 PC
	在线答疑	微信群
第六课 综合实战：020 优惠券使用预测		
05月29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 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 3&4 6.5 拼接测试及的特征 3&4 7.1 模型训练 7.2 预测 7.3 代码整理 7.4 结果提交	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
第七课 泰迪内推平台信息精准推荐应用（基于泰迪建模平台实现）		
05月30日 18:30-22:00	1 背景与目标 2 数据预处理 2.1 数据读取	泰迪云课堂

	2.2 数据预处理 2.3 数据拆分 3 模型构建 3.1 模型选型分析 3.2 模型构建 4 模型评价 4.1 模型性能评估 4.2 模型优化 5 部署设置	
	操作演练/作业	个人 PC
	在线答疑	微信群
第八课 在线考试		
05月31日 19:00-21:00	高级机器学习工程师职业资格证书在线考试	泰迪云课堂

附件三 自然语言处理实战（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	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	泰迪云课堂

	<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 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 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 	泰迪云课堂

- 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
 - 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		
05月21日 18:30-22:00	主讲：方海涛 主题内容：随机最优控制与人工智能	泰迪云课堂
第一课 Python 机器学习算法原理与实现		
05月21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 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	泰迪云课堂

	<p>4.8 网络输出的 Python 实现</p> <p>4.9 单样本网络训练的 Python 实现</p> <p>4.10 全样本网络训练的 Python 实现</p> <p>4.11 网络性能评价</p> <p>4.12 调用 sklearn 实现神经网络算法</p>	
	操作演练	个人 PC
	在线答疑	微信群
第二课 TensorFlow2 实战		
05 月 22 日 18:30-22:00	<p>1 任务 1: 构建一个线性模型</p> <p>1.1 tensorflow 介绍</p> <p>1.2 tensorflow2 常用数据类型和操作</p> <p>1.3 初始化模型</p> <p>1.4 构建损失函数</p> <p>1.5 模型训练及可视化</p> <p>1.6 使用高阶 API-keras</p> <p>2 任务 2: mnist 手写数字识别</p> <p>2.1 数据读取及探索</p> <p>2.2 交叉熵</p> <p>2.3 模型构建及训练</p> <p>2.4 调用保存好的模型对新样本进行预测</p>	泰迪云课堂
	操作演练	个人 PC
	在线答疑	微信群
第三课 TensorFlow 2 深度学习原理与实现		
05 月 23 日 18:30-22:00	<p>1.1 深度神经网络-引言</p> <p>2 卷积神经网络 CNN</p> <p>2.1 浅层神经网络的局限</p> <p>2.2 卷积操作</p> <p>2.3 卷积操作的优势</p> <p>2.4 池化及全连接</p> <p>2.5 高维输入及多 filter 卷积</p> <p>2.6 实现卷积操作</p> <p>2.7 实现池化操作</p>	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
05 月 24 日 18:30-22:00	<p>3 循环神经网络 RNN</p> <p>3.1 循环神经网络简介</p> <p>3.2 循环神经网络的常见结构</p> <p>4 长短时记忆网络 LSTM</p>	泰迪云课堂

	<ul style="list-style-type: none"> 4. 1LSTM 的三个门 4. 2LSTM 三个门的计算示例 4.3 利用 RNN&LSTM 实现 mnist 手写数字识别 5.1 生成对抗神经网络简介 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		
05月25日 18:30-22:00	<ul style="list-style-type: none"> 1 自然语言处理概述 1.1 自然语言处理概述 2NLP 的基本流程 2.1 NLP 的基本流程 2.2 语料字符处理 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 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 	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
05月26日 18:30-22:00	<ul style="list-style-type: none"> 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 2.6 jieba 分词 2.7 去停用词 	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
05月27日 18:30-22:00	<ul style="list-style-type: none"> 3 文本向量化 3.1.1 向量化与独热编码 	泰迪云课堂

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	
3.3 TF-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	
操作演练/作业	个人 PC
在线答疑	微信群

第五课 案例实战：公众健康问句分类

05月28日 18:30-22:00	<ul style="list-style-type: none"> 1 背景与挖掘目标 2 数据读取与探索 3.1 分词与去除停用词 3.2 词嵌入原理介绍 3.3 词嵌入实现 3.4 问句数据转换 4.1 LSTM 模型原理介绍 4.2 LSTM 模型构建 4.3 性能度量 4.4 模型评估 5 小结 	<p>泰迪云课堂</p>
	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

05月29日 18:30-22:00	<ul style="list-style-type: none"> 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 	<p>泰迪云课堂</p>
-----------------------	--	--------------

	<ul style="list-style-type: none"> 3.4 问答向量长度填充 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 	
	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

<p>05月30日 18:30-22:00</p>	<ul style="list-style-type: none"> 1 任务 1-了解背景与目标 2 数据探索 <ul style="list-style-type: none"> 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 2.3 数据获取与探索分析 3 获取用户单击流数据 <ul style="list-style-type: none"> 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 <ul style="list-style-type: none"> 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 <ul style="list-style-type: none"> 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 <ul style="list-style-type: none"> 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 <ul style="list-style-type: none"> 7.1 任务 7 说明-模型优化 	<p>泰迪云课堂</p>
-------------------------------	--	--------------

	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		
05月31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

附件四 计算机视觉处理实战（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	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	泰迪云课堂

	8.3 结语	
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> <p>3 Matplotlib 数据可视化基础</p> <p>3.1 掌握绘图基础语法与常用参数</p> <p>3.1.1 Matplotlib 介绍</p> <p>3.1.2 基础图形绘制</p> <p>3.1.3 常用参数设置</p> <p>3.2 分析特征间关系</p> <p>3.2.1 绘制散点图</p> <p>3.2.2 散点图参数设置</p> <p>3.2.3 绘制折线图</p> <p>3.3 分析特征内部数据分布与分散情况</p>	泰迪云课堂

- 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		
05月20日 18:30-22:00	主讲：方海涛 主题内容：随机最优控制与人工智能	泰迪云课堂
第一课 Python 机器学习算法原理与实现		
05月20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
05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 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

05月22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	泰迪云课堂	
	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

05月23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 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	泰迪云课堂	
	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	4 循环神经网络 RNN	泰迪云课堂
------	--------------	-------

	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	
	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		
05月24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	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
05月25日 18:30-22:00	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	泰迪云课堂
	操作演练/作业	个人 PC
	在线答疑	微信群
05月26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	泰迪云课堂

	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		
05月27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	
	4.3 模型训练	
	4.4 模型预测	
	5 模型验证及小结	
	操作演练/作业	个人 PC
	在线答疑	微信群
第六课 计算机视觉实战：基于深度学习的肝脏肿瘤分割		
05月28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	
	操作演练/作业	个人 PC
	在线答疑	微信群
05月29日 18:30-22:00	3.1 数据增强：定义图像生成器	泰迪云课堂
	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
第七课 实战案例：基于 GAN 的图像风格转换		
05 月 30 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		
05 月 31 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