

# 大数据分析挖掘与可视化实践培训班(面授/直播/录播)

产品名称	大数据分析挖掘与可视化实践培训班(面授/直播/录播)
公司名称	北京中培伟业管理咨询有限公司
价格	.00/件
规格参数	
公司地址	北京丰台夏家胡同育芳园东里3号楼B座
联系电话	400-808-2006 13910781835

## 产品详情

### 培训背景

在大数据时代，数据可视化技术在广泛应用的同时，也面临诸多新的挑战。大数据可视化是一个面向应用的研究领域，大数据可视化就是将抽象的“数据”以可见的形式表现出来，帮助人理解数据。大数据可视化相对传统的数据可视化，处理的数据对象有了本质不同，在已有的小规模或适度规模的结构化数据基础上，大数据可视化需要有效处理大规模、多类型、快速更新类型的数据。这给数据可视化研究与应用带来一系列新的挑战。

### 培训对象

从事大数据管理、大数据可视化分析和大数据建模应用的工作人员以及大数据爱好者、大中专在读学生。

### 培训特色

- 1.理论与实践相结合、案例分析与行业应用穿插进行；
- 2.专家精彩内容解析、学员专题讨论、分组研究；
- 3.通过全面知识理解、专题技能和实践结合的授课方式。

### 培训重点

- 1.数据分析实战

## 2.数据挖掘理论及核心技术

## 3.大数据算法原理及案例实现

## 4.Python应用实战

### 课程大纲（3天，每天6-7小时）

时间	知识模块	授课内容
第一天 上午	数据分析实战	第一讲 零基础学Python 讲解Python背景、国内发展状况、基础语法、数据结构及绘图操作等内容。算这块，着重介绍Python在这方面的优势及用法。
		第二讲 数据分析方法论 讲解统计分析基础，包括统计学基本概念，假设检验，置信区间等基础，并明其使用场景和运用方法。介绍数据分析流程和常见分析思路，并结合案例
第二天 下午	数据挖掘理论及核心技术	第三讲 数据处理技法 从数据接入、数据统计、数据转换等几个方面进行讲解。数据接入包含接入、Hadoop等常见数据库操作；数据统计包含Pandas包的具体用法和讲解；数据集的关联、合并、重塑等操作。此外，针对海量数据的情况下，介绍在数据处理技术，并结合真实环境进行操作讲解。
		第四讲 认识数据挖掘 讲解数据挖掘基本概念，细致讲解业务理解、数据理解、数据准备、建立模型部署各环节的工作内容及相关技术；结合业界经典场景，讲解数据挖掘法体系。
第二天	大数据算法原理及案例实现（1）	第五讲 数据挖掘核心技术 细致讲解抽样、分区、样本平衡、特征选择、训练模型、评估模型等数据挖掘，并结合案例讲解其具体实现和用法。尤其针对样本平衡，重点讲解人工合算法；针对特征选择，重点讲解特征选择的核心思路，并结合Python进行案
		第六讲 特征降维算法及Python实现 降维是大数据分析非常重要的算法，它可以在降低极少信息量的情况下，极模。主要讲解主成分、LDA以及t-SNE原理，并结合案例进行Python实现。特量数据情况下的应用场景，讲解实现思路和Python案例。
		第七讲 决策树算法及Python实现 决策树是非常经典的算法，一般常见于小数据的挖掘。由于决策树具有极强对海量数据仍然是非常重要的实用价值。主要讲解ID3、C4.5、C5.0以及CA实现原理，并结合案例进行Python实现。
		第八讲 好莱坞百万级影评数据分析与电影推荐实现 实战部分：基于好莱坞百万级的影评数据，对数据进行建模、清洗、透视表用户画像分析不同的用户喜好通过机器学习算法对不同性别、年龄阶段的用电影推荐，后把推荐的电影进行可视化的展示操作
		第十讲 深度学习算法及Python实现

对于大数据的建模任务，我们可以基于深度学习来实现，不仅能够针对海量数据进行建模，其效果也非常不错。本节主要讲解深度学习的发展历程，DBN、DNN等经典深度学习算法，深度学习优化算法以及一些技巧。同时，介绍Keras、OpenCV库的使用方法，并结合案例进行Python实现。

#### 第十一讲 采用OpenCV实现计算机视觉技术

实战部分：基于OpenCV面部模型，完成对图片和视频的人脸识别，实战中会讲解OpenCV的重要类和函数。主要内容包括OpenCV库的安装和部署、图像增强、像素操作、图形分析等各种技术，并且详细介绍了如何处理来自文件或摄像机的视频，以及如何检测和跟踪移动对象。

第三天	Python应用实战（一）	<p>第十二讲 Python自然语言处理原理及案例</p> <p>目前文档数据已经成为很多企业重要的资产，通过对文档数据进行解析、建模、可视化，我们能够发现不一样的洞察。本节主要讲解自然语言处理基本概念、分词、关键字提取、文摘提取、文本分类、主题模型、word2vec等内容。在AI加持下，与传统做法的区别，并使用Python进行案例讲解。</p> <p>第十三讲 数据分析图表及Python案例</p> <p>数据可视化是大数据分析的重要手段，通过合理地使用图表，不仅可以简洁明了，高效地发现问题，还可以为报告的编写以及数据分析web应用增色不少。常用的数据分析图表及其使用场景，介绍数据可视化的方法论，避免生搬硬套。针对不同的业务场景和需求，合理选择可视化方法。介绍的工具不限于matplotlib、ggplot、Bokeh、HoloViews、mpld3、plotly、pygal等常用可视化库。</p>
-----	---------------	---

#### 第十五讲 Seaborn可视化开发实战

Seaborn是一款不错的可视化框架，它和Pandas一样是建立在Matplotlib之上的。可以基于Seaborn快速开发一个轻量级的数据分析web应用。在网页中嵌入图表、数据以及分析的算法，非常适合打造企业内部的敏捷数据分析工具集。本节主要介绍Pie、Scatter、Radar等各种可视化解决方案，同时讲解一个用Seaborn实现数据分析功能（兼图表）的实际案例，搭建服务器，在企业内部实现轻量级数据分析应用。

#### 授课专家

刘老师 10多年的IT领域相关技术研究和项目开发工作，在长期软件领域工作过程中，对软件企业运作模式有深入研究，熟悉软件质量保障标准ISO9003和软件过程改进模型CMM/CMMI，在具体项目实施过程中总结经验，有深刻认识。通晓多种软件设计和开发工具。对软件开发整个流程非常熟悉，能根据项目特点定制具体软件过程，并进行项目管理和监控，有很强的软件项目组织管理能力。对C/C++、HTML5、python、Hadoop、java、java EE、android、IOS、大数据、云计算有比较深入的理解和应用，具有较强的移动互联网应用需求分析和系统设计能力，熟悉Android框架、IOS框架等技术，了解各种设计模式，能在具体项目中灵活运用。

#### 培训费用

线下面授培训费：7800元/人（含培训费、场地费、资料费、学习期间午餐及录播视频回放一年）。

网络直播培训费：6800元/人（含培训费、平台费、资料费及直播视频回放一年）。

## 结业证书

本课程由中国信息化培训中心颁发《大数据分析工程师》证书；

证书可作为技术人员职业能力考核的证明，以及技术人员岗位聘用、任职、定级和晋升职务的重要依据

。