《大数据分析与挖掘》课程实验教学大纲

一、课程基本信息

课程代码：16054103

课程名称：大数据分析与挖掘

英文名称: Big data analysis and mining 实验总学时：16

适用专业：软件工程、计算机科学与技术课程类别：专业选修课

先修课程：数理统计与概率论，算法设计，JAVA/Python程序设计

二、实验教学的总体目的和要求

实验教学目的：

**向学生教授经过实践检验的真理才是正确的，所有理论知识都应该经过实践检验，计算机科学这样，大数据分析与挖掘这门前沿信息技术尤其如此。**本课程实验是为了使学生在课程学习的同时，通过实验教学验证课堂教学的理论，理解和掌握大数据分析与挖掘中最基本、最广泛应用的概念、原理、理论和算法以及基本技术和方法，更好地掌握《大数据分析与挖掘》课程教学大纲要求的内容。

实验要求：

1. 对学生的要求：实验前要充分做好准备工作：
   * 复习和掌握与本实验有关的知识内容；
   * 预习、思考实验内容；
   * 对实验内容进行分析和设计。

实验过程中，实验者必须服从指导教师和实验室工作人员的安排，遵守纪律与实验制度，爱护设备及卫生。在指定的实验时间内，必须到实验室内做实验。

对于上机过程中出现的问题，尽量先独立思考和解决；对于难以解决的问题可以和同学交流或询问老师；对于同一个实验题目，可以考虑多种方法来实现，

然后比较并选择出一种较为有效的方法来实现。

对于设计型和验证型实验，实验时一人一组，独立上机。 2、对实验条件的要求：普通 PC 机房。

三、实验教学内容

实验项目一实验名称：数据分析与挖掘平台的搭建

实验内容：在 Windows（Linux,Mac）操作系统中安装 Python。可以使用官方下载 Python 的安装包安装，也可以使用 Anaconda 安装，还可以安装 Pycharm。安装完之后，进行入门操作，熟悉 Python 的使用。

实验性质：设计型实验学时：2

实验目的与要求：必修实验条件：实验机房，**阿里天池AI实训平台，数据分析平台。通过天池AI数据分析实训平台介绍与使用，使同学们掌握业内最先进的大数据分析挖掘开发平台，进行高质量的理论实践，提高动手能力，同时也可了解国内顶尖大数据分析技术发展现状，增强民族自豪感，激发学习热情。**

实验项目二实验名称：Python编程实践

实验内容：熟练掌握python编程技术，能将函数，代码复用，组合数据类型等熟练应用在编程实践中。

实验性质：设计型

实验学时：2

实验目的与要求：必修实验条件：实验机房

实验项目三实验名称：数据探索与数据预处理

实验内容：了解数据探索及预处理的常用经典方法及应用，并可编程用在实际问题领域。

实验性质：设计型

实验学时：4

实验目的与要求：必修实验条件：实验机房

实验项目四实验名称：电力窃漏电用户自动识别

实验内容：通过电力窃漏实际问题，掌握分类算法的原理以及应用。实验性质：设计型

实验学时：4

实验目的与要求：必修

实验条件：**阿里天池AI实训平台，数据分析平台。**实验机房。**通过天池AI数据分析实训平台使用，使同学们获得业内最先进的大数据分析挖掘开发平台技术，高质量的理论实践，提高动手能力，了解国内顶尖大数据分析技术发展现状，增强民族自豪感，激发学习热情。**

实验项目五实验名称： 航空公司客户价值分析

实验内容：通过航空公司的用户数据，掌握聚类算法的原理及使用，并完成客户流失问题的研究。

实验性质：设计型

实验学时：4

实验目的与要求：必修

实验条件：**阿里天池AI实训平台，数据分析平台。**实验机房。**通过天池AI数据分析实训平台使用，使同学们获得业内最先进的大数据分析挖掘开发平台技术，高质量的理论实践，提高动手能力，了解国内顶尖大数据分析技术发展现状，增强民族自豪感，激发学习热情。**

四、考核方式

1. 实验课的考核方式：每个实验单独计分。验证型和设计型实验依据源程序和运行结果评分，综合性实验要结合课程设计报告综合评分。
2. 实验课考核成绩的确定：实验课成绩计入课程总成绩，验证型和设计型实验成绩计为课程平时成绩。

**五、推荐实验教材和教学参考书**

推荐教材：

1.Python数据分析与挖掘实践（第二版），张良均，机械工业出版社，2019.3

2.数据挖掘导论（第二版）[Pang-Ning Tan](https://book.douban.com/search/Pang-Ning%20Tan) / [Michael Steinbach](https://book.douban.com/search/Michael%20Steinbach) / [Vipin Kumar](https://book.douban.com/search/Vipin%20Kumar)著，[范明](https://book.douban.com/search/%E8%8C%83%E6%98%8E) / [范宏建](https://book.douban.com/search/%E8%8C%83%E5%AE%8F%E5%BB%BA)译，人民邮电出版社，2018.10

3.Python语言程序设计基础(第二版)嵩天、礼欣、黄天羽著，高等教育出版社，2017.2

参考书目：

1. 机器学习，周志华，清华大学出版社

2. 利用Python进行数据分析，Wes Mckinney, 机械工业出版社

3. 高教版Python语言程序设计冲刺试卷(含线上题库)（第二版），黄天羽、李芬芬著，高等教育出版社，2019.14

**4. Python基础教程 Magnus Lie Hetland 人民邮电出版社  02/2018**

**5. 天池AI实训平台**

**https://tianchi.aliyun.com/education?spm=5176.15098393.J\_4642415020.12.529e5cd6UKUiAS&lang=zh-cn**

1. **其他需说明的**

无。

大纲修订人 ： 胡苏 修订日期 : 2022年1月

大纲审定者 ： 郭玉柱 审定日期 : 2022年1月

**注：在第二、三、五部分引入课程思政建设内容，其中第三部分的教学内容部分，引入思政教学知识点四个以上，均已标红加粗显示。**