

湖南省高等教育自学考试 课程考试大纲

智能数据处理

(理论部分)

(课程代码：12577)

湖南省教育考试院组编

2022年7月

高等教育自学考试课程考试大纲

课程名称：智能数据处理

课程代码：12577

第一部分 课程性质与目标

一、课程性质与特点

智能数据处理是高等教育自学考试物联网工程专业的专业核心课程。主要介绍了数据处理基础与人工智能等信息前沿技术、基础数据处理可视化操作方法、智能数据分析语言等内容。

二、课程目标与基本要求

通过本课程的学习，要求考生了解掌握数据处理基础及智能分析技术前沿、数据存储、数据存储应用（Access 2016）、以及通过 Python 程序设计语言完成简单的数据获取、数据智能分析和数据可视化展示等数据智能分析技术，培养和提高计算思维和智能计算思维。

三、与本专业其他课程的关系

本课程的先修课为《数据库原理》、《计算机技术基础》，学习本门课程为学习信息科学相关后续课程和利用信息科学的有关知识和工具解决本专业及相关领域的问题打下良好的基础。

第二部分 考核内容与考核目标

第一章 数据处理基础与人工智能

一、学习目的与要求

通过本章的学习，要求考生掌握信息、数据、大数据等基本概念和常用的数据智能处理和存储技术，了解智能数据分析的定义和步骤，并对数据挖掘、机器学习和区块链等智能科学前沿技术的应用领域和未来发展趋势有直观的认识。

二、考核知识点与考核目标

（一）信息、数据、大数据（重点）

识记：信息、数据、大数据的定义

理解：1.信息的特性

2.大数据的价值

应用：数据的结构、产生方式与阶段

（二）信息（数据）处理技术分类及发展（次重点）

识记：1.信息（数据）处理技术的定义和方式

2.信息（数据）存储技术的定义和数据存储方式

3.智能数据分析技术的定义和步骤

理解：1.数据储存的结构化数据储存—数据库技术

2.云储存技术

(三) 人工智能与信息技术前沿 (一般)

识记: 1.人工智能的定义

2.数据挖掘的定义

3.机器学习定义与类别

4.大数据处理与人工智能难题与对策

5.区块链的定义

6.智能科学发展新趋势中主要体现的八个方面

理解: 1.人工智能的研究与应用领域

2 数据挖掘步骤

3. 机器学习的应用

4. 区域连技术的核心技术和应用领域

第二章 数据存储基本理论 (关系数据库)

一、学习目的与要求

通过本章的学习, 要求考生掌握数据模型、数据库基本理论、数据库体系结构, 还主要介绍了数据库系统模式、内模式和外模式的三级体系结构和数据库管理系统的基本功能等内容。

二、考核知识点与考核目标

(一) 数据库实例与数据模型 (重点)

识记: 数据模型的定义

理解: 数据模型

应用: Access 数据库实例

(二) 关系数据模型的基本理论 (次重点)

识记: 1.关系数据模型的三要素

2.关系的定义、特点及相关概念

理解: 关系数据库的数据完整性约束

(三) 关系数据理论的进一步分析 (一般)

识记: 1.关系代数包括的一些运算

2.关系规范化的相关概念

理解: 关系模型与关系数据库

应用: 关系的并、交、差等运算

(四) 数据库体系结构 (一般)

识记: 三级模式和二级映射的相关知识

理解: 数据库管理系统概述

第三章 数据存储设计与 Access 数据库管理

一、学习目的与要求

通过本章的学习, 要求考生掌握数据存储中关系数据库的设计方法和详细步

骤，并以 Access2016 为例，详细地介绍了 Access 的特点、Access2016 的启动和工作界面，以及数据库的概念、存储、创建和管理等操作内容。

二、考核知识点与考核目标

- (一) 数据库设计方法（次重点）
 - 识记：1.数据库设计的定义
 - 2.数据库设计的步骤
- (二) 实体联系模型及转化（一般）
 - 识记：1.E-R 模型的基本概念
 - 2.E-R 图的定义
 - 3.E-R 模型向关系模型的转化方法
 - 4.设计 E-R 模型的方法
 - 5.术语对照表
 - 理解：需求调查与分析
- (三) 图书销售管理数据库设计（重点）
 - 识记：需求分析报告的内容以及相关概念
 - 理解：概念设计与逻辑设计
- (四) Access 概述（一般）
 - 识记：Access 的发展历程和优势
 - 应用：1.Access 的启动和退出
 - 2.Access 的用户界面
- (五) 创建 Access 数据库（重点）
 - 识记：Access 数据库对象和数据库存储的概念
 - 应用：创建数据库
- (六) Access 数据库管理（重点）
 - 应用：1.数据库的打开与关闭
 - 2.数据库管理

第四章 表与关系

一、学习目的与要求

通过本章的学习，要求考生掌握表结构的基本概念、设计视图创建表的方法、了解关系的定义方法及不同设置对操作数据的影响，还需要掌握对表的操作和设置以及表的导出等内容。

二、考核知识点与考核目标

- (一) Access 数据库的表对象及创建方法（一般）
 - 应用：建立表的几种方式
- (二) 数据类型、表的创建（次重点）
 - 识记：1.数据类型的定义
 - 2.数据库的物理设计

- 应用：1.应用设计视图创建表
2.用其他方法创建表
- (三) 建立表之间的关系（一般）
识记：表之间关系的类型以及表的创建
应用：对关系进行编辑
- (四) 表的操作（重点）
识记：表之间关系的类型及创建
应用：1.表记录的输入、修改和删除
2.对表的其他操作和导出
3.修改表结构和删除表

第五章 数据存储中的查询

一、学习目的与要求

通过本章的学习，要求考生掌握 Access 查询对象的意义、基础和用法、数据定义、数据查询、数据维护的命令及用法，和各种类型的可视交互查询设计视图的使用方法，通过本章的深入学习，要求考生能够对关系数据库和 SQL 的应用有深刻的认识，并能熟练地应用 Access 管理数据。

二、考核知识点与考核目标

- (一) 查询及查询对象（重点）
识记：1.查询的相关概念
2.SQL 的发展过程、基本功能、使用方式、以及特点
理解：1.查询的分类与查询对象
2.理解查询
应用：Access 查询的工作界面
- (二) SQL 查询（次重点）
识记：1.Access 数据运算、运算对象的内容以及表达式的定义
2.几种常用 SQL 查询的用法
理解：1.SQL 的追加功能
2.SQL 的更新功能
应用：1.SQL 的删除功能
2.SQL 的定义功能
- (三) 选择查询（一般）
识记：选择查询设置 DISTINCT、TOP 等子句的操作方法以及参数的概念
理解：汇总与分组统计查询设计
应用：1.创建选择查询
2.子查询设计与交叉表查询
- (四) 查询向导与动作查询（一般）
识记：1.简单查询向导的使用方法

2. 交叉表查询向导的使用方法
3. 查找重复项查询向导的使用方法
4. 查找不匹配项查询向导的使用方法

- 应用：1.生成表查询
2. 追加查询
3. 更新与删除查询

（五）SQL 特定查询（一般）

- 识记：1.联合查询的特点、功能
2.传递查询的特点、功能
理解：数据定义查询

第六章 智能数据分析语言——Python

一、学习目的与要求

通过本章的学习，要求考生掌握 Python 语言在进行程序设计时的语法基础和操作，主要介绍了搭建 Python 编程环境、Python 变量的使用、Python 中的不同数据类型的定义等内容，要求考生学习模仿一些已有的 Python 案例应用，逐步提高编程水平。

二、考核知识点与考核目标

（一）Python 语言概述（次重点）

- 识记：Python 流行的原因
理解：为何使用 Python 做数据分析

（二）Python 的安装与配置（重点）

- 理解：第一个 Python 程序
应用：Windows 系统中下载并安装 Python

（三）Python 语言基础（次重点）

- 识记：Python 语法特点
理解：1.变量
2. 常用数据类型
3. 运算符和表达式
4. 常用序列结构
5. 循环控制语句
6. 函数
7. 模块

应用：基本输入输出

（四）AI Studio 平台介绍（一般）

- 理解：运行一个简单的项目
应用：新建一个简单的项目

第七章 数值数据智能分析技术

一、学习目的与要求

通过本章的学习，要求考生掌握数值数据智能分析技术的基础知识、数据的创建、数据的导入和导出、数据的统计、数据的合并、连接和排序、数据的筛选和过滤等内容，通过大量相关实例分析，帮助考生巩固数据分析和可视化方法和技巧。

二、考核知识点与考核目标

（一）数值数据智能分析基础（重点）

识记：NumPy 数据处理的优势

理解：Pandas 库基础

（二）数值数据的导入和导出（重点）

识记：Python 与 Access 数据库的连接以及通过 sql 语句实现对数据库的基本操作

应用：1.导入、导出 CSV 文件

2. 导入、导出 Excel 文件

（三）数据统计、数据合并、连接和排序（次重点）

识记：基本和分组统计的相关概念

理解：1.Pandas 合并、连接

2.排序

应用：掌握 Pandas 合并、连接以及排序

（四）数据筛选和过滤功能（一般）

识记：1.筛选的方法

2.按筛选条件进行汇总的基本操作

理解：过滤

（五）数据科学制图（一般）

识记：Matplotlib 的功能特点

理解：1.折线图

2.散点图

3.柱状图

4.饼状图

5.雷达图

6.三维图形

第八章 文本数据智能分析技术

一、学习目的与要求

通过本章的学习，要求考生掌握文本数据智能分析基础知识，学习如何利用 Python 的工具库，通过自然语言处理技术进行文本分析，包括数据获取、文本数据的输输入和输出、中文分词技术、数据预处理技术和自然语言处理技术等。要求考生通过本章的学习，可以更加熟练地掌握文本数据智能分析的基本技术和理

论知识。

二、考核知识点与考核目标

(一) 数据获取及文本数据的输入和输出 (重点)

识记: 数据获取的定义

理解: 数据获取的方法

应用: 1. 文本数据的输入和输出
2. 导入、导出 TXT 文件

(二) 中文分词技术 (一般)

识记: 中文分词、精确模式、全模式、搜索引擎模式、jieba 分词的定义

理解: 中文分词、精确模式、全模式、搜索引擎模式、jieba 分词的特征

(三) 数据预处理技术 (一般)

识记: 1. 噪声的定义

2. 词性分析的定义

3. 停用词的定义

理解: 1. 回归和异常值检测

2. 停用词的分类

(四) 自然语言处理技术 (一般)

识记: 1. 词频统计的定义和功能

2. 词云分析的用途

第九章 人工智能分析方法

一、学习目的与要求

通过本章的学习, 要求考生掌握基于机器学习的智能数据分析方法, 本章主要涉及机器学习的相关概念、代表性方法、应用实例及其在 Python 中的实现过程等内容。

二、考核知识点与考核目标

(一) 机器学习简介 (重点)

识记: 1. 机器学习的基本概念

2. Python 机器学习库与学习平台

(二) 有监督学习 (一般)

识记: 1. 回归分析的定义

2. 决策树的定义以及构造决策树的核心问题

理解: 1. 支持向量机

2. KNN 算法

3. 人工神经网络

(三) 无监督学习 (一般)

识记: 无监督学习简介

理解: K-means 聚类

第十章 智能计算思维及其应用

一、学习目的与要求

通过本章的学习，要求考生掌握智能计算思维的概念、特征、结构、发展和价值，并通过 Python 算法案例展现了算法思维在求解问题中的应用。

二、考核知识点与考核目标

(一) 计算思维与人工智能（重点）

识记：1. 计算思维的概念和特征

2. 智能计算的定义和计算原则

理解：智能计算思维的应用

(二) 智能计算思维中的算法思维（次重点）

识记：1. 智能计算思维的组成

2. 算法思维的条件

理解：算法思维的表达和结构

应用：算法思维在求解问题中的应用

第三部分 有关说明与实施要求

一、考核的能力层次表述

本大纲在考核目标中，按照“识记”、“理解”、“应用”三个能力层次规定其应达到的能力层次要求。各能力层次为递进等级关系，后者必须建立在前者的基础上，其含义是：

识记：能知道有关的名词、概念、知识的含义，并能正确认识和表述，是低层次的要求。

理解：在识记的基础上，能全面把握基本概念、基本原理、基本方法，能掌握有关概念、原理、方法的区别与联系，是较高层次的要求。

应用：在理解的基础上，能运用基本概念、基本原理、基本方法联系学过的多个知识点分析和解决有关的理论问题和实际问题，是最高层次的要求。

二、教材

1. 指定教材

人工智能与数据处理基础，杨璠、张承德，清华大学出版社，2021 年版。

三、自学方法指导

1. 在开始阅读指定教材某一章之前，先翻阅大纲中有关这一章的考核知识点及对知识点的能力层次要求和考核目标，以便在阅读教材时做到心中有数，有的放矢。

2. 阅读教材时，要逐段细读，逐句推敲，集中精力，吃透每一个知识点，对基本概念必须深刻理解，对基本理论必须彻底弄清，对基本方法必须牢固掌握。

3. 在自学过程中，既要思考问题，也要做好阅读笔记，把教材中的基本概念、原理、方法等加以整理，这可从中加深对问题的认知、理解和记忆，以利于突出重点，并涵盖整个内容，可以不断提高自学能力。

4. 完成书后作业和适当的辅导练习是理解、消化和巩固所学知识，培养分析问题、解决问题及提高能力的重要环节，在做练习之前，应认真阅读教材，按考核目标所要求的不同层次，掌握教材内容，在练习过程中对所学知识进行合理的回顾与发挥，注重理论联系实际和具体问题具体分析，解题时应注意培养逻辑性，针对问题围绕相关知识点进行层次（步骤）分明的论述或推导，明确各层次（步骤）间的逻辑关系。

四、对社会助学的要求

1. 应熟知考试大纲对课程提出的总要求和各章的知识点。
2. 应掌握各知识点要求达到的能力层次，并深刻理解对各知识点的考核目标。
3. 辅导时，应以考试大纲为依据，指定的教材为基础，不要随意增删内容，以免与大纲脱节。
4. 辅导时，应对学习方法进行指导，宜提倡“认真阅读教材，刻苦钻研教材，主动争取帮助，依靠自己学通”的方法。
5. 辅导时，要注意突出重点，对考生提出的问题，不要有问即答，要积极启发引导。
6. 注意对考生能力的培养，特别是自学能力的培养，要引导考生逐步学会独立学习，在自学过程中善于提出问题，分析问题，做出判断，解决问题。
7. 要使考生了解试题的难易与能力层次高低两者不完全是一回事，在各个能力层次中会存在着不同难度的试题。
8. 助学学时：本课程共 4 学分，建议总课时 72 学时，其中助学课时分配如下：

章次	章节名称	学时
第一章	数据处理基础与人工智能	4
第二章	数据存储基本理论（关系数据库）	8
第三章	数据存储设计与 Access 数据库管理	8
第四章	表与关系	12
第五章	数据存储中的查询	12
第六章	智能数据分析语言——Python	12
第七章	数值数据智能分析技术	4
第八章	文本数据智能分析技术	4
第九章	人工智能分析方法	4
第十章	智能计算思维及其应用	4
合计		72

五、关于命题考试的若干规定

1. 本大纲为理论部分考试大纲，各章所提到的内容和考核目标都是考试内容。试题覆盖到章，适当突出重点。

2. 试卷中对不同能力层次的试题比例大致是：“识记”为 50%、“理解”为 40%、“应用”为 10%。
3. 试题难易程度应合理：易、较易、较难、难比例为 2：3：3：2。
4. 每份试卷中，各类考核点所占比例约为：重点占 60%，次重点占 30%，一般占 10%。
5. 试题类型一般分为：单项选择题、填空题、名词解释题、简答题。
6. 本课程由理论部分考试成绩和实操部分考核成绩两部分组成且缺一不可（缺少任一部分不予登分），两个部分的成绩分别占 60%和 40%，即课程成绩=理论部分考试成绩×60%+实操设计考核成绩×40%。成绩均当次有效。理论部分考试采用闭卷笔试，考试时间 90 分钟。实操部分考核由主考学校严格按照经我省审核通过的实操部分考核大纲组织实施。

六、题型示例（样题）

一、单项选择题（本大题共■小题，每小题■分，共■分）

在每小题列出的四个备选项中只有一个是符合题目要求的，请将其选出并将“答题卡”上的相应字母涂黑。错涂、多涂或未涂均无分。

1.以下算法属于相似重复记录清理常用算法的是

- A. 多趟近邻排序法 B. 冒泡排序法 C. 希尔排序法 D. 快速排序法

二、填空题（本大题共■小题，每小题■分，共■分）

1.大数据的储存和管理可以分为_____、_____和_____。

三、名词解释题（本大题共■小题，每小题■分，共■分）

1.决策树

四、简答题（本大题共■小题，每小题■分，共■分）

1.什么是数据？