

高纲 4268

江苏省高等教育自学考试大纲

# 13011 人工智能与大数据

南京理工大学编（2024 年）

# I 课程性质与课程目标

## 一、课程性质和特点

《人工智能与大数据》课程是江苏省高等教育自学考试信息管理与信息系统专业（专升本）的必修课，该课程是信息管理与信息系统专业课程体系中的基础课程之一。

本课程是一门非常实用的课程，既有理论基础又有实际应用价值。课程侧重于大数据和人工智能的理论和方法体系，同时兼顾利用机器学习和深度学习等方法解决实际应用能力的掌握。因此考生在掌握相关理论与方法的同时，一定要加强通过 Python 编程解决实际应用问题能力的培养，考生应当考察实际各类大数据和人工智能应用，注重解决实际问题，以便考生能把课程中所学的方法、原理应用到实际的大数据与人工智能中。

## 二、课程目标

《人工智能与大数据》课程是一门集知识和技能于一体，理论性和实践性都很强的课程，课程设置的目标是鼓励考生：

1. 掌握大数据理论和方法体系，包括大数据概述，大数据架构，大数据采集和预处理，大数据存储，大数据分析与挖掘，大数据可视化，以及大数据应用，为考生从事大数据挖掘和应用奠定基础，提高考生运用所学知识技能进行大数据分析处理的能力。

2. 掌握人工智能理论和方法体系，包括人工智能概论，机器学习，深度学习，以及人工智能应用，为考生使用机器学习和深度学习方法解决人工智能应用奠定基础。

3. 掌握 Python 基础操作，包括 Python 概述，Python 语法基础，列表与字典，Python 控制结构，文件与数据库，以及 Python 常用类库，具备 Python 的安装和基础操作能力。

## 三、课程的重点和难点

本课程的重点在于大数据的处理流程知识的掌握，次重点在于机器学习的几种典型算法的掌握，难点在于大数据分析挖掘算法和机器学习算法的掌握。

## II 考核目标

本大纲在考核目标中，按照识记、领会、简单应用和综合应用四个层次规定其应达到的能力层次要求。四个能力层次是递升的关系，后者必须建立在前者的基础上。各能力层次的含义是：

**识记：**要求考生能够识别和记忆本课程中有关大数据与人工智能的基本理论和基础知识（如定义、分类、技术、方法、技巧、模式、步骤、效果评价、特点等），并能够根据考核的不同要求，做正确的表述、选择和判断。

**领会：**要求考生能够领悟和理解本课程中有关大数据与人工智能基础知识的内涵及外延，理解大数据架构、大数据采集和预处理、大数据存储、大数据分析挖掘、大数据可视化、机器学习、深度学习；理解相关知识的区别和联系，并能根据考核的不同要求对大数据和人工智能问题进行分析，做出正确的判断、解释和说明。

**简单应用：**要求考生能够根据已知的知识和大数据与人工智能应用的需求，完成网络数据爬取、非结构化数据可视化、机器学习算法和 Python 控制结构的实现。

**综合应用：**要求考生能够面对具体、实际的人工智能应用问题，能探究解决问题的方法，完成分类算法的实现、使用 Python 实现控制结构经典算法、数据库应用，并使用 Keras 库实现相关应用。

## III 课程内容与考核要求

### 第一章 绪论

#### 一、学习目的与要求

通过本章学习，考生应当了解大数据相关的基础知识，掌握大数据的定义、产生、来源与作用，重点掌握大数据的特征。

#### 二、考核知识点与考核要求

##### （一）大数据的定义、产生及来源

**领会：**①大数据的定义；②大数据的产生；③大数据的来源。

(二) 大数据的特征

识记：①大数据的“4V”特征；②结构化数据。

领会：①半结构化数据；②非结构化数据。

(三) 大数据的作用

领会：①对政府治理的作用；②对经济治理的作用；③对公共服务的作用。

**三、本章的重点和难点**

重点：①大数据的特征；②结构化数据；③半结构化数据；④非结构化数据。

难点：①半结构化数据；②非结构化数据。

**第二章 大数据架构**

**一、学习目的与要求**

通过本章学习，考生应当掌握云计算的相关基础知识、大数据架构、Hadoop 体系架构，理解 MapReduce 计算的原理和过程。

**二、考核知识点与考核要求**

(一) 云计算的相关基础知识

识记：①云计算两大核心功能；②云计算的典型特征；③云计算的三种模式。

领会：①云计算的服务方式；②云计算数据中心；③云计算典型应用。

(二) 大数据架构及关键技术

识记：①大数据架构。

领会：①相关关键技术。

(三) Hadoop 体系架构

识记：①Hadoop 的来源、特点、架构、核心组件；②MapReduce 模型简介。

领会：①MapReduce 体系结构。

综合应用：①MapReduce 实例。

**三、本章的重点和难点**

重点：①云计算的三种服务方式、Hadoop 体系架构；②MapReduce 计算的

原理与过程。

难点：①MapReduce 计算的原理与过程。

### 第三章 大数据采集及预处理

#### 一、学习目的与要求

通过本章学习，考生应当了解大数据采集的概念、大数据采集方法，掌握大数据预处理和常用 ETL 工具。

#### 二、考核知识点与考核要求

##### （一）大数据采集

识记：①大数据采集的概念；②大数据采集设备。

领会：①大数据的采集方法。

##### （二）网络爬虫

识记：①网络爬虫的基本流程；

领会：①网络爬虫所需技术。

简单应用：①爬取网页标题；

综合应用：①使用 Scrapy 爬虫框架。

##### （三）数据预处理

识记：①数据预处理的原因。

领会：①数据预处理技术。

##### （四）常用 ETL 工具

识记：①ETL 技术。

领会：①常用 ETL 工具。

#### 三、本章的重点

重点：①使用 urllib 或 Requests 扩展库进行网络爬虫；②使用 Scrapy 框架进行网络爬虫。

难点：①使用 Scrapy 框架进行网络爬虫。

### 第四章 大数据存储

#### 一、学习目的与要求

通过本章学习，考生应当了解大数据存储面临的挑战、数据存储设备；掌

握传统数据库、NoSQL 数据库、NewSQL 数据库及数据仓库。

## 二、考核知识点与考核要求

### （一）大数据存储概述

识记：①大数据存储面临的挑战。

### （二）大数据的存储设备

识记：①数据存储设备；②数据存储模式。

### （三）传统数据库

识记：①传统数据库的概念；②传统数据库的分类。

领会：①主流的关系数据库。

### （四）大数据数据库

识记：①传统数据库的局限性；②NoSQL 数据库的概念和分类；③NewSQL 数据库的概念和特点。

### （五）数据仓库

识记：①数据仓库概念；②数据仓库的特征；③数据仓库的体系结构。

领会：①数据仓库的应用；②数据仓库工具 Hive。

简单应用：①数据库和数据仓库的区别。

## 三、本章的重点和难点

重点：①只能存储结构化数据的传统数据库；②可以存储半/非结构化的大数据的 NoSQL 数据库和 NewSQL 数据库；③数据库与数据仓库的区别。

难点：①NoSQL 数据库类型；②NewSQL 数据库特点。

## 第五章 大数据分析挖掘

### 一、学习目的与要求

通过本章学习，考生应当理解大数据分析的概念、目的和意义，掌握大数据分析的发展方向、研究方向和分类，掌握大数据分析的方法和挖掘的算法，理解数据分析与数据挖掘的关系。

### 二、考核知识点与考核要求

#### （一）大数据分析基础知识

识记：①大数据分析的概念；②大数据分析的目的；③大数据分析的意义。

领会：①大数据分析的发展方向；②大数据分析的研究方向；③大数据分析的分类。

#### （二）大数据分析的步骤与方法

识记：①大数据分析的步骤。

领会：①大数据分析的方法。

#### （三）大数据挖掘

识记：①数据挖掘的含义；②数据挖掘的详细步骤。

简单应用：①数据分析和数据挖掘的关系；②数据挖掘的主要算法。

### 三、本章的重点和难点

重点：①大数据分析与挖掘的方法和算法；②数据分析与数据挖掘的关系。

难点：①大数据分析与挖掘的算法；②数据分析与数据挖掘的关系。

## 第六章 大数据可视化

### 一、学习目的与要求

通过本章学习，考生应当理解数据可视化的基本特征、作用和流程，掌握大数据可视化的方法、大数据可视化的常用软件及工具。

### 二、考核知识点与考核要求

#### （一）数据可视化

识记：①数据可视化的基本特征；②数据可视化的作用；

领会：③数据可视化流程。

#### （二）大数据可视化的方法

识记：①科学可视化；②信息可视化。

简单应用：①用 Python 编程实现非结构化数据可视化。

#### （三）大数据可视化工具

识记：①入门级工具；②信息图表工具；③地图工具；④时间线工具。

领会：①简单工具的使用方法；②多维数据可视化方法。

综合应用：①高级可视化工具使用。

### 三、本章的重点和难点

重点：①文本可视化；②多维数据可视化方法；③高级可视化工具。

难点：①多维数据可视化方法；②高级可视化工具。

## 第七章 大数据应用

### 一、学习目的与要求

通过本章学习，考生应当了解国外大数据应用的经典案例和国内大数据的应用情况。

### 二、考核知识点与考核要求

#### （一）国外大数据应用

识记：①大数据的新价值。

#### （二）国内大数据应用

领会：①智慧城市应用；②保险大数据应用；③农业大数据应用。

## 第八章 人工智能概论

### 一、学习目的与要求

通过本章学习，考生应当理解人工智能的概念、分类与学派，了解人工智能的发展，掌握人工智能研究的基本内容、人工智能与大数据的关系，了解人工智能引发的机遇与挑战。

### 二、考核知识点与考核要求

#### （一）人工智能的概念

识记：①智能的定义；②智能的特征。

领会：①人工智能的定义；②图灵测试和中文屋子的概念。

#### （二）人工智能的分类和学派

领会：①弱人工智能；②强人工智能；③超人工智能；④符号主义；⑤连接主义；⑥行为主义。

简单应用：①人工智能的分类。

#### （三）人工智能的发展

领会：①人工智能的孕育和诞生；②人工智能的第一次热潮与低谷；③人工智能的第二次热潮与低谷；④人工智能的第三次热潮。

#### （四）人工智能研究的基本内容

识记：①认知建模的含义；②知识表示的含义；③知识推理的含义；④知

识应用的含义。

领会：①机器感知的含义；②机器思维的含义；③机器学习的含义；④机器行为的含义；⑤构建智能系统的含义。

简单应用：①人工智能的研究的内容

（五）人工智能与大数据的关系

领会：①大数据与人工智能互相促进发展；

简单应用：①大数据与人工智能的融合。

（六）人工智能引发的思考

识记：①奇点理论；②机器人定律

领会：①人工智能引发的挑战；②人工智能相关的伦理规范。

### 三、本章重点、难点

重点：①人工智能的分类与学派；②人工智能的基本研究内容。

难点：①人工智能研究的基本内容；②人工智能与大数据的关系。

## 第九章 机器学习

### 一、学习目的与要求

通过本章学习，考生应当理解机器学习的概念和应用范围，掌握机器学习的基本流程，掌握包括分类算法、聚类算法、关联规则分析算法及回归分析算法的几种典型机器学习算法。

### 二、考核知识点与考核要求

（一）机器学习的概念及应用范围

识记：①机器学习概念；②机器学习的应用范围。

领会：①机器学习与人类思考的类比。

（二）机器学习的基本流程

识记：①明确目标任务；②数据收集；③数据预处理与特征工程；④选择、训练和测试模型。

领会：①评估、优化模型；②应用模型。

（三）机器学习算法

识记：①监督算法的定义和性质；②无监督算法的定义和性质；③强化学习的定义；

领会：①典型的分类算法；②典型的聚类算法；③典型的关联规则分析算法；④典型的回归分析算法。

#### （四）决策树分类

识记：①决策树分类的概念。

综合应用：①用 ID3 算法从数据集中生成决策树。

#### （五）逻辑回归分类

识记：①逻辑回归的概念；②逻辑函数。

领会：①逻辑回归的一般流程。

综合应用：①使用逻辑回归分类算法完成分类。

#### （六）支持向量机

识记：①支持向量机的概念。

综合应用：①使用支持向量机进行数字模式识别。

#### （七）聚类

识记：①聚类的应用；②K-均值算法的原理。

简单应用：①使用 K-均值算法完成聚类。

#### （八）关联规则

识记：①关联规则分析的主要概念。

领会：①关联规则案例；②Apriori 算法的原理。

简单应用：①使用 Apriori 算法完成关联规则分析。

#### （九）神经网络

识记：①人工神经网络的定义；②神经网络模型。

领会：①前馈型神经网络；②反馈型神经网络；③神经网络工作原理。

简单应用：①使用神经网络算法识别手写数字。

### 三、本章的重点和难点

重点：①机器学习的基本流程；②决策树分类；③关联规则；④神经网络。

难点：①决策树分类；②神经网络。

## 第十章 深度学习

### 一、学习目的与要求

通过本章学习，考生应当理解深度学习的基本概念，掌握卷积神经网络、循环神经网络、生成对抗网络、强化学习、迁移学习和对偶学习等算法。

### 二、考核知识点与考核要求

#### （一）深度学习概述

识记：①深度学习的基本定义。

领会：①深度学习与机器学习的区别。

#### （二）卷积神经网络

识记：①卷积神经网络基本概念。

领会：①卷积神经网络模型结构；②卷积神经网络特点。

#### （三）循环神经网络

识记：①循环神经网络基本概念；②循环神经网络应用领域。

领会：①循环神经网络模型结构。

#### （四）生成对抗网络

识记：①生成对抗网络基本概念。

领会：①生成对抗网络框架模型；②生成对抗网络工作原理；③生成对抗网络典型事例。

#### （五）强化学习

识记：①强化学习基本定义。

领会：①强化学习特点。

#### （六）迁移学习

识记：①迁移学习基本定义。

领会：①迁移学习特点。

#### （七）对偶学习

领会：①对偶学习原理。

### 三、本章的重点和难点

重点：①卷积神经网络；②生成对抗网络；③强化学习。

难点：①卷积神经网络；②强化学习。

## 第十一章 人工智能应用

### 一、学习目的与要求

通过本章学习，考生应当了解人工智能的一些主要应用场景，包括图像识别与分类、语音识别、人脸识别和情感计算、自动驾驶、智能家居、专家系统、机器人、自然语言处理和其他 AI 应用。

### 二、考核知识点与考核要求

#### （一）图像识别与分类

识记：①图像识别流程；②图像识别技术发展。

领会：①图像识别技术的主要应用。

#### （二）语音识别

识记：①语音识别的步骤。

领会：①语音识别技术的应用。

#### （三）人脸识别和情感计算

识记：①人脸识别的定义。

领会：①人脸识别的一般流程；②情感计算。

#### （四）自动驾驶

识记：①人类与 AI 驾驶对比；②自动驾驶的 SAE 分类标准；③自动驾驶模块组成；④障碍物识别；⑤交通标志识别；⑥车道识别。

领会：①自动驾驶现状；②自动驾驶系统；③自动驾驶未来趋势。

#### （五）智能家居

识记：①智能家居的概念；②智能家居分类；③智能家居主要功能。

#### （六）专家系统

识记：①专家系统的基本概念；②专家系统组成；③专家系统的应用领域。

#### （七）机器人

识记：①机器人的定义；②机器人的发展；③机器人分类；④机器人的应用。

#### （八）自然语言处理

识记：①自然语言处理基本概念；②自然语言处理的发展；③自然语言处

理应用。

#### （九）其他 AI 应用

识记：①博弈上的应用；②机器视觉上的应用；③智慧医疗上的应用；④智能个性化推荐上的应用。

### 三、本章的重点和难点

重点：①图像识别与分类；②人脸识别和情感计算；③自动驾驶；④自然语言处理。

难点：①人脸识别和情感计算；②自动驾驶。

## 第十二章 Python 概述

### 一、学习目的与要求

通过本章学习，考生应当了解 Python 的发展简史，理解 Python 的特点，掌握 Python 的应用领域，学会开发环境安装配置和类库的导入。

### 二、考核知识点与考核要求

#### （一）Python 的发展简史

识记：①Python 的来源；②Python 各版本发行时间；Python 的广泛运用。

#### （二）Python 的特点

领会：①Python 简单易学、语法简单、开发速度快；②Python 跨平台、可移植、可扩展、交互式、解释型、面向对象、动态语言；③内置大量标准库和第三方库；④社区活跃、贡献者多、互帮互助；⑤开源语言、发展动力巨大；⑥发布的源代码不加密。

#### （三）Python 的应用领域

领会：①应用于人工智能；②应用于网络爬虫；③应用于数据分析；④应用于云计算；⑤应用于游戏开发；⑥应用于桌面软件及数据库程序开发；⑦应用于 Web 前后端开发；⑧应用于网络编程；⑨应用于自动化运维；⑩应用于文档处理。

#### （四）Python 开发环境安装配置

识记：①Anaconda 3；②Jupyter Notebook；③Spyder；④Anaconda Prompt；⑤pip 命令。

### （五）Python 类库的导入

领会：①导入方式 1: `import 模块名 [as 别名]`；②导入方式 2: `from 模块名 import 对象名 [as 别名]`；③导入方式 3: `from 模块名 import *`。

### 三、本章的重点

重点：①Python 开发环境安装配置；②pip 命令；③Python 标准库和扩展库的导入和使用。

难点：①Python 标准库和扩展库的导入和使用。

## 第十三章 Python 语法基础

### 一、学习目的与要求

通过本章学习，考生应当掌握 Python 的标识符、常量、变量、数据类型、运算符、内置函数和用户自定义函数等用法。

### 二、考核知识点与考核要求

#### （一）标识符、常量、变量

识记：①标识符的命名规则；②标识符的命名原则；③关键字的含义。

领会：①查看关键字的命令；②常量、变量的含义。

#### （二）数据类型、运算符

识记：①常用的数据类型；②字符串的定义。

领会：①数字的表示方法；②转义字符；③字符串的运算；④字符串的操作；⑤算数运算符；⑥赋值运算符；⑦关系（比较）运算符；⑧逻辑运算符；⑨成员测试运算符；⑩运算符的优先级。

#### （三）内置函数

领会：①`print()` 函数；②`input()` 函数；③`abs()` 函数；④`pow()` 函数；⑤`int()` 函数；⑥`float()` 函数；⑦`ord()`、`chr()` 函数；⑧`max()`、`min()` 函数；⑨`len()` 函数；⑩`type()` 函数。

简单应用：①内置函数。

#### （四）用户自定义函数

领会：①Python 定义函数；②无参函数；③有参函数。

综合应用：①用户自定义函数。

### 三、本章的重点和难点

重点：①数据类型；②字符串的操作；③运算符；④用户自定义函数。

难点：①字符串的操作；②用户自定义函数。

## 第十四章 列表与字典

### 一、学习目的与要求

通过本章学习，考生应当掌握 Python 中列表、字典、元组、集合等复合数据类型的特点、方法、函数，学会运用各种运算符对列表、字典、元组、集合中的数据进行操作，掌握复合数据类型的切片和索引。

### 二、考核知识点与考核要求

#### （一）列表

识记：①定义列表。

领会：①列表切片；②列表的其他操作；③列表相关函数；④列表相关方法。

#### （二）字典

识记：①定义字典。

领会：①访问字典中的数据；②修改字典；③删除字典元素。

#### （三）元组

识记：①元组的定义。

领会：①创建元组；②访问元组；③修改元组；④删除元组；⑤元组相关函数。

#### （三）集合

识记：①创建集合。

领会：①添加元素；②删除元素；③清空集合；④集合长度；⑤判断集合中是否存在元素；⑥集合相关方法。

### 三、本章的重点和难点

重点：①列表、字典、元组、集合等方法和函数；②运用各种运算符对列表、字典、元组、集合中的数据进行操作；③复合数据类型的切片与索引。

难点：①列表、字典、元组、集合等方法和函数；②复合数据类型的切片与索引。

## 第十五章 Python 控制结构

### 一、学习目的与要求

通过本章学习，考生应当掌握 Python 中顺序结构、选择结构和循环结构的执行流程。

### 二、考核知识点与考核要求

#### （一）顺序结构

识记：①顺序结构的流程。

#### （二）选择结构

识记：①选择结构的流程

领会：①单分支 if 语句的形式；②双分支 if 语句的形式；③多分支 if 语句的形式；④if 语句的常用嵌套形式。

简单应用：①使用单分支 if 语句、双分支 if 语句、多分支 if 语句和嵌套 if 语句实现选择结构。

#### （三）循环结构

识记：①while 循环结构的流程；②循环嵌套的概念。

领会：①while 循环结构的语法结构；②for 循环结构的语法格式；③range()函数；④break 语句的用法；⑤continue 语句的用法。

简单应用：①使用 while 语句、for 语句、break 语句和 continue 语句实现循环结构和循环嵌套结构。

#### （四）经典算法

综合应用：①掌握经典算法实现典型的应用。

### 三、本章的重点和难点

重点：①if 选择结构；②for 循环和 while 循环；③选择结构和循环结构的多层嵌套及执行过程。

难点：①for 循环和 while 循环；②选择结构和循环结构的多层嵌套及执行过程。

## 第十六章 文件与数据库

### 一、学习目的与要求

通过本章学习，考生应当掌握用 Python 编程访问数据文件和数据库的方法。

## 二、考核知识点与考核要求

### （一）文件的概念及分类

识记：①文件的基本概念；②ASCII 码文件；③二进制文件。

领会：①ASCII 码文件和二进制文件的优缺点。

### （二）文件的打开与关闭

识记：①open() 函数的格式；②close() 函数的格式。

领会：①open() 函数的用法。

### （三）文件的读/写

领会：①flush() 函数；②next() 函数；③read([size]) 函数；④readline([size])() 函数；⑤readlines([size])() 函数；⑥seek(offset[,whence]) 函数；⑦tell() 函数；⑧write(str)；⑨writelines(sequence)函数；⑩openpyxl 扩展库的用法。

简单应用：①掌握使用相关函数和扩展库实现文件的读/写操作。

### （四）数据库访问

综合应用：①使用 Python 访问 Access 数据库；②使用 Python 访问 MySQL 数据库；③使用 Python 访问 MongoDB 数据库。

## 三、本章的重点和难点

重点：①各种读/写文件的方法；②openpyxl 扩展库的用法；③数据库访问。

难点：①MySQL 数据库访问；②MongoDB 数据库访问。

## 第十七章 Python 常用类库

### 一、学习目的与要求

通过本章学习，考生应当学会安装 NumPy 库、Matplotlib 库、Pandas 库、Scikit-learn 库、Keras 库等一些常用的第三方扩展库。

## 二、考核知识点与考核要求

### （一）NumPy 库

识记：①NumPy 库的基本概念。

领会：①NumPy 库的常用方法。

简单应用：①实现 NumPy 库的主要功能。

#### （二）Matplotlib 库

识记：①Matplotlib 库的基本概念。

领会：①Matplotlib 库的常用方法。

简单应用：①实现 Matplotlib 库的主要功能。

#### （三）Pandas 库

识记：①Pandas 库的基本概念。

领会：①Pandas 库的常用方法。

简单应用：①实现 Pandas 库的主要功能。

#### （四）Scikit-learn 库

识记：①Scikit-learn 库的基本概念。

领会：①Scikit-learn 库的安装流程。

#### （五）Keras 库

识记：①Keras 库的基本概念。

领会：①Keras 库的常用方法。

综合应用：①使用 Keras 库实现相关应用。

### 三、本章的重点和难点

重点：①NumPy 库；②Matplotlib 库；③Scikit-learn 库。

难点：①Matplotlib 库；②Pandas 库；③Keras 库。

## IV 关于大纲的说明与考核实施要求

### 一、自学考试大纲的目的和作用

课程自学考试大纲是根据专业考试计划的要求，结合自学考试的特点而确定。其目的是对个人自学、社会助学和课程考试命题进行指导和规定。

课程自学考试大纲明确了课程学习的内容以及深广度，规定了课程自学考试的范围和标准。因此，它是编写自学考试教材和辅导书的依据，是社会助学组织进行自学辅导的依据，是考生学习教材、掌握课程内容知识范围和程度的依据，也是进行自学考试命题的依据。

### 二、课程自学考试大纲与教材的关系

课程自学考试大纲是进行学习和考核的依据，教材是学习掌握课程知识的基本内容与范围，教材的内容是大纲所规定的课程知识和内容的扩展与发挥。大纲与教材所体现的课程内容应基本一致；大纲里面的课程内容和考核知识点，教材里一般也要有。反过来教材里有的内容，大纲里就不一定体现。

### 三、关于自学教材

本课程使用教程为：《大数据与人工智能》，杨忠宝、佘向飞主编，北京大学出版社，2022年。

### 四、关于自学要求和自学方法的指导

本大纲的课程基本要求是依据专业考试计划和专业培养目标而确定的。课程基本要求明确了课程的基本内容，以及对基本内容掌握的程度。基本要求中的知识点构成了课程内容的主体部分。因此，课程基本内容掌握程度、课程考核知识点是高等教育自学考试考核的主要内容。

为有效地指导个人自学和社会助学，本大纲已指明了课程的重点和难点，在各章的基本要求中也指明了各章内容的重点和难点。

本课程是一门理论性和实践性都很强的课，重点在于对大数据和人工智能理论知识的掌握，以及应用大数据挖掘算法和机器学习算法解决实际的大数据应用和人工智能应用问题。自学考试主要是通过个人自学、教师辅导、社会助学和国家考试来考核考生掌握专业知识和能力的方法。考生应根据自己的特点，找出适合自己的学习方法，此外，考生在自学过程中，应注意以下几点：

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. 本课程内容涉及大数据与人工智能的各个方面，知识、范围比较广泛，全书是一个整体，但各章之间又有相对独立性。考生应首先全面系统地学习各章的内容，深刻领会大数据与人工智能中的理论知识；其次，要注意各章之间的联系；然后，在全面系统的基础上掌握重点，有目的地深入学习重点章节，但切忌在没有了解全貌的情况下孤立地去抓重点，押题目。

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。

5. 注意将大数据与人工智能理论与实践应用相结合。考生应明白，大数据与人工智能这门课是一门理论性和实践性都很强的课程，在学习过程中切忌死记硬背，而应当尽可能多上网，观察、分析和研究实际大数据与人工智能应用案例，将课程中所讲述的基本理论知识（如大数据架构、大数据采集与预处理、大数据分析挖掘、大数据可视化、机器学习、深度学习等），与一定的实际相联系，并通过实际操作（如实现经典的数据挖掘算法和机器学习算法）来更好地理解和掌握相关的理论知识。

## **五、应考指导**

### **1. 如何学习**

很好的计划和组织是你学习成功的法宝。如果你正在接受培训学习，一定要跟紧课程并完成作业。为了在考试中作出满意的回答，你必须对所学课程内容有很好的理解。使用“行动计划表”来监控你的学习进展。你阅读课本时可以做读书笔记。如有需要重点注意的内容，可以用彩笔来标注。如：红色代表重点；绿色代表需要深入研究的领域；黄色代表可以运用在工作之中。可以在空白处记录相关网站，文章。

### **2. 如何考试**

卷面整洁非常重要。书写工整，段落与间距合理，卷面赏心悦目有助于教师评分，教师只能为他能看懂的内容打分。回答所提出的问题。要回答所问的问题，而不是回答你自己乐意回答的问题！避免超过问题的范围。

### **3. 如何处理紧张情绪**

正确处理对失败的惧怕，要正面思考。如果可能，请教已经通过该科目考试的人，问他们一些问题。做深呼吸放松，这有助于使头脑清醒，缓解紧张情

绪。考试前合理膳食，保持旺盛精力，保持冷静。

#### 4. 如何克服心理障碍

这是一个普遍问题！如果你在考试中出现这种情况，试试下列方法：使用“线索”纸条。进入考场之前，将记忆“线索”记在纸条上，但你不能将纸条带进考场，因此当你阅读考卷时，一旦有了思路就快速记下。按自己的步调进行答卷。为每个考题或部分分配合理时间，并按此时间安排进行。

#### 六、对社会助学的要求

1. 社会助学者应根据本大纲规定的考试内容和考核目标，认真钻研指定教材，明确本课程与其它课程不同的特点和学习要求，对考生进行切实有效的辅导，引导他们防止自学中的各种偏向，把握社会助学的正确方向。

2. 要正确处理重点和一般的关系。课程内容有重点与一般之分，但考试内容是全面的，而且重点与一般是相互影响的，不是截然分开的，社会助学者应指导考生全面系统地学习教材，掌握全部考核内容和考核知识点，并在此基础上突出重点。总之，要把重点学习同兼顾一般结合起来，切勿孤立地抓重点，把考生引向猜题押题。

#### 七、对考核内容的说明

1. 本课程要求考生学习和掌握的知识点内容都作为考核的内容。课程中各章的内容均由若干知识点组成，在自学考试成为考核知识点。因此，课程自学考试大纲中所规定的考试内容是以分解为考核知识点的方式给出的。由于各知识点在课程中的地位、作用以及知识自身的特点不同，自学考试将对各知识点分别按四个能力层次确定其考核要求。

2. 本大纲在考核目标中，按照识记、领会、简单应用和综合应用四个层次要求考生掌握，四个能力层次是递进关系。

3. 课程分为十七部分，分别是绪论、大数据架构、大数据采集与预处理、大数据存储、大数据分析挖掘、大数据可视化、大数据应用、人工智能概论、机器学习、深度学习、人工智能应用、Python 概述、Python 语法基础、列表与字典、Python 控制结构、文件与数据库、Python 常用类库。

#### 八、关于考试命题的若干规定

本课程的命题考试，应根据本大纲规定的考试内容和考核目标来确定考试

范围和考核要求，按大纲规定试题中主观性题和客观性题的比例来组配试卷，适当掌握试题的内容覆盖面、能力层次和难易度。

1. 本大纲各章所规定的基本要求、知识点及知识点下的知识细目，都属于考核的内容。考试命题既要覆盖到章，又要避免面面俱到。要注意突出课程的重点、章节重点，加大重点内容的覆盖度。

2. 命题不应有超出大纲中考核知识点范围的题目，考核目标不得高于大纲中所规定的相应的最高能力层次要求。命题应着重考核考生对基本概念、基本知识和基本理论是否了解或掌握，对基本方法是否会用或熟练。不应出与基本要求不符的偏题或怪题。

3. 本课程在试卷中对不同能力层次要求的分数比例大致为：识记占 30%，领会占 30%，简单应用占 20%，综合应用占 20%。

4. 试题要合理安排难度结构。试题难易可分为易、较易、较难、难四个等级。试题的难易度与能力层次不是一个概念，在各能力层次上都会存在不同难度的问题，考生切勿混淆。

5. 试卷的题型有：单项选择题、多项选择题、填空题、简答题、综合题。在命题工作中必须按照本课程大纲中所规定的题型命制，考试试卷使用的题型可以略少，但不能超出本课程对题型规定。

6. 本课程考试方法采用闭卷、笔试，考试时间为 150 分钟，评分采用百分制，60 分为及格。考生只准携带 0.5 毫米黑色墨水的签字笔、铅笔、圆规、直尺、三角板、橡皮等必需的文具用品。不可携带计算器。

## 附录 题型举例

### 一、单项选择题

1. 不依赖外界交互、自动地利用未标记样本来提升学习性能的学习形式是（ ）

A. 有监督学习      B. 强化学习      C. 无监督学习      D. 半监督学习

参考答案： C

2. 下面不属于云计算技术的是（ ）

A. Hadoop      B. Spark      C. YARN      D. 集中式计算

参考答案： D

### 二、多项选择题

1. 对人工智能常见的误解有 ( )

- A. 机器学习只是人工智能中的一个方向      B. 人工智能就是机器学习  
C. 人工智能就是深度学习      D. 人工智能最近十年受到深度学习的驱动较多  
E. 人工智能可以解决一切问题

参考答案: BCE

2. 下面不属于相关分析的是 ( )

- A. 正相关    B. 负相关    C. 线性相关    D. 误差相关    E. 均值相关

参考答案: DE

### 三、填空题

1. 在支持向量机中, \_\_\_\_\_ 允许一些超平面附近的样本被错误分类, 从而提升泛化性能。

参考答案: 软间隔

### 四、简答题

1. 简述人工智能的三大学派。

参考答案: ①符号主义学派; ②连接主义学派; ③行为主义学派。

### 五、综合题

1. 请使用 while 语句完成 1~100 的偶数和。

参考答案:

```
i=2
sum=0
while i<=100:
    sum=sum+i
    i=i+2
print("2+4+6+...+100=", sum)
```