**人工智能训练师（三级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：数据清洗培训大纲编写

考核时间：10min

**1.**场地设备要求

（1）人工智能训练师主机；   
（2）Office软件。

**2.**工作任务

A企业是一家康复训练机构，需要利用人工智能技术为康复患者提供个性化的康复训练计划，并实时监控训练效果。为了提高人工智能技术的应用效果，其采集了大量的康复数据，现计划对新进技术人员进行数据清洗工具的使用培训。通过这次培训，将会使新进技术人员掌握4中基于Python的数据清洗工具的使用，能对大量的康复数据进行清洗，达到人工智能训练师四级/中级工的技能水平。  
请你根据要求补全素材4.1.3.docx中的培训大纲。

**3.**技能要求

（1）能编写初级培训讲义；   
（2）能对初级工、中级工开展知识和技术培训。

**4.**质量指标

（1）培训讲义编写逻辑合理；   
（2）培训资料正确有效；   
（3）语句组织通顺。

**人工智能训练师（三级）操作技能考核**

**试题评分表**

准考证号：

试题代码：

试题名称：数据清洗培训大纲编写

考核时间：10min

测量分评分表

| 细则编号 | 配分 | 评分细则描述 | 规定或  标称值 | 结果或  实际值 | 得分 |
| --- | --- | --- | --- | --- | --- |
| M1 | 5 | 回答内容与答案类似得1分；最多5分； | 根据数据 |  |  |
| 合计配分 | 5 | 合计得分 | |  |  |

**参考答案**（尽量将细则内容写在上面的表格内，写不下可另写，但要具体可评判）

**在下划线上填写答案。**

# 根据学习大纲补充学习目标

**引言**

**学习目标：** 理解人工智能在康复训练中的价值，掌握数据清洗对提升康复数据质量的核心作用。

内容：

内容：

介绍人工智能在康复训练中的应用，数据清洗在处理康复数据中的作用。

**数据清洗基础理论**

**学习目标：**  **掌握数据清洗的基本概念与核心任务，理解其在数据分析和模型训练中的关键作用**

内容：

什么是数据清洗？

数据清洗的常见任务：数据去重、缺失值处理、数据格式转换等。

数据清洗在数据分析和模型训练中的重要性。

**常用数据清洗工具简介**

**学习目标： 熟悉主流数据清洗工具的核心功能与适用场景，能够根据需求选择合适的工具**。

内容：

Pandas：强大的数据处理与分析工具。

NumPy：高性能科学计算和数据处理库。

OpenRefine：用于数据清洗的开源工具。

Dask：用于处理大规模数据的并行计算库。

**环境搭建与工具安装**

**学习目标： 掌握数据清洗工具的安装和环境配置**。

内容：

安装Python和pip包管理工具。

安装并配置Pandas、NumPy、OpenRefine、Dask。

**Pandas实战**

**学习目标：**  **能够使用Pandas完成数据导入、筛选、转换等核心清洗任务**。

内容：

数据导入与导出：读取和保存CSV、Excel等格式的数据。

数据筛选与过滤：条件筛选、去重、缺失值处理。

数据转换：数据类型转换、时间序列处理。

实践操作：使用Pandas清洗一个康复数据集。

**NumPy实战**

学习目标：综合运用所学知识进行数据清洗项目。

内容：

项目介绍：清洗一个多来源、多格式的康复数据集。

数据采集：导入多种格式的数据。

数据清洗：使用Pandas、NumPy、OpenRefine、Dask进行清洗。

项目评审：展示清洗成果，讲解实现思路和遇到的问题。