**第4部分**

操作技能复习题

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：自动驾驶系统的道路交通标识图数据质量检测

考核时间：20min

**1.**场地设备要求

（1）人工智能训练师主机；   
（2）Office软件；   
（3）1.1.1数据集。

**2.**工作任务

A公司开发的自动驾驶系统需要准确地识别和理解道路标识，以保证车辆行驶的安全性，减少违章行为。你作为自动驾驶技术项目组人工智能训练师，负责对预处理后的道路交通标识图数据集（1.1.1数据集）质量进行审核与评估，并梳理数据采集与处理规范。请完成以下任务：  
(1)请审核提供的图像数据集，找出两组重复或冗余的图片，将重复图片及编号截图记录到答题卷中。  
(2)请审核提供的图像数据集，找出两张道路标识图像模糊或者有严重遮挡等问题额图片，并将图片及编号截图记录到答题卷中。  
(3)撰写该系统数据集的数据质量评估报告到答题卷中，总结图像格式与分辨率、清晰度、数据重复、道路标识类别覆盖情况方面的发现。  
(4)请根据数据集的应用需求，提出采集范围和图片质量规范，写到答题卷中。  
(5)请根据数据集的应用需求，提出数据整理和分类规范各2条，写到答题卷中。  
请将答案写到答题卷文件中，答题卷文件命名为“1.1.1.docx”，保存到桌面考生文件夹，考生文件夹命名为“准考证号+身份证后6位”。

**3.**技能要求

（1）能对预处理后的业务数据进行审核；   
（2）能结合人工智能技术要求梳理业务数据采集规范；   
（3）能结合人工智能技术要求梳理业务数据处理规范。

**4.**质量指标

（1）业务数据采集梳理符合规范要求；   
（2）业务数据处理梳理符合规范要。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：自动驾驶系统的道路交通标识图数据质量检测

考核时间：20min

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 | 4 | 查找到重复或冗余的图片并截图，找到1组得2分， 2组及以上得4分； | 根据数据 |  |  |
| M2 | 4 | 查找到道路标识图像模糊或者有严重遮挡的图片并截图 。 找到1张得2分，2张及以上得4分； | 根据数据 |  |  |
| 合计配分 | 8 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 4 | 图片尺寸格式评估与答案类似：得1分； 清晰度评估与答案类似：得1分； 数据重复评估与答案类似：得1分； 道路标识类别覆盖评估与答案类似：得1分； |  |  |  |  |
| J2 | 4 | 回答采集规范要求与答案类似的每1个内容点可得1分，最多4分； |  |  |  |  |
| J3 | 4 | 数据整理规范要求与答案类似的每1个内容点可得1分，最多2分； 数据归类规范要求与答案类似的每1个内容点可得1分，最多2分； |  |  |  |  |
| 合计配分 | 12 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：公共交通工具识别图片数据集审核

考核时间：20min

**1.**场地设备要求

（1）人工智能训练师主机；   
（2）Office软件；   
（3）1.1.2数据集。

**2.**工作任务

A公司正在开发一种基于人工智能的公共交通工具识别系统，旨在提升城市交通管理效率。你作为该项目组的人工智能训练师，负责审核和评估预处理后的公共交通工具图像数据集质量，确保识别的准确性，并尽量减少分类错误。  
(1) 请审核提供的公共交通工具图像数据集，找出两组重复或冗余的图片，将重复图片及编号组截图记录到答题卷文件中。  
(2) 请审核提供的公共交通工具图像数据集，找出两张遮挡或者包含多种公共交通工具的图片，并将图片及其编号截图记录到答题卷文件中。  
(3) 对该系统数据集进行全面评估，总结图片尺寸与格式、清晰度、数据重复以及公共交通工具类别分布等方面的发现，写到答题卷文件中。  
(4) 根据公共交通工具识别系统的应用需求，提出4条适合的图像数据采集规范要求，说明数据采集过程中应考虑的关键因素及规范，如环境条件、拍摄设备、图像标注等，写到答题卷文件中  
(5) 针对公共交通工具识别系统的需求，提出4条适合的图像数据处理规范要求。详细描述数据预处理、图像增强和标准化的步骤和标准，以确保训练模型的稳定性和准确性，写到答题卷文件中。  
请将答案写到答题卷文件中，答题卷文件命名为“1.1.2.docx”，保存到桌面考生文件夹，考生文件夹命名为“准考证号+身份证后6位”。

**3.**技能要求

(1)能够对预处理后业务数据进行审核；   
(2)能够结合人工智能技术要求，梳理业务数据采集规范；   
(3)能够结合人工智能技术要求，梳理业务数据处理规范。

**4.**质量指标

(1)数据质量审核标准正确；   
(2)数据采集流程符合常规规范和可操作性；   
(3)对数据处理流程的要求清晰。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：公共交通工具识别图片数据集审核

考核时间：20min

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 | 4 | 找出每1组重复或冗余的图片，得2分，最多4分； | 根据数据 |  |  |
| M2 | 4 | 找出每1组遮挡或者包含多种公共交通工具的图片，得2分，最多4分； | 根据数据 |  |  |
| 合计配分 | 8 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 4 | (1) 回答图片尺寸格式评估与答案类似：得1分； (2) 回答清晰度评估与答案类似：得1分； (3) 回答数据重复评估与答案类似：得1分； (4) 回答公共交通工具类别分布与答案类似：得1分； |  |  |  |  |
| J2 | 4 | 回答图像数据采集中的环境条件、拍摄设备、图像标注方面的规范与答案类似：每条得1分；最多得4分 |  |  |  |  |
| J3 | 4 | 回答数据处理规范中的数据预处理、图像增强、标准化方面的规范与答案类似： 1条得1分；最多得4分 |  |  |  |  |
| 合计配分 | 12 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：人脸表情识别数据集审核

考核时间：20min

**1.**场地设备要求

（1）人工智能训练师主机；   
（2）Office软件；   
（3）1.1.3数据集。

**2.**工作任务

在科技浪潮中，情感智能已成为人工智能领域的一个重要分支，它不仅改变了人机交互的方式，还深刻影响了诸如心理健康评估、客户服务、娱乐和教育等多个行业。A公司是一家专注于AI情感分析的前沿企业，正致力于开发一款革命性的人脸表情识别系统，旨在通过精准分析人类的微表情，实现更深层次的情感理解和响应。你作为该项目组的人工智能训练师，负责审核和评估预处理后的人脸表情数据集质量，确保识别的准确性，并尽量减少分类错误。  
(1) 请审核提供的人脸表情数据集，找出两组重复或冗余的图片，并将重复图片及编号截图记录到答题卷文件中。  
(2) 请审核提供的人脸表情数据集，找出共有多少个人的表情和多少种表情，将人数和表情数记录到答题卷文件中。  
(3) 对该系统数据集进行全面评估，总结图片尺寸与格式、清晰度、数据重复以及日常生活图像类别分布等方面的发现，写到答题卷文件中。  
(4) 根据人脸表情识别系统的应用需求，提出4条适合的图像数据采集规范要求，说明数据采集过程中应考虑的关键因素及规范，如环境条件（光照、背景）、拍摄设备（相机规格）、图像标注等，写到答题卷文件中。  
(5) 针对人脸表情识别系统的需求，提出适合的图像数据处理规范要求，详细描述数据预处理、图像增强和标准化的步骤和标准，写到答题卷文件中。  
请将答案写到答题卷文件中，答题卷文件命名为“1.1.3.docx”，保存到桌面考生文件夹，考生文件夹命名为“准考证号+身份证后6位”。

**3.**技能要求

(1)能够对预处理后业务数据进行审核；   
(2)能够结合人工智能技术要求，梳理业务数据采集规范；   
(3)能够结合人工智能技术要求，梳理业务数据处理规范；

**4.**质量指标

(1)数据质量审核标准正确；   
(2)数据采集流程符合常规规范和可操作性；   
(3)对数据处理流程的要求清晰

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：人脸表情识别数据集审核

考核时间：20min

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 | 4 | 每找出每1组重复或冗余的图片，得2分。最多4分； | 根据数据 |  |  |
| M2 | 4 | 找出正确人数，得2分； 找出正确表情数，得2分； | 根据数据 |  |  |
| 合计配分 | 8 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 4 | 回答图片尺寸格式评估与答案类似：得1分； 回答清晰度评估与答案类似：得1分； 回答数据重复评估与答案类似：得1分； 回答日常生活图像类别分布与答案类似：得1分； |  |  |  |  |
| J2 | 4 | 回答图像数据采集中的环境条件、拍摄设备、图像标注规范与答案类似：1条得1分；最多得4分； |  |  |  |  |
| J3 | 4 | 回答数据处理规范中的数据预处理、图像增强、标准化方面的规范与答案类似：1条得1分，最多得4分； |  |  |  |  |
| 合计配分 | 12 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：化妆品销售数据质量检测

考核时间：20min

**1.**场地设备要求

（1）人工智能训练师主机   
（2）Office软件   
（3）1.1.4数据集

**2.**工作任务

化妆品行业呈现出稳步增长的趋势，A企业为优化化妆品销售策略，实现精准营销和个性化推荐，同时为企业产品研发和创新提供方向，收集了化妆品的销售数据（1.1.4数据集），用以利用人工智能算法进行分析。你作为一名人工智能训练师，请对该销售数据集质量进行审核，梳理数据采集和数据处理规范。请完成以下任务：  
(1)请根据数据质量审核遵循的核心纬度，对数据集的数据质量进行审核，记录发现的两个具体问题及所参考的纬度到答题卷文件。  
(2)对该数据集进行评估，总结数据集准确性、完整性、一致性、重复性方面的情况，写到到答题卷中。  
(3)请根据数据集的应用需求，描述数据采集规范中关于用户隐私保护、数据准确性、数据完整性、数据时效性的要求规范，写到答题卷中。  
(4)请根据数据集的应用需求，给出4条数据清洗和预处理规范，写到答题卷中。  
请将答案写到答题卷文件中，答题卷文件命名为“1.1.4.docx”，保存到桌面考生文件夹，考生文件夹命名为“准考证号+身份证后6位”。

**3.**技能要求

（1）能对预处理后的业务数据进行审核；   
（2）能结合人工智能技术要求梳理业务数据采集规范；   
（3）能结合人工智能技术要求梳理业务数据处理规范。

**4.**质量指标

（1）业务数据采集梳理符合规范要求；   
（2）业务数据处理梳理符合规范要。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：化妆品销售数据质量检测

考核时间：20min

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 | 4 | 发现的问题描述正确，并有具体举例说明。每发现一个问题得1分，最多得4分； | 根据数据 |  |  |
| 合计配分 | 4 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 8 | 总结报告涵盖准确性，表达与答案类似：得2分； 总结报告涵盖完整性，表达与答案类似：得2分； 总结报告涵盖一致性，表达与答案类似：得2分； 总结报告涵盖重复性，表达与答案类似：得2分； |  |  |  |  |
| J2 | 4 | 用户隐私表述正确，得1分； 数据准确性表述正确，得1分； 数据完整性表述正确，得1分； 数据时效性表述正确，得1分 |  |  |  |  |
| J3 | 4 | 回答与数据清洗、预处理答案类似，合理的每条得1分，最多得4分 |  |  |  |  |
| 合计配分 | 16 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：笔记本电脑评论数据质量检测

考核时间：20min

**1.**场地设备要求

（1）人工智能训练师主机；   
（2）Office软件；   
（3）1.1.5数据集。

**2.**工作任务

A企业采集了用户对该公司笔记本电脑的评论数据，通过对商品评论进行人工智能分析，判断用户对商品或服务的意见，从而帮助企业更好地了解用户需求和市场动态，提升产品竞争力和市场影响力。你作为一名人工智能训练师，请对预处理后的评论数据集（1.1.5数据集）质量进行审核，梳理数据采集和数据处理规范。请完成以下任务：  
(1)请根据数据质量审核遵循的核心纬度，对已采集的数据质量进行审核，请记录发现的两个问题及所参考的纬度到答题卷中。  
(2)对该系统数据集的数据质量进行评估，总结数据集准确性、完整性、平衡性、重复性方面的情况，写到答题卷中。  
(3)如果采用从电商平台网站爬取用户评论数据的方式采集数据，请梳理需要遵守的数据采集规范和要求，写到答题卷中。  
(4)请根据数据集的应用需求，给出4条数据清洗和预处理规范，写到答题卷中。  
请将答案写到答题卷文件中，答题卷文件命名为“1.1.5.docx”，保存到桌面考生文件夹，考生文件夹命名为“准考证号+身份证后6位”。

**3.**技能要求

（1）能对预处理后的业务数据进行审核；   
（2）能结合人工智能技术要求梳理业务数据采集规范；   
（3）能结合人工智能技术要求梳理业务数据处理规范。

**4.**质量指标

（1）业务数据采集梳理符合规范要求；   
（2）业务数据处理梳理符合规范要。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：笔记本电脑评论数据质量检测

考核时间：20min

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 | 4 | 问题的描述要正确，并有具体举例说明。写出1条得2分，2条及以上得4分； | 根据数据 |  |  |
| 合计配分 | 4 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 8 | 总结报告涵盖准确性、完整性、平衡性、重复性，表述正确，每给出1条得2分，给出4条及以上给8分； |  |  |  |  |
| J2 | 4 | 给出与答案类似的采集规范，每给出1条得1分，4条及以上得4分； |  |  |  |  |
| J3 | 4 | 与答案类似的数据清洗和预处理规范，每给出1条规范得1分，4条及以上得4分； |  |  |  |  |
| 合计配分 | 16 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：道路标识图数据处理方法优化

考核时间：15min

**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公司开发的自动驾驶系统需要准确地识别和理解道路标识，以保证车辆行驶的安全性，减少违章行为。道路标识图片数据的采集范围应该包括城市道路、高速公路、乡村道路、山区道路等多种道路类型，以确保系统能够识别各种道路环境下的交通标识；还需要在不同时间段（如白天、夜晚、黄昏、黎明）和天气条件下（晴天、雨天、雾天、雪天）采集数据，以应对复杂多变的驾驶环境。你作为自动驾驶技术项目组人工智能训练师，请对现有采集和处理流程进行优化。  
（1）根据系统需求，A公司已经制定出的道路交通标识图采集流程如下，请对该采集流程和方案提出优化建议。  
第一步：根据采集目标制定采集方案：通过车载摄像头拍摄沿途的交通标识，采集路线涵盖了主干道和部分次干道；采集时间8:00-20:00。  
第二步：设备准备：包括高清相机，设备电量充足、存储卡空间足够、镜头清洁无遮挡。   
第三步：实地采集：按照计划路线行进，确保不遗漏任何需要采集的标识。   
第四部：数据存储： 将采集到的图像数据及时备份至安全的存储设备中。  
第五步：数据交付：按照约定方式将数据包交付给客户或相关部门。  
（2）在数据质量审核时，发现数据集中有图片不清晰、图像失真、数据标注不一致、各类道路标识分布不均等情况，请选择其中2个问题分别给出两条优化建议，提升数据集质量。   
请将答案写到答题卷文件中，答题卷文件命名为“1.2.1.docx”，保存到桌面考生文件夹，考生文件夹命名为“准考证号+身份证后6位”。

**3.**技能要求

（1）能对业务数据采集流程提出优化建议；   
（2）能对业务数据处理流程提出优化建议。

**4.**质量指标

（1）对业务数据采集流程提出的优化建议合理，具有操作性；   
（2）对业务数据处理流程提出的优化建议合理，具有操作。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：道路标识图数据处理方法优化

考核时间：15min

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 6 | 提出增加数据质量监控（检查）环节，得2分； 提出采集路线优化建议，得2分； 提出采集时间优化建议，得2分； |  |  |  |  |
| J2 | 4 | 与参考答案中优化方案类似，每提出1点，得2分，最多得4分； |  |  |  |  |
| 合计配分 | 10 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：语音识别系统的数据采集与处理优化

考核时间：15min

**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公司开发的语音识别系统需要准确识别和理解不同方言、口音和语速的语音，以提高系统的用户体验和实用性。你作为语音识别技术项目组的人工智能训练师，将对语音数据集的采集和处理流程提出优化建议。  
(1)语音数据采集流程如下，请对采集流程和采集方案提出3条优化建议，写到答题卷文件中。  
第一步：根据采集目标制定采集方案：通过麦克风录制不同方言和口音的语音，采集样本涵盖不同年龄、性别、职业的人群。  
第二步：设备准备：选择麦克风和录音设备。  
第三步：实地采集：在选定的采集地点，按照采集方案要求，进行语音录制。  
第四步：数据存储：将采集到的语音数据及时存储到预先准备好的存储设备。  
第五步：数据交付：将数据交付给项目组的数据分析师或机器学习工程师。  
(2)在数据质量审核时，发现数据集中有背景噪音大、语音不清晰、采集样本方言和口音分布不均、部分数据标注不一致等问题，请制定一个数据处理优化方案，提升数据集质量。提出的数据处理优化方案写到答题卷文件中。  
请将答案写到答题卷文件中，答题卷文件命名为“1.2.2.docx”，保存到桌面考生文件夹，考生文件夹命名为“准考证号+身份证后6位”。

**3.**技能要求

(1) 能对业务数据采集流程提出优化建议；   
(2) 能对业务数据处理流程提出优化建议。

**4.**质量指标

(1)采集优化方案的建议具体、适用、可行；   
(2)数据处理优化方案具体、适用、可行。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：语音识别系统的数据采集与处理优化

考核时间：15min

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 6 | 回答采集方案优化建议或采集流程优化建议与答案类似：每条可得2分，最多6分； |  |  |  |  |
| J2 | 4 | 回答数据处理优化方案与答案类似：每条可得1分，最多4分； |  |  |  |  |
| 合计配分 | 10 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：智能家居系统的数据采集和处理优化

考核时间：15min

**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公司开发的智能家居系统需要通过传感器数据来自动控制家居设备，以提高居住环境的舒适性和安全性。你作为智能家居技术项目组的人工智能训练师，将对传感器数据集的采集和处理流程提出优化建议。  
(1)传感器数据采集流程如下，请对采集流程和采集方案提出3条优化建议，写到答题卷文件中。  
第一步：根据采集目标制定采集方案：通过智能家居传感器采集家庭环境数据，包括温度、湿度、光照强度、运动检测等；采集时间24小时连续采集。  
第二步：设备准备：准备所需传感器，如测量温度、湿度、光照强度、运动检测等指标的传感器。  
第三步：实地采集：确认所有设备正常工作后，启动数据采集程序，开始24小时连续采集。  
第四步：数据存储：根据数据量大小，选择本地服务器、云存储或混合存储方案。  
第五步：数据交付： 将数据及相关文档交付给项目团队或指定部门。  
(2)在数据质量审核时，发现数据集中有数据缺失、传感器数据噪音大、不同传感器数据时间不同步、部分数据标注不准确等问题，请制定一个数据处理优化方案，提升数据集质量。数据处理优化方案写到答题卷文件中。  
请将答案写到答题卷文件中，答题卷文件命名为“1.2.3.docx”，保存到桌面考生文件夹，考生文件夹命名为“准考证号+身份证后6位”。

**3.**技能要求

(1) 能对业务数据采集流程提出优化建议；   
(2) 能对业务数据处理流程提出优化建议。

**4.**质量指标

(1)采集优化方案的建议具体、适用、可行 ；   
(2)数据处理优化方案具体、适用、可行。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：智能家居系统的数据采集和处理优化

考核时间：15min

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 6 | 回答采集方案优化建议或采集流程优化建议与答案类似：每条可得2分，最多6分； |  |  |  |  |
| J2 | 4 | 回答数据处理优化方案与答案类似：每条可得2分，最多4分； |  |  |  |  |
| 合计配分 | 10 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：化妆品销售数据处理方法优化

考核时间：15min

**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

化妆品行业呈现出稳步增长的趋势，A企业为优化化妆品销售策略，实现精准营销和个性化推荐，同时为产品研发和创新提供方向，需要收集化妆品销售数据并开发人工智能系统进行分析。你作为人工智能训练师，请对数据采集及处理方法进行优化。  
（1）目前已经制定出的化妆品销售数据采集流程如下，请对该数据采集流程提出优化建议。  
第一步：制定采集方案：使用电商平台公开的API接口爬取数据。  
第二步： 环境准备：安装并配置爬虫所需的编程环境（如Python环境）、IDE（如PyCharm）、爬虫框架和库。  
第三步：编写爬虫代码：根据采集方案编写爬虫脚本，包括API接口调用、数据解析、数据存储、异常处理等。  
第四步：数据采集与存储：按照计划启动爬虫，将采集到的数据存储到本地数据库。  
第五步：数据交付：整理好的数据交付给客户或相关部门。  
（2）来自不同电商平台的数据存在格式不一致，并发现有数值不正确和缺失的问题，请给出优化建议。  
请将答案写到答题卷文件中，答题卷文件命名为“1.2.4.docx”，保存到桌面考生文件夹，考生文件夹命名为“准考证号+身份证后6位”。

**3.**技能要求

（1）能对业务数据采集流程提出优化建议；   
（2）能对业务数据处理流程提出优化建议。

**4.**质量指标

（1）对业务数据采集流程提出的优化建议合理，具有操作性；   
（2）对业务数据处理流程提出的优化建议合理，具有操作性。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：化妆品销售数据处理方法优化

考核时间：15min

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 6 | 回答数据采集流程提出优化建议与答案类似：每条可得2分，最多6分； |  |  |  |  |
| J2 | 4 | 回答针对数据问题的优化建议与答案类似：每条可得2分，最多4分； |  |  |  |  |
| 合计配分 | 10 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：推荐系统的数据采集策略和流程优化

考核时间：15min

**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公司开发的推荐系统需要高质量的用户行为数据集，以提高推荐的准确性和用户满意度。你作为推荐系统技术项目组的人工智能训练师，将对用户行为数据集的采集流程和策略提出优化建议。  
(1)用户行为数据采集流程如下，请对采集流程和采集方案提出优化建议，写到答题卷文件中。  
第一步：根据采集目标制定采集方案：通过网站和移动应用收集用户浏览、点击、购买等行为数据，样本涵盖不同年龄、性别、地域的用户；采集时间为一个月。  
第二步：工具准备  
第三步：数据采集  
第四步：数据存储  
第五步：数据交付  
(2)在数据采集策略审核时，发现当前的采集方案未能充分覆盖不同用户群体，缺乏多样化的行为数据，请制定优化建议以改进数据采集策略，确保更全面的数据覆盖和准确性。采集优化建议写到答题卷文件中。  
请将答案写到答题卷文件中，答题卷文件命名为“1.2.5.docx”，保存到桌面考生文件夹，考生文件夹命名为“准考证号+身份证后6位”。

**3.**技能要求

(1) 能对业务数据采集流程提出优化建议；   
(2) 能对业务数据采集策略提出优化建议。

**4.**质量指标

(1)采集优化方案的建议具体、适用、可行 ；   
(2)数据采集策略优化方案具体、适用、可行。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：推荐系统的数据采集策略和流程优化

考核时间：15min

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 6 | 回答采集方案优化建议或采集流程优化建议与答案类似：每条可得2分，最多6分； |  |  |  |  |
| J2 | 4 | 回答数据处理优化方案与答案类似：每条可得1分，最多4分； |  |  |  |  |
| 合计配分 | 10 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：基于人类照片的性别的数据归类和定义

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机；   
（2）labelImg；   
（3）Office软件；   
（4）Python 编译环境。

**2.**工作任务

人类性别智能分类系统是一种集成的一种基于人工智能技术的系统，旨在根据人的外貌特征和其他相关信息，自动识别和分类人的性别。在本项目中，作为一名人工智能训练师，请对已经清洗完成的数据使用labelImg工具进行数据标注，并对数据进行处理。  
（1）使用labelImg对2.1.1素材文件夹中“2.1.1pic”文件夹中的15张图片进行性别分类标注，性别枚举值: [Male,Female]，标注结果文件命名为“2.1.1性别.xml”。  
（2）打开2.1.1素材文件夹中的程序文件“2.1.1.ipynb”，按照下列要求补全代码，对“2.1.1pic”文件夹中的15张图片数据进行以下操作。  
1）将“2.1.1pic”文件夹中图片样本数据格式修改为“jpg”格式，将此部分代码截图保存至“2.1.1程序1.docx”文件中。  
2）将上步骤“2.1.1pic”文件夹中已经被修改为“jpg”格式的图片样本以数字编号重命名，数据从“0000”开始计数，顺序累加，如0000.jpg，0001.jpg，0002.jpg….以此类推直至结束。将此部分代码截图保存至“2.1.1程序2.docx”文件中。下载处理好的数据集，将其压缩打包保存至“2.1.1pic-更新.zip”。  
所有操作提交结果（“2.1.1性别.xml”、“2.1.1程序1.docx”、“2.1.1程序2.docx”、“2.1.1pic-更新.zip”）保存在桌面新建的一个文件夹中，文件夹命名为“准考证号+身份号后六位”。

**3.**技能要求

（1）能够运用工具，对杂乱数据进行分析，输出内在关联及特征；   
（2）能够运用工具，对标注后的数据进行提取归档；   
（3）能根据数据内在关联和特征进行数据定义。

**4.**质量指标

（1）代码书写规范；   
（2）代码正常运行；   
（3）数据标注正确，输出归档文件正确。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：基于人类照片的性别的数据归类和定义

考核时间：30min

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 | 15 | 按照指定格式进行性别分类标注：每标注1张得1分，总共15分； | 根据数据 |  |  |
| M2 | 6 | 正确补全代码，每空得2分，总共6分； | 根据数据 |  |  |
| M3 | 6 | 正确补全代码，每空得2分，总共6分； | 根据数据 |  |  |
| M4 | 3 | 下载处理好的数据集，结果正确得3分，结果不正确得0分。 | 根据数据 |  |  |
| 合计配分 | 30 | 合计得分 | |  |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：基于人类照片的肤色的数据归类和定义

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机；   
（2）labelImg；   
（3）Office软件；   
（4）Python 编译环境。

**2.**工作任务

人类肤色智能分类系统是一种集成的一种基于人工智能技术的系统，旨在根据人的外貌特征和其他相关信息，自动识别和分类人的肤色。在本项目中，作为一名人工智能训练师，将会对已经清洗完成的数据使用labelImg工具进行数据标注，并对数据进行处理。  
（1）使用labelImg对2.1.2素材文件夹中“2.1.2pic”文件夹中的15张图片进行肤色分类标注，肤色枚举值: [White race,Yellow race，Black race]，标注结果文件命名为“2.1.2肤色.xml”。  
（2）打开素材文件夹中的程序文件 “2.1.2.ipynb”，按照下列要求补全代码，对“2.1.2pic”文件夹中的15张图片数据进行操作。  
1）将“2.1.2pic”文件夹中图片样本数据格式修改为“jpg”格式，将此部分代码截图保存至“提交/2.1.2程序1.docx”文件中。  
2）将上步骤“2.1.2pic”文件夹中已经被修改为“jpg”格式的图片样本名称重新定义，并利用数字编号。数据从“0000”开始计数，顺序累加，如0000.jpg，0001.jpg，0002.jpg….以此类推直至结束。将此部分代码截图保存至“2.1.2程序2.docx”文件中。下载处理好的数据集，将其压缩打包保存至“2.1.2pic-更新.zip”。  
所有操作提交结果（“2.1.2肤色.xml”、“2.1.2程序1.docx”、“2.1.2程序2.docx”、“2.1.2pic-更新.zip”）保存在桌面新建的一个文件夹中，文件夹命名为“准考证号+身份号后六位”。

**3.**技能要求

（1）能够运用工具，对杂乱数据进行分析，输出内在关联及特征；   
（2）能够运用工具，对标注后的数据进行提取归档；   
（3）能根据数据内在关联和特征进行数据定义。

**4.**质量指标

（1）代码书写规范；   
（2）代码正常运行；   
（3）数据标注正确，输出归档文件正确。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：基于人类照片的肤色的数据归类和定义

考核时间：30min

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 | 15 | 按照指定格式进行性别分类标注：每标注1张得1分，总共15分； | 根据数据 |  |  |
| M2 | 6 | 正确补全代码，每空得2分，总共6分； | 根据数据 |  |  |
| M3 | 6 | 正确补全代码，每空得2分，总共6分； | 根据数据 |  |  |
| M4 | 3 | 下载处理好的数据集，结果正确得3分，结果不正确得0分。 | 根据数据 |  |  |
| 合计配分 | 30 | 合计得分 | |  |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：基于运输工具图像的数据归类和定义

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机；   
（2）labelImg；   
（3）Office软件；   
（4）Python 编译环境。

**2.**工作任务

运输工具类型识别系统是一种集成的基于AI视觉技术的系统，旨在根据运输工具的视觉特征和其他相关信息，自动识别和分类不同类型的运输工具。在本项目中，作为一名人工智能训练师，将会对已经清洗完成的数据使用labelImg工具对已经预处理的运输工具图像进行标注，并对标注后的数据进行特征提取和归档。  
（1）使用labelImg工具对题目下dataset文件夹中的所有图片进行标注，并对每张图片生成标注的结果导出并命名为“图片名.xml”文件保存在相同的目录下，具体标注要求如下：  
1）使用方形圈出图片中的运输工具；  
2）对图片进行标注，标注的格式为: 运输工具类型；其中预设的运输工具类型见本题文件夹中的“predefined\_classes.txt”文件。  
（2）打开文件夹中的程序文件 2.1.3.py，按照下列要求及代码中的注释补全代码，对上题每张图片标注的结果进行提取，将提取到的结果进行归档合并。  
1）遍历dataset文件夹中所有的xml文件，提取标注内容。  
2）参考sample.csv的格式，将最终的归档结果提取到 result.csv中。  
所有操作结果文件储存在桌面新建的一个文件夹中，文件夹命名为“准考证号+身份号后六位”。

**3.**技能要求

（1）能够运用工具，对杂乱数据进行分析，输出内在关联及特征；   
（2）能够运用工具，对标注后的数据进行提取归档；   
（3）能根据数据内在关联和特征进行数据定义

**4.**质量指标

（1）代码书写规范；   
（2）代码正常运行；   
（3）数据标注正确，输出归档文件正确

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：基于运输工具图像的数据归类和定义

考核时间：30min

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 | 15 | 按照指定格式标注，并将其标注结果正确保存为相应xml，每标注1张正确的图片，得1分，保存成1个xml文件得0.5分，总共15分； | 根据数据 |  |  |
| M2 | 10 | 正确打开文件补全代码，每空得1分，总共10分； | 根据数据 |  |  |
| M3 | 5 | 每张图片归档结果提取正确，得0.5分，总共5分； | 根据数据 |  |  |
| 合计配分 | 30 | 合计得分 | |  |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：智能相册系统的图像数据定义和分类

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机；   
（2）labelImg；   
（3）Office软件；   
（4）Python 编译环境。

**2.**工作任务

相册分类系统是一种基于人工智能技术的集成系统，旨在根据照片中的人物、风景和物品的外观特征及相关信息，自动识别和分类不同的内容类型。你作为人工智能训练师，需要对数据进行定义，并使用labelImg工具对已经清洗完成的数据进行标注，并对标注后的数据进行提取和归档。  
（1）使用labelImg工具对题目下dataset文件夹中的所有图片进行标注，并对每张图片生成标注的结果导出并命名为“图片名.xml”文件保存在相同的目录下，具体标注要求如下：  
1）根据图片特征，使用方形圈出图片中的各类；  
2）对图片进行标注，标注的格式为: 分类类型；其中分类类型的名称来源于“predefined\_classes.txt”文件。  
（2）打开文件夹中的程序文件 2.1.4.py，按照下列要求及代码中的注释补全代码，对上题每张图片标注的结果进行提取，将提取到的结果进行归档合并。  
1）遍历dataset文件夹中所有的xml文件，提取标注内容。  
2）参考sample.csv的格式，将最终的归档结果提取到 result.csv中。  
所有操作结果文件储存在桌面新建的一个文件夹中，文件夹命名为“准考证号+身份号后六位”。

**3.**技能要求

（1）能够运用工具，对杂乱数据进行分析，输出内在关联及特征；   
（2）能够运用工具，对标注后的数据进行提取归档；   
（3）能根据数据内在关联和特征进行数据定义。

**4.**质量指标

（1）代码书写规范；   
（2）代码正常运行；   
（3）数据标注正确，输出归档文件正确。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：智能相册系统的图像数据定义和分类

考核时间：30min

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 | 15 | 按照指定格式标注，并将其标注结果正确保存为相应xml，每标注1张正确的图片，得1分，保存成1个xml文件得0.5分，总共15分； | 根据数据 |  |  |
| M2 | 10 | 正确打开文件补全代码，每空得1分，总共10分； | 根据数据 |  |  |
| M3 | 5 | 每张图片归档结果提取正确，得0.5分，总共5分； | 根据数据 |  |  |
| 合计配分 | 30 | 合计得分 | |  |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：动物分类识别系统的图像数据定义和分类

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机；   
（2）labelImg；   
（3）Office软件；   
（4）Python 编译环境。

**2.**工作任务

动物分类识别系统是一种集成的基于人工智能技术的系统，旨在根据动物的外貌特征和其他相关信息，自动识别和分类不同种类的动物。在本项目中，作为一名人工智能训练师，将会对已经清洗完成的数据使用labelImg工具进行数据标注，并对标注后的数据进行提取和归档。  
（1）使用labelImg工具对题目下dataset文件夹中的所有图片进行标注，并对每张图片生成标注的结果导出并命名为“图片名.xml”文件保存在相同的目录下，具体标注要求如下：  
1）根据图片特征，使用方形圈出图片中动物的分类；  
2）对图片进行标注，标注的格式为: 动物分类类型；其中动物分类类型的名称来源于“predefined\_classes.txt”文件。  
（2）打开文件夹中的程序文件 2.1.5.py，按照下列要求及代码中的注释补全代码，对上题每张图片标注的结果进行提取，将提取到的结果进行归档合并。  
1）遍历dataset文件夹中所有的xml文件，提取标注内容。  
2）参考sample.csv的格式，将最终的归档结果提取到 result.csv中。  
所有操作结果文件储存在桌面新建的一个文件夹中，文件夹命名为“准考证号+身份号后六位”。

**3.**技能要求

（1）能够运用工具，对杂乱数据进行分析，输出内在关联及特征；   
（2）能够运用工具，对标注后的数据进行提取归档；   
（3）能根据数据内在关联和特征进行数据定义。

**4.**质量指标

（1）代码书写规范；   
（2）代码正常运行；   
（3）数据标注正确，输出归档文件正确。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：动物分类识别系统的图像数据定义和分类

考核时间：30min

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 | 15 | 按照指定格式标注，并将其标注结果正确保存为相应xml，每标注1张正确的图片，得1分，保存成1个xml文件得0.5分，总共15分； | 根据数据 |  |  |
| M2 | 10 | 正确打开文件补全代码，每空得1分，总共10分； | 根据数据 |  |  |
| M3 | 5 | 每张图片归档结果提取正确，得0.5分，总共5分； | 根据数据 |  |  |
| 合计配分 | 30 | 合计得分 | |  |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：基于人类照片的性别年龄标注数据审核

考核时间：25min

**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

人类性别年龄智能分类系统是一种集成的一种基于人工智能技术的系统，旨在根据人的外貌特征和其他相关信息，自动识别和分类人的性别和年龄段。作为一名人工智能训练师，请对该系统中已标注的性别、年龄段数据进行审核，并完成以下任务。  
（1）文件《2.2.1标注结果\_待审核.xlsx》的数据是对“2.2.1pic”文件夹中15张图片的标注结果，其中label\_age是对年龄段的标注，label\_gender是对性别的标注，请从数据的完整性和准确性两方面进行审核，填写审核报告，保存在“2.2.1审核报告.docx”文件中。  
（2）请纠正审核过程中发现的错误，将结果保存至“2.2.1审核纠正.xlsx”文件中。  
（3）请根据纠正后的结果，为模型训练按照“7：3”的比例分配训练数据和测试数据，分配数据打包压缩保存至“2.2.1分配结果.zip”文件中。  
所有操作提交结果（“2.2.1审核报告.docx”、“2.2.1审核纠正.xlsx”、“2.2.1分配结果.zip”）储存在桌面新建的一个文件夹中，文件夹命名为“准考证号+身份号后六位”。

**3.**技能要求

（1）能够完成对标注数据准确性和完整性审核，输出审核报告；   
（2）能够对审核过程中发现的错误进行纠正；   
（3）能够根据审核结果完成数据筛选。

**4.**质量指标

（1）错误描述或纠正正确；   
（2）label书写正确；   
（3）数据保存格式正确。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：基于人类照片的性别年龄标注数据审核

考核时间：25min

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 | 4 | 对数据的完整性进行审核，每审核出1处错误，得2分，总共4分； | 根据数据 |  |  |
| M2 | 4 | 对数据的准确性进行审核，每审核出1处错误，得2分，总共4分； | 根据数据 |  |  |
| M3 | 2 | 对标注错误的数据修改正确，每正确修改1处，得1分，总共2分； | 根据数据 |  |  |
| M4 | 2 | 对未标注的数据进行标注，每正确标注1处，得1分，总共2分； | 根据数据 |  |  |
| M5 | 8 | 训练数据数量在范围内，正确得3分，范围外得0分； 测试数据数量在范围内，正确得3分，范围外得0分； 数据与测试数据数量之和为15条，正确得3分，不正确得0分； | 根据数据 |  |  |
| 合计配分 | 20 | 合计得分 | |  |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：基于人类照片的性别肤色标注数据审核

考核时间：25min

**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

人类性别肤色智能分类系统是一种集成的一种基于人工智能技术的系统，旨在根据人的外貌特征和其他相关信息，自动识别和分类人的性别和肤色。作为一名人工智能训练师，请对该系统中已标注的性别、肤色数据进行审核，并完成以下任务。  
（1）文件《2.2.2标注结果\_待审核.xlsx》为对目录下的15张图片的标注结果，其中label\_skin colour为对肤色的标注，label\_gender为对性别的筛选，请从数据的完整性和准确性两方面进行审核，填写审核报告，保存在“2.2.2审核报告.docx”文件中。  
（2）请纠正审核过程中发现的错误，将结果保存至“2.2.2审核纠正.xlsx”文件中。  
（3）请根据纠正后的结果，为模型训练按照“7：3”的比例分配训练数据和测试数据，分配数据打包压缩保存至“2.2.2分配结果.zip”文件中。  
所有操作提交结果（“2.2.2审核报告.docx”、“2.2.2审核纠正.xlsx”、“2.2.2分配结果.zip”）储存在桌面新建的一个文件夹中，文件夹命名为“准考证号+身份号后六位”。

**3.**技能要求

（1）能够完成对标注数据准确性和完整性审核，输出审核报告；   
（2）能够对审核过程中发现的错误进行纠正；   
（3）能够根据审核结果完成数据筛选、

**4.**质量指标

（1）错误描述或纠正正确；   
（2）label书写正确；   
（3）数据保存格式正确。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：基于人类照片的性别肤色标注数据审核

考核时间：25min

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 | 4 | 对数据的完整性进行审核，每审核出1处错误，得2分，总共4分； | 根据数据 |  |  |
| M2 | 4 | 对数据的准确性进行审核，每审核出1处错误，得2分，总共4分； | 根据数据 |  |  |
| M3 | 2 | 对标注错误的数据修改正确，每正确修改1处，得1分，总共2分； | 根据数据 |  |  |
| M4 | 2 | 对未标注的数据进行标注，每正确标注1处，得1分，总共2分； | 根据数据 |  |  |
| M5 | 8 | 训练数据数量在范围内，正确得3分，范围外得0分； 测试数据数量在范围内，正确得3分，范围外得0分； 数据与测试数据数量之和为15条，正确得3分，不正确得0分； | 根据数据 |  |  |
| 合计配分 | 20 | 合计得分 | |  |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：智能面部表情识别系统的标注数据审核

考核时间：25min

**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

智能面部表情识别系统是一种基于人工智能技术的应用，旨在自动识别和分类人的面部表情，以便在情感分析、用户体验优化等领域进行应用。在本项目中，作为一名人工智能训练师，你将对已标注的面部图像数据进行审核，并进行错误纠正及数据筛选，确保其准确性和完整性。  
（1）文件《表情标注结果\_待审核.xlsx》包含对dataset文件夹下面部图像的标注结果，面部情绪包括：快乐（Happiness）、悲伤（Sadness）、惊讶（Surprise）、愤怒（Anger）、厌恶（Disgust）、恐惧（Fear），面部情绪每项评分最高5分，最低1分。请从数据的完整性和准确性两方面进行审核，输出审核报告。  
（2）请列出审核过程中发现的错误，并进行纠正，将结果保存至《标注结果\_已审核.xlsx》。  
（3）请根据审核结果，对数据进行筛选，选出适合模型训练的测试数据，结果保存至《标注结果\_train.xlsx》。  
所有操作结果文件储存在桌面新建的一个文件夹中，文件夹命名为“准考证号+身份号后六位”。

**3.**技能要求

（1）能够完成对标注数据准确性和完整性审核，输出审核报告；   
（2）能够对审核过程中发现的错误进行纠正；   
（3）能够根据审核结果完成数据筛选。

**4.**质量指标

（1）审核报告清晰、正确、全面；   
（2）审核过程中纠错正确；   
（3）数据筛选正确。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：智能面部表情识别系统的标注数据审核

考核时间：25min

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 | 4 | 从数据的完整性进行审核，输出报告：其中每分析一条结果得1分，最多得3分；对完整性审核的总结与答案类似得1分，本项合计最高得4分 | 根据数据 |  |  |
| M2 | 3 | 从数据的准确性进行审核，输出报告：其中每分析一条结果的1分，最多得2分；对准确性审核的总结与答案类似得1分，本项合计最高得3分； | 根据数据 |  |  |
| M3 | 2 | 对多余标注数据删除正确：每条得1分，最多得2分； | 根据数据 |  |  |
| M4 | 2 | 对每个错误标注的数据进行修改：得1分，最高得2分； | 根据数据 |  |  |
| M5 | 1 | 标注结果\_已审核.xlsx正确保存：得1分； | 根据数据 |  |  |
| M6 | 5 | 标注结果\_已审核.xlsx内容正确：每4条得1分，最高得5分 | 根据数据 |  |  |
| M7 | 1 | 标注结果\_train.xlsx正确保存：得1分 | 根据数据 |  |  |
| M8 | 2 | 数据筛选结果文件（标注结果\_train.xlsx）内容正确：每10条得1分，最高得2分； | 根据数据 |  |  |
| 合计配分 | 20 | 合计得分 | |  |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：产品评论情感分析数据审核与筛选

考核时间：25min

**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

在自然语言处理（NLP）领域，情感分析是一项重要的任务，它涉及到识别和提取文本中表达的情感或情绪。一家电子商务公司希望改进其产品评论的情感分析算法，以便更准确地了解顾客反馈并提高客户满意度。为此，他们收集了大量的产品评论数据，并进行了初步的情感标注。在本项目中，作为一名人工智能训练师，你将对已标注的数据进行审核，并进行错误纠正及数据筛选，确保其准确性和完整性，以确保模型训练的质量。  
（1）文件《产品评论\_待审核.xlsx》包含了22条产品评论及其初步的情感标注结果。标注包括正面（Positive）、负面（Negative）和中立（Neutral）。审核标准如下：  
完整性：检查每条评论是否都有情感标注。  
准确性：确认情感标注是否正确反映评论的真实情感倾向。  
编写一份审核报告，详细说明数据的完整性与准确性情况。请从数据的完整性和准确性两方面进行审核，输出审核报告。  
（2）在审核过程中，记录并纠正所有发现的标注错误。将纠正后的数据保存至《产品评论\_已审核.xlsx》。  
（3）对数据进行筛选，移除任何有争议的、模糊不清或信息不足的评论，保留高质量的标注数据用于模型训练。将筛选后的数据保存至《产品评论\_train.xlsx》。  
所有操作结果文件储存在桌面新建的一个文件夹中，文件夹命名为“准考证号+身份号后六位”。

**3.**技能要求

（1）能够完成对标注数据准确性和完整性审核，输出审核报告；   
（2）能够对审核过程中发现的错误进行纠正；   
（3）能够根据审核结果完成数据筛选。

**4.**质量指标

（1）审核报告清晰、正确、全面；   
（2）审核过程中纠错正确；   
（3）数据筛选正确。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：产品评论情感分析数据审核与筛选

考核时间：25min

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 | 3 | 从数据的完整性进行审核，输出报告：其中每分析一条结果得1分，最多得2分；对完整性审核的总结与答案类似得1分，本项合计最高得3分 | 根据数据 |  |  |
| M2 | 2 | 从数据的准确性进行审核，输出报告：其中每分析一条结果的1分，最多得1分；对准确性审核的总结与答案类似得1分，本项合计最高得2分； | 根据数据 |  |  |
| M3 | 3 | 对标注数据错误发现1条得1分，最多得3分； | 根据数据 |  |  |
| M4 | 3 | 对每个错误标注的数据进行修改：得1分，最高得3分； | 根据数据 |  |  |
| M5 | 1 | 标注结果\_已审核.xlsx正确保存：得1分； | 根据数据 |  |  |
| M6 | 4 | 标注结果\_已审核.xlsx内容正确：每5条得1分，最高得4分 | 根据数据 |  |  |
| M7 | 1 | 数据进行筛选结论与答案类似，得1分 | 根据数据 |  |  |
| M8 | 1 | 标注结果\_train.xlsx正确保存：得1分 | 根据数据 |  |  |
| M9 | 2 | 数据筛选结果文件（标注结果\_train.xlsx）内容正确：每10条得1分，最高得2分； | 根据数据 |  |  |
| 合计配分 | 20 | 合计得分 | |  |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：智能企业沟通分析系统数据审核与筛选

考核时间：25min

**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

随着远程工作和在线协作的兴起，企业内部沟通平台上的聊天记录成为了解员工互动模式、团队氛围和工作效率的重要数据源。这些数据不仅可用于分析和改进工作流程，还能帮助HR部门识别潜在的团队冲突或协作问题。为此，A公司决定建立一个自动分析工具，该工具能够识别聊天记录中的关键事件和情绪动态。本项目中，作为一名人工智能训练师，为了训练这个工具，需要一个精心标注的聊天记录数据集，确保其准确性和完整性，以确保模型训练的质量。  
（1）文件《待审核.xlsx》包含了20条企业内部聊天记录及其初步的事件和情绪标注结果。标注包括事件类别（如会议安排、任务分配、问题解决、个人成就、团队冲突等）和情绪标签（如积极、消极、中立）。审核标准如下：  
完整性：检查每条聊天记录是否都有事件类别和情绪标签。  
准确性：确认事件和情绪的标注是否正确反映聊天内容。  
编写一份审核报告，详细说明数据的完整性与准确性情况。请从数据的完整性和准确性两方面进行审核，输出审核报告。  
（2）在审核过程中，记录并纠正所有发现的标注错误。将纠正后的数据保存至《已审核.xlsx》。  
（3）对数据进行筛选，移除任何有争议的、模糊不清或信息不足的评论，保留高质量的标注数据用于模型训练。将筛选后的数据保存至《train.xlsx》。  
所有操作结果文件储存在桌面新建的一个文件夹中，文件夹命名为“准考证号+身份号后六位”。

**3.**技能要求

（1）能够完成对标注数据准确性和完整性审核，输出审核报告；   
（2）能够对审核过程中发现的错误进行纠正；   
（3）能够根据审核结果完成数据筛选。

**4.**质量指标

（1）审核报告清晰、正确、全面；   
（2）审核过程中纠错正确；   
（3）数据筛选正确。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：智能企业沟通分析系统数据审核与筛选

考核时间：25min

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 | 3 | 从数据的完整性进行审核，输出报告：其中每分析一条结果得1分，最多得2分；对完整性审核的总结与答案类似得1分，本项合计最高得3分； | 根据数据 |  |  |
| M2 | 2 | 从数据的准确性进行审核，输出报告：其中每分析一条结果的1分，最多得1分；对准确性审核的总结与答案类似得1分，本项合计最高得2分； | 根据数据 |  |  |
| M3 | 3 | 对标注数据错误发现1条得1分，最多得3分； | 根据数据 |  |  |
| M4 | 3 | 对每个错误标注的数据进行修改：得1分，最高得3分； | 根据数据 |  |  |
| M5 | 1 | 标注结果\_已审核.xlsx正确保存：得1分； | 根据数据 |  |  |
| M6 | 4 | 标注结果\_已审核.xlsx内容正确：每4条得1分，最高得4分； | 根据数据 |  |  |
| M7 | 1 | 数据进行筛选结论与答案类似，得1分； | 根据数据 |  |  |
| M8 | 1 | 标注结果\_train.xlsx正确保存：得1分； | 根据数据 |  |  |
| M9 | 2 | 数据筛选结果文件（标注结果\_train.xlsx）内容正确：每9条得1分，最高得2分； | 根据数据 |  |  |
| 合计配分 | 20 | 合计得分 | |  |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：超市智能销售系统运维

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机1台；   
（2）开发工具：Python编译工具；   
（3）标注工具：Microsoft Excel；   
（4）Word。

**2.**工作任务

为支持A超市的备货决策，找出销售爆款品种。请完成以下任务。  
（1）使用Excel加载“A超市的销售数据.csv”，找到小类名称为“不锈钢碗”的数据，并对缺失值进行填充，输出填充后小类名称为“不锈钢碗”的数据。提供截图，以jpg的格式保存并命名为3.1.1-1.jpg。  
（2）补全3.1.1.ipynb代码完成以下统计分析：  
1、加载数据集“A超市的销售数据.csv”打印结果，统计数据集大类的名称个数是多少。  
2、计算不同“大类名称”下的总销售金额，并按照总销售金额从大到小排序。  
3、选择销售金额最大的“大类名称”分析在该“大类”下销售总金额最大的前15个小类的分布情况，以柱状图的方式输出。  
将以上代码以及运行结果，以html格式保存并命名为3.1.1-2.html。  
（3）分析出当前最热大类下的前五个小类产品，用于支持下个季度的备货决策。将分析结果填入答题文件3.1.1-3.docx。

**3.**技能要求

（1）能够维护智能系统所需的知识库；   
（2）能维护智能系统所需的数据；   
（3）能为单一的产品找到合适的应用场景。

**4.**质量指标

（1）数据完整性和准确性；   
（2）数据分析的准确性；   
（3）输出结果可视化。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：超市智能销售系统运维

考核时间：30min

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 | 4 | 正确填充1条缺失数据：+2分，总共4分； | 根据数据 |  |  |
| M2 | 2 | 正确加载数据集并输出：+2分； | 根据数据 |  |  |
| M3 | 1 | 正确统计出大类个数：+2分； | 根据数据 |  |  |
| M4 | 2 | 正确统计出按大类分布的销售金额情况：+2； | 根据数据 |  |  |
| M5 | 1 | 以上数据从大到小降序排序：+2分； | 根据数据 |  |  |
| M6 | 2 | 筛选出销售金额最大的大类名称：+2分； | 根据数据 |  |  |
| M7 | 2 | 计算该大类下各小类的销售总金额：+2分； | 根据数据 |  |  |
| M8 | 1 | 筛选出各小类的销售总金额前15：+1分； | 根据数据 |  |  |
| M9 | 1 | 绘制柱状图：+1分； | 根据数据 |  |  |
| M10 | 2 | 分析出当前销售金额最大的前五个小类：+2分； | 根据数据 |  |  |
| 合计配分 | 18 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 2 | 根据分析结果给出合理的备货建议，涵盖品类，资金，备货时间，备货量等方面中的一个：+2分； |  |  |  |  |
| 合计配分 | 2 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：气象分析智能系统运维

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机1台；   
（2）开发工具：Python编译工具；   
（3）标注工具：Microsoft Excel；   
（4）Word。

**2.**工作任务

为了评估气压变化对天气条件的影响，你需要对天气数据进行分析。气压的变化可能对天气条件有显著影响，进而影响农业。例如，剧烈的气压变化可能会导致极端天气事件，如暴风雨或干旱，这对作物的生长和收成有直接影响。完成以下任务：  
（1）使用Excel加载“A地区的天气数据.csv”，查找所有记录中“压力差”为缺失值的数据，并使用“最高压力”减去“最低压力”填充。输出填充后的数据，提供截图，以jpg的格式保存并命名为3.1.2-1.jpg。  
（2）补全3.1.2.ipynb代码完成以下统计分析：  
1、加载数据集“A地区的天气数据.csv”，并用print函数输出加载后的数据的头部信息。  
2、计算2009年1-11月份的每月的平均压力差，并按照平均压力差从大到小排序。  
3、给出2009年压力差平均值最大的月份，以折线图的方式展示该月份每天的压力差。  
将以上代码以及运行结果，以html格式保存并命名为3.1.2-2.html。  
（3）基于所有年份压力差的统计数据，分析哪些月份的天气条件可能最不稳定，并讨论这些不稳定性可能对当地的农业活动有何影响。将答案填入word文件保存并命名为3.1.2-3.docx。

**3.**技能要求

（1）能够维护智能系统所需的知识库   
（2）能维护智能系统所需的数据   
（3）能为单一的产品找到合适的应用场景

**4.**质量指标

（1）数据完整性和准确性   
（2）数据分析的准确性   
（3）输出结果可视化

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：气象分析智能系统运维

考核时间：30min

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 | 2 | 正确输出2条填充后的数据：+2分； | 根据数据 |  |  |
| M2 | 1 | 正确加载数据集：+1分； | 根据数据 |  |  |
| M3 | 1 | 正确输出数据的头部信息：+1分； | 根据数据 |  |  |
| M4 | 1 | 正确格式化筛选日期：+1分； | 根据数据 |  |  |
| M5 | 1 | 正确筛选2009年：+1分 | 根据数据 |  |  |
| M6 | 2 | 正确计算出2009年的每月平均压力差：+2分； | 根据数据 |  |  |
| M7 | 2 | 输出的月份平均压力差从大到小到小排序：+2分； | 根据数据 |  |  |
| M8 | 2 | 正确给出2009年压力差平均值最大的月份：+2分； | 根据数据 |  |  |
| M9 | 1 | 正确筛选月数据：+1分； | 根据数据 |  |  |
| M10 | 2 | 正确绘制折线图：+2分； | 根据数据 |  |  |
| M11 | 1 | 正确显示折线图：+1分； | 根据数据 |  |  |
| M12 | 2 | 分析过程采用所有年份数据，正确列举出前2个月份：+2分； | 根据数据 |  |  |
| 合计配分 | 18 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 2 | 给出合理的数据分析，分析出天气条件最不稳定的月份，给出合理理由：+2分； |  |  |  |  |
| 合计配分 | 2 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：餐饮智能系统运维

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机1台；   
（2）开发工具：Python编译工具；   
（3）标注工具：Microsoft Excel；   
（4）Word。

**2.**工作任务

为了更好地评估上海各行政区餐厅的服务质量和消费者评价，你需要对“上海餐饮数据.csv”进行分析。完成以下任务：  
（1）使用Excel加载“上海餐饮数据.csv”，查找所有记录中行政区缺失的数据，并使用“未知”填充“行政区”字段，用“其他”填充“类别”字段。输出填充后的数据，提供截图，以jpg的格式保存并命名为3.1.3-1.jpg。  
（2）补全3.1.3.ipynb代码完成以下统计分析：  
1、加载数据集“上海餐饮数据.csv”，并用print函数输出加载后“行政区”字段为“浦东新区”的数据头部信息。  
2、计算每个行政区的平均“口味”、“环境”和“服务”评分（不包含“未知”行政区），并按照综合评分从高到低排序。综合评分可由这三个评分的平均值计算得出。  
3、选择综合评分最高的行政区，以人均消费最高的前15个类别并以条形图的方式展示。  
将以上代码以及运行结果，以html格式保存并命名为3.1.3-2.html。  
（3）综合评分最高的行政区和该区的人均消费数据，分析消费者的偏好及其对餐饮业的潜在影响。讨论这些因素如何影响餐厅的服务提升和定价策略。将答案填入word文件保存并命名为3.1.3-3.docx。

**3.**技能要求

（1）能够维护智能系统所需的知识库；   
（2）能维护智能系统所需的数据；   
（3）能为单一的产品找到合适的应用场景。

**4.**质量指标

（1）数据完整性和准确性；   
（2）数据分析的准确性；   
（3）输出结果可视化。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：餐饮智能系统运维

考核时间：30min

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 | 4 | 正确填充1条缺失数据：+1分，总共4分； | 根据数据 |  |  |
| M2 | 1 | 正确加载数据集：+1分； | 根据数据 |  |  |
| M3 | 1 | 正确输出数据的头部信息：+1分； | 根据数据 |  |  |
| M4 | 1 | 正确过滤未知区域：+1分； | 根据数据 |  |  |
| M5 | 1 | 正确计算综合评分：+1分； | 根据数据 |  |  |
| M6 | 1 | 正确计算每个区域的口味”、“环境”和“服务”评分：+1分； | 根据数据 |  |  |
| M7 | 1 | 正确输出综合评分最高的区域：+1分； | 根据数据 |  |  |
| M8 | 1 | 筛选出综合评分最高区域的数据：+1分； | 根据数据 |  |  |
| M9 | 2 | 正确计算人均消费的均值：+2分； | 根据数据 |  |  |
| M10 | 1 | 正确取前15个类别：+1分； | 根据数据 |  |  |
| 合计配分 | 14 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 2 | 合理列出潜在影响：+2分； |  |  |  |  |
| J2 | 2 | 合理描述服务提升：+2分； |  |  |  |  |
| J3 | 2 | 合理描述定价策略：+2分； |  |  |  |  |
| 合计配分 | 6 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：房屋挂牌智能系统运维

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机1台；   
（2）开发工具：Python编译工具；   
（3）标注工具：Microsoft Excel；   
（4）Word。

**2.**工作任务

为了更好地评估各地区的房地产市场表现，你需要对“各地区房屋挂牌数据.csv”进行分析。完成以下任务：  
（1）使用Excel加载“各地区房屋挂牌数据.csv”，查找所有记录中“地区”字段缺失的数据，并使用“未知”填充。输出填充后的数据，提供截图，以jpg的格式保存并命名为3.1.4-1.jpg。  
（2）补全3.1.4.ipynb代码完成以下统计分析：  
1、加载数据集“各地区房屋挂牌数据.csv”，输出加载后“地区”字段为“黄浦”的数据的头部信息。  
2、计算“城市”为“上海”的每个地区的平均“单价(每平方米)”，并按照从高到低排序。  
3、计算“城市”为“上海”的所有地区中平均“单价(每平方米)”最高的地区。同时分析该地区不同“装修风格”的平均单价分布，并柱状图的方式展示“平均单价”最高的前10组数据。  
将以上代码以及运行结果，以html格式保存并命名为3.1.4-2.html。  
（3）根据“城市”为“上海”地区的平均“单价(每平方米)”最高的地区的房产数据，分析消费者的购买偏好及其对房地产市场的潜在影响。讨论价格如何影响房地产的营销策略和定价策略。将答案填入word文件保存并命名为3.1.4-3.docx。

**3.**技能要求

（1）能够维护智能系统所需的知识库；   
（2）能维护智能系统所需的数据；   
（3）能为单一的产品找到合适的应用场景。

**4.**质量指标

（1）数据完整性和准确性；   
（2）数据分析的准确性；   
（3）输出结果可视化。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：房屋挂牌智能系统运维

考核时间：30min

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 | 4 | 正确填充1条缺失数据：+2分，总共4分； | 根据数据 |  |  |
| M2 | 1 | 正确加载数据集：+1分； | 根据数据 |  |  |
| M3 | 1 | 正确筛选地区是黄浦的数据集：+1分； | 根据数据 |  |  |
| M4 | 1 | 正确输出数据的头部信息：+1分； | 根据数据 |  |  |
| M5 | 1 | 正确筛选城市是上海的数据集：+1分； | 根据数据 |  |  |
| M6 | 1 | 正确计算上海每个区域的单价：+1分； | 根据数据 |  |  |
| M7 | 1 | 以上数据集按照从高到低排序输出：+1分； | 根据数据 |  |  |
| M8 | 1 | 正确筛选上海平均单价最高的区域：+1分； | 根据数据 |  |  |
| M9 | 2 | 正确计算该地区不同“装修风格”的平均单价分布：+1分； | 根据数据 |  |  |
| M10 | 1 | 以上数据取前10：+1分； | 根据数据 |  |  |
| M11 | 1 | 正确设置为柱图：+1分； | 根据数据 |  |  |
| M12 | 1 | 显示柱图：+1分； | 根据数据 |  |  |
| 合计配分 | 16 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 2 | 合理分析出消费者的购买偏好和对房地产市场的影响：+2分； |  |  |  |  |
| J2 | 2 | 合理探讨价格因素对房地产的营销策略和定价策略的影响：+2分； |  |  |  |  |
| 合计配分 | 4 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：电影院销售智能系统运维

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机1台；   
（2）开发工具：Python编译工具；   
（3）标注工具：Microsoft Excel；   
（4）Word。

**2.**工作任务

为了帮助电影院运营者更好地掌握各地区电影院的票房表现和观众行为，使用提供的“电影院销售数据.csv”进行以下任务：  
（1）使用Excel加载“电影院销售数据.csv”，查找所有记录中“电影院代码”字段缺失的数据，并使用“未知”填充。输出填充后的数据，提供截图，以jpg的格式保存并命名为3.1.5-1.jpg。  
（2）补全3.1.5.ipynb代码完成以下统计分析：  
1、加载数据集“电影院销售数据.csv”，并用print函数输出加载后“电影代码”为“33”的数据的头部信息。  
2、按照“电影院代码”计算不同电影院平均“总售票金额”和平均“售票数”，并按照平均“总售票金额”从高到低排序，输出平均“总售票金额”最高的前15条记录。  
3、计算“总售票金额”最高的电影院代码。统计平均“总售票金额”最高的电影院在不同“开场时间”的总售票金额分布，并以柱状图的方式展示。  
将以上代码以及运行结果，以html格式保存并命名为3.1.5-2.html。  
（3）综合评估平均“总售票金额”最高的电影院的销售数据，讨论票价、开场时间如何影响电影院的销售情况。后期如何在开场时间对电影的排片做优化。将答案填入word文件保存并命名为3.1.5-3.docx。

**3.**技能要求

（1）能够维护智能系统所需的知识库；   
（2）能维护智能系统所需的数据；   
（3）能为单一的产品找到合适的应用场景。

**4.**质量指标

（1）数据完整性和准确性；   
（2）数据分析的准确性；   
（3）输出结果可视化。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：电影院销售智能系统运维

考核时间：30min

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 | 4 | 正确填充1条缺失数据：+2分，总共4分； | 根据数据 |  |  |
| M2 | 1 | 正确加载数据集：+1分； | 根据数据 |  |  |
| M3 | 1 | 正确筛选电影代码是33的数据集：+1分； | 根据数据 |  |  |
| M4 | 1 | 输出以上数据集的头部信息：+1分； | 根据数据 |  |  |
| M5 | 2 | 正确筛选按照“电影院代码”分组，计算平均“总售票金额”和平均“售票数”：+2分； | 根据数据 |  |  |
| M6 | 1 | 以上数据集按照“总售票金额”从高到低排序：+1分； | 根据数据 |  |  |
| M7 | 1 | 以上数据集输出前15条：+1分； | 根据数据 |  |  |
| M8 | 1 | 筛选“总售票金额”最高的电影院代码：+1分； | 根据数据 |  |  |
| M9 | 1 | 获取以上电影院代码的数据：+1分； | 根据数据 |  |  |
| M10 | 2 | 正确筛选以上电影院按照“开场时间”分组，计算总售票金额：+2分； | 根据数据 |  |  |
| M11 | 1 | 输出柱状图：+1分； | 根据数据 |  |  |
| 合计配分 | 16 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 2 | 结合“总售票金额最高”的电影院销售数据，分析出开场时间、票价对电影院销售情况的影响：+2分； |  |  |  |  |
| J2 | 2 | 探讨如何合理利用开场时间拍片来增加电影院的收益：+2分； |  |  |  |  |
| 合计配分 | 4 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：超市智能销售系统优化

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机1台；   
（2）开发工具：Python编译工具；   
（3）标注工具：Microsoft Excel；   
（4）Word。

**2.**工作任务

为帮助财务人员了解A超市的经营状况，发掘当前的热度产品，以制定更好的经营策略。完成下列任务。  
（1）使用Excel，从“A超市的销售数据.csv”数据集中删减小类名称为“杂粮”、销售日期为“20150106”条件下的所有重复的记录，将去重后的数据截图，以jpg的格式保存并命名为3.2.1-1.jpg。  
（2）补全3.2.1.ipynb代码完成以下统计分析：  
1、加载数据集“A超市的销售数据.csv”打印结果。  
2、计算大类为“家居”条件下每个月的总销售金额，打印结果。  
3、生成一份大类为“家居”条件下“销售月份”为201501的日销售趋势图，以折线图的形式呈现。  
4、分析“家居”产品的销售趋势，给出家居产品的最佳表现月份。  
将以上代码以及运行结果，以html格式保存并命名为3.2.1-2.html。  
（3）现需要以“A超市的销售数据.csv”为训练数据集，使用一种AI模型构建销售趋势预测模型，请简述该预测模型部署的流程及优化方法，写入word文件保存并命名为3.2.1-3.docx。

**3.**技能要求

（1）能利用分析工具进行数据分析，输出分析报告；   
（2）能根据数据分析结论对智能产品的单一功能提出优化需求。

**4.**质量指标

（1）数据完整性和准确性；   
（2）数据分析的准确性；   
（3）输出结果的可视化的可解释度。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：超市智能销售系统优化

考核时间：30min

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 | 3 | 正确删减重复数据1条：+1分，总共3分； | 根据数据 |  |  |
| M2 | 1 | 正确加载数据集并输出：+1分； | 根据数据 |  |  |
| M3 | 2 | 正确筛选数据为“家居”大类：+1分； | 根据数据 |  |  |
| M4 | 2 | 正确给出“家居”月度销售总金额汇总：+2分； | 根据数据 |  |  |
| M5 | 2 | 正确筛选大类为“家居”且销售月份为201501的数据：+2分； | 根据数据 |  |  |
| M6 | 2 | 正确计算每天的销售金额：+2分； | 根据数据 |  |  |
| M7 | 1 | 正确绘制折线图：+1分； | 根据数据 |  |  |
| M8 | 1 | 正确显示折线图：+1分； | 根据数据 |  |  |
| M9 | 2 | 正确给出家居产品表现最好的月份：+2分； | 根据数据 |  |  |
| 合计配分 | 16 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 2 | 给出部署流程：+2分； |  |  |  |  |
| J2 | 2 | 给出至少一个优化方法：+2分； |  |  |  |  |
| 合计配分 | 4 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：气象分析智能系统优化

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机1台；   
（2）开发工具：Python编译工具；   
（3）标注工具：Microsoft Excel；   
（4）Word。

**2.**工作任务

气候变化对农业活动的影响显著，例如，极端高温可能导致作物减产，而降水模式的变化会影响灌溉需求和收成时间。为了帮助当地政府和居民更好地了解气候变化和其可能的影响，使用提供的“A地区的天气数据.csv”进行以下任务：。  
（1）使用Excel，筛选出所有“最高温度”超过90°F的记录，并检查这些记录中是否存在重复数据。找到并删除所有重复记录。将去重后的数据截图，以jpg的格式保存并命名为3.2.2-1.jpg。  
（2）补全3.2.2.ipynb代码完成以下统计分析：  
1、加载数据集“A地区的天气数据.csv”打印结果。  
2、计算2010年每个月的平均降水量，打印结果。  
3、以2010年的数据生成一个显示每月最低平均湿度变化的折线图。  
将以上代码以及运行结果，以html格式保存并命名为3.2.2-2.html。  
（3）利用历史天气数据构建一个简单的人工智能预测模型以预测下个月的平均降水量。选择适合的AI模型，描述模型的基本部署过程。根据数据分析的结论，提出针对智能产品的预测功能的优化方法。写入word文件保存并命名为3.2.2-3.docx。

**3.**技能要求

（1）能利用分析工具进行数据分析，输出分析报告；   
（2）能根据数据分析结论对智能产品的单一功能提出优化需求。

**4.**质量指标

（1）数据完整性和准确性；   
（2）数据分析的准确性；   
（3）输出结果的可视化的可解释度。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：气象分析智能系统优化

考核时间：30min

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 | 4 | 正确去重1条数据；+2分，总共4分； | 根据数据 |  |  |
| M2 | 1 | 正确加载数据集：+1分； | 根据数据 |  |  |
| M3 | 1 | 正确输出数据的头部信息：+1分； | 根据数据 |  |  |
| M4 | 2 | 正确格式化筛选日期：+1分； | 根据数据 |  |  |
| M5 | 2 | 正确筛选2010年：+1分 | 根据数据 |  |  |
| M6 | 2 | 正确计算2010年每个月的平均降水量：+2分； | 根据数据 |  |  |
| M7 | 2 | 正确计算2010年每个月的最低平均湿度：+2分； | 根据数据 |  |  |
| M8 | 1 | 正确绘制折线图x轴数据：+1分； | 根据数据 |  |  |
| M9 | 1 | 正确绘制折线图y轴数据：+1分； | 根据数据 |  |  |
| M10 | 1 | 正确显示折线图：+1分； | 根据数据 |  |  |
| 合计配分 | 17 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 1 | 给出一个机器学习模型：+1分； |  |  |  |  |
| J2 | 1 | 给出部署流程：+1分； |  |  |  |  |
| J3 | 1 | 给出至少一个优化方法：+1分； |  |  |  |  |
| 合计配分 | 3 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：餐饮智能系统优化

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机1台；   
（2）开发工具：Python编译工具；   
（3）标注工具：Microsoft Excel；   
（4）Word。

**2.**工作任务

为了帮助当地政府和商家更好地了解上海餐饮市场的表现和消费者偏好，使用提供的“上海餐饮数据.csv”进行以下任务：。  
（1）使用Excel，筛选出所有“人均消费”超过2000元的记录，并检查这些记录中是否存在重复数据。找到并删除所有重复记录。将去重后的数据截图，以jpg的格式保存并命名为3.2.3-1.jpg。  
（2）补全3.2.3.ipynb代码完成以下统计分析：  
1、加载数据集“上海餐饮数据.csv”，过滤掉“未知”行政区，打印结果。  
2、计算每个行政区的平均服务评分，按照从高到低的顺序排序。  
3、选择“类别”字段为“川菜”的数据，分析不同行政区的平均服务分汇总数据，以条形图的形式绘制。  
4、根据各个行政区的“人均消费数据”，分析出餐饮消费能力最高和最低的行政区。  
将以上代码以及运行结果，以html格式保存并命名为3.2.3-2.html。  
（3）如果要利用“上海餐饮数据”分析口味、环境、服务与人均消费的关系，选择一个适合的模型（如线性回归模型），并描述如何处理数据，部署模型，以及如何优化模型性能，将答案写入word文件保存并命名为3.2.3-3.docx。

**3.**技能要求

（1）能利用分析工具进行数据分析，输出分析报告   
（2）能根据数据分析结论对智能产品的单一功能提出优化需求

**4.**质量指标

（1）数据完整性和准确性   
（2）数据分析的准确性   
（3）输出结果的可视化的可解释度

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：餐饮智能系统优化

考核时间：30min

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 | 4 | 正确去重1条数据：+2分，总共4分； | 根据数据 |  |  |
| M2 | 1 | 正确加载数据集：+1分； | 根据数据 |  |  |
| M3 | 1 | 正确过滤未知区域：+1分； | 根据数据 |  |  |
| M4 | 2 | 计算每个行政区的平均服务评分：+2分 | 根据数据 |  |  |
| M5 | 1 | 正确筛选类别是四川的数据：+1分； | 根据数据 |  |  |
| M6 | 2 | 正确计算川菜在不同行政区的服务平均分：+2分； | 根据数据 |  |  |
| M7 | 1 | 正确设置图形为柱状图：+1分； | 根据数据 |  |  |
| M8 | 1 | 正确显示图像：+1分； | 根据数据 |  |  |
| M9 | 2 | 正确生成餐饮消费能力最高的行政区：+2分； | 根据数据 |  |  |
| M10 | 2 | 正确生成餐饮消费能力最低的行政区：+2分； | 根据数据 |  |  |
| 合计配分 | 17 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 1 | 合理描述选择的模型如何处理数据：+1分； |  |  |  |  |
| J2 | 1 | 合理描述选择的模型如何部署：+1分； |  |  |  |  |
| J3 | 1 | 合理描述选择的模型如何优化：+1分； |  |  |  |  |
| 合计配分 | 3 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：房屋挂牌智能系统优化

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机1台；   
（2）开发工具：Python编译工具；   
（3）标注工具：Microsoft Excel；   
（4）Word。

**2.**工作任务

为了帮助政府和商家更好地了解各个地区房地产市场的表现和消费者购房偏好，使用提供的“各地区房屋挂牌数据.csv”进行以下任务：。  
（1）使用Excel，筛选出所有“总价(元)”超过100,000,000元的记录，并检查这些记录中是否存在重复数据。找到并删除所有重复记录。将去重后的数据截图，以jpg的格式保存并命名为3.2.4-1.jpg。  
（2）补全3.2.4.ipynb代码完成以下统计分析：  
1、加载数据集“各地区房屋挂牌数据.csv”，打印结果。  
2、计算“城市”为“长沙”的每个地区的平均“单价(每平方米)”，按照从高到低的顺序排序。  
3、选择“房屋类型”为“3室1厅2卫”的数据，分析不同城市的平均单价数据，以柱状图的形式绘制。  
4、按照城市统计房屋单价的平均值，找出房屋平均单价最高的城市，同时找出该城市中房屋单价最高的地区。  
将以上代码以及运行结果，以html格式保存并命名为3.2.4-2.html。  
（3）如果要利用“各地区房屋挂牌数据.csv”分析同一个城市中面积、装修风格、朝向、楼层、单价和房屋总价的关系，选择一个适合的AI模型，并描述如何处理数据，部署模型，以及如何优化模型性能，将答案写入word文件保存并命名为3.2.4-3.docx。

**3.**技能要求

（1）能利用分析工具进行数据分析，输出分析报告；   
（2）能根据数据分析结论对智能产品的单一功能提出优化需求。

**4.**质量指标

（1）数据完整性和准确性；   
（2）数据分析的准确性；   
（3）输出结果的可视化的可解释度。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：房屋挂牌智能系统优化

考核时间：30min

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 | 4 | 正确去重1条数据：+2分，总共4分； | 根据数据 |  |  |
| M2 | 1 | 正确加载数据集：+1分； | 根据数据 |  |  |
| M3 | 1 | 正确筛选城市是长沙的数据集：+1分； | 根据数据 |  |  |
| M4 | 1 | 正确计算长沙每个地区的平均单价：+1分； | 根据数据 |  |  |
| M5 | 1 | 以上数据按照从高到低输出：+2分； | 根据数据 |  |  |
| M6 | 1 | 正确筛选房屋类型是3室1厅2卫的数据集：+1分； | 根据数据 |  |  |
| M7 | 2 | 对以上数据，筛选不同城市的平均单价数据：+2分； | 根据数据 |  |  |
| M8 | 1 | 图形设置成柱形：+1分； | 根据数据 |  |  |
| M9 | 1 | 显示柱状图：+1分； | 根据数据 |  |  |
| M10 | 2 | 筛选单价最高的城市：+2分； | 根据数据 |  |  |
| M11 | 2 | 筛选单价最高城市的区域：+2分； | 根据数据 |  |  |
| 合计配分 | 17 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 1 | 合理描述选择的模型如何处理数据：+1分； |  |  |  |  |
| J2 | 1 | 合理描述选择的模型如何部署：+1分； |  |  |  |  |
| J3 | 1 | 合理描述选择的模型如何优化：+1分； |  |  |  |  |
| 合计配分 | 3 | 合计得分 | | | |  |

考评员： 日期：

**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）

**人工智能训练师（四级）操作技能考核**

**试题单**

准考证号：

试题代码：

试题名称：电影院销售智能系统优化

考核时间：30min

**1.**场地设备要求

（1）人工智能训练师主机1台；   
（2）开发工具：Python编译工具；   
（3）标注工具：Microsoft Excel；   
（4）Word。

**2.**工作任务

为了帮助电影院运营者更好地掌握各地区电影院的票房表现和观众行为，使用提供的“电影票销售数据集.csv”进行以下任务：  
（1）使用Excel，筛选出所有“总销售金额”低于30,000元（包括30,000元）的记录，并检查这些记录中是否存在重复数据。找到并删除所有重复记录。将去重后的数据截图，以jpg的格式保存并命名为3.2.5-1.jpg。  
（2）补全3.2.5.ipynb代码完成以下统计分析：  
1、加载数据集“电影票销售数据集.csv”，打印结果。  
2、已知“电影院代码”与电影院一一对应，利用“电影院代码”计算每家电影院的平均“票价”，按照从高到低的顺序排序输出平均票价最高的前15组数据。  
3、选择“开场时间”为“9”的数据，分析不同电影院在“9”点的“总售票金额”，以柱状图的形式绘制“总售票金额”最高的前15组数据。  
将以上代码以及运行结果，以html格式保存并命名为3.2.5-2.html。  
（3）如果要利用“电影票销售数据集.csv”分析同一个城市中的电影类型、开场时间、票价与总售票金额的关系，选择一个适合的模型，并描述如何部署模型，以及如何优化模型性能，将答案写入word文件保存并命名为3.2.5-3.docx。

**3.**技能要求

（1）能利用分析工具进行数据分析，输出分析报告；   
（2）能根据数据分析结论对智能产品的单一功能提出优化需求。

**4.**质量指标

（1）数据完整性和准确性；   
（2）数据分析的准确性；   
（3）输出结果的可视化的可解释度。

**人工智能训练师（四级）操作技能考核**

**试题评分表及答案（主客观评分方式）**

准考证号：

试题代码：

试题名称：电影院销售智能系统优化

考核时间：30min

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 | 4 | 正确去重1条数据：+2分，总共4分； | 根据数据 |  |  |
| M2 | 1 | 正确加载数据集：+1分； | 根据数据 |  |  |
| M3 | 1 | 正确输出数据的头部信息：+1分； | 根据数据 |  |  |
| M4 | 2 | 正确计算每家电影院的平均票价：+2分； | 根据数据 |  |  |
| M5 | 1 | 以上数据集按照降序排序：+1分； | 根据数据 |  |  |
| M6 | 1 | 以上数据集取前10：+1分； | 根据数据 |  |  |
| M7 | 1 | 正确筛选开场为9的数据：+1分； | 根据数据 |  |  |
| M8 | 2 | 正确计算不同电影院在“9”点的“总售票金额”：+2分； | 根据数据 |  |  |
| M9 | 1 | 以上结果降序排序：+1分； | 根据数据 |  |  |
| M10 | 1 | 以上结果取前15：+1分； | 根据数据 |  |  |
| M11 | 1 | 图形设置为柱状图：+1分； | 根据数据 |  |  |
| M12 | 1 | 显示图形：+1分； | 根据数据 |  |  |
| 合计配分 | 17 | 合计得分 | |  |  |

考评员： 日期：

评价分评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 细则编号 | 配分 | 评分细则描述 | 考评员评分 | | | 得分 |
| 1 | 2 | 3 |
| J1 | 1 | 合理描述选择的模型如何处理数据：+1分； |  |  |  |  |
| J2 | 1 | 合理描述选择的模型如何部署：+1分； |  |  |  |  |
| J3 | 1 | 合理描述选择的模型如何优化：+1分； |  |  |  |  |
| 合计配分 | 3 | 合计得分 | | | |  |

考评员： 日期：

**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）