

第 4 部分

操作技能复习题

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：1.1.1

试题名称：自动化营销模型业务流程设计

考核时间：25 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

自动化营销可以利用技术代替重复性的人力劳动，规模化地降低营销运营成本和采购成本，提高投资回报率，在构建商业体系、提升商业价值中都占有重要的作用。A 公司作为一家电商平台的卖家，计划开发一个营销自动化模型。该模型基于用户数据（基本信息、订单信息、消费行为等），能够通过分析潜在用户的行为，刻画用户画像，为销售人员提供销售线索，对不同阶段的客户进行个性化营销。你将作为一名人工智能训练师，为自动化营销模型的训练提供可用的数据，成功训练出模型。

- (1) 假设你需要收集客户的行为数据和客户基本信息以支持自动化营销模型，你将如何设计整个业务数据采集流程？请给出详细步骤。（5 分）
- (2) 在自动化营销模型中，如何确保采集到的数据准确、一致和可用？请给出

你的务数据处理流程设计。(5分)

(3) 数据质量对于 AI 模型的性能至关重要。请描述你将如何实施数据审核流程以保证数据质量，包括使用哪些审核指标、采取哪些审核方法，以及如何处理审核结果。(5分)

3、技能要求

- (1) 能够结合人工智能技术要求和业务特征，设计整套业务数据采集流程
- (2) 能够结合人工智能技术要求和业务特征，设计整套业务数据处理流程
- (3) 能够结合人工智能技术要求和业务特征，设计整套业务数据审核流程

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	5分	采集流程给出类似参考答案中的 1 条：+1 分 采集流程给出类似参考答案中的 3 条及以上：+3 分 采集的数据种类给出类似参考答案中的 1 类：+1 分 采集的数据种类给出类似参考答案中的 2 类及以上： +2 分 加总得分在 0 到 5 分之间，超过 5 分计 5 分				
(2)	5分	正确给出类似参考答案中的 1 条：+2 分 正确给出类似参考答案中的 2 条：+4 分 正确给出类似参考答案中的 3 条及以上：+5 分				
(3)	5分	正确给出类似参考答案中的 1 条：+2 分 正确给出类似参考答案中的 2 条：+4 分 正确给出类似参考答案中的 3 条及以上：+5 分				

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：1.1.2

试题名称：自动化营销模型业务模块效果优化

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

自动化营销可以利用技术代替重复性的人力劳动，规模化地降低营销运营成本和采购成本，提高投资回报率，在构建商业体系、提升商业价值中都占有重要的作用。A 公司作为一家电商平台的卖家，计划开发一个营销自动化模型。该模型基于用户数据（基本信息、订单信息、消费行为等），能够通过分析潜在用户的行为，刻画用户画像，为销售人员提供销售线索，对不同阶段的客户进行个性化营销。你将作为一名人工智能训练师，为自动化营销模型的训练提供可用的数据，成功训练出模型。

(1) 在模型运行过程中，你可能会遇到各种问题，例如数据采集困难、模型准确度低等。请结合你的业务知识，识别可能的问题，并分析可能的原因。（4 分）

(2) 针对你在（1）中识别的问题，你会如何设计一套业务模块优化方案以提升模型的性能？请详细描述你的方案，包括优化的目标、方法，以及预期的结果。（6 分）

3、技能要求

- (1) 能够结合业务知识，识别业务流程中单一模块的问题
- (2) 能够结合人工智能技术设计业务模块优化方案并推动实现

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	4分	正确给出类似参考答案中的 1 条：+2 分 正确给出类似参考答案中的 2 条及以上：+4 分				
(2)	6分	给出类似优化方案 1 条：+2 分 给出类似优化方案 2 条及以上：+4 分 给出类似的优化后预期目标：+1 分				

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：1.2.1

试题名称：顾客评价情感识别数据处理规范制定

考核时间：30 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

在电商平台激烈竞争的背景下，除了提高商品质量，压低商品价格外，了解更多的消费者心声对于商家来说至关重要。了解消费者心声的一种方式就是针对消费者的购物评价数据进行内在信息的数据挖掘和分析。你作为一名人工智能训练师，将对买家的购物评价数据进行情感识别数据处理规范的制定。

(1) 基于标注后用于模型训练测试的需求和原始数据 (data.csv)，为购物评价数据制定情感识别数据清洗标准，并保存清洗后的数据为“data_cleaned.csv”。

(10 分)

(2) 业务需求为识别“积极(用数字 1 代表)”和“消极(用数字 0 代表)”两种情感标签。请基于标注后用于模型训练测试的需求，为购物评价数据制定情感识别标注规范。(5 分)

3、技能要求

- (1) 能够结合人工智能技术要求和业务特征，设计数据清洗和标注流程
- (2) 能够结合人工智能技术要求和业务特征，制定数据清洗和标注规范

4、质量指标

无

主观评分表/过程评分表

序号	配 分	评分细则描述	考评员评分			得 分
			1	2	3	
(1)	10 分	正确给出类似参考答案标准中的一条：+1 分 正确给出类似参考答案标准中的两条：+3 分 正确给出类似参考答案标准中的三条：+5 分 若给出一条及以上不合理的数据清洗标准：-1 分 需要删除的数据每删除一条：+0.1 分 不需要删除的数据每删除一条：-0.1 分 加总得分在 0 到 10 分之间。				
(2)	5 分	正确给出标注内容与范围：+0.5 分 正确给出标注格式：+0.5 分 正确给出数据交付格式：+0.5 分 积极的标注标准，每给出 4 条常见表述中类似的 1 条 得 0.5 分，最多得 1 分 消极的标注标准，每给出 4 条常见表述中类似的 1 条 得 0.5 分，最多得 1 分 正确给出特殊情况反馈途径：+1 分 加总得分在 0 到 5 分之间。				

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：1.2.2

试题名称：顾客评价情感识别算法测试

考核时间：60 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

在电商平台激烈竞争的背景下，除了提高商品质量，压低商品价格外，了解更多的消费者心声对于商家来说至关重要。了解消费者心声的一种方式就是针对消费者的购物评价数据进行内在信息的数据挖掘和分析。你作为一名人工智能训练师，将对买家的购物评价数据进行简单的情感识别算法训练及测试。

请先打开 环境准备.ipynb，运行代码。

(1)请打开“data_split.ipynb”，按照要求补齐相关代码，基于“raw_data.csv”对训练集（保存为 train.csv）和测试集（保存为 test.csv）进行划分。（2分）

(2)请打开“train_test.ipynb”，按照要求补齐相关代码，基于已划分好的训练集和测试集对模型进行训练。（2分）

(3)请打开“train_test.ipynb”，按照要求补齐相关代码，使用“prediction_test.csv”数据集对模型进行测试，输出模型性能指标及混淆矩阵。（4分）

(4)请对测试结果进行分析，并编写测试报告。（6分）

(5)请找出至少一个错误案例，分析错误产生的原因并指出可能的纠正方法。（6分）

3、技能要求

- (1) 能够维护日常训练集与测试集
- (2) 能使用工具对算法进行训练
- (3) 能够使用测试工具对人工智能产品的使用进行测试
- (4) 能够对测试结果进行分析，编写测试报告
- (5) 能够运用工具，分析算法中错误案例产生的原因并进行纠正

4、质量指标

- (1) 代码书写规范
- (2) 代码正常运行

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
(1)	2分	正确填写一处代码：+1分 汇总得分在0到2分之间		
(2)	2分	正确填写一处代码：+1分 汇总得分在0到2分之间		
(3)	4分	正确填写一处代码：+1分 正确输出模型性能指标及混淆矩阵：+2分 汇总得分在0到4分之间		

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(4)	6分	对模型指标做出类似参考答案一条至多条合理分析：+4分 对混淆矩阵做出类似参考答案分析说明：+2分 得分为6分、4分、2分或0分				
(5)	6分	正确指出至少一个错误案例：+2分 正确给出可能原因分析及纠正方案：+4分 加总得分为6分、4分、2分或0分				

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：1.3.1

试题名称：客户体验管理智能系统监控和优化

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

客户体验管理（Customer Experience Management, CEM）智能系统是一种集成的解决方案，可帮助企业跟踪、分析和管理工作与客户的所有交互，从而提高客户体验和满意度。乐言科技的 CEM 产品是一款面向电商客服主管及质检团队的产品。你作为一名人工智能训练师，将参考该系统的操作手册（CEM 操作手册.pdf），对该智能系统使用的数据进行全面分析，并输出分析报告，基于分析提出优化需求，并为其应用设计智能解决方案。

- (1) 请对该智能产品的评价分析数据（评价分析.xlsx）进行分析，并输出分析报告。（4分）
- (2) 请对该智能产品提出优化需求。（4分）
- (3) 请基于提出的优化需求，为该智能产品的应用设计智能解决方案。（7分）

3、技能要求

- (1) 能够对单一智能产品使用的数据进行全面分析，输出分析报告
- (2) 能够对单一智能产品提出优化需求
- (3) 能够为单一智能产品的应用设计智能解决方案

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	4分	正确给出标签统计或占比最高的标签说明：+2分 正确给出类似参考答案分析中的一条：+2分 正确给出类似参考答案分析中的两条及以上：+4分 给出错误分析：-2分 加总得分在0到4分之间				
(2)	4分	正确给出类似参考答案优化需求中的一条：+2分 正确给出类似参考答案优化需求中的两条及以上：+4分 给出错误优化需求：-2分 加总得分在0到4分之间				
(3)	7分	正确给出类似参考答案中的解决方案：+7分 得分为7分或0分				

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：1.3.2

试题名称：客户体验管理智能系统人机交互流程设计

考核时间：15 min

1、场地设备要求

(1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）

(2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

客户体验管理（Customer Experience Management, CEM）智能系统是一种集成的解决方案，可帮助企业跟踪、分析和管理客户与公司的所有交互，从而提高客户体验和满意度。乐言科技的 CEM 产品是一款面向电商客服主管及质检团队的产品。你作为一名人工智能训练师，将参考该系统的操作手册（CEM 操作手册.pdf），为该智能系统设计人工与智能交互的最优方式及流程。

(1) 请为该智能产品找到人工与智能交互的最优方式。（10 分）

(2) 请为该智能产品设计人工与智能交互的最优流程。（10 分）

3、技能要求

(1) 能够通过数据分析，找到单一场景下人工和智能交互的最优方式

(2) 能够通过数据分析，设计单一场景下人工和智能交互的最优流程

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	

(1)	10分	提出类似参考答案的最优方式: +10分 得分为10分或0分				
(2)	10分	提出类似参考答案的最优流程: +10分 得分为10分或0分				

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：1.4.1

试题名称：数据采集和处理培训

考核时间：10 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

- (1) 请提供数据采集和处理培训讲义的目录，并对培训内容做简单描述。（1分）
- (2) 请针对数据采集和处理培训，从下述方法中选择合适的培训方法，并简述原因：

讲授法、演示法、研讨法、视听法、角色扮演法（2分）

3、技能要求

- (1) 能够编写初级培训讲义
- (2) 能够对五级/初级工、四级/中级工开展知识和技术培训

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	1分	目录设计合理，循序渐进由浅入深：+1分 得分为1分或0分				
(2)	2分	正确选择一到多种培训方法并给出原因（可与参考答案不同）：+2分 得分为2分或0分				

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：1.4.2

试题名称：数据采集和处理指导

考核时间：10 min

1、场地设备要求

(1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）

(2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

(1) 请简述数据采集和处理的常见问题的解决方案内容（至少提供教材中五个中的一个部分名称及简要说明）。（1 分）

(2) 请简述数据标注流程中可指导的常见问题指导方向（至少提供教材中七个中的一个部分名称及简要说明）。（1 分）

3、技能要求

(1) 能够指导五级/初级工、四级/中级工解决数据采集、处理问题

(2) 能够指导五级/初级工、四级/中级工解决数据标注问题

4、质量指标

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	1分	正确提供至少一个名称及简要说明：+1 分 错误不扣分 得分为 1 分或 0 分				
(2)	1分	正确提供至少一个名称及简要说明：+1 分 错误不扣分 得分为 1 分或 0 分				

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：2.1.1

试题名称：商品推荐系统业务流程设计

考核时间：25 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

在当前竞争激烈的电商环境中，提供个性化的用户体验已成为提升用户满意度和保持用户忠诚度的关键。A 公司作为一家电商企业计划开发一个商品推荐系统。该系统旨在通过精准推荐，优化用户的购物体验，提升用户满意度和购买转化率，最终提高公司销售额。该商品推荐系统将利用大数据和人工智能技术，根据用户的历史购买记录、浏览行为、用户属性等信息，向用户推荐他们可能感兴趣的商品。你将作为一名人工智能训练师为商品推荐模型的训练提供可用的数据，成功训练出模型。

(1) 请根据用户行为数据和商品信息，设计一套完整的业务数据采集流程，以便为个性化推荐提供支持。（5 分）

(2) 在电商推荐系统中，数据处理是非常重要的一环。请设计一套完整的业务数据处理流程，请详细描述您的流程设计。（5 分）

(3) 为了保证数据质量，数据审核是不可缺少的一步。请设计一套完整的业务数据审核流程，以便对处理后的数据进行质量控制和异常检测。请详细描述流程设计，包括但不限于审核的指标、方法以及处理审核结果的策略。（5 分）

3、技能要求

- (1) 能够结合人工智能技术要求和业务特征，设计整套业务数据采集流程
- (2) 能够结合人工智能技术要求和业务特征，设计整套业务数据处理流程
- (3) 能够结合人工智能技术要求和业务特征，设计整套业务数据审核流程

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	5分	数据需求和数据来源明确合理：+1分 数据收集方法的选择合理，考虑到了数据的实时性和完整性：+1分 数据存储的设计考虑到了数据的安全性和便利性：+1分 对数据更新机制的描述合理，考虑到了数据的实时性：+1分 确定数据需求、数据来源、数据采集方式、数据存储、数据更新机制等流程内容合理，涉及4个：+5分				
(2)	5分	数据清洗方法的选择合理，能够有效处理缺失值、异常值、重复值等问题：+2分 数据转换设计合理，满足模型的需求：+2分 特征工程的设计考虑到了业务的特性和需求：+2分 数据切分方法的选择合理，适用于模型的训练和评估：+2分 涉及到3个流程：+5分				
(3)	5分	审核指标的设置合理：+1分 审核方法的设计科学：+1分 异常处理策略的设计合理，能够有效处理数据质量问题：+1分 考虑到了审核结果的反馈和使用，为数据采集和处理的改进提供支持：+1分 涉及4个步骤：+5分				

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：2.1.2

试题名称：商品推荐系统业务模块效果优化

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

在当前竞争激烈的电商环境中，提供个性化的用户体验已成为提升用户满意度和保持用户忠诚度的关键。A 公司作为一家电商企业计划开发一个商品推荐系统。该系统旨在通过精准推荐，优化用户的购物体验，提升用户满意度和购买转化率，最终提高公司销售额。该商品推荐系统将利用大数据和人工智能技术，根据用户的历史购买记录、浏览行为、用户属性等信息，向用户推荐他们可能感兴趣的物品。你将作为一名人工智能训练师为商品推荐模型的训练提供可用的数据，成功训练出模型。

- (1) 在实际运行中，推荐系统可能会遇到各种问题，例如用户对推送的商品反馈较少、推荐的准确率较低等。请针对识别可能存在的问题，分析其可能的原因（4 分）
- (2) 针对在（1）中识别的问题，设计一套业务模块优化方案，以提高电商推荐系统的性能。请详细描述您的方案设计，包括但不限于优化的目标、优化的方法以及预期的结果等。（6 分）

3、技能要求

- (1) 能够结合业务知识，识别业务流程中单一模块的问题
- (2) 能够结合人工智能技术设计业务模块优化方案并推动实现

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	4分	可能存在的问题及原因合理即可，每个+2分				
(2)	6分	优化目标的设定明确、合理：+1分 优化方法的设计科学，考虑到了数据、模型和用户反馈等多方面的因素，一个优化方法+2分，两个优化方法+3分 对优化方法的实施有明确的计划，如定期进行实验和评估等：+1分 对预期结果的描述准确，反映了优化的效果：+1分				

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：2.2.1

试题名称：新闻文本分类数据处理规范制定

考核时间：30 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

新闻文本分类是指将大量的新闻文本按照其内容进行自动分类，使得用户可以更快速地查找到自己所需要的新闻信息。目前新闻文本分类已被广泛应用于新闻门户网站、社交媒体、搜索引擎等领域，成为提高用户粘性和体验的一项重要技术。本项目中，你作为一名人工智能训练师，将进行新闻文本分类数据的清洗与标注。

(1) 请根据项目背景描述，为文件(data.csv，数据来源今日头条)中提供的新闻文本分类数据进行清洗，并制定相应的清洗标准，保存清洗后的数据为(dataCleaned.csv)。(10分)

(2) 请基于标注后用于模型训练测试的需求，从以下内容方面为新闻文本分类数据制定分类标注流程及规范。(5分)

标注需求：从内容方面标注 15 种新闻分类标签：民生故事(news_story)、文化(news_culture)、娱乐(news_entertainment)、体育(news_sports)、财经(news_finance)、房产(news_house)、汽车(news_car)、教育(news_edu)、科技(news_tech)、军事(news_military)、旅游(news_travel)、国际(news_world)、证券(stock)、农业(news_agriculture)、游戏(news_game)

3、技能要求

(1) 能够结合人工智能技术要求和业务特征，设计数据清洗和标注流程

(2) 能够结合人工智能技术要求和业务特征，制定数据清洗和标注规范

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	10分	清洗标准占5分： 正确给出参考答案标准中的一条：+2分 正确给出参考答案标准中的三条及以上：+5分 若给出一条及以上不合理的数据清洗标准：-2分 加总得分在0到5分之间。 正确保存清洗后的数据占5分： 保存文件格式正确：+1分 保存文件格式不正确：-1分 正确删除特殊字符■：+1分 正确删除纯英文，纯数字，连续标点符号的文本：+1分 正确删除空值：+1分 正确删除24条重复数据：+1分 加总得分在0到5分之间。				
(2)	5分	正确给出标注内容与范围：+1分 正确给出标注规则1：+1分 正确给出标注规则2或规则3中任一点：+1分 正确给出特殊情况反馈途径：+1分 正确给出标注与交付格式：+1分 加总得分在0到5分之间。				

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：2.2.2

试题名称：新闻文本分类算法测试

考核时间：60 min

1、场地设备要求

(1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）

(2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

新闻文本分类是指将大量的新闻文本按照其内容进行自动分类，使得用户可以更快速地查找到自己所需要的新闻信息。目前新闻文本分类已被广泛应用于新闻门户网站、社交媒体、搜索引擎等领域，成为提高用户粘性和体验的一项重要技术。本项目中，你作为一名人工智能训练师，将对新闻文本数据分类任务建模进行训练及测试验证。

请先打开 环境准备.ipynb，运行代码。

(1) 请打开 data_split.ipynb，按照要求补齐相关代码，基于 raw_data.csv 对训练集（保存为 train.csv）、测试集（保存为 test.csv）和验证集（保存为 dev.csv）进行划分。（2 分）

(2) 请打开“train_test.ipynb”，按照要求补齐相关代码，基于已划分好的训练集、测试集和验证集对模型进行训练。（2 分）

(3) 请打开“train_test.ipynb”，按照要求补齐相关代码，使用“prediction_test.csv”数据集对模型进行测试，输出模型性能指标及混淆矩阵。（4 分）

(4) 请对测试结果进行分析，并编写测试报告。（6 分）

(5) 请找出至少一个错误案例，分析错误产生的原因并指出可能的纠正方法。

(6分)

3、技能要求

- (1) 能够维护日常训练集与测试集
- (2) 能使用工具对算法进行训练
- (3) 能够使用测试工具对人工智能产品的使用进行测试
- (4) 能够对测试结果进行分析，编写测试报告
- (5) 能够运用工具，分析算法中错误案例产生的原因并进行纠正

4、质量指标

- (1) 代码书写规范
- (2) 代码正常运行

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
(1)	2分	正确填写一处代码：+1分 汇总得分在0到2分之间		
(2)	2分	正确填写一处代码：+1分 汇总得分在0到2分之间		
(3)	4分	正确填写一处代码：+1分 正确输出模型性能指标及混淆矩阵：+2分 汇总得分在0到4分之间		

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(4)	6分	对模型指标做出类似参考答案一条至多条合理分析：+4分 对混淆矩阵做出类似参考答案分析说明：+2分 得分为6分、4分、2分或0分				
(5)	6分	正确指出至少一个错误案例：+2分 正确给出可能原因分析及纠正方案：+4分 加总得分为6分、4分、2分或0分				

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：2.3.1

试题名称：智能回评系统监控和优化

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

在电商领域，评价是顾客与商家沟通及反馈的一个重要模块，订单评价的维护也是客服或运营的重要工作之一，因此部分商家要求客服/运营对买家针对订单的评价进行回评，并进行回评话术模版的持续维护。传统的回评方式需要客服人工逐条查看买家评价，并针对性选择模版中合适的回评话术复制粘贴发送。乐言科技基于大语言模型（LLM）打造了智能回评产品，利用 AI 文本生成能力进行回评话术智能生成。本项目中，你作为人工智能训练师，将对该智能产品进行分析和优化设计。请阅读产品操作手册（AI 智能回评.pdf）后进行以下试题的作答。

- (1) 请对该智能产品的回评生成数据（智能回评生成.csv）进行分析，并输出分析报告。（4 分）
- (2) 请对该智能产品提出优化需求。（4 分）
- (3) 请基于提出的优化需求，为该智能产品的应用设计智能解决方案。（7 分）

3、技能要求

- (1) 能够对单一智能产品使用的数据进行全面分析，输出分析报告
- (2) 能够对单一智能产品提出优化需求
- (3) 能够为单一智能产品的应用设计智能解决方案

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	4分	给出话术风格模版统计数据：+1分 指出同一话术风格模版被用于多评价类型的回评：+2分 指出相同评价生成回评不同：+2分 指出回评话术生成错误：+2分 指出回评话术长度问题：+2分 错误描述不扣分 汇总得分为0到4分之间，超过4分的得分为4分				
(2)	4分	正确描述类似参考答案优化需求中的一至多条：+4分 错误描述不扣分 汇总得分为4分或0分				
(3)	7分	针对优化需求设计出可行的智能解决方案：+7分 得分为7分或0分				

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：2.3.2

试题名称：智能回评系统人机交互流程设计

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

在电商领域，评价是顾客与商家沟通及反馈的一个重要模块，订单评价的维护也是客服或运营的重要工作之一，因此部分商家要求客服/运营对买家针对订单的评价进行回评，并进行回评话术模版的持续维护。传统的回评方式需要客服人工逐条查看买家评价，并针对性选择模版中合适的回评话术复制粘贴发送。乐言科技基于大语言模型（LLM）打造了智能回评产品，利用 AI 文本生成能力进行回评话术智能生成。本项目中，你作为人工智能训练师，将对该智能产品进行分析和优化设计。请阅读产品操作手册（AI 智能回评.pdf）后进行以下试题的作答。

- (1) 请为该智能产品找到人工与智能交互的最优方式。（10 分）
- (2) 请为该智能产品设计人工与智能交互的最优流程。（10 分）

3、技能要求

- (1) 能够通过数据分析，找到单一场景下人工和智能交互的最优方式
- (2) 能够通过数据分析，设计单一场景下人工和智能交互的最优流程

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	10分	提出类似参考答案的最优方式：+10分 得分为10分或0分				
(2)	10分	提出类似参考答案的最优流程：+10分 得分为10分或0分				

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：2.4.1

试题名称：数据清洗培训

考核时间：10 min

1、场地设备要求

(1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）

(2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

(1) 请提供数据清洗培训讲义的目录，并对培训内容做简单描述。（1分）

(2) 请针对数据清洗培训，从下述方法中选择合适的培训方法，并简述原因：
讲授法、演示法、研讨法、视听法、角色扮演法（2分）

3、技能要求

(1) 能够编写初级培训讲义

(2) 能够对五级/初级工、四级/中级工开展知识和技术培训

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	1分	目录设计合理，循序渐进由浅入深：+1分 得分为1分或0分				
(2)	2分	正确选择一到多种培训方法并给出原因（可与参考答案不同）：+2分 得分为2分或0分				

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：2.4.2

试题名称：数据采集和处理指导

考核时间：10 min

1、场地设备要求

(1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）

(2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

(1) 请简述数据采集和处理的常见问题的解决方案内容（至少提供教材中五个中的一个部分名称及简要说明）。（1 分）

(2) 请简述数据标注流程中可指导的常见问题指导方向（至少提供教材中七个中的一个部分名称及简要说明）。（1 分）

3、技能要求

(1) 能够指导五级/初级工、四级/中级工解决数据采集、处理问题

(2) 能够指导五级/初级工、四级/中级工解决数据标注问题

4、质量指标

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	1分	正确提供至少一个名称及简要说明：+1 分 错误不扣分 得分为 1 分或 0 分				
(2)	1分	正确提供至少一个名称及简要说明：+1 分 错误不扣分 得分为 1 分或 0 分				

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：3.1.1

试题名称：政务服务智能问答业务流程设计

考核时间：25 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

智能问答服务是“互联网+政务服务”技术体系建设的要求之一，也是政府网站及新媒体考核要求之一。随着移动互联网的发展，政务服务从线下向线上的转移带来巨大的咨询服务量。传统的政府网上公开的政务办事信息，多按政府部门、专业的政务办事术语等为分类制定检索条件，对普通群众查询办事信息形成认知门槛，人们无法根据自己的需求快速找到相应的政务办事项目和信息。A 公司是一家电子政务领域的科技公司，计划开发一套政务服务智能问答模型，根据政务服务事项的相关信息，解答用户关于办理政务服务的相关问题。

你将作为一名人工智能训练师，为政务服务智能问答提供可用的数据，成功训练出模型。

- (1) 假设你需要设计政务服务智能问答模型，为了构建政务服务问答知识库，需要采集什么样的数据？请列出采集数据的种类和要求。（5 分）
- (2) 为了构建政务服务问答知识库，可以通过哪些采集的方法获取数据？（5 分）
- (3) 数据质量对于 AI 模型的性能至关重要。当清洗数据时，发现疑似重复数据该怎么处理？（5 分）

3、技能要求

- (1) 能够结合人工智能技术要求和业务特征，设计整套业务数据采集流程
- (2) 能够结合人工智能技术要求和业务特征，设计整套业务数据处理流程
- (3) 能够结合人工智能技术要求和业务特征，设计整套业务数据审核流程

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	5分	采集数据的种类给出类似参考答案中的 2 个：+1 分 采集数据的种类给出类似参考答案中的 4 个：+2 分 采集数据的种类给出类似参考答案中的 6 个及以上：+3 分 采集数据的要求给出类似参考答案中的 1 种：+1 分 采集数据的要求给出类似参考答案中的 2 种及以上：+2 分				
(2)	5分	给出类似参考答案中的 1 种采集方法：+2 分 给出类似参考答案中的 2 种采集方法：+4 分 给出类似参考答案中的 3 种及以上采集方法：+5 分				
(3)	5分	给出类似参考答案中的 1 个：+2 分 给出类似参考答案中的 2 个：+4 分 给出类似参考答案中的 3 个及以上：+5 分				

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：3.1.2

试题名称：政务服务智能问答业务模块效果优化

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

智能问答服务是“互联网+政务服务”技术体系建设的要求之一，也是政府网站及新媒体考核要求之一。随着移动互联网的发展，政务服务从线下向线上的转移带来巨大的咨询服务量。传统的政府网上公开的政务办事信息，多按政府部门、专业的政务办事术语等为分类制定检索条件，对普通群众查询办事信息形成认知门槛，人们无法根据自己的需求快速找到相应的政务办事项目和信息。A 公司是一家电子政务领域的科技公司，计划开发一套政务服务智能问答模型，根据政务服务事项的相关信息，解答用户关于办理政务服务的相关问题。

你将作为一名人工智能训练师，为政务服务智能问答提供可用的数据，成功训练出模型。

- (1) 用户通过政务服务咨询平台咨询问题时，咨询了关于 A 事项的问题，但系统未识别该问题，可能的原因有哪些？（4 分）
- (2) 哪些情况下需要对数据库进行优化，请结合你的业务知识，识别可能的问题，并分析可能的原因？（6 分）

3、技能要求

- (1) 能够结合业务知识，识别业务流程中单一模块的问题
- (2) 能够结合人工智能技术设计业务模块优化方案并推动实现

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	4分	给出类似参考答案中的 1 个原因：+2 分 给出类似参考答案中的 2 个及以上原因：+4 分				
(2)	6分	给出类似参考答案中的 1 个：+2 分 给出类似参考答案中的 2 个：+4 分 给出类似参考答案中的 3 个及以上：+6 分				

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：3.2.1

试题名称：命名实体识别数据处理规范制定

考核时间：30 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

命名实体识别（Named Entity Recognition，简称 NER）（也称为实体识别、实体分块和实体提取）是 NLP 中一项非常基础的任务，旨在将文本中的命名实体定位并分类为预先定义类别，如人员、组织、位置、时间表达式、数量、货币值、百分比等。你作为一名人工智能训练师，将进行命名实体识别任务的数据准备。

(1) 基于下述实体识别数据集格式描述和原始数据“data.tsv”，为命名实体识别数据制定数据清洗标准，并保存清洗后的待标注数据为“data_cleaned.tsv”。

(10 分)

input format:

Input format (prefer BIO tag scheme), with each character its label for one line. Sentences are split with a null line.

Example:

美 B-LOC
国 I-LOC
的 O
华 B-NAME
莱 I-NAME
士 I-NAME

我 O
跟 O
他 O

(2) 基于下述标注需求和 2.1.1 中的格式示例，为命名实体识别数据制定完整的标注规范。(5 分)

标注数据需求:

按 BIO schema 标注 8 种实体: 国籍 (CONT), 学历学位 (EDU), 地区 (LOC), 人名 (NAME), 组织 (ORG), 专业 (PRO), 民族 (RACE), 职位头衔 (TITLE)。

3、技能要求

(1) 能够结合人工智能技术要求和业务特征，设计数据清洗和标注流程

(2) 能够结合人工智能技术要求和业务特征，制定数据清洗和标注规范

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	10分	正确给出类似参考答案标准中的一条: +2 分 正确给出类似参考答案标准中的两条: +4 分 若给出一条及以上不合理的数据清洗标准: -1 分 data_cleaned. tsv 文件数据完整且中每行存在字符不多于 1 个: +3 分 data_cleaned. tsv 文件数据完整且中不存在连续空行: +3 分 加总得分在 0 到 10 分之间。				

(2)	5分	正确给出标注内容与范围：+0.5分 正确给出标注格式：+0.3分 正确给出数据交付格式：+0.3分 标注规则每1个标签给出完整定义：+0.2分 正确给出特殊情况反馈途径：+0.5分 加总得分在0到5分之间。				
-----	----	--	--	--	--	--

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：3.2.2

试题名称：命名实体识别算法测试

考核时间：60 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

命名实体识别（Named Entity Recognition，简称 NER）（也称为实体识别、实体分块和实体提取）是 NLP 中一项非常基础的任务，旨在将文本中的命名实体定位并分类为预先定义的类别，如人员、组织、位置、时间表达式、数量、货币值、百分比等。你作为一名人工智能训练师，将进行命名实体识别任务的模型训练，并对结果进行测试与分析。

请先打开 环境准备.ipynb，运行代码。

(1) 请打开 “data_split.ipynb”，按照要求补齐相关代码，基于 “raw.tsv” 对训练集（保存为 train.tsv）、测试集（保存为 test.tsv）和验证集（保存为 dev.tsv）进行划分。（2 分）

(2) 请打开 “train_test.ipynb”，按照要求补齐相关代码，基于已划分好的训练集、测试集和验证集对模型进行训练。（2 分）

(3) 请打开 “train_test.ipynb”，按照要求补齐相关代码，使用 “prediction_test.tsv” 数据集对模型进行测试，输出模型性能指标及混淆矩阵。（4 分）

(4) 请对测试结果进行分析，并编写测试报告。（6 分）

(5) 请找出至少一个错误案例，分析错误产生的原因并指出可能的纠正方法。

(6分)

3、技能要求

- (1) 能够维护日常训练集与测试集
- (2) 能使用工具对算法进行训练
- (3) 能够使用测试工具对人工智能产品的使用进行测试
- (4) 能够对测试结果进行分析，编写测试报告
- (5) 能够运用工具，分析算法中错误案例产生的原因并进行纠正

4、质量指标

- (1) 代码书写规范
- (2) 代码正常运行

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
(1)	2分	正确填写一处代码：+1分 汇总得分在0到2分之间		
(2)	2分	正确填写一处代码：+1分 汇总得分在0到2分之间		
(3)	4分	正确填写一处代码：+1分 正确输出模型性能指标及混淆矩阵：+2分 汇总得分在0到4分之间		

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(4)	6分	对模型指标做出类似参考答案一条至多条合理分析：+4分 对混淆矩阵做出类似参考答案分析说明：+2分 得分为6分、4分、2分或0分				
(5)	6分	正确指出至少一个错误案例：+2分 正确给出可能原因分析及纠正方案：+4分 加总得分为6分、4分、2分或0分				

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：3.3.1

试题名称：智能卖点生成系统监控和优化

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

在电商领域，当消费者进店、进行商品咨询或主动要求推荐时，客服人员通常会介绍商品卖点，以期望能打动消费者促成下单购买。因此，商品卖点的文案编写维护是客服/营销的重要工作之一。传统的卖点文案编写费时费力，且非常考验客服/营销人员的文案编写能力。为帮助降低商品卖点文案编写难度，乐言科技基于大语言模型（LLM）打造了智能卖点生成产品，利用 AI 文本生成能力进行商品卖点话术的智能生成。本项目中，你作为人工智能训练师，将对该智能产品进行分析和优化设计。请阅读产品操作手册“AI 智能卖点生成.pdf”后进行以下试题的作答。

- (1) 请对该智能产品的智能卖点生成数据“智能卖点生成.csv”进行分析，并输出分析报告。（4分）
- (2) 请对该智能产品提出优化需求。（4分）
- (3) 请基于提出的优化需求，为该智能产品的应用设计智能解决方案。（7分）

3、技能要求

- (1) 能够对单一智能产品使用的数据进行全面分析，输出分析报告
- (2) 能够对单一智能产品提出优化需求

(3) 能够为单一智能产品的应用设计智能解决方案

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	4分	给出话术风格模版统计数据：+1分 指出同一话术风格模版被用于多家店铺：+2分 指出话术风格模版中的商品与被应用于进行卖点生成的商品可以有很大差异：+1分 指出卖点话术风格继承问题：+2分 错误描述不扣分 加总得分为0到4分之间，超过4分的得分为4分				
(2)	4分	正确描述类似参考答案优化需求中的一至多条：+4分 错误描述不扣分 汇总得分为4分或0分				
(3)	7分	针对优化需求设计出可行的智能解决方案：+7分 得分为7分或0分				

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：3.3.2

试题名称：智能卖点生成系统人机交互流程设计

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

在电商领域，当消费者进店、进行商品咨询或主动要求推荐时，客服人员通常会介绍商品卖点，以期望能打动消费者促成下单购买。因此，商品卖点的文案编写维护是客服/营销的重要工作之一。传统的卖点文案编写费时费力，且非常考验客服/营销人员的文案编写能力。为帮助降低商品卖点文案编写难度，乐言科技基于大语言模型（LLM）打造了智能卖点生成产品，利用 AI 文本生成能力进行商品卖点话术的智能生成。本项目中，你作为人工智能训练师，将对该智能产品进行分析和优化设计。请阅读产品操作手册“AI 智能卖点生成.pdf”后进行以下试题的作答。

- (1) 请为该智能产品找到人工与智能交互的最优方式。（10 分）
- (2) 请为该智能产品设计人工与智能交互的最优流程。（10 分）

3、技能要求

- (1) 能够通过数据分析，找到单一场景下人工和智能交互的最优方式
- (2) 能够通过数据分析，设计单一场景下人工和智能交互的最优流程

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	10分	提出类似参考答案的最优方式：+10分 得分为10分或0分				
(2)	10分	提出类似参考答案的最优流程：+10分 得分为10分或0分				

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：3.4.1

试题名称：数据标注培训

考核时间：10 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

- (1) 请提供数据标注培训讲义的目录，并对培训内容做简单描述。（1 分）
- (2) 请针对数据标注培训，从下述方法中选择合适的培训方法，并简述原因：
讲授法、演示法、研讨法、视听法、角色扮演法（2 分）

3、技能要求

- (1) 能够编写初级培训讲义
- (2) 能够对五级/初级工、四级/中级工开展知识和技术培训

4、质量指标

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	1分	目录设计合理，循序渐进由浅入深：+1 分 得分为 1 分或 0 分				
(2)	2分	正确选择一到多种培训方法并给出原因（可与参考答案不同）：+2 分 得分为 2 分或 0 分				

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：3.4.2

试题名称：数据采集和处理指导

考核时间：10 min

1、场地设备要求

(1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）

(2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

(1) 请简述数据采集和处理的常见问题的解决方案内容（至少提供教材中五个中的一个部分名称及简要说明）。（1 分）

(2) 请简述数据标注流程中可指导的常见问题指导方向（至少提供教材中七个中的一个部分名称及简要说明）。（1 分）

3、技能要求

(1) 能够指导五级/初级工、四级/中级工解决数据采集、处理问题

(2) 能够指导五级/初级工、四级/中级工解决数据标注问题

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	1分	正确提供至少一个名称及简要说明：+1 分 错误不扣分 得分为 1 分或 0 分				
(2)	1分	正确提供至少一个名称及简要说明：+1 分 错误不扣分				

	得分为 1 分或 0 分				
--	--------------	--	--	--	--

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：4.1.1

试题名称：便携式智能取色仪业务流程设计

考核时间：25min

1、场地设备要求

人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；
开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，
标注工具：Microsoft Excel），流程图管理软件 1 个，可绘制流程图

2、工作任务

为了解决在高端布料的生产加工中，颜色抽样评估依赖于“行业老师傅”的问题，B 公司计划设计一款便携式智能取色仪来代替“行业老师傅”进行颜色的抽样评估。在此背景下，您作为人工智能专家，需要应用深度学习技术，设计一个光谱值-Lab 色值预测模型。

数据标注背景知识：

“raw_xyz”代表设备传感器采集到的原始光谱值，这是在采集过程中直接得到的数据；而“lab”则代表 Lab 色值，这是通过人工进行标注的数据，也是我们深度学习模型需要进行预测的目标数据。

数据标注前，我们的数据格式如下所示：我们可以看到一系列的数据对象，每个对象都包含一个“raw_xyz”键，其对应的值为一个字符串，字符串中包含三个数字，以逗号分隔，分别代表设备采集到的光谱值。

[

```
  {"raw_xyz": "272, 328, 119"},
```

```
  {"raw_xyz": "176, 216, 81"},
```

```
  {"raw_xyz": "75, 91, 35"},
```

```
  {"raw_xyz": "23, 27, 10"}]
```

]

经过人工标注后，我们的数据格式则发生了改变：现在每个数据对象除了包含“raw_xyz”键之外，还多出了一个“lab”键，其对应的值为一个字符串，字符串中包含三个数字，以逗号分隔，这三个数字代表通过人工标注得到的 Lab 色值。

[

```
    {"raw_xyz": "272, 328, 119", "lab": "94.88, -0.51, 1.58"},  
    {"raw_xyz": "176, 216, 81", "lab": "81.62, -1.34, 0.4"},  
    {"raw_xyz": "75, 91, 35", "lab": "58.57, -0.99, -0.32"},  
    {"raw_xyz": "23, 27, 10", "lab": "35.22, 1.14, -1.67"}
```

]

我们的目标是，通过训练深度学习模型，使其能够根据输入的“raw_xyz”光谱值，准确预测出对应的“lab”色值。

(1) 请结合人工智能技术要求和业务特征，设计一套完整的业务数据采集流程，包括设备调试，场地布置，布料采集，数据标注等环节，并详细描述每个环节的操作步骤和注意事项。(5分)

(2) 在布料颜色预测模型中，数据处理是非常重要的环节。请设计一套完整的业务数据处理流程，包括数据清洗，数据预处理，数据分割，数据增强等环节，并详细描述每个环节的操作步骤和注意事项。(5分)

(3) 为了保证数据质量，数据审核是不可缺少的一步。请设计一套完整的业务数据审核流程，以便对处理后的数据进行质量控制和异常检测。请详细描述流程设计，包括但不限于审核的指标，方法以及处理审核结果的策略。(5分)

3、技能要求

- (1) 能够结合人工智能技术要求和业务特征，设计整套业务数据采集流程
- (2) 能够结合人工智能技术要求和业务特征，设计整套业务数据处理流程
- (3) 能够结合人工智能技术要求和业务特征，设计整套业务数据审核流程

4、质量指标

无

主观评分表/过程评分表

序号	配	评分细则描述	考评员评分	得分
----	---	--------	-------	----

	分		1	2	3	
(1)	5分	数据需求和数据来源明确合理：+1分 数据收集方法的选择合理，考虑到了数据的实时性和完整性：+1分 数据存储的设计考虑到了数据的安全性和便利性：+1分 对数据更新机制的描述合理，考虑到了数据的实时性：+1分 确定数据需求、数据来源、数据采集方式、数据存储、数据更新机制等流程内容合理，涉及4个：+5分				
(2)	5分	数据清洗方法的选择合理，能够有效处理缺失值、异常值、重复值等问题：+2分 数据转换设计合理，满足模型的需求：+2分 特征工程的设计考虑到了业务的特性和需求：+2分 数据切分方法的选择合理，适用于模型的训练和评估：+2分 涉及到3个流程：+5分				
(3)	5分	审核指标的设置合理：+1分 审核方法的设计科学：+1分 异常处理策略的设计合理，能够有效处理数据质量问题：+1分 考虑到了审核结果的反馈和使用，为数据采集和处理的改进提供支持：+1分 涉及4个步骤：+5分				

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：4.1.2

试题名称：便携式智能取色仪业务模块效果优化

考核时间：15min

1、场地设备要求

人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）；流程图管理软件 1 个，可绘制流程图

2、工作任务

你将作为一个人工智能训练师参与到便携式智能取色仪的业务模块优化工作中。在这个过程中，你需要处理各种可能出现的问题，例如取色精度不高、用户反馈色彩不准确等。你的目标是利用人工智能和大数据技术优化这些问题，提高用户满意度，最终提高便携式智能取色仪的市场竞争力。

（1）请识别出便携式智能取色仪可能存在的问题，并分析其可能的原因（4 分）

（2）针对在（1）中识别的问题，设计一套业务模块优化方案，以提高便携式智能取色仪的性能。请详细描述您的方案设计，包括但不限于优化的目标、优化的方法以及预期的结果等。（6 分）

3、技能要求

（1）能够结合业务知识，识别业务流程中单一模块的问题

（2）能够结合人工智能技术设计业务模块优化方案并推动实现

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	4分	可能存在的问题及原因合理即可，每个+2分				

(2)	6分	优化目标的设定明确、合理：+1分 优化方法的设计科学，考虑到了数据、模型和用户反馈等多方面的因素，一个优化方法+2分，两个优化方法+3分 对优化方法的实施有明确的计划，如定期进行实验和评估等：+1分 对预期结果的描述准确，反映了优化的效果： +1分				
-----	----	--	--	--	--	--

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：4.2.1

试题名称：便携式智能取色仪数据处理规范制定

考核时间：30min

1、场地设备要求

人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练）

2、工作任务

作为一个数据处理专家，你的工作是为便携式智能取色仪项目进行数据准备。这个项目需要大量的图像数据和色彩数据。你将处理从光谱传感器采集到的数据，并为这些数据进行色值标注。

项目使用的光谱数据是基于 CIE 1931 XYZ 颜色空间的，这是一个基于人类视觉感知的颜色空间。在这个颜色空间中，颜色被表示为 X、Y、Z 三个参数，这三个参数在理论上没有明确的上下限，它们的具体取值取决于颜色刺激的强度。然而，在实际应用中，人们通常会选择一个白点作为参考，然后将所有的颜色的 XYZ 值都相对于这个白点进行标定。在此任务中，选择的白点是 D65 白点，它的 XYZ 值被定义为 (95.047, 100.000, 108.883)。因此，在这种情况下，对于一般的颜色刺激，X、Y、Z 的值通常都会在 0 和这个白点的 XYZ 值之间。除非是非常强烈的颜色刺激，XYZ 值才可能会超过这个范围。

你的任务包括：

(1) 基于下述原始数据“data.json”，为光谱数据制定数据清洗标准，并保存清洗后的待标注数据为“data_cleaned.json”。(10 分)

数据格式：

```
[  
  {"raw_xyz": "272,328,119"},  
  {"raw_xyz": "176,216,81"},  
  {"raw_xyz": "75,91,35"},  
  {"raw_xyz": "23,27,10"},  
  {"raw_xyz": "320,364,132"},  
  {"raw_xyz": "220,250,90"},  
  {"raw_xyz": "156,178,65"},
```

```
 {"raw_xyz": "86,98,35"},
```

]

注：以上数字均为示例，实际的光谱值和对应的 Lab 色值需要依据具体设备和环境测量得出。

(2) 基于实际需求，为色值标注制定完整的标注规范。(5 分)

注：

- 标注的 Lab 色值应符合 CIE Lab 颜色空间的标准。其中，L 值（亮度）应在 0 到 100 之间，a 值和 b 值（色度）应在-128 到 127 之间。
- 实际标注过程是由调色师使用肉眼与色卡进行对，为每个光谱数据标注对应的 Lab 色值。此处只需要模拟生成相应的数据即可，并制定数据标注规范。
- 标注的数据应保存为“data_labeled.json”，格式如下：

[

```
 {"raw_xyz": "272,328,119", "lab": "94.88,-0.51,1.58"},
```

```
 {"raw_xyz": "176,216,81", "lab": "81.62,-1.34,0.4"},
```

...

]

3、技能要求

(1) 能够结合人工智能技术要求和业务特征，设计数据清洗和标注流程

(2) 能够结合人工智能技术要求和业务特征，制定数据清洗和标注规范

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	10分	正确给出类似参考答案标准中的一条：+1 分 正确给出类似参考答案标准中的两条：+3 分 正确给出类似参考答案标准中的三条：+5 分 若给出一条及以上不合理的数据清洗标准：-1 分 需要删除的数据每删除一条：+0.1 分 不需要删除的数据每删除一条：-0.1 分 加总得分在 0 到 10 分之间。				

(2)	5分	<p>正确给出标注内容与范围：+0.5分</p> <p>正确给出标注格式：+0.5分</p> <p>正确给出数据交付格式：+0.5分</p> <p>积极的标注标准，每给出4条常见表述中类似的1条得0.5分，最多得1分</p> <p>消极的标注标准，每给出4条常见表述中类似的1条得0.5分，最多得1分</p> <p>正确给出特殊情况反馈途径：+1分</p> <p>加总得分在0到5分之间。</p>				
-----	----	--	--	--	--	--

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：4.2.2

试题名称：便携式智能取色仪模型训练评估

考核时间：60min

1、场地设备要求

(1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练

(2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

在面辅料激烈竞争的背景下，对于面辅料的颜色感知还是通过人工感知的非标准化的方式识别，除了培养颜色感知能力或者寻找颜色感知力天生强的人以外，没有更好的方式来将非标的过程标准化。了解整个过程，对操作流程的掌握后，你作一名人工智能训练师，将对这样非标的操作流程要标准化。因此将使用设备将颜色量化，并训练出符合标准的设备模型。

(1) 请打开“train.ipynb”，按照要求补齐相关代码，基于“xxx.json”文件，对训练集进行训练。（4分）

(2) 请打开“tests.ipynb”，按照要求补齐相关代码，使用“[106, 186, 53]”单个色点对每个设备模型进行评测，按照“0.85”输出设备达标的占比。（4分）

(3) 请对测试结果分析，并编写测试报告（6分）

(4) 分析不合格设备的原因，并指出可能得纠正方法（6分）

3、技能要求

(1) 能够维护日常训练集和生成相关测试集

(2) 能够使用工具对算法进行训练

(3) 能够使用测试工具对人工智能产品的使用进行测试

(4) 可以对测试结果进行分析，编写测试包裹

(5) 能够运用工具，分析算法，优化算法

4、质量指标

(1) 代码书写规范

(2) 代码正常运行

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01	4分	正确填写一处代码：+1分 汇总得分在0到4分之间		
02	4分	正确填写一处代码：+1分 汇总得分在0到4分之间		

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
03	6分	对模型指标做出类似参考 答案一条至多条合理分析： +4分 对模型提出改善方案：+2 得分为6分、4分、2分、 或0分				
04	6分	正确的指出至少一个错误 案例：+2分 正确的给出可能得原因分 析及纠正方案：+4分 得分为6分、4分、2分、 或0分				

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：4.3.1

试题名称：便携式智能取色仪系统监控和优化

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

智能取色仪在高端布料的生产加工中发挥了重要的作用，但在取色过程以及数据标注和比对中可能存在一些操作问题，影响了最终的结果。您作为数据分析专家，需要进行以下任务：

- (1) 请基于已有数据，进行数据分析，并输出分析报告，指出可能存在的问题。（3分）
- (2) 基于数据分析的结果，提出优化需求，并为这些需求设计智能解决方案。（4分）
- (3) 描述设计的解决方案的实施步骤，并指出预期效果和可能的挑战。（3分）

3、技能要求

- (1) 能够对单一智能产品使用的数据进行全面分析，输出分析报告
- (2) 能够对单一智能产品提出优化需求
- (3) 能够为单一智能产品的应用设计智能解决方案

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	3分	正确给出类似参考答案问题分析中的一条：+2分 正确给出类似参考答案问题分析中的两条及以上：+3分 给出错误问题分析：-1分 加总得分在0到3分之间				
(2)	4分	正确给出类似参考答案优化需求中的一条：+2分 正确给出类似参考答案优化需求中的两条及以上：+4分 给出错误优化需求：-2分 加总得分在0到4分之间				
(3)	3分	正确给出类似参考答案中的实施方案：酌情给分，不超过3分 加总得分在0到4分之间				

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：4.3.2

试题名称：便携式智能取色仪系统人机交互流程设计

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

智能取色仪在高端布料的生产加工中发挥了重要的作用，但在取色过程以及数据标注和比对中可能存在一些操作问题，影响了最终的结果。作为交互设计专家，您需要综合考虑智能取色仪的使用场景，对人机交互流程进行设计。请完成以下任务：

- (1) 通过数据分析，找到高端布料生产加工场景下人工和智能交互的最优方式，并解释为什么这是最优的。（3分）
- (2) 设计高端布料生产加工场景下人工和智能交互的最优流程，并制作相关的流程图。（3分）
- (3) 结合你的交互流程设计，解释如何解决取色过程中可能出现的问题以及提高数据标注和比对的效率。（4分）

3、技能要求

- (1) 能够通过数据分析，找到单一场景下人工和智能交互的最优方式
- (2) 能够通过数据分析，设计单一场景下人工和智能交互的最优流程

4、质量指标

无

主观评分表/过程评分表

序号	配 分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	3 分	提出类似参考答案的最优方式：+3分 得分为3分或0分				
(2)	4 分	提出类似参考答案的最优流程：+4分 得分为4分或0分				
(3)	3 分	提出类似参考答案的解释：+3分 得分为3分或0分				

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：4.4.1

试题名称：数据采集和处理培训

考核时间：10 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

为了解决在高端布料的生产加工中，颜色抽样评估依赖于“行业老师傅”的问题，B 公司计划设计一款便携式智能取色仪来代替“行业老师傅”进行颜色的抽样评估。在此背景下，您作为人工智能专家，需要组建团队，应用深度学习技术，设计一个光谱值-Lab 色值预测模型。

首先，你需要对数据采集和处理团队进行岗前培训：

- (1) 请提供相应的数据采集和处理培训讲义的目录，并对培训内容做简单描述。（5 分）
- (2) 请针对数据采集和处理培训，从下述方法中选择合适的培训方法，并简述原因：

讲授法、演示法、研讨法、视听法、角色扮演法（5 分）

讲授法、演示法、研讨法、视听法、角色扮演法（5 分）

3、技能要求

- (1) 能够编写初级培训讲义
- (2) 能够对五级/初级工、四级/中级工开展知识和技术培训

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	5分	目录设计合理，循序渐进由浅入深：酌情给分，不超过5分				
(2)	5分	正确选择一到多种培训方法并给出原因（可与参考答案不同）：酌情给分，不超过5分				

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：4.4.2

试题名称：数据采集和处理指导

考核时间：10 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

为了解决在高端布料的生产加工中，颜色抽样评估依赖于“行业老师傅”的问题，B 公司计划设计一款便携式智能取色仪来代替“行业老师傅”进行颜色的抽样评估。在此背景下，您作为人工智能专家，需要带领团队，应用深度学习技术，设计一个光谱值-Lab 色值预测模型。

在数据采集和处理过程中，不可避免的会出现一些问题，作为专家需要为提供输出相应的解决方案：

- (1) 请简述数据采集和处理的常见问题的解决方案内容。（1 分）
- (2) 请简述数据标注流程中可指导的常见问题指导方向（1 分）

3、技能要求

- (1) 能够指导五级/初级工、四级/中级工解决数据采集、处理问题
- (2) 能够指导五级/初级工、四级/中级工解决数据标注问题

4、质量指标

无

主观评分表/过程评分表

序号	配	评分细则描述	考评员评分	得分
----	---	--------	-------	----

	分		1	2	3	
(1)	5分	正确提供至少一个数据采集和处理的常见问题及解决方案简要说明：酌情给分，不超过+5分				
(2)	5分	正确提供至少一个数据标注流程中的问题及指导方向简要说明：酌情给分，不超过+5分				

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：5.1.1

试题名称：电商评论情感分析系统的流程设计

考核时间：20min

1、场地设备要求

(1) 人工智能训练师主机：CPU（intel i5 及以上）、内存（不少于 16GB）、操作系统（windows10）、支持深度学习训练

(2) 图像文本标注工具、python 开发语言支持工具

2、工作任务

业务背景：

当今社交媒体和数字平台的迅猛发展，使得大量的用户生成了海量的文本数据，其中包含了丰富的情感信息。情感分析任务是一种重要的自然语言处理技术，它的主要目标是自动识别和分析文本中的情感倾向和情感极性。情感分析在多个领域中都具有广泛的应用，包括社交媒体监测、品牌管理、市场调研、舆情分析等。

某平台作为一个电子商务平台，希望通过情感分析技术来获取用户评论的情感极性，以帮助平台更好地运营和提升用户体验。随着数字平台的迅速发展和用户规模的不断扩大，该平台已经积累了大量的用户评论和反馈。这些评论包含了用户对平台上的产品、服务、功能和用户体验等方面的意见和情感表达。平台意识到，通过深入了解用户的情感反应，可以获得宝贵的洞察，帮助平台更好地了解用户需求、改善产品和提升服务质量。

为了更好地设计情感分析系统，请针对该系统进行流程设计。请根据需求分析需要采集什么样的数据，分析维度包括采集种类、采集规则和采集要求。内容保存在”提交资料/业务分析.docx“中

3、技能要求

能够结合人工智能技术要求和业务特征，设计整套业务数据采集流程、处理流程、审核流程。

4、质量指标

流程设计要完整全面，可得 5 分

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分	

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1	1	流程设计里面包含采集数据的种类				
2	2	流程设计里面包含采集规则的描述				
3	2	流程设计里面包含采集的要求说明				

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：5.1.2

试题名称：电商评论情感分析系统的模块效果优化

考核时间： 20min

1、场地设备要求

(1) 人工智能训练师主机：CPU (intel i5 及以上)、内存 (不少于 16GB)、操作系统 (windows10)、支持深度学习训练

(2) 图像文本标注工具、python 开发语言支持工具

2、工作任务

业务背景：

当今社交媒体和数字平台的迅猛发展，使得大量的用户生成了海量的文本数据，其中包含了丰富的情感信息。情感分析任务是一种重要的自然语言处理技术，它的主要目标是自动识别和分析文本中的情感倾向和情感极性。情感分析在多个领域中都具有广泛的应用，包括社交媒体监测、品牌管理、市场调研、舆情分析等。

某平台作为一个电子商务平台，希望通过情感分析技术来获取用户评论的情感极性，以帮助平台更好地运营和提升用户体验。随着数字平台的迅速发展和用户规模的不断扩大，该平台已经积累了大量的用户评论和反馈。这些评论包含了用户对平台上的产品、服务、功能和用户体验等方面的意见和情感表达。平台意识到，通过深入了解用户的情感反应，可以获得宝贵的洞察，帮助平台更好地了解用户需求、改善产品和提升服务质量。

为了更好地设计情感分析系统，请针对该系统进行模块效果优化。

(1) 采集的数据会包含各种各样的问题，如污染文本信息的 html 符号等，请再列出两种采集后可能存在的数据噪音并给出清洗代码。

(2) 给定标注好(data.csv)的数据，判断数据是否符合要求，请从几种不同的方向进行分析，并给出分析方法，若需要代码实现则提供相应代码。

3、技能要求

能够结合人工智能技术要求和业务特征，识别单一模块中的问题，设计业务模块优化方案并推动实现。

4、质量指标

数据清洗和优化要正确，可得 10 分

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
1	5	清洗代码可跑通		

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1	2	数据噪声罗列至少两个				
2.1	1	给出数据是否符合要求的结论				
2.2	2	针对是否符合要求进一步说明				

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：5.2.1

试题名称：电商评论情感分析系统的数据处理规范制定

考核时间：20min

1、场地设备要求

(1) 人工智能训练师主机：CPU（intel i5 及以上）、内存（不少于 16GB）、操作系统（windows10）、支持深度学习训练

(2) 图像文本标注工具、python 开发语言支持工具

2、工作任务

业务背景：

当今社交媒体和数字平台的迅猛发展，使得大量的用户生成了海量的文本数据，其中包含了丰富的情感信息。情感分析任务是一种重要的自然语言处理技术，它的主要目标是自动识别和分析文本中的情感倾向和情感极性。情感分析在多个领域中都具有广泛的应用，包括社交媒体监测、品牌管理、市场调研、舆情分析等。

某平台作为一个电子商务平台，希望通过情感分析技术来获取用户评论的情感极性，以帮助平台更好地运营和提升用户体验。随着数字平台的迅速发展和用户规模的不断扩大，该平台已经积累了大量的用户评论和反馈。这些评论包含了用户对平台上的产品、服务、功能和用户体验等方面的意见和情感表达。平台意识到，通过深入了解用户的情感反应，可以获得宝贵的洞察，帮助平台更好地了解用户需求、改善产品和提升服务质量。

为了更好地设计情感分析系统，请针对该系统进行数据处理规范制定

(3) 请根据需求完成标注规范的制定。标注规范保存在“提交资料”目录下，并命名为“情感分析任务标注规范”。标注规范需要包括以下内容：

a. 任务背景：介绍数据标注任务的业务背景，以及情感分析任务的任务介绍；

- b. 标注规则：给出那些情况不进行标注，进行标注的情况应该以什么样的标注将文本情感分别标注为**积极、消极**，并给出标注样例；
- c. 特殊标注情况：一些难以标注的情况应该如何标注，以及给出样例；
- d. 特殊情况反馈：遇到无法确定标注标签的情况，应该如何进行反馈；

3、技能要求

能够结合业务知识，设计数据清洗和标注流程，并制定完整的规范标准。

4、质量指标

标注流程设计要规范并且正确，可得 10 分

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1	2	答案中包含任务背景描述				
2	3	答案包含标注规则				
3	3	答案包含困难标注样本的处理方法				
4	2	答案包含一些特殊情况的案例说明				

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：5.2.2

试题名称：电商评论情感分析系统的算法测试

考核时间：40min

1、场地设备要求

(1) 人工智能训练师主机：CPU (intel i5 及以上)、内存 (不少于 16GB)、操作系统 (windows10)、支持深度学习训练

(2) 图像文本标注工具、python 开发语言支持工具

2、工作任务

业务背景：

当今社交媒体和数字平台的迅猛发展，使得大量的用户生成了海量的文本数据，其中包含了丰富的情感信息。情感分析任务是一种重要的自然语言处理技术，它的主要目标是自动识别和分析文本中的情感倾向和情感极性。情感分析在多个领域中都具有广泛的应用，包括社交媒体监测、品牌管理、市场调研、舆情分析等。

某平台作为一个电子商务平台，希望通过情感分析技术来获取用户评论的情感极性，以帮助平台更好地运营和提升用户体验。随着数字平台的迅速发展和用户规模的不断扩大，该平台已经积累了大量的用户评论和反馈。这些评论包含了用户对平台上的产品、服务、功能和用户体验等方面的意见和情感表达。平台意识到，通过深入了解用户的情感反应，可以获得宝贵的洞察，帮助平台更好地了解用户需求、改善产品和提升服务质量。

为了更好地设计情感分析系统，请针对该系统进行算法测试相关操作。数据集存在于 (data.csv)。请完成一下操作：

(1) 划分训练集、验证集和测试集。创建 “data_split.ipynb” 文件，使用代码将标注好的数据按照 8:1:1 的比例分割为训练集、验证集和测试集。将其按

照“train.csv”、“valid.csv”和“test.csv”的名称保存到“提交资料”目录下。

(2) 配置训练环境、完成模型训练代码，并在训练集上进行模型训练，模型保存到“提交资料/models”文件下。

(3) 在测试集上测试模型效果，需要给出 f1 值，并将测试结果保存到“提交资料”文件下。

3、技能要求

能够正确拆分训练集、测试集、验证集，能够使用工具对算法进行训练，能够针对训练过程进行测试报告撰写。

4、质量指标

3.2.2.1 电商评论情感分析系统的算法测试

算法开发过程要按照指导进行正确训练，并且训练结果保证测试集准确率在 50%以上可得 25 分

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
1	5	data_split.ipynb 文件可以跑通		
2	10	测试集评测结果准确率达到 50%		

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1	3	数据集划分分成三份				
2	3	‘提交资料/models’里				

		面有输出结果				
3	4	模型代码有提交				

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：5.3.1

试题名称：电商评论情感分析系统的系统监控和优化

考核时间：40min

1、场地设备要求

- (1) 人工智能训练师主机：CPU (intel i5 及以上)、内存 (不少于 16GB)、操作系统 (windows10)、支持深度学习训练
- (2) 图像文本标注工具、python 开发语言支持工具

2、工作任务

业务背景：

当今社交媒体和数字平台的迅猛发展，使得大量的用户生成了海量的文本数据，其中包含了丰富的情感信息。情感分析任务是一种重要的自然语言处理技术，它的主要目标是自动识别和分析文本中的情感倾向和情感极性。情感分析在多个领域中都具有广泛的应用，包括社交媒体监测、品牌管理、市场调研、舆情分析等。

某平台作为一个电子商务平台，希望通过情感分析技术来获取用户评论的情感极性，以帮助平台更好地运营和提升用户体验。随着数字平台的迅速发展和用户规模的不断扩大，该平台已经积累了大量的用户评论和反馈。这些评论包含了用户对平台上的产品、服务、功能和用户体验等方面的意见和情感表达。平台意识到，通过深入了解用户的情感反应，可以获得宝贵的洞察，帮助平台更好地了解用户需求、改善产品和提升服务质量。

为了更好地设计情感分析系统，请针对该系统进行系统监控和优化。

- (1) 如何对情感分析数据集标注结果进行质检，至少描述 2 种质检方法和流程，并应描述每种方法的优劣势。

3、技能要求

能够针对输出数据结果进行全面分析，输出分析报告，够针对模块进行优化，能够为单一智能产品的应用设计智能解决方案

4、质量指标

至少描述 2 种质检方法和流程，并应完整描述每种方法的优劣势。可得 10 分

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1	5	答案有两种质检方法				
2	5	答案包含每种质检方法的流程说明				

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：5.3.2

试题名称：电商评论情感分析系统的人机交互流程设计

考核时间：20min

1、场地设备要求

(1) 人工智能训练师主机：CPU（intel i5 及以上）、内存（不少于 16GB）、操作系统（windows10）、支持深度学习训练

(2) 图像文本标注工具、python 开发语言支持工具

2、工作任务

业务背景：

当今社交媒体和数字平台的迅猛发展，使得大量的用户生成了海量的文本数据，其中包含了丰富的情感信息。情感分析任务是一种重要的自然语言处理技术，它的主要目标是自动识别和分析文本中的情感倾向和情感极性。情感分析在多个领域中都具有广泛的应用，包括社交媒体监测、品牌管理、市场调研、舆情分析等。

某平台作为一个电子商务平台，希望通过情感分析技术来获取用户评论的情感极性，以帮助平台更好地运营和提升用户体验。随着数字平台的迅速发展和用户规模的不断扩大，该平台已经积累了大量的用户评论和反馈。这些评论包含了用户对平台上的产品、服务、功能和用户体验等方面的意见和情感表达。平台意识到，通过深入了解用户的情感反应，可以获得宝贵的洞察，帮助平台更好地了解用户需求、改善产品和提升服务质量。

为了更好地设计情感分析系统，请通过对数据进行分析，结合数据特征选择合适的情感分析模型，并且通过实验对比选择最优模型。

(2) 给出至少两种情感分析模型试验结果。

(3) 分析实验结果，给出分析过程及结论，选择最合适的情感分析模型

3、技能要求

能够通过数据分析，找到场景下人工和智能交互的最优解，通过数据分析设计交互的最优流程。

4、质量指标

人机交互流程设计要全面并且清晰，可得 25 分

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
1	15	答案有两种方案的代码		

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
	5	答案包含分析过程和结果				
	5	有最优的情感分析模型结果说明				

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：5.4.1

试题名称： 电商评论情感分析系统的培训

考核时间： 10min

1、场地设备要求

(1) 人工智能训练师主机：CPU（intel i5 及以上）、内存（不少于 16GB）、操作系统（windows10）、支持深度学习训练

(2) 图像文本标注工具、python 开发语言支持工具

2、工作任务

业务背景：

当今社交媒体和数字平台的迅猛发展，使得大量的用户生成了海量的文本数据，其中包含了丰富的情感信息。情感分析任务是一种重要的自然语言处理技术，它的主要目标是自动识别和分析文本中的情感倾向和情感极性。情感分析在多个领域中都具有广泛的应用，包括社交媒体监测、品牌管理、市场调研、舆情分析等。

某平台作为一个电子商务平台，希望通过情感分析技术来获取用户评论的情感极性，以帮助平台更好地运营和提升用户体验。随着数字平台的迅速发展和用户规模的不断扩大，该平台已经积累了大量的用户评论和反馈。这些评论包含了用户对平台上的产品、服务、功能和用户体验等方面的意见和情感表达。平台意识到，通过深入了解用户的情感反应，可以获得宝贵的洞察，帮助平台更好地了解用户需求、改善产品和提升服务质量。

为了更好地设计情感分析系统，请针对该系统进行培训相关知识学习。分析人工智能训练师需要掌握的工作内容有哪些。

业务背景：

3、技能要求

能够编写初级培训讲义，针对理论知识进行讲解。

4、质量指标

培训技巧的知识分享符合要求即可，可得2分

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1	2	给出的方案 有两种即可				

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：5.4.2

试题名称：电商评论情感分析系统的指导

考核时间：10min

1、场地设备要求

(1) 人工智能训练师主机：CPU（intel i5 及以上）、内存（不少于 16GB）、操作系统（windows10）、支持深度学习训练

(2) 图像文本标注工具、python 开发语言支持工具

2、工作任务

业务背景：

当今社交媒体和数字平台的迅猛发展，使得大量的用户生成了海量的文本数据，其中包含了丰富的情感信息。情感分析任务是一种重要的自然语言处理技术，它的主要目标是自动识别和分析文本中的情感倾向和情感极性。情感分析在多个领域中都具有广泛的应用，包括社交媒体监测、品牌管理、市场调研、舆情分析等。

某平台作为一个电子商务平台，希望通过情感分析技术来获取用户评论的情感极性，以帮助平台更好地运营和提升用户体验。随着数字平台的迅速发展和用户规模的不断扩大，该平台已经积累了大量的用户评论和反馈。这些评论包含了用户对平台上的产品、服务、功能和用户体验等方面的意见和情感表达。平台意识到，通过深入了解用户的情感反应，可以获得宝贵的洞察，帮助平台更好地了解用户需求、改善产品和提升服务质量。

为了更好地设计情感分析系统，请针对该系统进行指导相关知识学习。分析人工智能训练师在数据标注的时候需要采用的工具有什么，有什么需要注意的事项。

3、技能要求

能够针对初级工数据采集、处理问题进行指导。

4、质量指标

针对数据标注和清晰有自己的认知即可，可得 3 分

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1	3	给出的工具至少两种				

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：6.1.1

试题名称：实体链接-业务流程设计

考核时间： 20 min

1、场地设备要求

(1) 电脑主机：设备指标满足 CPU：建议英特尔 i5 以上；内存：建议不少于 16GB；操作系统：Windows 10；开发语言：python；开发工具：python 编译工具；深度学习训练环境：支持主流深度学习框架训练

(2) 额外设备：鼠标，键盘等配置设备

2、工作任务

(1) 随着网络应用技术的飞速发展，多元化、低密度数据的急剧膨胀对人们获取正确信息带来巨大挑战，大量冗余信息出现的根源在于自然语言表达的多样性，即一词多义和多词同义。例如，“苹果”在不同语境下既可以表示蔷薇科苹果属植物又可以表示苹果产品公司；“李明”在同一城市街道既可以表示正在中学上学的李明同学又可以表示快递员李明；“申城”和“魔都”尽管字面完全不同，却都是上海市的别称。因此，伴随着大数据时代的到来，实现对海量 Web 数据的高效处理，理解用户意图，降低信息过载，是实体链接追求的目标。在政企领域，用户关注的信息通常是事件的时间、地点、人物和事件本身等四个核心维度。随着时间的积累，事件无论从概念范畴还是信息纵深上都有了长足发展，已成长为信息空间的参天大树，针对政务类中某领导的一年的事件都长达 GB 维度的数据，在众多人物和事件信息中，确保我们能精准直达对特定人物肖像的描绘显得十分关键，因此如何精确识别用户所指显得尤为重要。构建政企知识库为基础的实体链接服务，提升政企搜索、政企问答、政企知识挖掘和抽取等工作的效果，对改善政企用户体验有重要意义。

(2) 你将作为一名人工智能训练师为实体链接的训练提供可用的文本数据，成

功训练出模型。根据业务特征与业务需求，选择采集方法并设计采集流程；根据业务需求，设计处理流程和处理方法，并利用 python 完成数据清洗和处理；根据业务需求，设计标注与质检规范与流程。

3、技能要求

(1) 业务流程设计

基本情况：

- ① 本体分值：10 分
- ② 考核时间：20min
- ③ 考核形式：实操

1.1、请根据实体链接需求，在给定数据集基础上分析训练需要什么样的数据集，分析构造所需要训练数据集的理由并说明原因。内容提交至“提交资料/业务分析.docx”文件中；

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1.1	10	1、提到构建文本相似度用二分类的思想去构建数据集即可				

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：6.1.2

试题名称：实体链接-业务模块效果优化

考核时间： 40 min

1、场地设备要求

(1) 电脑主机：设备指标满足 CPU：建议英特尔 i5 以上；内存：建议不少于 16GB；操作系统：Windows 10；开发语言：python；开发工具：python 编译工具；深度学习训练环境：支持主流深度学习框架训练

(2) 额外设备：鼠标，键盘等配置设备

2、工作任务

(1) 随着网络应用技术的飞速发展，多元化、低密度数据的急剧膨胀对人们获取正确信息带来巨大挑战，大量冗余信息出现的根源在于自然语言表达的多样性，即一词多义和多词同义。例如，“苹果”在不同语境下既可以表示蔷薇科苹果属植物又可以表示苹果产品公司；“李明”在同一城市街道既可以表示正在中学上学的李明同学又可以表示快递员李明；“申城”和“魔都”尽管字面完全不同，却都是上海市的别称。因此，伴随着大数据时代的到来，实现对海量 Web 数据的高效处理，理解用户意图，降低信息过载，是实体链接追求的目标。在政企领域，用户关注的信息通常是事件的时间、地点、人物和事件本身等四个核心维度。随着时间的积累，事件无论从概念范畴还是信息纵深上都有了长足发展，已成长为信息空间的参天大树，针对政务类中某领导的一年的事件都长达 GB 维度的数据，在众多人物和事件信息中，确保我们能精准直达对特定人物肖像的描绘显得十分关键，因此如何精确识别用户所指显得尤为重要。构建政企知识库为基础的实体链接服务，提升政企搜索、政企问答、政企知识挖掘和抽取等工作的效果，对改善政企用户体验有重要意义。

(2) 你将作为一名人工智能训练师为实体链接的训练提供可用的文本数据，成

功训练出模型。根据业务特征与业务需求，选择采集方法并设计采集流程；根据业务需求，设计处理流程和处理方法，并利用 python 完成数据清洗和处理；根据业务需求，设计标注与质检规范与流程。

3、技能要求

(1) 业务模块效果优化

基本情况：

- ① 本体分值：20 分
- ② 考核时间：40min
- ③ 考核形式：实操

1.1、请根据实体链接需求，请给出给定数据转化为实体链接训练数据的转化流程--脚本参考” code/data_process.py”，内容提交至“提交资料/数据转化.docx”；

1.2、完成对转化的实体链接数据集基本信息统计，包括训练和测试的数据量大小，内容提交至“提交资料/数据分析.docx”；

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
1.2	10	返回比例值，形式不固定	638:100	

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1.1	10	提到用 kb_id 链接到数据库的数据为正样本，其余为负样本，至于负样本个数不做限制，有相关论述即可				

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：6.2.1

试题名称：实体链接-数据处理规范制定

考核时间： 20min

1、场地设备要求

(1) 电脑主机：设备指标满足 CPU：建议英特尔 i5 以上；内存：建议不少于 16GB；操作系统：Windows 10；开发语言：python；开发工具：python 编译工具；深度学习训练环境：支持主流深度学习框架训练

(2) 额外设备：鼠标，键盘等配置设备

2、工作任务

(1) 随着网络应用技术的飞速发展，多元化、低密度数据的急剧膨胀对人们获取正确信息带来巨大挑战，大量冗余信息出现的根源在于自然语言表达的多样性，即一词多义和多词同义。例如，“苹果”在不同语境下既可以表示蔷薇科苹果属植物又可以表示苹果产品公司；“李明”在同一城市街道既可以表示正在中学上学的李明同学又可以表示快递员李明；“申城”和“魔都”尽管字面完全不同，却都是上海市的别称。因此，伴随着大数据时代的到来，实现对海量 Web 数据的高效处理，理解用户意图，降低信息过载，是实体链接追求的目标。在政企领域，用户关注的信息通常是事件的时间、地点、人物和事件本身等四个核心维度。随着时间的积累，事件无论从概念范畴还是信息纵深上都有了长足发展，已成长为信息空间的参天大树，针对政务类中某领导的一年的事件都长达 GB 维度的数据，在众多人物和事件信息中，确保我们能精准直达对特定人物肖像的描绘显得十分关键，因此如何精确识别用户所指显得尤为重要。构建政企知识库为基础的实体链接服务，提升政企搜索、政企问答、政企知识挖掘和抽取等工作的效果，对改善政企用户体验有重要意义。

(2) 你将作为一名人工智能训练师为实体链接的训练提供可用的文本数据，成

功训练出模型。根据业务的需求，利用 python 划分训练集与测试集，随后利用工具完成模型训练，测试模型，对测试结果进行分析。

3、技能要求

(1) 数据处理规范制定

基本情况：

- ① 本体分值：10 分
- ② 考核时间：20min
- ③ 考核形式：实操

1.1、划分训练集与测试集，打开“提交资料/data”项目，需要完成如下任务：说明转化后的训练集和测试集的比例是多少，造成这样结果的原因是什么，内容提交至“提交资料/数据集分析.docx”文件；

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
1.1	10	提到动态负采样即可	动态负采样操作	

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：6.2.2

试题名称：实体链接-算法测试

考核时间： 40 min

1、场地设备要求

(1) 电脑主机：设备指标满足 CPU：建议英特尔 i5 以上；内存：建议不少于 16GB；操作系统：Windows 10；开发语言：python；开发工具：python 编译工具；深度学习训练环境：支持主流深度学习框架训练

(2) 额外设备：鼠标，键盘等配置设备

2、工作任务

(1) 随着网络应用技术的飞速发展，多元化、低密度数据的急剧膨胀对人们获取正确信息带来巨大挑战，大量冗余信息出现的根源在于自然语言表达的多样性，即一词多义和多词同义。例如，“苹果”在不同语境下既可以表示蔷薇科苹果属植物又可以表示苹果产品公司；“李明”在同一城市街道既可以表示正在中学上学的李明同学又可以表示快递员李明；“申城”和“魔都”尽管字面完全不同，却都是上海市的别称。因此，伴随着大数据时代的到来，实现对海量 Web 数据的高效处理，理解用户意图，降低信息过载，是实体链接追求的目标。在政企领域，用户关注的信息通常是事件的时间、地点、人物和事件本身等四个核心维度。随着时间的积累，事件无论从概念范畴还是信息纵深上都有了长足发展，已成长为信息空间的参天大树，针对政务类中某领导的一年的事件都长达 GB 维度的数据，在众多人物和事件信息中，确保我们能精准直达对特定人物肖像的描绘显得十分关键，因此如何精确识别用户所指显得尤为重要。构建政企知识库为基础的实体链接服务，提升政企搜索、政企问答、政企知识挖掘和抽取等工作的效果，对改善政企用户体验有重要意义。

(2) 你将作为一名人工智能训练师为实体链接的训练提供可用的文本数据，成

功训练出模型。根据业务的需求，利用 python 划分训练集与测试集，随后利用工具完成模型训练，测试模型，对测试结果进行分析。

3、技能要求

(1) 算法测试

基本情况：

- ① 本体分值：40 分
- ② 考核时间：40min
- ③ 考核形式：实操

1.1、给出实体链接的基础模型，放在 ‘code/train.py’ 下，用于指示实体链接所用到的深度学习模型，将 train.py 的损失函数换成 Focal-loss，源码已给出在 ‘code/focal_loss.py’，请说明替换的过程，内容提交至“提交资料/Loss 替换.docx”

1.2、按照实体链接模型需要自行搭建好训练环境，利用 ‘code/train.py’ 完成整个实体链接的训练过程；在训练过程中需要记录训练的参数以及训练过程的 Loss，F1 值等关键信息，这些运行过程产生的信息保存在 ‘提交资料//训练信息.docx’ 下；

1.3、将训练所得到的最优模型保存在 ‘提交代码/checkpoint’ 下，用于后续的模式预测；

1.4、测试模型，运行 ‘code/infer.py’ 文件，能够完成整个模型的测试过程，需要给出测试数据集的 F1 值保存在 ‘提交资料/测试信息.docx’ 下；

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
1.1	10	替换正确即可	将 code/confi.py 文件的 criterion='CE' 改为 criterion='FL'	
1.2	10	见图	见图 ‘提交资料//训练信息.docx’ 下	
1.3	10	有数据模型即可		
1.4	10	见图	见图 ‘提交资料/测试信息.docx’ 下	

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：6.3.1

试题名称：智能系统监控和优化

考核时间：20 min

1、场地设备要求

(1) 电脑主机：设备指标满足 CPU：建议英特尔 i5 以上；内存：建议不少于 16GB；操作系统：Windows 10；开发语言：python；开发工具：python 编译工具；深度学习训练环境：支持主流深度学习框架训练

(2) 额外设备：鼠标，键盘等配置设备

2、工作任务

(1) 随着网络应用技术的飞速发展，多元化、低密度数据的急剧膨胀对人们获取正确信息带来巨大挑战，大量冗余信息出现的根源在于自然语言表达的多样性，即一词多义和多词同义。例如，“苹果”在不同语境下既可以表示蔷薇科苹果属植物又可以表示苹果产品公司；“李明”在同一城市街道既可以表示正在中学上学的李明同学又可以表示快递员李明；“申城”和“魔都”尽管字面完全不同，却都是上海市的别称。因此，伴随着大数据时代的到来，实现对海量 Web 数据的高效处理，理解用户意图，降低信息过载，是实体链接追求的目标。在政企领域，用户关注的信息通常是事件的时间、地点、人物和事件本身等四个核心维度。随着时间的积累，事件无论从概念范畴还是信息纵深上都有了长足发展，已成长为信息空间的参天大树，针对政务类中某领导的一年的事件都长达 GB 维度的数据，在众多人物和事件信息中，确保我们能精准直达对特定人物肖像的描绘显得十分关键，因此如何精确识别用户所指显得尤为重要。构建政企知识库为基础的实体链接服务，提升政企搜索、政企问答、政企知识挖掘和抽取等工作的效果，对改善政企用户体验有重要意义。

(2) 你将作为一名人工智能训练师为实体链接的训练提供可用的文本数据，成

功训练出模型。通过对数据的分析，结合数据特征，选择一些合适的目标检测模型，并通过实现对比选择最优模型。

3、技能要求

(1) 智能系统监控和优化

基本情况：

- ① 本体分值：10分
- ② 考核时间：20min
- ③ 考核形式：实操

1.1、通过对实体链接数据进行分析，至少描述两种优化策略跟方法，并描述该方法的优劣势，内容提交至“提交资料/智能系统设计.docx”；

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1.1	10	1. 提到文本想改改为实体向量匹配、负采样个数、loss函数替换等两点即可				

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：6.3.2

试题名称：实体链接-人机交互流程设计

考核时间： 20 min

1、场地设备要求

(1) 电脑主机：设备指标满足 CPU：建议英特尔 i5 以上；内存：建议不少于 16GB；操作系统：Windows 10；开发语言：python；开发工具：python 编译工具；深度学习训练环境：支持主流深度学习框架训练

(2) 额外设备：鼠标，键盘等配置设备

2、工作任务

(1) 随着网络应用技术的飞速发展，多元化、低密度数据的急剧膨胀对人们获取正确信息带来巨大挑战，大量冗余信息出现的根源在于自然语言表达的多样性，即一词多义和多词同义。例如，“苹果”在不同语境下既可以表示蔷薇科苹果属植物又可以表示苹果产品公司；“李明”在同一城市街道既可以表示正在中学上学的李明同学又可以表示快递员李明；“申城”和“魔都”尽管字面完全不同，却都是上海市的别称。因此，伴随着大数据时代的到来，实现对海量 Web 数据的高效处理，理解用户意图，降低信息过载，是实体链接追求的目标。在政企领域，用户关注的信息通常是事件的时间、地点、人物和事件本身等四个核心维度。随着时间的积累，事件无论从概念范畴还是信息纵深上都有了长足发展，已成长为信息空间的参天大树，针对政务类中某领导的一年的事件都长达 GB 维度的数据，在众多人物和事件信息中，确保我们能精准直达对特定人物肖像的描绘显得十分关键，因此如何精确识别用户所指显得尤为重要。构建政企知识库为基础的实体链接服务，提升政企搜索、政企问答、政企知识挖掘和抽取等工作的效果，对改善政企用户体验有重要意义。

(2) 你将作为一名人工智能训练师为实体链接的训练提供可用的文本数据，成

功训练出模型。通过对数据的分析，结合数据特征，选择一些合适的目标检测模型，并通过实现对比选择最优模型。

3、技能要求

(1) 人机交互流程设计

基本情况：

- ① 本体分值：10分
- ② 考核时间：20min
- ③ 考核形式：实操

1.1、通过对实体链接相关参数的微调，至少描绘一种参数的变化对最终模型的影响曲线，内容提交至“提交资料/智能系统设计.docx”；

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
1.1	10	随意改变config.py文件参数训练，有相关F1曲线图即可				

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：6.4.1

试题名称：实体链接-培训

考核时间： 10 min

1、场地设备要求

（1）电脑主机：设备指标满足 CPU：建议英特尔 i5 以上；内存：建议不少于 16GB；操作系统：Windows 10；开发语言：python；开发工具：python 编译工具；深度学习训练环境：支持主流深度学习框架训练

（2）额外设备：鼠标，键盘等配置设备

2、工作任务

（1）随着网络应用技术的飞速发展，多元化、低密度数据的急剧膨胀对人们获取正确信息带来巨大挑战，大量冗余信息出现的根源在于自然语言表达的多样性，即一词多义和多词同义。例如，“苹果”在不同语境下既可以表示蔷薇科苹果属植物又可以表示苹果产品公司；“李明”在同一城市街道既可以表示正在中学上学的李明同学又可以表示快递员李明；“申城”和“魔都”尽管字面完全不同，却都是上海市的别称。因此，伴随着大数据时代的到来，实现对海量 Web 数据的高效处理，理解用户意图，降低信息过载，是实体链接追求的目标。在政企领域，用户关注的信息通常是事件的时间、地点、人物和事件本身等四个核心维度。随着时间的积累，事件无论从概念范畴还是信息纵深上都有了长足发展，已成长为信息空间的参天大树，针对政务类中某领导的一年的事件都长达 GB 维度的数据，在众多人物和事件信息中，确保我们能精准直达对特定人物肖像的描绘显得十分关键，因此如何精确识别用户所指显得尤为重要。构建政企知识库为基础的实体链接服务，提升政企搜索、政企问答、政企知识挖掘和抽取等工作的效果，对改善政企用户体验有重要意义。

(2) 你将作为一名人工智能训练师为实体链接的训练提供可用的文本数据，成功训练出模型。培训与指导：能够对五级、四级/中级工开展知识和技术培训；能够指导五级/初级工、四级/中级工解决数据标注问题。

3、技能要求

(1) 培训

基本情况：

- ① 本体分值：10分
- ② 考核时间：10min
- ③ 考核形式：实操

1.1、人工智能训练师五级/初级工、四级/中级工需要掌握的工作内容有哪些？

主观评分表/过程评分表

序号	配分	评分细则描述	规定或标称值	得分
1.1	10	内容描述正确即可	<p>人工智能训练师五级/初级工需要掌握的工作内容包括：</p> <ol style="list-style-type: none"> 1. 熟悉机器学习的基础知识，包括常见的模型、算法和数据预处理技术。 2. 具备编程能力，能够使用Python等编程语言实现机器学习算法。 3. 能够根据业务需求，设计、构建和优化机器学习模型，并对模型进行评估和调优。 4. 具备数据分析和数据可视化能力，能够对数据进 	

			<p>行清洗、挖掘和分析，提取有价值的信息。</p> <p>5. 能够将机器学习模型部署到生产环境中，并进行效果监测和升级维护。</p> <p>人工智能训练师四级/中级工需要在此基础上扩展以下工作内容：</p> <p>1. 深入了解自然语言处理、计算机视觉、推荐系统等人工智能领域的专业知识，并能够应用到具体业务场景中。</p> <p>2. 具备多任务学习、强化学习等高级机器学习算法的掌握和应用能力。</p> <p>3. 能够对超大规模数据进行分布式处理和机器学习模型训练。</p> <p>4. 具备团队管理和项目管理能力，能够协调项目进度、质量和人员，对团队成员进行指导和培训。</p>	
--	--	--	---	--

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：6.4.2

试题名称：实体链接-指导

考核时间： 10 min

1、场地设备要求

(1) 电脑主机：设备指标满足 CPU：建议英特尔 i5 以上；内存：建议不少于 16GB；操作系统：Windows 10；开发语言：python；开发工具：python 编译工具；深度学习训练环境：支持主流深度学习框架训练

(2) 额外设备：鼠标，键盘等配置设备

2、工作任务

(1) 随着网络应用技术的飞速发展，多元化、低密度数据的急剧膨胀对人们获取正确信息带来巨大挑战，大量冗余信息出现的根源在于自然语言表达的多样性，即一词多义和多词同义。例如，“苹果”在不同语境下既可以表示蔷薇科苹果属植物又可以表示苹果产品公司；“李明”在同一城市街道既可以表示正在中学上学的李明同学又可以表示快递员李明；“申城”和“魔都”尽管字面完全不同，却都是上海市的别称。因此，伴随着大数据时代的到来，实现对海量 Web 数据的高效处理，理解用户意图，降低信息过载，是实体链接追求的目标。在政企领域，用户关注的信息通常是事件的时间、地点、人物和事件本身等四个核心维度。随着时间的积累，事件无论从概念范畴还是信息纵深上都有了长足发展，已成长为信息空间的参天大树，针对政务类中某领导的一年的事件都长达 GB 维度的数据，在众多人物和事件信息中，确保我们能精准直达对特定人物肖像的描绘显得十分关键，因此如何精确识别用户所指显得尤为重要。构建政企知识库为基础的实体链接服务，提升政企搜索、政企问答、政企知识挖掘和抽取等工作的效果，对改善政企用户体验有重要意义。

(2) 你将作为一名人工智能训练师为实体链接的训练提供可用的文本数据，成

功训练出模型。培训与指导：能够对五级、四级/中级工开展知识和技术培训；能够指导五级/初级工、四级/中级工解决数据标注问题。

3、技能要求

(1) 指导

基本情况：

- ① 本体分值：10分
- ② 考核时间：10min
- ③ 考核形式：实操

1.1、人工智能训练师五级/初级工、四级/中级工针对数据训练一般的运行流程是什么过程？

主观评分表/过程评分表

序号	配分	评分细则描述	规定或标称值	得分
1.1	10	内容描述正确即可	<p>对于数据训练一般的运行流程，可以概括为以下几个步骤：</p> <ol style="list-style-type: none"> 1. 数据获取和清洗：根据业务需求，获取相关的数据集，并进行数据清洗和预处理，剔除无效数据和异常值。 2. 特征工程：根据问题类型和数据特点，进行特征工程的处理，包括特征选择、特征变换和特征生成等操作，提取有意义的特征信息。 3. 模型选择和训练：根据任务需求，选择适当的机器学习模型，并使用训练集对模型进行训练和调优，以提高 	

			<p>模型的精度和性能。</p> <p>4. 模型评估和验证：使用测试集对模型进行评估和验证，计算各种指标并对模型进行比较和选择。</p> <p>5. 模型部署和监测：将训练好的模型部署到生产环境中，并进行在线监测和性能调优，确保模型的稳定性和可用性。</p> <p>以上是数据训练一般的运行流程，实际操作中，可能会涉及到更多细节和问题的解决。人工智能训练师需要熟练掌握各种工具和技术，才能在实际应用中迭代优化模型，提高数据分析和决策的效率和准确率。</p>	
--	--	--	---	--

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：7.1.1

试题名称：自然场景吸烟检测业务流程设计

考核时间：25 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

吸烟作为一种常见的社会问题，已经成为全球面临的公共卫生挑战之一。据统计，吸烟是导致许多健康问题和死亡的主要原因之一，其中包括肺癌、慢性阻塞性肺病和心血管疾病等。为此，在大部分的公共场所，都禁止吸烟的行为。随着智能化技术的不断发展，吸烟检测的新的方法和技术也将不断出现。其中，机器学习等人工智能技术为吸烟检测带来更多的创新和可能性。

你将作为一名人工智能训练师为自然场景吸烟检测模型的训练提供可用的数据，成功训练出模型。

- (1) 请根据自然场景下吸烟行为分析，设计一套完整的业务数据采集流程，以便针对模型训练提供支持。（5分）
- (2) 在检测系统中，图像数据前处理是非常重要的环节。请设计一套完整的业务数据处理流程，请详细描述您的流程设计。（5分）
- (3) 为了保证数据质量，数据审核是不可缺少的一步。请设计一套完整的数据审核流程，以便对处理后的数据进行质量控制和异常检测。请详细描述流程设计，包括但不限于审核的指标、方法以及处理审核结果的策略。（5分）

3、技能要求

- (1) 能够结合人工智能技术要求和业务特征，设计整套业务数据采集流程
- (2) 能够结合人工智能技术要求和业务特征，设计整套业务数据处理流程
- (3) 能够结合人工智能技术要求和业务特征，设计整套业务数据审核流程

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	5分	数据需求和数据来源明确合理：+1分 数据收集方法的选择合理，考虑到了数据的实时性和完整性：+1分 数据存储的设计考虑到了数据的安全性和便利性：+1分 对数据更新机制的描述合理，考虑到了数据的实时性：+1分 确定数据需求、数据来源、数据采集方式、数据存储、数据更新机制等流程内容合理，涉及4个：+5分				
(2)	5分	数据清洗方法的选择合理，能够有效处理无效图像等问题：+2分 数据转换设计合理，满足模型的需求：+2分 数据切分方法的选择合理，适用于模型的训练和评估：+2分 涉及到3个流程：+5分				
(3)	5分	审核指标的设置合理：+1分 审核方法的设计科学：+1分 异常处理策略的设计合理，能够有效处理数据质量问题：+1分 考虑到了审核结果的反馈和使用，为数据采集和处理的改进提供支持：+1分 涉及4个步骤：+5分				

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：7.1.2

试题名称：自然场景吸烟检测模块效果优化

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

吸烟作为一种常见的社会问题，已经成为全球面临的公共卫生挑战之一。据统计，吸烟是导致许多健康问题和死亡的主要原因之一，其中包括肺癌、慢性阻塞性肺病和心血管疾病等。为此，在大部分的公共场所，都禁止吸烟的行为。随着智能化技术的不断发展，吸烟检测的新的方法和技术也将不断出现。其中，机器学习等人工智能技术为吸烟检测带来更多的创新和可能性。你将作为一名人工智能训练师为吸烟检测模型的训练提供可用的数据，成功训练出模型。

- (1) 在实际运行中，吸烟检测系统可能会遇到各种问题，例如误检、漏检等。请针对检测可能存在的问题，分析其可能的原因（4分）
- (2) 针对在（1）中识别的问题，设计一套业务模块优化方案，以提高吸烟检测系统的性能。请详细描述您的方案设计，包括但不限于优化的目标、优化的方法以及预期的结果等。（6分）

3、技能要求

- (1) 能够结合业务知识，识别业务流程中单一模块的问题

(2) 能够结合人工智能技术设计业务模块优化方案并推动实现

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	4分	可能存在的问题及原因合理即可，每个+2分				
(2)	6分	优化目标的设定明确、合理：+1分 优化方法的设计科学，考虑到了数据、模型和用户反馈等多方面的因素，一个优化方法+2分，两个优化方法+3分 对优化方法的实施有明确的计划，如定期进行实验和评估等：+1分 对预期结果的描述准确，反映了优化的效果：+1分				

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：7.2.1

试题名称：自然场景吸烟检测数据处理规范制定

考核时间：30 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

吸烟作为一种常见的社会问题，已经成为全球面临的公共卫生挑战之一。据统计，吸烟是导致许多健康问题和死亡的主要原因之一，其中包括肺癌、慢性阻塞性肺病和心血管疾病等。为此，在大部分的公共场所，都禁止吸烟的行为。随着智能化技术的不断发展，吸烟检测的新的方法和技术也将不断出现。其中，机器学习等人工智能技术为吸烟检测带来更多的创新和可能性。

本项目中，你作为一名人工智能训练师，将进行自然场景吸烟检测数据的采集与标注。

- (1) 题目：根据项目需求，确定采集什么样的数据，分析维度需包含采集种类，采集规则和采集要求。（15 分）
- (2) 至少写出两种采集方法，并描述如何通过该采集方法采集需要的数据。（15 分）

3、技能要求

- (1) 能够结合人工智能技术要求和业务特征，设计数据采集与标注流程
- (2) 能够结合人工智能技术要求和业务特征，制定数据采集与标注规范

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	10分	正确给出类似参考答案标准中的一条：+1分 正确给出类似参考答案标准中的两条：+3分 正确给出类似参考答案标准中的三条：+5分 若给出一条及以上不合理的数据清洗标准：-1分 需要删除的数据每删除一条：+0.1分 不需要删除的数据每删除一条：-0.1分 加总得分在0到10分之间。				
(2)	5分	正确给出标注内容与范围：+0.5分 正确给出标注格式：+0.5分 正确给出数据交付格式：+0.5分 积极的标注标准，每给出4条常见表述中类似的1条得0.5分，最多得1分 消极的标注标准，每给出4条常见表述中类似的1条得0.5分，最多得1分 正确给出特殊情况反馈途径：+1分 加总得分在0到5分之间。				

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：7.2.2

试题名称：自然场景吸烟检测算法测试

考核时间：60 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

吸烟作为一种常见的社会问题，已经成为全球面临的公共卫生挑战之一。据统计，吸烟是导致许多健康问题和死亡的主要原因之一，其中包括肺癌、慢性阻塞性肺病和心血管疾病等。为此，在大部分的公共场所，都禁止吸烟的行为。随着智能化技术的不断发展，吸烟检测的新的方法和技术也将不断出现。其中，机器学习等人工智能技术为吸烟检测带来更多的创新和可能性。

本项目中，你作为一名人工智能训练师，将对吸烟检测模型进行训练。

请先打开 环境准备.ipynb，运行代码。

- (1) 请创建 train.ipynb，修改 train.py 中的训练集路径、验证集路径，调用 train.py，基于已划分好的训练集、测试集和验证集对模型进行训练。（4 分）
- (2) 请创建“test.ipynb”，修改 test.py 中的测试集路径，调用 test.py，基于已划分好的训练集、测试集和验证集对模型进行测试。（4 分）
- (4) 请对测试结果进行分析，计算 mAP 值，并编写测试报告。（6 分）
- (5) 请找出至少一个错误案例，分析错误产生的原因并指出可能的纠正方法。（6 分）

3、技能要求

- (1) 能使用工具对算法进行训练
- (2) 能够使用测试工具对人工智能产品的使用进行测试
- (3) 能够对测试结果进行分析，编写测试报告
- (4) 能够运用工具，分析算法中错误案例产生的原因并进行纠正

4、质量指标

- (1) 代码书写规范
- (2) 代码正常运行

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
(1)	2分	正确填写一处代码：+1分 汇总得分在0到2分之间		
(2)	2分	正确填写一处代码：+1分 汇总得分在0到2分之间		
(3)	4分	正确填写一处代码：+1分 正确输出模型性能指标及混淆矩阵：+2分 汇总得分在0到4分之间		

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(4)	6分	对模型指标做出类似参考答案一条至多条合理分析：+4分 对混淆矩阵做出类似参考答案分析说明：+2分 得分为6分、4分、2分或0分				
(5)	6分	正确指出至少一个错误案例：+2分 正确给出可能原因分析及纠正方案：+4分 加总得分为6分、4分、2分或0分				

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：7.3.1

试题名称：吸烟检测系统监控和优化

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

吸烟作为一种常见的社会问题，已经成为全球面临的公共卫生挑战之一。据统计，吸烟是导致许多健康问题和死亡的主要原因之一，其中包括肺癌、慢性阻塞性肺病和心血管疾病等。为此，在大部分的公共场所，都禁止吸烟的行为。随着智能化技术的不断发展，吸烟检测的新的方法和技术也将不断出现。其中，机器学习等人工智能技术为吸烟检测带来更多的创新和可能性。

本项目中，你作为人工智能训练师，将对该智能产品进行分析和优化设计。

- (1) 通过对数据的分析，结合数据特征，提出优化方法。（5分）
- (2) 结合上述统计的比例数据分布情况，分析一下目前方案的缺点。（5分）
- (3) 结合吸烟情况下，香烟与人之间的联系，提出优化方案。内容提交至“提交资料/题目 3/智能系统设计.docx”文件中。（5分）

3、技能要求

- (1) 能够对单一智能产品使用的数据进行全面分析，输出分析报告
- (2) 能够对单一智能产品提出优化需求
- (3) 能够为单一智能产品的应用设计智能解决方案

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	10分	正确描述数据分布 + 2分 将数据归一化 + 2分 提出优化方法 + 1分				
(2)	1分	正确描述类似参考答案优化需求中的一至多条: +4分 错误描述不扣分 汇总得分为4分或0分				

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：7.3.2

试题名称：自然场景吸烟检测系统人机交互流程设计

考核时间：15 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

吸烟作为一种常见的社会问题，已经成为全球面临的公共卫生挑战之一。据统计，吸烟是导致许多健康问题和死亡的主要原因之一，其中包括肺癌、慢性阻塞性肺病和心血管疾病等。为此，在大部分的公共场所，都禁止吸烟的行为。随着智能化技术的不断发展，吸烟检测的新的方法和技术也将不断出现。其中，机器学习等人工智能技术为吸烟检测带来更多的创新和可能性。

本项目中，你作为人工智能训练师，将对该智能产品进行分析和优化设计。

- (1) 请为该智能产品找到人工与智能交互的最优方式。（10 分）
- (2) 请为该智能产品设计人工与智能交互的最优流程。（10 分）

3、技能要求

- (1) 能够通过数据分析，找到单一场景下人工和智能交互的最优方式
- (2) 能够通过数据分析，设计单一场景下人工和智能交互的最优流程

4、质量指标

无

主观评分表/过程评分表

序号	配	评分细则描述	考评员评分	得
----	---	--------	-------	---

	分		1	2	3	分
(1)	10分	提出类似参考答案的最优方式: +10分 得分为10分或0分				
(2)	10分	提出类似参考答案的最优流程: +10分 得分为10分或0分				

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：7.4.1

试题名称：数据清洗培训

考核时间：10 min

1、场地设备要求

(1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）

(2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

(1) 请提供数据清洗培训讲义的目录，并对培训内容做简单描述。（1分）

(2) 请针对数据清洗培训，从下述方法中选择合适的培训方法，并简述原因：
讲授法、演示法、研讨法、视听法、角色扮演法（2分）

3、技能要求

(1) 能够编写初级培训讲义

(2) 能够对五级/初级工、四级/中级工开展知识和技术培训

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	1分	目录设计合理，循序渐进由浅入深：+1分 得分为1分或0分				
(2)	2分	正确选择一到多种培训方法并给出原因（可与参考答案不同）：+2分 得分为2分或0分				

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：7.4.2

试题名称：数据采集和处理指导

考核时间：10 min

1、场地设备要求

- (1) 人工智能训练师主机 1 台（CPU：建议英特尔 i5 以上，操作系统：Windows 10；开发工具：Python 编译工具，深度学习训练环境：支持主流深度学习框架训练，标注工具：Microsoft Excel）
- (2) 流程图管理软件 1 个，可绘制流程图

2、工作任务

- (1) 请简述数据采集和处理的常见问题的解决方案内容（至少提供教材中五个中的一个部分名称及简要说明）。（1 分）
- (2) 请简述数据标注流程中可指导的常见问题指导方向（至少提供教材中七个中的一个部分名称及简要说明）。（1 分）

3、技能要求

- (1) 能够指导五级/初级工、四级/中级工解决数据采集、处理问题
- (2) 能够指导五级/初级工、四级/中级工解决数据标注问题

4、质量指标

无

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	1分	目录设计合理，循序渐进由浅入深：+1 分 得分为 1 分或 0 分				
(2)	2分	正确选择一到多种培训方法并给出原因（可与参考答案不同）：+2 分 得分为 2 分或 0 分				

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：8.1.1

试题名称：人工智能训练师业务流程设计能力考核

考核时间：15min

1、场地设备要求

(1) 指定考试场地

(2)

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 4GB 操作系统：Windows10 python 版本：不低于 3.7 版本	台	1	

2、工作任务

(1) 业务背景

大型企业在各种业务过程中，总是会涉及到各种类型的合同文档。公司相关的法务、财务人员进行合同审核时，经常都是一行行的条款内容进行人工审核，整个过程，耗时耗力，还容易出现遗漏的风险。

公司的相关审核人员，希望通过人工智能技术帮忙自动提取合同当中具备风险可能的一些关键内容，基于人工智能技术抽取的结果辅助人工进行内容的复核。

(2) 业务目标

你将作为一名人工智能训练师，针对私募基金合同的审核场景，定义出其中需要的关键信息，清洗出适合训练的数据集，并成功训练提取相关字段的信息抽取模型。本次需要抽取的字段包括，基金名称、存续期限、募集总额、托管机构这 4 个字段。

业务分析阶段：需要理解业务的含义。根据关键信息的特征，对标注数据进行处理，使后续模型训练可以正常进行。

(3) 考试内容

- ① 请根据人工智能中自然语言处理技术的相关内容，写出信息抽取任务的标注数据，需要包含哪些关键内容。
- ② 客户的业务人员给出了一批标注样例数据，但标注的数据中只有部分内容。请你运行以下程序“data_label.py”生成带有更多信息的标注数据，并截图存在“业务分析.docx”文档中。

3、技能要求

- (1) 能够结合人工智能技术要求和业务特征，设计整套业务数据采集流程。
- (2) 能够结合人工智能技术要求和业务特征，设计整套业务数据处理流程。
- (3) 能够结合人工智能技术要求和业务特征，设计整套业务数据审核流程。

4、质量指标

- (1) 理解自然语言处理领域中信息抽取的基本概念。
- (2) 能够掌握自然语言处理实操的具体步骤。

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：8.1.2

试题名称：人工智能训练师业务模块效果优化能力考核

考核时间：10min

1、场地设备要求

(1) 指定考试场地

(2)

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 4GB 操作系统：Windows10 python 版本：不低于 3.7 版本	台	1	

2、工作任务

(1) 业务背景

大型企业在各种业务过程中，总是会涉及到各种类型的合同文档。公司相关的法务、财务人员进行合同审核时，经常都是一行行的条款内容进行人工审核，整个过程，耗时耗力，还容易出现遗漏的风险。

公司的相关审核人员，希望通过人工智能技术帮忙自动提取合同当中具备风险可能的一些关键内容，基于人工智能技术抽取的结果辅助人工进行内容的复核。

(2) 业务目标

你将作为一名人工智能训练师，针对私募基金合同的审核场景，定义出其中需要的关键信息，清洗出适合训练的数据集，并成功训练提取相关字段的信息抽取模型。本次需要抽取的字段包括，基金名称、存续期限、募集总额、托管机构这 4 个字段。

业务分析阶段：需要理解业务的含义。根据关键信息的特征，对标注数据进行处理，使后续模型训练可以正常进行。

（3）考试内容

请你阅读“data_label.py”代码，理解代码和结果的具体含义。可以看出有一些标注内容，在原文中出现了多次，请你删除错误标注的相关内容，将重复内容中正确的结果粘贴在“业务分析.docx”文档中，不重复的内容不必保留。并尝试从数据标注的角度解释说明，你保留这些结果的原因。

3、技能要求

（1）能够结合人工智能技术要求和业务特征，设计整套业务数据采集流程。

（2）能够结合人工智能技术要求和业务特征，设计整套业务数据处理流程。

（3）能够结合人工智能技术要求和业务特征，设计整套业务数据审核流程。

4、质量指标

（1）理解自然语言处理领域中信息抽取的基本概念。

（2）掌握自然语言处理领域数据标注处理的核心要点。

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： 8.2.1

试题名称：人工智能训练师数据处理规范制定能力考核

考核时间： 15min

1、场地设备要求

(1) 指定考试场地

(2)

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 4GB 操作系统：Windows10 python 版本：不低于 3.7 版本	台	1	

2、工作任务

(1) 业务背景

很多企业会在自己的 app 应用上接入各类舆情数据，在企业处理舆情新闻时，需要对新闻进行分类，以方便对不同领域感兴趣的用户能够快速阅读到相关领域的新闻。

(2) 业务目标

你将作为人工智能训练师，针对新闻分类这个场景训练一个文本分类模型，类别总共有 15 类，包括文化、体育、财经、娱乐等多个领域，详细标签内容参见“labels.json”文件。

(3) 考试内容

- ① 海量新闻资讯的采集一般通过什么方法？新闻采集一般至少要包含哪些内容。
- ② 采集过程中，不同新闻网站经常会互相转载新闻，这些重复新闻对用户

的体验影响很大。请针对海量新闻去重场景，说明至少一种新闻去重算法。

3、技能要求

(1) 能够结合人工智能技术要求和业务特征，设计数据清洗和标注流程。

(2) 能够结合人工智能技术要求和业务特征，制定数据清洗和标注规范。

4、质量指标

(1) 深入理解业务场景下的数据采集、预处理方法

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：8.2.2

试题名称：人工智能训练师模型训练能力考核

考核时间：30min

1、场地设备要求

(1) 指定考试场地

(2)

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 4GB 操作系统：Windows10 python 版本：不低于 3.7 版本	台	1	

2、工作任务

(1) 业务背景

很多企业会在自己的 app 应用上接入各类舆情数据，在企业处理舆情新闻时，需要对新闻进行分类，以方便对不同领域感兴趣的用户能够快速阅读到相关领域的新闻。

(2) 业务目标

你将作为人工智能训练师，针对新闻分类这个场景训练一个文本分类模型，类别总共有 15 类，包括文化、体育、财经、娱乐等多个领域，详细标签内容参见“labels.json”文件。

(3) 考试内容

- ① 请你针对待分类的文本数据进行数据清洗和预处理，需要至少包含去除特殊符号和停用词等可能影响文本分类效果的内容，并对文本进行分词处理。请将清洗过程在“data_process.py”文件中的 clean_text 方法

内进行实现，并运行该文件。截图保存至“智能训练.docx”中。

- ② 请将上题中实现的 `clean_text` 方法复制到“`classifier_svm_trainer.py`”中的同名方法中，并运行 `svm` 模型训练和评估程序，将 `accuracy` 结果保存到“智能训练.docx”中。

3、技能要求

- (1) 能够维护日常训练集与测试集。
- (2) 能使用工具对算法进行训练。
- (3) 能够使用测试工具对人工智能产品的使用进行测试。
- (4) 能够对测试结果进行分析，编写测试报告。
- (5) 能够运用工具，分析算法中错误案例产生的原因并进行纠正。

4、质量指标

- (1) 深入理解业务场景下的数据清洗方法。
- (2) 深入理解业务场景下的模型训练和评估方法。

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：8.3.1

试题名称：人工智能训练师智能系统和监控能力考核

考核时间：25 min

1、场地设备要求

(1) 指定考试场地

(2)

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 4GB 操作系统：Windows10 python 版本：不低于 3.7 版本	台	1	

2、工作任务

(1) 业务背景

很多企业会在自己的 app 应用上接入各类舆情数据，在企业处理舆情新闻时，需要对新闻进行分类，以方便对不同领域感兴趣的用户能够快速阅读到相关领域的新闻。

(2) 业务目标

你将作为人工智能训练师，针对新闻分类这个场景训练一个文本分类模型，类别总共有 15 类，包括文化、体育、财经、娱乐等多个领域，详细标签内容参见“labels.json”文件。

(3) 考试内容

- ① 标注质量对模型训练的结果好坏有着根本性的影响，请你设计一套流程，针对分类模型的标注质量进行质检。请描述至少 2 种质检方法，并说明其优劣势。

- ② 请结合 `train.json` 训练数据，还有哪些特征没有使用到。尝试在“`classifier_svm.py`”的训练和评估过程中加入该特征，使得准确率可以大幅提升。

3、技能要求

- (1) 能够对单一智能产品使用的数据进行全面分析，输出分析报告。
- (2) 能够对单一智能产品提出优化需求。
- (3) 能够为单一智能产品的应用设计智能解决方案。

4、质量指标

- (1) 深入理解业务场景下的数据预测和特征提取的方法

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：8.3.2

试题名称：人工智能训练师人机交互流程设计考核

考核时间：20 min

1、场地设备要求

(1) 指定考试场地

(2)

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 4GB 操作系统：Windows10 python 版本：不低于 3.7 版本	台	1	

2、工作任务

(1) 业务背景

很多企业会在自己的 app 应用上接入各类舆情数据，在企业处理舆情新闻时，需要对新闻进行分类，以方便对不同领域感兴趣的用户能够快速阅读到相关领域的新闻。

(2) 业务目标

你将作为人工智能训练师，针对新闻分类这个场景训练一个文本分类模型，类别总共有 15 类，包括文化、体育、财经、娱乐等多个领域，详细标签内容参见“labels.json”文件。

(3) 考试内容

① 根据 classifier_svm.py 中的程序，以下样例的分类结果如下：

- 1) 【人类象棋高手为什么不敌阿法狗？】news_culture，正确分类应该是 news_tech

2) 【企业同时属于高新企业和科技型中小企业，研发费用如何扣除？】

news_tech，正确分类应该是 news_finance

请尝试分析模型分类错误的原因。

3、技能要求

- (1) 能够通过数据分析，找到单一场景下人工和智能交互的最优方式。
- (4) 能够通过数据分析，设计单一场景下人工和智能交互的最优流程。

4、质量指标

- (1) 深入理解业务场景下数据建模原理，并具备分析取得效果原因的能力。

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：8.4.1

试题名称：人工智能训练师培训能力考核

考核时间：12min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

(1) 业务背景

随着人工智能技术在各领域的应用不断深入，人工智能训练师这一职业不断发展壮大，作为一名三级/高级人工智能训练师，既需要熟练使用各种工具完成数据采集、标注、训练和分析，也需要参与人工智能训练师五级/初级工、四级/中级培训工作。

(2) 业务目标

您将作为一名三级/高级人工智能训练师，参与人工智能训练师五级/初级工、四级/中级培训工作。有能力完成培训讲义编写、开展知识和技术培训、指导解决数据采集/处理和数据标准问题等多项工作。

(3) 考试内容

能够编写初级培训讲义；能够对五级/初级工、四级/中级工开展知识和技术培训。

①人工智能训练师五级/初级工、四级/中级工需要掌握的工作内容有哪些？

3、技能要求

- (1) 能够编写初级培训讲义
- (2) 能够对五级/初级工、四级/中级工开展知识和技术培训

4、质量指标

- (1) 包括但不限于：数据采集和处理、数据标注、智能系统运维

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01	1分	数据采集和处理，回答正确得1分	回答内容相似即可得分，不要求答案表述一模一样	1分
02	1分	数据标注，回答正确得1分	回答内容相似即可得分，不要求答案表述一模一样	1分
03	1分	智能系统运维，回答正确得1分	回答内容相似即可得分，不要求答案表述一模一样	1分

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：8.4.2

试题名称：人工智能训练师指导能力考核

考核时间：8min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

(1) 业务背景

随着人工智能技术在各领域的应用不断深入，人工智能训练师这一职业不断发展壮大，作为一名三级/高级人工智能训练师，既需要熟练使用各种工具完成数据采集、标注、训练和分析，也需要参与人工智能训练师五级/初级工、四级/中级指导工作。

(2) 业务目标

您将作为一名三级/高级人工智能训练师，参与人工智能训练师五级/初级工、四级/中级指导工作。有能力完成培训讲义编写、开展知识和技术培训、指导解决数据采集/处理和数据标准问题等多项工作。

(3) 考试内容

能够指导五级/初级工、四级/中级工解决数据采集/处理问题；能够指导五级/初级工、四级/中级工解决数据标注问题。

①人工智能训练师五级/初级工、四级/中级工解决数据采集、处理问题、解决数据标注问题使用的工具有哪些？

3、技能要求

- (1) 能够指导五级/初级工、四级/中级工解决数据采集/处理问题
- (2) 能够指导五级/初级工、四级/中级工解决数据标注问题

4、质量指标

(1) 一类是公开通用的标注工具，如：NumPy、Pandas、Excel、Word；还有一类是企业开发的标注工具：达观研发标注文档的 IDP 智能文档审阅系统，还有标注图片的 OCR 标注器

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01	1分	一类是公开通用的标注工具，如：NumPy、Pandas、Excel、Word。回答出通用标注工具得1分	回答结果符合实际要求即可，不要求和标准答案一模一样	1分
02	1分	还有一类是企业开发的标注工具：达观研发标注文档的 IDP 智能文档审阅系统，还有标注图片的 OCR 标注器。回答出企业开发标注工具得1分	回答结果符合实际要求即可，不要求和标准答案一模一样	1分
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：9.1.1

试题名称：业务和数据分析采集

考核时间：20min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10 开发语言：Python 开发工具：Python 编译工具； 深度学习训练环境：支持主流深度学习框架训练 配件：鼠标、键盘及显示	台	1	

2、工作任务

OCR 文字识别技术是搭建视觉图像和计算机文本操作的桥梁，它基于计算机视觉和人工智能技术，可以将手写或印刷的字符图像转换成可编辑或搜索的文本。OCR 技术已经成功应用于各个领域。在社会业务中，被广泛的用于企业的文档处理和管理。现代企业的日常工作涉及到大量的文档处理，如果采用传统的手写或印刷方式，不仅效率低下，还容易出现错误。采用 OCR 技术，可以极大地提高文档处理效率和准确性。

某 RPA（机器人流程自动化）系统可以帮助业务人员进行网页表单数据的自动填写，但是每次填写都需要先输入数字验证码，这个操作会使流程造成一定的阻塞

和效率下降。因此，系统开发人员想要利用人工智能技术准确识别验证码数字，通过识别验证码图片得到验证码数字，将人工填写验证码的步骤优化为自动填写。

你将作为一名人工智能训练师为验证码识别模型的训练采集可用的图像数据，并对图像数据进行清洗。

根据业务特征与业务需求，选择采集方法并设计采集流程；根据业务需求，设计处理流程与处理方法，并利 Python 完成数据清洗与处理；根据业务需求，设计标注与质检规范与流程。

(1) 请根据此需求分析需要采集什么样的数据，分析维度需包含采集种类、采集规则和采集要求。内容提交至“提交资料/题目 1/业务分析.docx”文件中。

(2) 至少写出两种采集方法，并描述如何通过该采集方法采集需要的数据。内容提交至“提交资料/题目 1/业务分析.docx”文件中。

3、技能要求

(1) 业务分析，数据理解

(2) 数据采集能力

4、质量指标

(1) 业务和数据分析方法的综合全面性

(2) 采集方法的多样性

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：9.1.2

试题名称：数据清理和处理

考核时间：25min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10 开发语言：Python 开发工具：Python 编译工具； 深度学习训练环境：支持主流深度学习框架训练 配件：鼠标、键盘及显示	台	1	

2、工作任务

OCR 文字识别技术是搭建视觉图像和计算机文本操作的桥梁，它基于计算机视觉和人工智能技术，可以将手写或印刷的字符图像转换成可编辑或搜索的文本。OCR 技术已经成功应用于各个领域。在社会业务中，被广泛的用于企业的文档处理和管理。现代企业的日常工作涉及到大量的文档处理，如果采用传统的手写或印刷方式，不仅效率低下，还容易出现错误。采用 OCR 技术，可以极大地提高文档处理效率和准确性。

某 RPA（机器人流程自动化）系统可以帮助业务人员进行网页表单数据的自动填写，但是每次填写都需要先输入数字验证码，这个操作会使流程造成一定的

阻塞和效率下降。因此，系统开发人员想要利用人工智能技术准确识别验证码数字，通过识别验证码图片得到验证码数字，将人工填写验证码的步骤优化为自动填写。

你将作为一名人工智能训练师为验证码识别模型的训练采集可用的图像数据，并对图像数据进行清洗。

根据业务特征与业务需求，选择采集方法并设计采集流程；根据业务需求，设计处理流程与处理方法，并利 Python 完成数据清洗与处理；根据业务需求，设计标注与质检规范与流程。

(1) 经过数据采集已经采集了一批用于验证码识别的图像数据，需要对数据进行全面处理与清洗。打开“业务分析”项目中的 processing_cleaning.ipynb 文件，考生在规定时间内完成以下任务：待清洗的数据存储在‘captcha’文件夹中，清洗数据样本内像素过小图像，删除短边小于 10 像素的图像，在 notebook 文件中完成程序，返回删除图像的名称和数量，将其记录在“提交资料/题目 1/业务分析.docx”文件中；

3、技能要求

- (1) 技术开发能力
- (2) 数据清洗能力

4、质量指标

- (1) 数据处理结果的准确性

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：9.2.1

试题名称：数据标注和训练数据划分

考核时间：25min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10 开发语言：Python 开发工具：Python 编译工具； 深度学习训练环境：支持主流深度学习框架训练 配件：鼠标、键盘及显示	台	1	

2、工作任务

OCR 文字识别技术是搭建视觉图像和计算机文本操作的桥梁，它基于计算机视觉和人工智能技术，可以将手写或印刷的字符图像转换成可编辑或搜索的文本。OCR 技术已经成功应用于各个领域。在社会业务中，被广泛的用于企业的文档处理和管理。现代企业的日常工作涉及到大量的文档处理，如果采用传统的手写或印刷方式，不仅效率低下，还容易出现错误。采用 OCR 技术，可以极大地提高文档处理效率和准确性。

某 RPA（机器人流程自动化）系统可以帮助业务人员进行网页表单数据的自动填写，但是每次填写都需要先输入数字验证码，这个操作会使流程造成一定的

阻塞和效率下降。因此，系统开发人员想要利用人工智能技术准确识别验证码数字，通过识别验证码图片得到验证码数字，将人工填写验证码的步骤优化为自动填写。

为了能训练出验证码识别模型，需要对采集好的验证数据进行标注，要求制定合理的标注方法和数据格式，以满足验证码识别模型的训练需求。在完成数据标注和处理后，基于标注数据实现验证码识别模型的训练和测试。

为了能训练出验证码识别模型，需要对处理好的数据进行标注，要求标注出数字验证码图片的验证码文字内容。

(1) 请根据此需求制定标注规范，标注规范需考虑如下内容。内容提交至“提交资料/ 题目 2/模型训练.docx”文件中。

- 标注方法：采用什么标注方法；
- 标注范围：哪些情况需要标注，哪些情况不需要标注；
- 标注规则：如标注哪些验证码数据，样式不一致的数据，验证码严重遮挡的情况等；
- 特殊情况反馈途径：若遇到举棋不定的标注情况如何进行反馈；
- 交付的数据格式：限定数据交付数据。

(2) 划分训练集与测试集，打开“划分数据”项目中 data_spilt.ipynb 文件，需要完成如下任务：

待划分的数据集存储在“captcha”中，包含图像文件夹 image 和标注文件 label.txt

1 按照 8:2 的比例将数据集划分为训练集与测试集，以如下格式进行存储；image 文件夹存储图像文件，label.txt 存储标注文件。

2 将训练集和测试集的数量记录在“提交资料/题目 2 模型训练.docx”文件中；将此部分代码截图保存至“提交资料/题目 2 模型训练阶段.docx”文件中。下载划分好的数据集，将其保存至“提交资料/题目 2/划分数据集”中

3、技能要求

- (1) 标注数据理解
- (2) 标注情况分析
- (3) 标注方法掌握
- (4) 编码能力

(5) 数据理解能力

4、质量指标

(1) 标注相关问题的回答全面性

(2) 数据集划分的准确性

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：9.2.2

试题名称：模型训练和测试验证

考核时间：35min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10 开发语言：Python 开发工具：Python 编译工具； 深度学习训练环境：支持主流深度学习框架训练 配件：鼠标、键盘及显示	台	1	

2、工作任务

OCR 文字识别技术是搭建视觉图像和计算机文本操作的桥梁，它基于计算机视觉和人工智能技术，可以将手写或印刷的字符图像转换成可编辑或搜索的文本。OCR 技术已经成功应用于各个领域。在社会业务中，被广泛的用于企业的文档处理和管理。现代企业的日常工作涉及到大量的文档处理，如果采用传统的手写或印刷方式，不仅效率低下，还容易出现错误。采用 OCR 技术，可以极大地提高文档处理效率和准确性。

某 RPA（机器人流程自动化）系统可以帮助业务人员进行网页表单数据的自动填写，但是每次填写都需要先输入数字验证码，这个操作会使流程造成一定的

阻塞和效率下降。因此，系统开发人员想要利用人工智能技术准确识别验证码数字，通过识别验证码图片得到验证码数字，将人工填写验证码的步骤优化为自动填写。

(1) 训练模型，打开 “模型训练与测试” 项目中 `train.ipynb` 文件，需要完成如下任务：

1 按照 notebook 上的说明运行代码准备好训练环境，利用 `python config` 目录下的 `train.py` 文件进行模型训练，等待模型训练完成，同时需输出训练时长（单位为秒）。

2 将此部分代码截图保存至 “提交资料/题目 2 模型训练阶段.docx” 文件中。同时下载模型训练第 3 轮的结果 (`epoch_3.pth`)，保存至 “提交资料/题目 2/模型训练结果” 中。

(2) 测试模型，运行 `train.ipynb` 文件中的模型测试部分程序，完成如下任务：

1 按照 notebook 上的说明加载 “model 文件夹” 中的第 3 轮训练参数，测试模型是否能准确识别 `images` 文件夹中验证码图像。

2 请将测试结果记录在 “提交资料/题目 2 模型训练.docx” 文件中，需将图像测试结果粘贴到文档中，若测试不准确，请分析原因并提出解决办法。

3、技能要求

- (1) 模型训练能力
- (2) 模型分析管理能力
- (3) 编码验证模型能力
- (4) 模型结果分析能力

4、质量指标

- (1) 训练任务的成功运行
- (2) 模型结果的准确性

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：9.3.1

试题名称：数据质检和流程设计

考核时间：20min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10 开发语言:Python 开发工具：Python 编译 配件：鼠标、键盘及显示	台	1	

2、工作任务

OCR 文字识别技术是搭建视觉图像和计算机文本操作的桥梁，它基于计算机视觉和人工智能技术，可以将手写或印刷的字符图像转换成可编辑或搜索的文本。OCR 技术已经成功应用于各个领域。在社会业务中，被广泛的用于企业的文档处理和管理。现代企业的日常工作涉及到大量的文档处理，如果采用传统的手写或印刷方式，不仅效率低下，还容易出现错误。采用 OCR 技术，可以极大地提高文档处理效率和准确性。

某 RPA（机器人流程自动化）系统可以帮助业务人员进行网页表单数据的自动填写，但是每次填写都需要先输入数字验证码，这个操作会使流程造成一定的阻塞和效率下降。因此，系统开发人员想要利用人工智能技术准确识别验证码数字，通过识别验证码图片得到验证码数字，将人工填写验证码的步骤优化为自动

填写。

(1) 如何对验证码数据集标注结果进行质检，至少描述 2 种质检方法和流程，并应描述 每种方法的优劣势。内容提交至“提交资料/题目 3/智能系统设计.docx”文件中。

3、技能要求

- (1) 样本质检方法掌握
- (2) 样本质检流程掌握
- (3) 标注方法掌握
- (4) 编码能力
- (5) 数据理解能力

4、质量指标

- (1) 质检方法回答全面性
- (2) 质检流程的正确性

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：9.3.2

试题名称：数据处理

考核时间：40min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10 开发语言：Python 开发工具：Python 编译工具； 配件：鼠标、键盘及显示	台	1	

2、工作任务

OCR 文字识别技术是搭建视觉图像和计算机文本操作的桥梁，它基于计算机视觉和人工智能技术，可以将手写或印刷的字符图像转换成可编辑或搜索的文本。OCR 技术已经成功应用于各个领域。在社会业务中，被广泛的用于企业的文档处理和管理。现代企业的日常工作涉及到大量的文档处理，如果采用传统的手写或印刷方式，不仅效率低下，还容易出现错误。采用 OCR 技术，可以极大地提高文档处理效率和准确性。

某 RPA（机器人流程自动化）系统可以帮助业务人员进行网页表单数据的自

动填写，但是每次填写都需要先输入数字验证码，这个操作会使流程造成一定的阻塞和效率下降。因此，系统开发人员想要利用人工智能技术准确识别验证码数字，通过识别验证码图片得到验证码数字，将人工填写验证码的步骤优化为自动填写。

(1) 打开“智能系统设计”文件夹，将‘captcha’文件夹中图像样本数据格式修改为“jpg”。将此部分代码截图保存至“提交资料/题目 3/ 智能系统设计.docx”文件中。

(2) 打开“智能系统设计”文件夹，将‘captcha’文件夹中图像样本名称重新定义，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“提交资料 /题目 2 业务分析.docx”文件中。下载处理好的数据集，将其保存至“提交资料/题目 3/智能系统设计”文件夹中。

3、技能要求

- (1) 编码能力
- (2) 数据处理能力

4、质量指标

- (1) 代码的正确性
- (2) 数据处理结果的准确性

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：9.4.1

试题名称：人工智能训练师培训能力考核

考核时间：12min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

(1) 业务背景

随着人工智能技术在各领域的应用不断深入，人工智能训练师这一职业不断发展壮大，作为一名三级/高级人工智能训练师，既需要熟练使用各种工具完成数据采集、标注、训练和分析，也需要参与人工智能训练师五级/初级工、四级/中级培训工作。

(2) 业务目标

您将作为一名三级/高级人工智能训练师，参与人工智能训练师五级/初级工、四级/中级培训工作。有能力完成培训讲义编写、开展知识和技术培训、指导解决数据采集/处理和数据标准问题等多项工作。

(3) 考试内容

能够编写初级培训讲义；能够对五级/初级工、四级/中级工开展知识和技术培训。

①如果安排你撰写人工智能训练师初级培训讲义，你觉得讲义应该重点讲解哪些方面的知识？

3、技能要求

- (1) 能够编写初级培训讲义
- (2) 能够对五级/初级工、四级/中级工开展知识和技术培训

4、质量指标

- (1) 包括但不限于：数据采集和处理、数据标注、智能系统运维

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01	1分	数据采集和处理，回答正确得1分	回答内容相似即可得分，不要求答案表述一模一样	1分
02	1分	数据标注，回答正确得1分	回答内容相似即可得分，不要求答案表述一模一样	1分
03	1分	智能系统运维，回答正确得1分	回答内容相似即可得分，不要求答案表述一模一样	1分

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：9.4.2

试题名称：人工智能训练师指导能力考核

考核时间：8min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

(1) 业务背景

随着人工智能技术在各领域的应用不断深入，人工智能训练师这一职业不断发展壮大，作为一名三级/高级人工智能训练师，既需要熟练使用各种工具完成数据采集、标注、训练和分析，也需要参与人工智能训练师五级/初级工、四级/中级指导工作。

(2) 业务目标

您将作为一名三级/高级人工智能训练师，参与人工智能训练师五级/初级工、四级/中级指导工作。有能力完成培训讲义编写、开展知识和技术培训、指导解决数据采集/处理和数据标准问题等多项工作。

(3) 考试内容

能够指导五级/初级工、四级/中级工解决数据采集/处理问题；能够指导五级/初级工、四级/中级工解决数据标注问题。

①人工智能训练师五级/初级工、四级/中级工解决数据采集、处理问题、解决数据标注问题使用的工具有哪些？

3、技能要求

- (1) 能够指导五级/初级工、四级/中级工解决数据采集/处理问题
- (2) 能够指导五级/初级工、四级/中级工解决数据标注问题

4、质量指标

(1) 一类是公开通用的标注工具，如：NumPy、Pandas、Excel、Word；还有一类是企业开发的标注工具：达观研发标注文档的 IDP 智能文档审阅系统，还有标注图片的 OCR 标注器

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01	1分	一类是公开通用的标注工具，如：NumPy、Pandas、Excel、Word。回答出通用标注工具得1分	回答结果符合实际要求即可，不要求和标准答案一模一样	1分
02	1分	还有一类是企业开发的标注工具：达观研发标注文档的 IDP 智能文档审阅系统，还有标注图片的 OCR 标注器。回答出企业开发标注工具得1分	回答结果符合实际要求即可，不要求和标准答案一模一样	1分
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

人工智能训练师（3级）操作技能考核

试题单

准考证号：

试题代码：10.1.1

试题名称：人工智能训练师数据分析处理能力考核

考核时间：24 min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

(1) 业务背景

互联网金融飞速发展，使得个人金融理财变得越来越容易。按照艾瑞《2022年中国消费金融行业研究报告》的统计口径，消费信贷市场（不含信用卡透支余额）放款规模达到 20.2 万亿元，余额达到 5.8 万亿元。其中，互联网消费贷款余额约为 4 万亿元，消费贷市场线上化渗透率为 69.2%。

信用评级技术是一种对贷款申请人（信用卡申请人）做风险评估分值的统计模型。信用评级卡模型是一种成熟的预测方法，其在信用风险评估以及金融风险控制领域得到广泛应用。信用评级卡可以根据客户提供的资料、客户的历史数据、第三方平台数据（芝麻分、京东、微信等），对客户的信用进行评估。信用评级卡的建立是以对大量数据的统计分析结果为基础，具有较高的准确性和可靠性。

(2) 业务目标

你将作为一名人工智能训练师，需要为根据提供的数据集进行处理，使得数据集能够满足信用卡评分模型的训练需求。

需要理解业务的含义，根据关键信息的特征，对原始数据进行处理，使后续模型训练可以正常进行。

(3) 考试内容

数据集为训练集，目的是开发一个申请的评分卡模型，对未来一段时间内借贷人出现违约的概率进行预测，对客户信用进行评估打分。

训练集样本数据一共 150000 条，10 个自变量，1 个因变量 (SeriousDlqin2yrs)。详细说明如下：

变量名	描述	类型
SeriousDlqin2yrs	是否有超过90天或更长时间逾期未还的行为	Y/N
RevolvingUtilizationOfUnsecuredLines	可用额度比值，信用卡和个人信用额度（不包括房地产和分期付款债务，如汽车贷款）的总余额除以信用额度总和。	percentage
age	借款人年龄	integer
NumberOfTime30-59DaysPastDueNotWorse	借款人历史有出现逾期30~59天记录的次数，但两年内无更严重逾期	integer
DebtRatio	每月偿还债务（比如，赡养费，生活费）除以月总收入	percentage
MonthlyIncome	月收入	real
NumberOfOpenCreditLinesAndLoans	未偿还贷款数量（汽车贷款或抵押贷款等分期付款）和信贷额度（如信用卡）	integer
NumberOfTimes90DaysLate	借款人历史有出现逾期90天或以上的记录的次数	integer
NumberRealEstateLoansOrLines	抵押贷款和房地产贷款数量，包括房屋净值贷款额度	integer
NumberOfTime60-89DaysPastDueNotWorse	借款人历史有出现逾期60~89天记录的次数，但两年内无更严重逾期	integer
NumberOfDependents	家庭中的家属人数（配偶、子女等）	integer

数据探索分析，获取样本总体的大概情况，分析以下三个问题，为后续分析提供目标和方向。

(a) 借款人逾期超过 90 天的比例，即借款客户的整体质量情况？

我们定义逾期超过 90 天的为坏客户，反之为好客户，统计好客户与坏客户数量，及坏客户占比。

(b) 借款人月收入分布对违约率的影响，收入区间设置为：低于 5000，5000~10000，10000~15000，高于 15000，并完成以下表格。

	违约率	客户数
[0, 5000)		
[5000, 10000)		
[10000, 15000]		
[15000, +∞)		

(c) 年龄对违约率的影响，年龄区间设置为：低于 25 岁，26 岁-35 岁，36 岁-55 岁，46 岁-55 岁，56 岁-65 岁，高于 65 岁，并完成以下表格。

	违约率	客户数
低于 25 岁		
26 岁-35 岁		
36 岁-55 岁		
46 岁-55 岁		
56 岁-65 岁		
高于 65 岁		

3、技能要求

- (1) 能够结合人工智能技术要求和业务特征，设计整套业务数据采集流程。
- (2) 能够结合人工智能技术要求和业务特征，设计整套业务数据处理流程。
- (3) 能够结合人工智能技术要求和业务特征，设计整套业务数据审核流程。

4、质量指标

- (1) 深刻理解数据，细致分析。

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

人工智能训练师（3级）操作技能考核

试题单

准考证号：

试题代码：10.1.2

试题名称：人工智能训练师数据分析处理能力考核

考核时间：16 min

5、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

6、工作任务

(1) 业务背景

互联网金融飞速发展，使得个人金融理财变得越来越容易。按照艾瑞《2022年中国消费金融行业研究报告》的统计口径，消费信贷市场（不含信用卡透支余额）放款规模达到20.2万亿元，余额达到5.8万亿元。其中，互联网消费贷款余额约为4万亿元，消费贷市场线上化渗透率为69.2%。

信用评级技术是一种对贷款申请人（信用卡申请人）做风险评估分值的统计模型。信用评级卡模型是一种成熟的预测方法，其在信用风险评估以及金融风险控制领域得到广泛应用。信用评级卡可以根据客户提供的资料、客户的历史数据、第三方平台数据（芝麻分、京东、微信等），对客户的信用进行评估。信用评级卡的建立是以对大量数据的统计分析结果为基础，具有较高的准确性和可靠性。

(2) 业务目标

你将作为一名人工智能训练师，需要为根据提供的数据集进行处理，使得数据集能够满足信用卡评分模型的训练需求。

需要理解业务的含义，根据关键信息的特征，对原始数据进行处理，使后续模型训练可以正常进行。

(3) 考试内容

数据集为训练集，目的是开发一个申请的评分卡模型，对未来一段时间内借贷人出现违约的概率进行预测，对客户信用进行评估打分。

训练集样本数据一共 150000 条，10 个自变量，1 个因变量 (SeriousDlqin2yrs)。详细说明如下：

变量名	描述	类型
SeriousDlqin2yrs	是否有超过90天或更长时间逾期未还的行为	Y/N
RevolvingUtilizationOfUnsecuredLines	可用额度比值，信用卡和个人信用额度（不包括房地产和分期付款债务，如汽车贷款）的总余额除以信用额度总和。	percentage
age	借款人年龄	integer
NumberOfTime30-59DaysPastDueNotWorse	借款人历史有出现逾期30~59天记录的次数，但两年内无更严重逾期	integer
DebtRatio	每月偿还债务（比如，赡养费，生活费）除以月总收入	percentage
MonthlyIncome	月收入	real
NumberOfOpenCreditLinesAndLoans	未偿还贷款数量（汽车贷款或抵押贷款等分期付款）和信贷额度（如信用卡）	integer
NumberOfTimes90DaysLate	借款人历史有出现逾期90天或以上的记录的次数	integer
NumberRealEstateLoansOrLines	抵押贷款和房地产贷款数量，包括房屋净值贷款额度	integer
NumberOfTime60-89DaysPastDueNotWorse	借款人历史有出现逾期60~89天记录的次数，但两年内无更严重逾期	integer
NumberOfDependents	家庭中的家属人数（配偶、子女等）	integer

数据预处理，包括异常值处理、缺失值处理等，将原始数据转化为可用作模型开发的格式化数据：

异常值处理，根据第一步的数据分布，我们已经发现了一些特征里面的异常值，请剔除以下异常值，并给出剔除后训练数据还剩多少条：

- (a) 可用额度比值大于 1
- (b) 年龄小于 18 岁
- (c) 逾期 30~59 天次数大于等于 80
- (d) 逾期 60~89 天次数大于等于 80
- (e) 逾期超过 90 天次数大于等于 80

缺失值处理，

(a) 月收入数据缺失达到了 20%，如果直接丢弃，可能会导致模型效果受影响，请使用随机森林模型对其进行填充。

(b) 其余有缺失数据的直接丢弃，给出最后剩余训练数据条数。

7、技能要求

- (1) 能够结合业务知识，识别业务流程中单一模块的问题。
- (2) 能够结合人工智能技术设计业务模块优化方案并推动实现。

8、质量指标

(1) 深刻理解数据，细致分析。

人工智能训练师（3级）操作技能考核

试题单

准考证号：

试题代码：10.1.2

试题名称：人工智能训练师数据分析处理能力考核

考核时间：16 min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

(1) 业务背景

互联网金融飞速发展，使得个人金融理财变得越来越容易。按照艾瑞《2022年中国消费金融行业研究报告》的统计口径，消费信贷市场（不含信用卡透支余额）放款规模达到 20.2 万亿元，余额达到 5.8 万亿元。其中，互联网消费贷款余额约为 4 万亿元，消费贷市场线上化渗透率为 69.2%。

信用评级技术是一种对贷款申请人（信用卡申请人）做风险评估分值的统计模型。信用评级卡模型是一种成熟的预测方法，其在信用风险评估以及金融风险控制领域得到广泛应用。信用评级卡可以根据客户提供的资料、客户的历史数据、第三方平台数据（芝麻分、京东、微信等），对客户的信用

进行评估。信用评分卡的建立是以对大量数据的统计分析结果为基础，具有较高的准确性和可靠性。

(2) 业务目标

你将作为一名人工智能训练师，需要为根据提供的数据集进行处理，使得数据集能够满足信用卡评分模型的训练需求。

需要理解业务的含义，根据关键信息的特征，对原始数据进行处理，使后续模型训练可以正常进行。

(3) 考试内容

数据集为训练集，目的是开发一个申请的评分卡模型，对未来一段时间内借贷人出现违约的概率进行预测，对客户信用进行评估打分。

训练集样本数据一共 150000 条，10 个自变量，1 个因变量 (SeriousDlqin2yrs)。详细说明如下：

变量名	描述	类型
SeriousDlqin2yrs	是否有超过90天或更长时间逾期未还的行为	Y/N
RevolvingUtilizationOfUnsecuredLines	可用额度比值，信用卡和个人信用额度（不包括房地产和分期付款债务，如汽车贷款）的总余额除以信用额度总和。	percentage
age	借款人年龄	integer
NumberOfTime30-59DaysPastDueNotWorse	借款人历史有出现逾期30~59天记录的次数，但两年内无更严重逾期	integer
DebtRatio	每月偿还债务（比如，赡养费，生活费）除以月收入	percentage
MonthlyIncome	月收入	real
NumberOfOpenCreditLinesAndLoans	未偿还贷款数量（汽车贷款或抵押贷款等分期付款）和信贷额度（如信用卡）	integer
NumberOfTimes90DaysLate	借款人历史有出现逾期90天或以上的记录的次数	integer
NumberRealEstateLoansOrLines	抵押贷款和房地产贷款数量，包括房屋净值贷款额度	integer
NumberOfTime60-89DaysPastDueNotWorse	借款人历史有出现逾期60~89天记录的次数，但两年内无更严重逾期	integer
NumberOfDependents	家庭中的家属人数（配偶、子女等）	integer

数据预处理，包括异常值处理、缺失值处理等，将原始数据转化为可用作模型开发的格式化数据；

异常值处理，根据第一步的数据分布，我们已经发现了一些特征里面的异常值，请剔除以下异常值，并给出剔除后训练数据还剩多少条：

- (a) 可用额度比值大于 1
- (b) 年龄小于 18 岁
- (c) 逾期 30~59 天次数大于等于 80
- (d) 逾期 60~89 天次数大于等于 80
- (f) 逾期超过 90 天次数大于等于 80

缺失值处理，

- (a) 月收入数据缺失达到了 20%，如果直接丢弃，可能会导致模型效果受影响，请使用随机森林模型对其进行填充。
- (b) 其余有缺失数据的直接丢弃，给出最后剩余训练数据条数。

3、技能要求

- (1) 能够结合业务知识，识别业务流程中单一模块的问题。
- (2) 能够结合人工智能技术设计业务模块优化方案并推动实现。

4、质量指标

- (1) 深刻理解数据，细致分析。

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

人工智能训练师（3级）操作技能考核

试题单

准考证号：

试题代码： 10.2.1

试题名称：人工智能训练师模型训练能力考核

考核时间： 25 min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

业务背景

互联网金融飞速发展，使得个人金融理财变得越来越容易。按照艾瑞《2022年中国消费金融行业研究报告》的统计口径，消费信贷市场（不含信用卡透支余额）放款规模达到20.2万亿元，余额达到5.8万亿元。其中，互联网消费贷款余额约为4万亿元，消费贷市场线上化渗透率为69.2%。

信用评分技术是一种对贷款申请人（信用卡申请人）做风险评估分值的统计模型。信用评分卡模型是一种成熟的预测方法，其在信用风险评估以及金融风险控制领域得到广泛应用。信用评分卡可以根据客户提供的资料、客户的历史数据、第三方平台数据（芝麻分、京东、微信等），对客户的信用进行评估。信用评分卡的建立是以对大量数据的统计分析结果为基础，具有较高的准确性和可靠性。

业务目标

你将作为一名人工智能训练师，希望你能够根据提供的训练数据（已经完成预处理，可用直接使用），选择合适的特征。

考试内容

数据集为训练集（已经完成预处理，可用直接使用），目的是开发一个申请的评分卡模型，对未来一段时间内借贷人出现违约的概率进行预测，对客户信用进行评估打分。

训练集样本数据一共一万多条，10个自变量，1个因变量（SeriousDlqin2yrs）。详细说明如下：

变量名	描述	类型
SeriousDlqin2yrs	是否有超过90天或更长逾期未还的行为	Y/N
RevolvingUtilizationOfUnsecuredLines	可用额度比值，信用卡和个人信用额度（不包括房地产和分期付款债务，如汽车贷款）的总余额除以信用额度总和。	percentage
age	借款人年龄	integer
NumberOfTime30-59DaysPastDueNotWorse	借款人历史有出现逾期30~59天记录的次数，但两年内无更严重逾期	integer
DebtRatio	每月偿还债务（比如，赡养费，生活费）除以月总收入	percentage
MonthlyIncome	月收入	real
NumberOfOpenCreditLinesAndLoans	未偿还贷款数量（汽车贷款或抵押贷款等分期付款）和信贷额度（如信用卡）	integer
NumberOfTimes90DaysLate	借款人历史有出现逾期90天或以上的记录的次数	integer
NumberRealEstateLoansOrLines	抵押贷款和房地产贷款数量，包括房屋净值贷款额度	integer
NumberOfTime60-89DaysPastDueNotWorse	借款人历史有出现逾期60~89天记录的次数，但两年内无更严重逾期	integer
NumberOfDependents	家庭中的家属人数（配偶、子女等）	integer

变量选择，好的特征选择能够提升模型的性能，更能帮助我们理解数据的特点、底层结构，这对进一步改善模型、算法都有着重要作用。

(a) **分箱处理**，变量分箱（binning）是对连续变量离散化（discretization）的一种称呼。信用评分卡开发中一般有常用的等距分段、等深分段、最优分段。其中等距分段（Equal length intervals）是指分段的区间是一致的，比如年龄以十年作为一个分段；等深分段（Equal frequency intervals）是先确定分段数量，然后令每个分段中数据数量大致相等；最优分段（Optimal Binning）又叫监督离散化

（supervised discretization），使用递归划分（Recursive Partitioning）将连续变量分为分段，背后是一种基于条件推断查找较佳分组的算法。

在这里，我们选择对连续变量进行最优分段，在连续变量的分布不满足最优分段的要求时，再考虑对连续变量进行等距分段。

先介绍一下 WOE (Weight of Evidence) 和 IV (information value)

WOE (Weight of Evidence)

对分箱后的每组数据求 woe。公式如下：

$$W_i = \ln\left(\frac{p(y_i)}{p(n_i)}\right)$$

其中， $p(y_i)$ 表示当前条件下风险用户占总风险用户的比例， $p(n_i)$ 表示当前条件下无风险用户占无风险用户的比例。

对上述公式变换后得到：

$$W_i = \ln\left(\frac{p(y_i)}{p(n_i)}\right) = \ln\left(\frac{y_i/y_t}{n_i/n_t}\right) = \ln\left(\frac{y_i/n_i}{y_t/n_t}\right)$$

WOE 表示当前分组里风险用户和无风险用户的比值，和所有样本中风险用户和无风险用户的比值的差异。WOE 越大，说明这种差异越大，这个分组里样本有风险的可能性就越大。

IV (information value)

衡量某一个变量的信息量。

计算公式为：

$$W_i = \ln\left(\frac{p(y_i)}{p(n_i)}\right) = \ln\left(\frac{y_i/y_t}{n_i/n_t}\right) = \ln\left(\frac{y_i/n_i}{y_t/n_t}\right)$$

IV 值对应的预测能力

IV	预测能力
<0.03	无预测能力
0.03~0.09	低
0.1~0.29	中
0.3~0.49	高
>=0.5	极高

最优分箱会根据 IV 值来调整分箱结构并重新计算 WOE 和 IV，直到 IV 达到最大值，此时的分箱效果最好。

请对 `RevolvingUtilizationOfUnsecuredLines`、`age`、`DebtRatio` 和 `MonthlyIncome` 使用最优分段，剩余变量使用等距分段。

(b) 相关性分析

用经过清洗后的数据查看变量间的相关性。相关程度用相关系数 r 表示， $-1 \leq r < 1$ ， r 的绝对值越大，表示变量之间的相关程度越高， r 为负数时，表示一个变量的增加可能引起另一个变量的减少，此时，叫做负相关。

请分析训练数据中，给出各变量间的相关性，并画出热力图，剔除相关系数大于 0.6 的变量，如果相关性较小，则不需要操作。

(c) IV 筛选

根据自动分箱函数中提供的 `iv` 值，剔除显著较低（小于 0.1）的变量，请给出需要剔除哪些变量？

3、技能要求

(1) 能够结合人工智能技术要求和业务特征，设计数据清洗和标注流程。

(2) 能够结合人工智能技术要求和业务特征，制定数据清洗和标注规范。

4、质量指标

(1) 深入理解业务，训练符合业务需求的模型

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

人工智能训练师（3级）操作技能考核

试题单

准考证号：

试题代码：10.2.2

试题名称：人工智能训练师模型训练能力考核

考核时间：35 min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

业务背景

互联网金融飞速发展，使得个人金融理财变得越来越容易。按照艾瑞《2022年中国消费金融行业研究报告》的统计口径，消费信贷市场（不含信用卡透支余额）放款规模达到20.2万亿元，余额达到5.8万亿元。其中，互联网消费贷款余额约为4万亿元，消费贷市场线上化渗透率为69.2%。

信用评分技术是一种对贷款申请人（信用卡申请人）做风险评估分值的统计模型。信用评分卡模型是一种成熟的预测方法，其在信用风险评估以及金融风险控制领域得到广泛应用。信用评分卡可以根据客户提供的资料、客户的历史数据、第三方平台数据（芝麻分、京东、微信等），对客户的信用进行评估。信用评分卡的建立是以对大量数据的统计分析结果为基础，具有较高的准确性和可靠性。

业务目标

你将作为一名人工智能训练师，希望你能够根据提供的训练数据（已经完成预处理，可用直接使用），训练一个信用卡评分模型。

考试内容

数据集为训练集（已经完成预处理，可用直接使用），目的是开发一个申请的评分卡模型，对未来一段时间内借贷人出现违约的概率进行预测，对客户信用进行评估打分。

训练集样本数据一共一万多条，10个自变量，1个因变量（SeriousDlqin2yrs）。详细说明如下：

变量名	描述	类型
SeriousDlqin2yrs	是否有超过90天或更长时间逾期未还的行为	Y/N
RevolvingUtilizationOfUnsecuredLines	可用额度比值，信用卡和个人信用额度（不包括房地产和分期付款债务，如汽车贷款）的总余额除以信用额度总和。	percentage
age	借款人年龄	integer
NumberOfTime30-59DaysPastDueNotWorse	借款人历史有出现逾期30~59天记录的次数，但两年内无更严重逾期	integer
DebtRatio	每月偿还债务（比如，赡养费，生活费）除以月总收入	percentage
MonthlyIncome	月收入	real
NumberOfOpenCreditLinesAndLoans	未偿还贷款数量（汽车贷款或抵押贷款等分期付款）和信贷额度（如信用卡）	integer
NumberOfTimes90DaysLate	借款人历史有出现逾期90天或以上的记录的次数	integer
NumberRealEstateLoansOrLines	抵押贷款和房地产贷款数量，包括房屋净值贷款额度	integer
NumberOfTime60-89DaysPastDueNotWorse	借款人历史有出现逾期60~89天记录的次数，但两年内无更严重逾期	integer
NumberOfDependents	家庭中的家属人数（配偶、子女等）	integer

请使用 Logistic 模型，选取所需特征进行模型训练，为了方便起见，这里可以不采用 WOE 转换，直接使用原始数据进行训练。

要求在训练时，剔除以下特征：

- DebtRatio
- MonthlyIncome
- NumberOfOpenCreditLinesAndLoans
- NumberRealEstateLoansOrLines
- NumberOfDependents

3、技能要求

- （1）能够维护日常训练集与测试集。
- （2）能使用工具对算法进行训练。
- （3）能够使用测试工具对人工智能产品的使用进行测试。
- （4）能够对测试结果进行分析，编写测试报告。
- （5）能够运用工具，分析算法中错误案例产生的原因并进行纠正。

4、质量指标

- （1）深入理解业务，训练符合业务需求的模型

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

工智能训练师（3级）操作技能考核

试题单

准考证号：

试题代码：10.3.1

试题名称：人工智能训练师模型评估能力考核

考核时间：25 min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

业务背景

互联网金融飞速发展，使得个人金融理财变得越来越容易。按照艾瑞《2022年中国消费金融行业研究报告》的统计口径，消费信贷市场（不含信用卡透支余额）放款规模达到20.2万亿元，余额达到5.8万亿元。其中，互联网消费贷款余额约为4万亿元，消费贷市场线上化渗透率为69.2%。

信用评分技术是一种对贷款申请人（信用卡申请人）做风险评估分值的统计模型。信用评分卡模型是一种成熟的预测方法，其在信用风险评估以及金融风险控制领域得到广泛应用。信用评分卡可以根据客户提供的资料、客户的历史数据、第三方平台数据（芝麻分、京东、微信等），对客户的信用进行评估。信用评分卡的建立是以对大量数据的统计分析结果为基础，具有较高的准确性和可靠性。

业务目标

你将作为人工智能训练师，我们会提供一个测试集和一个已经训练好的模型，你需要评估这个模型的好坏。

考试内容

数据集分为训练集（已经完成预处理，可用直接使用）和评估集，目的是开发一个申请的评分卡模型，对未来一段时间内借贷人出现违约的概率进行预测，对客户信用进行评估打分。

训练集样本数据一共一万多条，10个自变量，1个因变量（SeriousDlqin2yrs）。详细说明如下：

变量名	描述	类型
SeriousDlqin2yrs	是否有超过90天或更长时间逾期未还的行为	Y/N
RevolvingUtilizationOfUnsecuredLines	可用额度比值，信用卡和个人信用额度（不包括房地产和分期付款债务，如汽车贷款）的总余额除以信用额度总和。	percentage
age	借款人年龄	integer
NumberOfTime30-59DaysPastDueNotWorse	借款人历史有出现逾期30~59天记录的次数，但两年内无更严重逾期	integer
DebtRatio	每月偿还债务（比如，赡养费，生活费）除以月总收入	percentage
MonthlyIncome	月收入	real
NumberOfOpenCreditLinesAndLoans	未偿还贷款数量（汽车贷款或抵押贷款等分期付款）和信贷额度（如信用卡）	integer
NumberOfTimes90DaysLate	借款人历史有出现逾期90天或以上的记录的次数	integer
NumberRealEstateLoansOrLines	抵押贷款和房地产贷款数量，包括房屋净值贷款额度	integer
NumberOfTime60-89DaysPastDueNotWorse	借款人历史有出现逾期60~89天记录的次数，但两年内无更严重逾期	integer
NumberOfDependents	家庭中的家属人数（配偶、子女等）	integer

(a) Logistics 模型的常用评估指标有哪些？

(b) 逻辑回归为什么要对特征进行离散化？

3、技能要求

- (1) 能够对单一智能产品使用的数据进行全面分析，输出分析报告。
- (2) 能够对单一智能产品提出优化需求。
- (3) 能够为单一智能产品的应用设计智能解决方案。

4、质量指标

- (1) 深入理解模型和业务，合理评估模型

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

人工智能训练师（3级）操作技能考核

试题单

准考证号：

试题代码：10.3.2

试题名称：人工智能训练师模型评估能力考核

考核时间：35 min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

业务背景

互联网金融飞速发展，使得个人金融理财变得越来越容易。按照艾瑞《2022年中国消费金融行业研究报告》的统计口径，消费信贷市场（不含信用卡透支余额）放款规模达到20.2万亿元，余额达到5.8万亿元。其中，互联网消费贷款余额约为4万亿元，消费贷市场线上化渗透率为69.2%。

信用评分技术是一种对贷款申请人（信用卡申请人）做风险评估分值的统计模型。信用评分卡模型是一种成熟的预测方法，其在信用风险评估以及金融风险控制领域得到广泛应用。信用评分卡可以根据客户提供的资料、客户的历史数据、第三方平台数据（芝麻分、京东、微信等），对客户的信用进行评估。信用评分卡的建立是以对大量数据的统计分析结果为基础，具有较高的准确性和可靠性。

业务目标

你将作为人工智能训练师，我们会提供一个测试集和一个已经训练好的模型，你需要评估这个模型的好坏。

考试内容

数据集分为训练集（已经完成预处理，可用直接使用）和评估集，目的是开发一个申请的评分卡模型，对未来一段时间内借贷人出现违约的概率进行预测，对客户信用进行评估打分。

训练集样本数据一共一万多条，10个自变量，1个因变量（SeriousDlqin2yrs）。详细说明如下：

变量名	描述	类型
SeriousDlqin2yrs	是否有超过90天或更长时间逾期未还的行为	Y/N
RevolvingUtilizationOfUnsecuredLines	可用额度比值，信用卡和个人信用额度（不包括房地产和分期付款债务，如汽车贷款）的总余额除以信用额度总和。	percentage
age	借款人年龄	integer
NumberOfTime30-59DaysPastDueNotWorse	借款人历史有出现逾期30~59天记录的次数，但两年内无更严重逾期	integer
DebtRatio	每月偿还债务（比如，赡养费，生活费）除以月总收入	percentage
MonthlyIncome	月收入	real
NumberOfOpenCreditLinesAndLoans	未偿还贷款数量（汽车贷款或抵押贷款等分期付款）和信贷额度（如信用卡）	integer
NumberOfTimes90DaysLate	借款人历史有出现逾期90天或以上的记录的次数	integer
NumberRealEstateLoansOrLines	抵押贷款和房地产贷款数量，包括房屋净值贷款额度	integer
NumberOfTime60-89DaysPastDueNotWorse	借款人历史有出现逾期60~89天记录的次数，但两年内无更严重逾期	integer
NumberOfDependents	家庭中的家属人数（配偶、子女等）	integer

已经训练了一个Logistics回归模型，训练代码见train.py

请通过AUC来评估信用卡评分模型的拟合能力，

- 画出ROC曲线
- 计算AUC
- 简单评估模型质量
- 评估代码请在evaluate.py中补充。

3、技能要求

- 能够通过数据分析，找到单一场景下人工和智能交互的最优方式。
- 能够通过数据分析，设计单一场景下人工和智能交互的最优流程。

4、质量指标

- 深入理解模型和业务，合理评估模型

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01				
02				
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描	考评员评分	得分
----	----	-------	-------	----

		述	1	2	3	
S1						
S2						
S3						
S4						

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：10.4.1

试题名称：人工智能训练师培训能力考核

考核时间：12min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

(1) 业务背景

随着人工智能技术在各领域的应用不断深入，人工智能训练师这一职业不断发展壮大，作为一名三级/高级人工智能训练师，既需要熟练使用各种工具完成数据采集、标注、训练和分析，也需要参与人工智能训练师五级/初级工、四级/中级培训工作。

(2) 业务目标

您将作为一名三级/高级人工智能训练师，参与人工智能训练师五级/初级工、四级/中级培训工作。有能力完成培训讲义编写、开展知识和技术培训、指导解决数据采集/处理和数据标准问题等多项工作。

(3) 考试内容

能够编写初级培训讲义；能够对五级/初级工、四级/中级工开展知识和技术培训。

①一名合格的中级人工智能训练师，应该具备哪些能力？

3、技能要求

- (1) 能够编写初级培训讲义
- (2) 能够对五级/初级工、四级/中级工开展知识和技术培训

4、质量指标

- (1) 具备的能力包括但不限于：数据采集和处理、数据标注、智能系统运维

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01	1分	数据采集和处理，回答正确得1分	回答内容相似即可得分，不要求答案表述一模一样	1分
02	1分	数据标注，回答正确得1分	回答内容相似即可得分，不要求答案表述一模一样	1分
03	1分	智能系统运维，回答正确得1分	回答内容相似即可得分，不要求答案表述一模一样	1分

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

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：10.4.2

试题名称：人工智能训练师指导能力考核

考核时间：8min

1、场地设备要求

(1) 场地要求：指定考试场地

(2) 设备要求：

序号	名称	规格	单位	数量	备注
1	人工智能训练师主机	CPU：建议英特尔 i5 以上 内存：建议不少于 16GB 操作系统：Windows10	台	1	

2、工作任务

(1) 业务背景

随着人工智能技术在各领域的应用不断深入，人工智能训练师这一职业不断发展壮大，作为一名三级/高级人工智能训练师，既需要熟练使用各种工具完成数据采集、标注、训练和分析，也需要参与人工智能训练师五级/初级工、四级/中级指导工作。

(2) 业务目标

您将作为一名三级/高级人工智能训练师，参与人工智能训练师五级/初级工、四级/中级指导工作。有能力完成培训讲义编写、开展知识和技术培训、指导解决数据采集/处理和数据标准问题等多项工作。

(3) 考试内容

能够指导五级/初级工、四级/中级工解决数据采集/处理问题；能够指导五级/初级工、四级/中级工解决数据标注问题。

①常见的数据标注工具有哪些？

3、技能要求

- (1) 能够指导五级/初级工、四级/中级工解决数据采集/处理问题
- (2) 能够指导五级/初级工、四级/中级工解决数据标注问题

4、质量指标

(1) 一类是公开通用的标注工具，如：NumPy、Pandas、Excel、Word；还有一类是企业开发的标注工具：达观研发标注文档的 IDP 智能文档审阅系统，还有标注图片的 OCR 标注器

客观评分表/结果评分表

序号	配分	评分细则描述	规定或标称值	得分
01	1分	一类是公开通用的标注工具，如：NumPy、Pandas、Excel、Word。回答出通用标注工具得1分	回答结果符合实际要求即可，不要求和标准答案一模一样	1分
02	1分	还有一类是企业开发的标注工具：达观研发标注文档的 IDP 智能文档审阅系统，还有标注图片的 OCR 标注器。回答出企业开发标注工具得1分	回答结果符合实际要求即可，不要求和标准答案一模一样	1分
03				

注：如有分解为子项目评分表的，请分列客观评分表，并注明子评分表名称。

主观评分表/过程评分表

序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
S1						
S2						
S3						
S4						

