

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：1.1.1

试题名称：客服会话业务数据质量检测

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

智能客服机器人能通过识别理解用户意图，将客户的需求进行合理分类、传递，从知识库中寻求答案，或者直接回答客户问题，或者辅助人工客服智能应答。本项目中，你作为一名人工智能训练师，将对预处理后的客服会话数据进行审核，并制定相应的采集和处理规范。

- (1) 请对文件(dataForReview.csv)中的预处理后会话数据进行审核，指出其中存在的问题。（4分）
- (2) 请基于标注后用于模型训练测试的需求，为会话数据制定采集规范。（8分）
- (3) 请基于标注后用于模型训练测试的需求，为会话数据制定处理规范。（8分）

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

无

人工智能训练师（四级）操作技能考核

试题评分表及答案

准考证号：

试题代码：1.1.1

试题名称：客服会话业务数据质量检测

考核时间：30 min

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	正确指出上述存在问题中的一个：+1分 正确指出上述存在问题中的三个及以上：+4分 答错不扣分				
(2)	8分	提供一个合理的采集规范：+2分 提供四个及以上合理的采集规范：+8分 提供一个及以上不合理的采集规范：-2分 加总得分在0到8分之间				
(3)	8分	提供一个合理的处理规范：+2分 提供三个及以上合理的处理规范：+8分 提供一个及以上不合理的处理规范：-2分 加总得分在0到8分之间				

参考答案（尽量将细则内容写在上面的表格内，写不下的可另写，但要具体可评判）

(1) 请对文件(dataForReview.csv)中的预处理后会话数据进行审核，指出其中存在的问题。(4分)

参考答案：

1. 存在系统消息。

由 服务助手 转交给***

帮您转接到售后（相关）客服

2. 数据分列异常。第13414行的message可能存在空格。如下图所示：

1	dt	store_id	assistant_id	buyer_id	direction	message	message_time
13414		20230501sid_005	assistant_011	user_002861	ASSISTANT_TO_BUYER	亲亲，这款宝贝口碑非常好，希望你能喜欢 [url]	2023/5/1 16:03

3. 存在缺失值。第73行存在空白行，第13715行message数据缺失。如下图所示

示：

1	dt	store_id	assistant_id	buyer_id	direction	message	message_time
73							
13715		20230501 sid_005	assistant_011	user_002934	ASSISTANT_TO_BUYER		2023/5/1 10:39

4. 存在串行问题。第 16487 和第 16488 行串行。如下图所示：

1	dt	store_id	assistant_id	buyer_id	direction	message	message_time
16487		20230501 sid_005	assistant_012	user_003501	ASSISTANT_TO_BUYER		
16488	您可以先拍70A的回去对	2023/5/1 17:13					

5. 存在数据重复。

(2) 请基于标注后用于模型训练测试的需求，为会话数据制定采集规范。(8分)

参考答案：

1. 目的性原则（又称针对性原则）：需要明确采集数据的目的和用途，比如训练集用于模型训练，验证集用于模型调参，测试集用于模型评估，只有明确了目的，才能有针对性地采集数据，避免浪费时间和资源；
2. 可靠性原则：信息必须是真实对象或环境所产生的，必须保证信息来源是可靠的，必须保证采集的信息能反映真实的状况；
3. 完整性原则（又称全面性原则）：必须按照一定的标准要求，尽可能全面地采集符合需求的信息；
4. 准确性原则：数据不存在错误或缺失，属于采集目的范畴内，对于模型训练测试有价值；
5. 有效性原则：数据采集是否符合逻辑，根据线上流量分布采集；
6. 及时性原则（又称实时性原则）：采集到的信息能够反映当前线上真实情况；
7. 安全性和隐私性原则：采集数据时，应保护数据的安全性和隐私性。

(3) 请基于标注后用于模型训练测试的需求，为会话数据制定处理规范。(8分)

处理规范：整理、归类、整合、汇总等，目的在于减少后续的数据清洗和标注的工作量。

1. 整理，对数据进行格式和名称的规范，保障后续的归类和汇总的有效性。；
2. 归类，把性质或参数相同类的数据划分在一起。归类原则：第一，所有的数据都必须涵盖全，不能遗留；第二，分类之间不允许重复和交叉；第三，同级次分类的维度要统一，颗粒度要一致；

3. 整合，按照归类的类别，把同类数据合并在一起；
4. 汇总，按照类别，对数据进行排列组合，获取有价值的信息。

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：1.1.2

试题名称：客服会话数据处理方法优化

考核时间：30 min

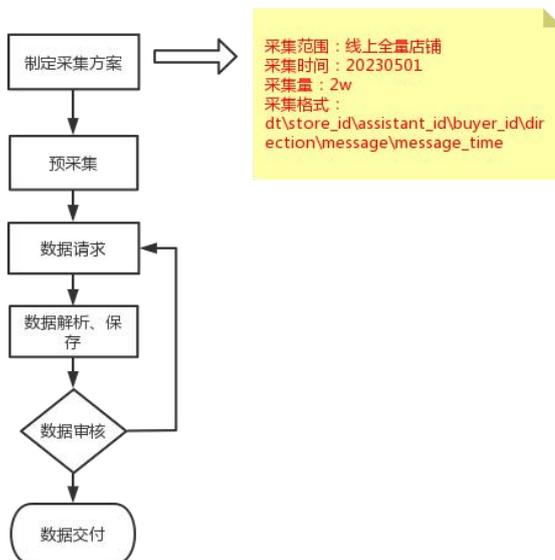
1、场地设备要求

- (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

智能客服机器人能通过识别理解用户意图，将客户的需求进行合理分类、传递，从知识库中寻求答案，或者直接回答客户问题，或者辅助人工客服智能应答。本项目中，你作为一名人工智能训练师，将对客服会话数据的采集处理流程提出优化建议。

- (1) 请阅读如下的会话数据采集流程，对采集流程提出优化建议。（5 分）



(2) 请阅读如下的会话数据处理流程，对处理流程提出优化建议。

会话数据处理流程：

第一步，将数据按 store_id 进行归类；

第二步，对 message 进行 unique 处理，得到 unique_count；

第三步，将采集的会话数据进行撞库处理，得到模型预测识别结果；

第四步，汇总时筛选出 direction 为 buyer_to_assistant 的数据。(5分)

3、技能要求

(1) 能够对业务数据采集流程提出优化建议

(2) 能够对业务数据处理流程提出优化建议

4、质量指标

无

人工智能训练师（四级）操作技能考核

试题评分表及答案

准考证号：

试题代码：1.1.2

试题名称：客服会话数据处理方法优化

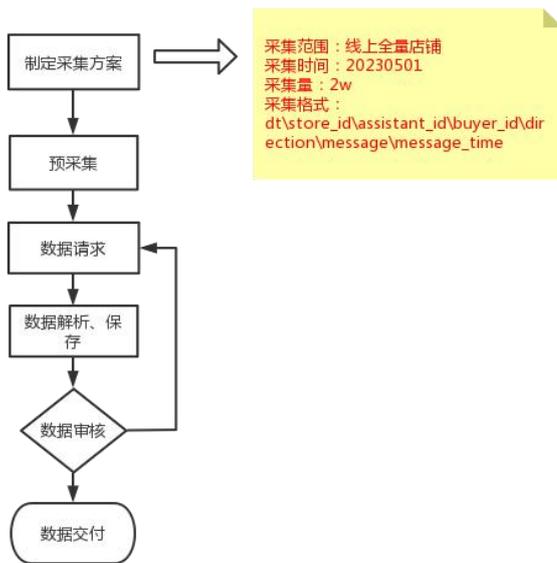
考核时间：30 min

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	正确给出类似上述优化需求中的一条：+2分 正确给出类似上述优化需求中的两条及以上：+5分 若给出一个及以上不合理的优化需求：-2分 (优化需求合理即可) 加总得分在0到5分之间				
(2)	5分	正确给出类似上述优化需求中的一条：+2分 正确给出类似上述优化需求中的两条及以上：+5分 若给出一个及以上不合理的优化需求：-2分 (优化需求合理即可) 加总得分在0到5分之间				

参考答案（尽量将细则内容写在上面的表格内，写不下的可另写，但要具体可评判）

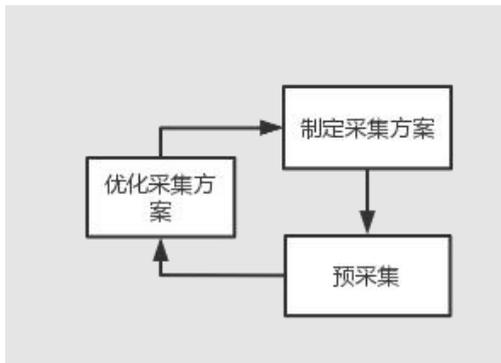
(1) 请阅读如下的会话数据采集流程，对采集流程提出优化建议。(5分)



参考答案:

可以有以下优化方向:

1. 制定采集方案后可进行预采集, 便于及时调整优化采集方案, 如下图所示:



2. 可按照线上流量分布采集, 使得采集到的数据接近线上真实分布;
3. 适当扩大数据采集的时间, 避免采集量不够, 初步制定采集时间为近一周;
4. 做好数据维护, 注意数据的时效性, 按月/按季更新采集会话数据。

(2) 请阅读如下的会话数据处理流程, 对处理流程提出优化建议。

会话数据处理流程:

第一步, 将数据按 store_id 进行归类;

第二步, 对 message 进行 unique 处理, 得到 unique_count;

第三步, 将采集的会话数据进行撞库处理, 得到模型预测识别结果;

第四步, 汇总时筛选出 direction 为 buyer_to_assistant 的数据。(5分)

参考答案：

可以有以下优化方向：

1. 首先需要进行数据整理，避免存在变量数据的格式不符合计算需求；
2. 按照 store_id 和 message 进行归类，便于后续标注以及汇总；
3. 按照模型预测结果过滤出无需标注的部分，例如系统消息等；
4. 按业务需求考虑过滤掉 unique_count 低且字符数过长（超过 50 字符）或者过短（少于 5 个字符）的数据。

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：1.2.1

试题名称：客服会话数据归类和定义

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

智能客服机器人能通过识别理解用户意图，将客户的需求进行合理分类、传递，从知识库中寻求答案，或者直接回答客户问题，或者辅助人工客服智能应答。本项目中，你作为一名人工智能训练师，将对已清洗完的数据进行分析，并根据分析结果进行数据归类和数据定义。

(1) 打开文件夹中的程序文件 2.1_clustering.ipynb，按下列要求及代码中的注释补全代码，并以原文件名保存在文件夹中。代码实现对已清洗完并获得相应 embedding 的会话数据进行 K-means 聚类分析。（10 分）

1. 导入 sikit-learn 包里的聚类方法
2. 读取 bert_embedding.csv 文件，查看数据
3. 使用 K-means 聚类算法，将数据聚为 3 类。

(2) 打开文件夹中的程序文件 2.1_clustering.ipynb，按下列要求及代码中的注释补全代码，找出最佳聚类分组，进行数据分类，提交分类后的数据。（10 分）

1. 计算 CH 系数，确认最佳 K 值。
2. 参考 dataSample.csv 文件，保存最佳聚类结果到 cluster_result.csv

(3) 请根据上述结果，对数据进行定义，以指导后续的数据标注；选取其中一

簇判断聚类结果是否合理，给出理由。（10分）

聚类结果定义示例：

请求改价

定义：买家请求卖家改价，或者询问能否改价，不包括咨询卖家是否已经改好价格

例句：那你给我改价格吧

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）能够根据数据内在关联和特征进行数据归类
- （3）能够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

人工智能训练师（四级）操作技能考核

试题评分表及答案

准考证号：

试题代码：1.2.1

试题名称：客服会话数据归类和定义

考核时间：30 min

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)	10分	正确补全代码，每空格2分		
(2)	10分	正确补全代码，每空格2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	
(3)	10分	每个簇给出合适名称及定义各2分。 聚类结果判断为合理或不合理均正确，给出的合理理由即可：+2分				

参考答案（尽量将细则内容写在上面的表格内，写不下的可另写，但要具体可评判）

(1) 打开文件夹中的程序文件 2.1_clustering.ipynb，按下列要求及代码中的注释补全代码，并以原文件名保存在文件夹中。代码实现对已清洗完并获得相应 embedding 的会话数据进行 K-means 聚类分析。（10分）

1. 导入 sikit-learn 包里的聚类方法
2. 读取 bert_embedding.csv 文件，查看数据
3. 使用 K-means 聚类算法，将数据聚为 3 类。

参考答案：

请查看 2.1_clustering[answer].ipynb

(2) 打开文件夹中的程序文件 2.1_clustering.ipynb, 按下列要求及代码中的注释补全代码, 找出最佳聚类分组, 进行数据分类, 提交分类后的数据。(10分)

1. 计算 CH 系数, 确认最佳 K 值。
2. 参考 dataSample.csv 文件, 保存最佳聚类结果到 cluster_result.csv

参考答案:

请查看 2.1_clustering[answer].ipynb

(3) 请根据上述结果, 对数据进行定义, 以指导后续的数据标注; 选取其中一簇判断聚类结果是否合理, 给出理由。(10分)

聚类结果定义示例:

请求改价

定义: 买家请求卖家改价, 或者询问能否改价, 不包括咨询卖家是否已经改好价格

例句: 那你给我改价格吧

参考答案:

簇 0: 请求多送: 买家请求多送赠品, 询问能否多送。

簇 1: 陈述中奖了: 买家陈述自己中奖了, 中了几等奖, 中奖渠道等

簇 2: 陈述没收到: 买家陈述自己没收到

簇 4: 议价: 买家询问能不能便宜, xx 钱买不买

簇 1 中的各问题聚为同一类是合理的, 因为都是陈述中奖相关问题。

簇 1 中的各问题聚为同一类是不合理的, 因为里面有陈述自己中奖的问题, 也有中奖要求免单的问题, 诉求不一样。

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：1.2.2

试题名称：客服会话标注数据审核

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

智能客服机器人能通过识别理解用户意图，将客户的需求进行合理分类、传递，从知识库中寻求答案，或者直接回答客户问题，或者辅助人工客服智能应答。本项目中，你作为一名人工智能训练师，对已标注的数据进行审核，并进行错误纠正及数据筛选。

- (1) 文件“标注结果_待审核.xlsx”包含两个 label 的标注结果，请对这份标注数据进行准确性和完整性审核，输出审核报告。（10 分）
- (2) 请列出审核过程中发现的三类错误并进行纠正，结果保存至“标注结果_已审核.xlsx”文件。（5 分）
- (3) 请根据审核结果，对数据进行筛选，选出适合用于模型训练测试的数据。结果保存至“标注结果_train.xlsx”文件。（5 分）

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

准考证号：

试题代码：1.2.2

试题名称：客服会话标注数据审核

考核时间：60 min

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)	10分	答对一点得5分，答错不扣分。		
(2)	5分	5条未标注数据填写正确的label：+2分 3条标签有误的数据更改为正确的label：+3分		
(3)	5分	删除错误的系统消息、客服回复数据，每条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。

参考答案（尽量将细则内容写在上面的表格内，写不下的可另写，但要具体可评判）

(1) 文件“标注结果_待审核.xlsx”包含两个label的标注结果，请对这份标注数据进行准确性和完整性审核，输出审核报告。（10分）

参考答案：

文件里的标注结果不符合准确性原则，含有错误标签名称，共3条；不符合完整性原则，有未标注数据，共5条

(2) 请列出审核过程中发现的三类错误并进行纠正，结果保存至“标注结果_已审核.xlsx”文件。（5分）

参考答案：

请查看 标注结果_已审核.xlsx

(3) 请根据审核结果，对数据进行筛选，选出适合用于模型训练测试的数据。

结果保存至“标注结果_train.xlsx”文件。(5分)

参考答案:

请查看 标注结果_train.xlsx

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：1.3.1

试题名称：客户体验管理系统维护

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。本项目中，作为一名人工智能训练师，你将进行该智能系统的知识与数据维护，并为其单一智能功能找到合适的应用场景。请参考该智能系统的操作手册（CEM 操作手册.pdf）进行以下试题的作答。

- (1) 请对质检规则进行维护，添加一条自定义质检规则，检测客服是否有仅发送单个标点符号。请用文字描述如何操作来添加这条规则。（3 分）
- (2) 客服培训案例的维护需要会话数据，请用文字描述添加为客服培训案例的会话数据的判断标准。（3 分）
- (3) 请为该智能系统的客服对话质检功能找到合适的应用场景，并用文字描述应用场景。（4 分）

3、技能要求

- (1) 能够维护智能系统所需知识
- (2) 能够维护智能系统所需数据

(3) 能够为单一智能产品找到合适应用场景

4、质量指标

无

人工智能训练师（四级）操作技能考核

试题评分表及答案

准考证号：

试题代码：1.3.1

试题名称：客户体验管理系统维护

考核时间：15 min

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	正确说明选择检测对象：+1分 说明质检关键词添加整句且正确描述至少一个标点符号：+2分				
(2)	3分	提供一个合理标准：+1分 提供两个及以上合理标准：+3分 提供一个及以上不合理标准：-1分 加总得分在0到3分之间				
(3)	4分	提供至少一个合理应用场景：+3分 提供两个及以上合理应用场景：+4分 提供一个及以上不合理应用场景：-2分 加总得分在0到4分之间				

参考答案（尽量将细则内容写在上面的表格内，写不下的可另写，但要具体可评判）

(1) 请对质检规则进行维护，添加一条自定义质检规则，检测客服是否有仅发送单个标点符号。请用文字描述如何操作来添加这条规则。（3分）

参考答案：

从“设置->规则设置”进入，选择“自定义规则”Tab页，点击“新建规则”，并选择检测对象为“只检测客服”，在质检关键词下添加若干整句，例如：

“？”“！”“……”“。”“，”“，”

(2) 客服培训案例的维护需要会话数据，请用文字描述添加为客服培训案例的会话数据的判断标准。（3分）

参考答案：

可以参考以下几个判断标准选择合适的会话添加为客服培训案例：

1. 会话轮数超过两轮；
2. 会话涉及内容较为典型且客服回复符合 SOP 标准，例如催发货，申请退货等常见客服流程，或是客服 SOP 中关注的重点对话场景；
3. 会话中客服的接待较为优秀，例如安抚好客户情绪并获得客户好评或致谢，或成功引导顾客完成下单购买等；
4. 会话中客服回答没有被质检为异常，否则提供的客服答案不正确会误导接受培训的客服。

（3）请为该智能系统的客服对话质检功能找到合适的应用场景，并用文字描述应用场景。（4分）

参考答案：

该功能可用于客服主管离线质检客服每日接待会话，将质检及复检结果计入绩效考核的一部分；

该功能也可用于有质检异常的客服检查其异常接待会话及异常原因，加深对客服 SOP 的学习理解；

该功能也可作为客服接待时的实时督导，即使提示预警客服接待异常，让客服可以在发出异常消息后及时撤回，或在发出消息前预审告警，以减少不符合规范的接待应答；

该功能也可用于买家情绪预警，当买家对话被质检为高危负面情绪（例如表达强烈不满甚至骂人）时，可以主动预警并通知相关主管介入主动人工干预。

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：1.3.2

试题名称：客户体验管理系统优化

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。本项目中，作为一名人工智能训练师，你将对其输出的数据进行分析，根据数据分析结论提出优化需求。请参考该智能系统的操作手册（CEM 操作手册.pdf）进行以下试题的作答。

- (1) 请对评价分析结果（评价分析.xlsx）进行数据分析，输出分析报告。（4 分）
- (2) 请根据数据分析的结论，对评价分析提出优化需求。（6 分）

3、技能要求

- (1) 能够利用分析工具进行数据分析，输出分析报告
- (2) 能够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

无

人工智能训练师（四级）操作技能考核

试题评分表及答案

准考证号：

试题代码：1.3.2

试题名称：客户体验管理系统优化

考核时间：15 min

主观评分表/过程评分表

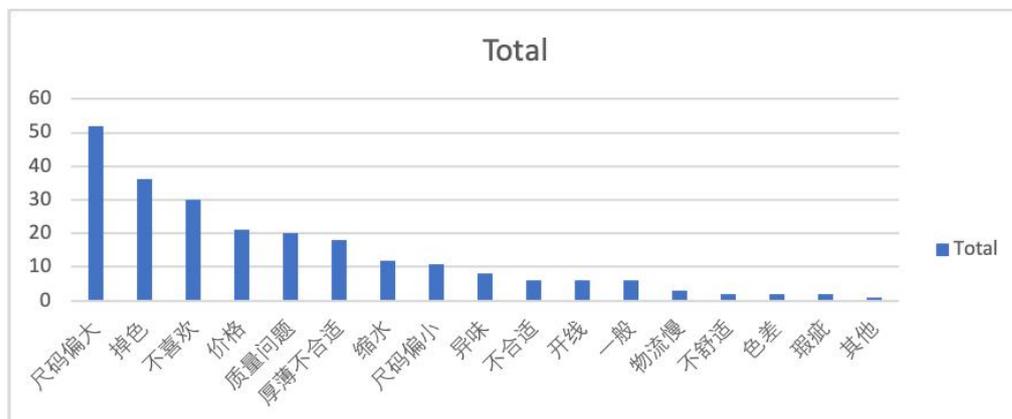
序号	配分	评分细则描述	考评员评分			得分
			1	2	3	
(1)	4分	正确给出标签统计：+2分 正确给出占比最高的标签说明：+2分 错误不扣分 加总得分在0到4分之间				
(2)	6分	正确给出类似参考答案优化需求中的一条：+3分 正确给出类似参考答案优化需求中的两条及以上：+6分 错误不扣分 加总得分在0到6分之间				

参考答案（尽量将细则内容写在上面的表格内，写不下的可另写，但要具体可评判）

(1) 请对评价分析结果（评价分析.xlsx）进行数据分析，输出分析报告。（4分）

参考答案：

评价分析输出的负面标签的统计如下：



从该统计可看出，评价日期（5月10日至5月16日）内，顾客反映最多的问题是“尺码偏大”、“掉色”和“不喜欢”。

由于每日订单评价量在评价分析结果中未提供，所以不需要进行每日负面评价的数量对比，也无法统计负面评价的占比。

（2）请根据数据分析的结论，对评价分析提出优化需求。（6分）

参考答案：

可以有以下一些优化方向：

1. 根据业务需求新增或拆分更多有意义的负面评价标签，并实现其识别；
2. 提供正面评价的好评标签识别，以及未评价的订单数量（及订单相关商品），与目前负面评价的中差评标签进行汇总分析。