

一种基于多模态模型的数据处理方法、装置、设备及介质

技术领域

[01]本申请涉及人工智能技术领域，尤其是涉及一种基于多模态模型的数据处理方法、装置、设备及介质。

背景技术

[02]随着网络技术的快速发展，网络数据越来越多，为了从大量网络数据中检索到用户所需数据，需要由用户提供关键词，服务器基于关键词从大量网络数据中检索到用户所需数据。然而，由于网络数据的数据量很大，服务器需要消耗大量计算资源且需要花费大量时间才能够检索到用户所需数据，从而导致服务器的大量资源消耗，且用户等待很长时间才能够获知所需数据。比如说，在实际场景中会部署大量摄像机，这些摄像机会采集大量图像，服务器为了从这些图像中分析用户所需数据，需要消耗大量计算资源，花费大量时间。

发明内容

[03]本申请提供一种基于多模态模型的数据处理方法，所述方法包括：基于原始输入问题确定与已配置的问题模板匹配的初始问题特征；基于所述初始问题特征获取与所述初始问题特征关联的多个附加问题特征；基于所述初始问题特征和所述多个附加问题特征生成多个目标问题，每个目标问题包括所述初始问题特征和至少一个附加问题特征；针对每个目标问题，将所述目标问题输入给已训练的多模态目标模型，由所述多模态目标模型输出所述目标问题对应的候选答案；基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答案。

[04]本申请提供一种基于多模态模型的数据处理装置，所述装置包括：确定模块，用于基于原始输入问题确定与问题模板匹配的初始问题特征；获取模块，用于获取与所述初始问题特征关联的多个附加问题特征；生成模块，用于基于所述初始问题特征和所述多个附加问题特征生成多个目标问题，每个目标问题包括所述初始问题特征和至少一个附加问题特征；所述获取模块，还用于针对每个目标问题，将所述目标问题输入给已训练的多模态目标模型，由所述多模态目标模型输出所述目标问题对应的候选答案；所述确定模块，还用于基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答

案，所述目标答案是所述原始输入问题的输出结果。

[05]本申请提供一种电子设备，包括：处理器和机器可读存储介质，所述机器可读存储介质存储有能够被所述处理器执行的机器可执行指令；所述处理器用于执行机器可执行指令，以实现上述示例的基于多模态模型的数据处理方法。

[06]由以上技术方案可见，本申请实施例中，可以基于初始问题特征和多个附加问题特征生成多个目标问题，并通过多模态目标模型得到每个目标问题对应的候选答案，基于每个目标问题对应的候选答案确定原始输入问题对应的目标答案，从而借助多模态目标模型得到目标答案，且通过多个目标问题得到准确可靠的目标答案，只需要消耗少量计算资源就能够从大量数据中查询到原始输入问题对应的目标答案，获取目标答案的时间更短，从而节省服务器的计算资源，并节省时间开销，能够快速定位原始输入问题对应的目标答案，提供异常逻辑分析推理能力，实现自动异常诊断，利用多模态目标模型的分析能力给出准确可靠的目标答案，综合考虑图像、视频、语音等各种维度的可能性。

附图说明

[07]为了更加清楚地说明本申请实施例或者现有技术中的技术方案，下面将对本申请实施例或者现有技术描述中所需要使用的附图作简单地介绍。下面描述中的附图仅仅是本申请中记载的一些实施例，对于本领域技术人员来讲，还可以根据本申请实施例的这些附图获得其他的附图。

[08]图 1 是一种实施方式中的基于多模态模型的数据处理方法的流程示意图。

[09]图 2 是一种实施方式中的基于多模态模型的数据处理方法的流程示意图。

[10]图 3 是一种实施方式中的基于多模态模型的自推理方法的示意图。

[11]图 4 是一种实施方式中的基于多模态模型的数据处理装置的结构示意图。

[12]图 5 是一种实施方式中的电子设备的硬件结构图。

具体实施方式

[13]在本申请实施例中使用的术语仅仅是出于描述特定实施例的目的，而非限制本申请。本申请的说明书和权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其它含义。还应当理解，本文中使用的术语“和/或”是指包含一个或多个相关联的列出项目的任何或所有可能组合。

[14]应当理解，尽管在本申请实施例中可能采用术语第一、第二、第三等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本申请范围的情况下，第一信息也可以被称为第二信息，类似地，第二信息也可以被称为第一信息。取决于语境，此外，所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[15]本申请实施例提出一种基于多模态模型的数据处理方法，该方法可以应用于任意设备，如服务器等，参见图 1 所示，该方法可以包括：

[16]步骤 101、基于原始输入问题确定与已配置的问题模板匹配的初始问题特征。

[17]步骤 102、基于初始问题特征获取与初始问题特征关联的多个附加问题特征。

[18]步骤 103、基于初始问题特征和多个附加问题特征生成多个目标问题；其中，每个目标问题包括该初始问题特征和至少一个附加问题特征。

[19]步骤 104、针对每个目标问题，将该目标问题输入给已训练的多模态目标模型，由多模态目标模型输出该目标问题对应的候选答案。

[20]步骤 105、基于每个目标问题对应的候选答案确定原始输入问题对应的目标答案，该目标答案可以是针对原始输入问题的最终输出结果。

[21]示例性的，基于原始输入问题确定与已配置的问题模板匹配的初始问题特征，可以包括但不限于：若问题模板包括特征类型，则确定该原始输入问题中是否存在与该特征类型匹配的关键值；若存在与该特征类型匹配的关键值，则可以基于该关键值确定初始问题特征。若不存在与该特征类型匹配的关键值，则可以对该原始输入问题进行分析，得到与该特征类型匹配的初始问题特征；或者，可以向用户发出针对该特征类型的提示信息，由用户基于该提示信息提供新的原始输入问题，若新的原始输入问题中存在与该特征类型匹配的关键值，则可以基于该关键值确定初始问题特征。

[22]示例性的，基于初始问题特征获取与初始问题特征关联的多个附加问题特征，可以包括但不限于：将该初始问题特征输入给已训练的特征分析目标模型，由特征分析目标模型输出与该初始问题特征关联的多个附加问题特征。或者，将该初始问题特征输入给多模态目标模型，由多模态目标模型输出与该初始问题特征关联的多个附加问题特征。或者，基于该初始问题特征分析与该初始问题特征关联的多个附加问题特征（即对该初始问题特征进行分析，以得到与该初始问题特征关联的多个附加问题特征）。

[23]示例性的，特征分析目标模型可以是与目标场景类型匹配的特征分析目标模型；问

题模板可以是与目标场景类型匹配的问题模板；其中，该目标场景类型可以为行为异常类型、车辆异常类型、或工业异常类型。

[24]示例性的，基于初始问题特征和多个附加问题特征生成多个目标问题，可以包括但不限于：从多个附加问题特征中选取候选附加问题特征；在将初始问题特征和候选附加问题特征组合之后，针对每一个组合，为该初始问题特征和/或该候选附加问题特征添加上下文信息，得到候选附加问题特征对应的目标问题。

[25]示例性的，基于每个目标问题对应的候选答案确定原始输入问题对应的目标答案，可以包括但不限于：基于每个目标问题对应的候选答案，从所有候选答案中过滤内容重复的候选答案，得到至少一个剩余候选答案，并确定每个剩余候选答案在所有候选答案中的出现次数；将最大出现次数的剩余候选答案确定为原始输入问题对应的目标答案；或者，若最大出现次数对应的剩余候选答案在所有候选答案中的出现比例值大于预设阈值，则将最大出现次数的剩余候选答案确定为原始输入问题对应的目标答案。

[26]示例性的，在基于每个目标问题对应的候选答案确定原始输入问题对应的目标答案之后，可以输出该目标答案对应的目标数据，该目标数据可以包括目标答案，或者，该目标数据可以包括目标答案和目标答案对应的附加问题特征；其中，目标答案可以包括但不限于图像、文本、声音和视频至少一种维度的答案；其中，目标答案可以对应多个候选答案，且目标答案对应的附加问题特征可以包括目标答案对应的每个候选答案对应的附加问题特征。

[27]示例性的，上述执行顺序只是为了方便描述给出的示例，在实际应用中，还可以改变步骤之间的执行顺序，对此执行顺序不做限制。而且，在其它实施例中，并不一定按照本申请中示出和描述的顺序来执行相应方法的步骤，其它实施例中的方法所包括的步骤可以比本申请中所描述的更多或更少。此外，本申请中所描述的单个步骤，在其它实施例中可能被分解为多个步骤进行描述；本申请中所描述的多个步骤，在其它实施例中也可能被合并为单个步骤进行描述。

[28]由以上技术方案可见，本申请实施例中，可以基于初始问题特征和多个附加问题特征生成多个目标问题，并通过多模态目标模型得到每个目标问题对应的候选答案，基于每个目标问题对应的候选答案确定原始输入问题对应的目标答案，从而借助多模态目标模型得到目标答案，且通过多个目标问题得到准确可靠的目标答案，只需要消耗少量计算资源就能够从大量数据中查询到原始输入问题对应的目标答案，获取目标答案的时间更短，从而节省服务器的计算资源，并节省时间开销，能够快速定位原始输入问题对应

的目标答案，提供异常逻辑分析推理能力，实现自动异常诊断，利用多模态目标模型的分析能力给出准确可靠的目标答案，综合考虑图像、视频、语音等各种维度的可能性。

[29]以下结合具体应用场景，对本申请实施例中的上述技术方案进行说明。

[30]由于网络数据（如文本数据、视频数据、音频数据、图像数据等）的数据量很大，服务器需要消耗大量计算资源且需要花费大量时间才能够检索到用户所需数据，从而导致服务器的大量资源消耗，且用户等待很长时间才能够获知所需数据。比如说，中实际场景中会部署大量摄像机，这些摄像机采集大量图像，为了从这些图像中分析用户所需数据，需要消耗大量计算资源，花费大量时间。

[31]针对上述发现，本申请实施例提出一种基于多模态模型的数据处理方法。基于该方法，借助多模态目标模型得到目标答案，且通过多个目标问题得到准确可靠的目标答案，只需要消耗少量计算资源就能够从大量数据中查询到原始输入问题对应的目标答案，获取目标答案的时间更短，从而能够节省服务器的计算资源，并节省时间开销，能够快速定位原始输入问题对应的目标答案。

[32]示例性的，可以预先配置问题模板，该问题模板可以包括至少一个特征类型，表示原始输入问题应该提供与该特征类型匹配的关键值。其中，问题模板可以是与某个场景类型匹配的问题模板，比如说，场景类型可以包括但不限于行为异常类型、车辆异常类型、工业异常类型等，因此，可以预先配置与行为异常类型匹配的问题模板 a1、与车辆异常类型匹配的问题模板 a2、与工业异常类型匹配的问题模板 a3 等，当然，上述类型只是场景类型的示例，对此不做限制。

[33]比如说，问题模板 a1 可以包括但不限于以下至少一种：时间类型、位置类型、行为事件类型。该行为事件类型可以是某种特定行为，例如打架行为、抢劫行为、遗失物品行为等，后续将该行为事件类型记为行为 X，表示发生行为 X 的事件类型，对此行为 X 不做限制。

[34]问题模板 a2 可以包括但不限于以下至少一种：时间类型、位置类型、行为事件类型。该行为事件类型可以是某种特定车辆行为，如“闯红灯”行为、“超速”行为等，表示发生该特定车辆行为的事件类型，对此不做限制。

[35]问题模板 a3 可以包括但不限于以下至少一种：声音类型、振动类型、加速度类型、角速度类型、温度类型、湿度类型，对此不做限制。

[36]当然，上述模板只是问题模板的几个示例，对此问题模板不做限制。

[37]示例性的，针对每个场景类型，可以训练与该场景类型对应的特征分析目标模型，比如说，可以训练与行为异常类型对应的特征分析目标模型 b1、与车辆异常类型对应的特征分析目标模型 b2、与工业异常类型对应的特征分析目标模型 b3，当然，也可以为所有场景类型训练同一个特征分析目标模型。

[38]为了训练特征分析目标模型 b1，可以收集与“行为异常类型”有关的样本数据，该样本数据包括与时间类型有关的信息、与位置类型有关的信息、与行为 X 有关的信息，除了上述信息，该样本数据还可以包括其它类型的信息，基于大量样本数据就可以执行训练，以得到特征分析目标模型 b1，对此训练过程不做限制。

[39]为了训练特征分析目标模型 b2，可以收集与“车辆异常类型”有关的样本数据，该样本数据包括与时间类型有关的信息、与位置类型有关的信息、与特定车辆行为有关的信息，除了上述信息，该样本数据还可以包括其它类型的信息，基于大量样本数据就可以执行训练，以得到特征分析目标模型 b2，对此训练过程不做限制。

[40]为了训练特征分析目标模型 b3，可以收集与“工业异常类型”有关的样本数据，该样本数据包括与声音类型有关的信息、与振动类型有关的信息、与加速度类型有关的信息、与角速度类型有关的信息、与温度类型有关的信息、与湿度类型有关的信息，除了上述信息，该样本数据还可以包括其它类型的信息，基于大量样本数据就可以执行训练，以得到特征分析目标模型 b3，对此训练过程不做限制。

[41]当然，上述模型只是特征分析目标模型的几个示例，对此不做限制。

[42]在上述应用场景下，本申请实施例提出一种基于多模态模型的数据处理方法，该方法可以应用于服务器，参见图 2 所示，该方法可以包括：

[43]步骤 201、获取原始输入问题。比如说，服务器可以接收用户输入的原始输入问题，也可以采用其它方式获取原始输入问题，对此原始输入问题的来源不做限制。

[44]比如说，针对行为异常类型的应用场景，原始输入问题的一个例子可以如“帮我搜一下 2022 年 3 月 21 日下午 3 点左右在创新路十字路口出现的行人，这些行人中谁最有可能是这个时间点在该十字路口发生行为 X 的目标”。

[45]比如说，针对车辆异常类型的应用场景，原始输入问题的一个例子可以如“帮我搜一下 2022 年 3 月 21 日下午 3 点左右在创新路十字路口出现的车辆，这些车辆中谁最有可能是这个时间点在该十字路口发生闯红灯行为的目标”。

[46]当然，上述问题只是原始输入问题的两个示例，对此原始输入问题不做限制。

[47]步骤 202、基于原始输入问题确定与问题模板匹配的初始问题特征。

[48]示例性的，服务器可以基于或者不基于原始输入问题确定目标场景类型（即当前应用场景的场景类型）。例如，服务器可以基于原始输入问题确定目标场景类型并确定与目标场景类型匹配的问题模板，基于原始输入问题确定与该问题模板匹配的初始问题特征。比如说，若基于原始输入问题确定目标场景类型为行为异常类型，则服务器基于原始输入问题确定与问题模板 a1 匹配的初始问题特征。服务器还可以不基于原始输入问题确定目标场景类型；例如，服务器可以先确定目标场景类型（可以是用户提前确定好的），并确定与目标场景类型匹配的问题模板，此时输入的原始输入问题需要符合当前应用场景的场景类型，然后服务器确定与该问题模板匹配的初始问题特征。比如说，若目标场景类型为车辆异常类型，则服务器基于原始输入问题确定与问题模板 a2 匹配的初始问题特征。若目标场景类型为工业异常类型，则服务器基于原始输入问题确定与问题模板 a3 匹配的初始问题特征。

[49]比如说，与目标场景类型匹配的问题模板可以是问题模板 a1，且问题模板 a1 可以包括时间类型、位置类型、行为事件类型。服务器从原始输入问题中提取与时间类型匹配的关键值“2022 年 3 月 21 日下午 3 点”，从原始输入问题中提取与位置类型匹配的关键值“创新路十字路口”，从原始输入问题中提取与行为事件类型匹配的关键值“行为 X”，服务器根据这些关键值确定与目标场景类型匹配的问题模板是问题模板 a1，并且将上述关键值作为初始问题特征，即初始问题特征包括“2022 年 3 月 21 日下午 3 点”、“创新路十字路口”、“行为 X”。

[50]又例如，与目标场景类型匹配的问题模板可以是问题模板 a2，且问题模板 a2 可以包括时间类型、位置类型、行为事件类型。服务器从原始输入问题中提取与时间类型匹配的关键值“2022 年 3 月 21 日下午 3 点”，从原始输入问题中提取与位置类型匹配的关键值“创新路十字路口”，从原始输入问题中提取与行为事件类型匹配的关键值“闯红灯行为”，服务器根据这些关键值确定与目标场景类型匹配的问题模板是问题模板 a2，可以将上述关键值作为初始问题特征。

[51]综上所述可以看出，基于与目标场景类型匹配的问题模板，可以从原始输入问题中提取出初始问题特征，对此初始问题特征的提取方式不做限制。

[52]在一种可能的实施方式中，为了确定初始问题特征，可以采用如下方式：

[53]方式 1：针对问题模板中的每个特征类型（如时间类型、位置类型、行为事件类型

等), 确定原始输入问题中是否存在与该特征类型匹配的关键值; 若存在与该特征类型匹配的关键值, 则基于该关键值确定初始问题特征, 如将该关键值作为初始问题特征。若不存在与该特征类型匹配的关键值, 则对该原始输入问题进行分析, 得到与该特征类型匹配的初始问题特征。

[54]比如说, 若原始输入问题中存在与该特征类型匹配的关键值, 则可以使用键值映射将该关键值替换为变量, 从而将该关键值作为初始问题特征。

[55]比如说, 若原始输入问题中不存在与该特征类型匹配的关键值, 则可以对该原始输入问题进行分析, 得到与该特征类型匹配的关键值, 并将该关键值作为初始问题特征, 本实施例中对此原始输入问题的分析过程不做限制。

[56]显然, 在问题模板对应多个特征类型时, 需要确定每个特征类型匹配的关键值, 即将每个关键值作为初始问题特征, 每个特征类型对应一个初始问题特征。

[57]方式 2: 针对问题模板中的每个特征类型, 确定原始输入问题中是否存在与该特征类型匹配的关键值; 若存在与该特征类型匹配的关键值, 则基于该关键值确定初始问题特征, 如将该关键值作为初始问题特征。若不存在与该特征类型匹配的关键值, 则向用户发出针对该特征类型的提示信息(该提示信息表示原始输入问题中缺少与该特征类型匹配的关键值), 由用户基于该提示信息提供新的原始输入问题(即在之前的原始输入问题中添加与该特征类型匹配的关键值, 得到新的原始输入问题), 若新的原始输入问题中存在与该特征类型匹配的关键值, 则可以基于该关键值确定初始问题特征。

[58]方式 3: 针对问题模板中的每个特征类型, 确定原始输入问题中是否存在与该特征类型匹配的关键值; 若存在与该特征类型匹配的关键值, 则基于该关键值确定初始问题特征, 如将该关键值作为初始问题特征。若不存在与该特征类型匹配的关键值, 则对该原始输入问题进行分析, 若分析到与该特征类型匹配的关键值, 则将该关键值作为与该特征类型匹配的初始问题特征。若未分析到与该特征类型匹配的关键值, 则向用户发出针对该特征类型的提示信息, 由用户基于该提示信息提供新的原始输入问题, 若新的原始输入问题中存在与该特征类型匹配的关键值, 则基于该关键值确定初始问题特征。

[59]当然, 上述三种方式只是示例, 对此初始问题特征的获取方式不做限制。

[60]综上所述, 在步骤 202 中, 可以基于原始输入问题确定与问题模板匹配的初始问题特征, 后续将初始问题特征记为特征 A、特征 B 和特征 C。

[61]步骤 203、基于初始问题特征获取与初始问题特征关联的多个附加问题特征。

[62]示例性的，在得到初始问题特征之后，还可以基于初始问题特征获取多个附加问题特征，附加问题特征是在初始问题特征基础上额外补充的特征，附加问题特征可以变化，与初始问题特征有关即可。比如说，针对行为异常类型的应用场景，初始问题特征包括“行为 X”，附加问题特征是与“行为 X”有关的特征，如行人速度特征、行人穿戴特征、行人表情特征等。针对车辆异常类型的应用场景，初始问题特征包括“闯红灯行为”，附加问题特征是与“闯红灯行为”有关的特征，如车辆速度特征、车牌标识特征、车辆颜色特征等。

[63]在一种可能的实施方式中，可以采用如下方式获取多个附加问题特征：

[64]方式 1：将初始问题特征输入给已训练的特征分析目标模型，由特征分析目标模型输出与该初始问题特征关联的多个附加问题特征。

[65]示例性的，特征分析目标模型是与目标场景类型匹配的特征分析目标模型，该目标场景类型可以为行为异常类型、车辆异常类型、工业异常类型。比如说，若目标场景类型为行为异常类型，则将初始问题特征输入给特征分析目标模型 b1，由特征分析目标模型 b1 对初始问题特征进行处理，得到与该初始问题特征关联的多个附加问题特征，输出多个附加问题特征。若目标场景类型为车辆异常类型，则将初始问题特征输入给特征分析目标模型 b2，由特征分析目标模型 b2 对初始问题特征进行处理，得到与该初始问题特征关联的多个附加问题特征，输出多个附加问题特征。若目标场景类型为工业异常类型，则将初始问题特征输入给特征分析目标模型 b3，由特征分析目标模型 b3 对初始问题特征进行处理，得到与该初始问题特征关联的多个附加问题特征，输出多个附加问题特征。

[66]示例性的，在基于样本数据执行训练以得到特征分析目标模型时，样本数据可以包括与初始问题特征有关的信息（如与时间类型有关的信息、与位置类型有关的信息、与行为 X 有关的信息），且样本数据还可以包括其它类型的信息（其它类型的信息可以作为初始问题特征的标签信息，即反映初始问题特征的信息），因此，通过不断学习优化特征分析目标模型，特征分析目标模型能够反映初始问题特征与附加问题特征的关系，在将初始问题特征输入给已训练的特征分析目标模型之后，可以得到与初始问题特征关联的多个附加问题特征。

[67]方式 2：将初始问题特征输入给多模态目标模型（即多模态目标大模型），由多模态目标模型输出与该初始问题特征关联的多个附加问题特征。

[68]示例性的，多模态目标模型是基于大数据不断学习优化得到的，具有强大的处理能力，可以适用于所有场景类型的数据处理，如行为异常类型、车辆异常类型、工业异常类型等，对此多模态目标模型的来源不作限制。在将初始问题特征输入给多模态目标模型之后，由多模态目标模型对初始问题特征进行处理，得到与该初始问题特征关联的多个附加问题特征，输出多个附加问题特征。

[69]方式 3：基于初始问题特征分析与该初始问题特征关联的多个附加问题特征。

[70]比如说，服务器可以获取初始问题特征对应的原始数据，基于原始数据分析与该初始问题特征关联的多个附加问题特征。例如，初始问题特征包括“2022 年 3 月 21 日下午 3 点”、“创新路十字路口”、“行为 X”，服务器可以获取 2022 年 3 月 21 日下午 3 点在创新路十字路口产生的图像，基于图像分析与行为 X 关联的附加问题特征，如行人速度特征、行人穿戴特征、行人表情特征等。当然，上述方式只是基于原始数据分析得到多个附加问题特征的示例，对此不作限制，只要能够分析得到与初始问题特征关联的多个附加问题特征即可。

[71]综上所述，可以得到与初始问题特征关联的多个附加问题特征，当然，方式 1、方式 2 和方式 3 只是示例，对此附加问题特征的获取方式不作限制。

[72]步骤 204、基于初始问题特征和多个附加问题特征生成多个目标问题；其中，每个目标问题可以包括该初始问题特征和至少一个附加问题特征。

[73]比如说，服务器可以从多个附加问题特征中选取至少一个附加问题特征作为候选附加问题特征，并基于初始问题特征和候选附加问题特征生成目标问题。若多个附加问题特征包括附加问题特征 c1、附加问题特征 c2、附加问题特征 c3，那么，服务器可以基于初始问题特征和附加问题特征 c1（即将附加问题特征 c1 作为候选附加问题特征）生成目标问题 d1，基于初始问题特征和附加问题特征 c2 生成目标问题 d2，基于初始问题特征和附加问题特征 c3 生成目标问题 d3，基于初始问题特征、附加问题特征 c1 和附加问题特征 c2 生成目标问题 d4，基于初始问题特征、附加问题特征 c1 和附加问题特征 c3 生成目标问题 d5，以此类推。

[74]示例性的，基于初始问题特征和候选附加问题特征生成目标问题，可以包括但不限于：将初始问题特征和候选附加问题特征组合，在将初始问题特征和候选附加问题特征组合之后，针对每一个组合，为该初始问题特征和/或该候选附加问题特征添加上下文信息，得到该候选附加问题特征对应的目标问题。比如说，服务器可以为该初始问题特征

添加上下文信息，即在该初始问题特征的前面添加上下文信息，和/或，在该初始问题特征的后面添加上下文信息，对此上下文信息的内容不作限制。又例如，服务器可以为该候选附加问题特征添加上下文信息。又例如，服务器可以为该初始问题特征添加上下文信息，并为该候选附加问题特征添加上下文信息。

[75]对问题特征添加上下文信息意味着将与问题相关的环境、条件、先前信息或语境纳入考虑范围，以增强模型对问题的理解和解决能力。具体而言，这里的上下文信息可以包括：前文信息、后文信息、周围环境和先前知识。

[76]前文信息指前面提到的与当前问题相关的文本或数据。在自然语言处理中，这可能是前一句话、前一段落或前一篇文档中的信息。

[77]后文信息指与问题有关的后续文本或数据。这可以帮助模型预测未来事件或更好地理解当前问题的语境。

[78]周围环境指问题发生的背景或环境信息。在视觉任务中，这可能是与图像相关的其他对象、场景或背景信息。在自然语言处理中，这可能是与当前文本相关的外部信息，例如时间、地点或社会背景等。

[79]先前知识指模型已经学习或训练过的先前知识。这可以包括模型在训练过程中接触到的语料库、数据集或经验。综合来看，添加上下文信息意味着将问题嵌入到更广泛的语境中，以便模型能够更全面地理解和解决问题。

[80]步骤 205、针对每个目标问题，将该目标问题输入给已训练的多模态目标模型（多模态目标大模型），由多模态目标模型输出该目标问题对应的候选答案。

[81]示例性的，多模态目标模型是基于大数据不断学习优化得到的，具有强大的处理能力，可以适用于所有场景类型的数据处理，对此多模态目标模型的来源不作限制。多模态指除了文本数据外，还有图像、视频、语音等其他模态数据，例如，一个具有声音和视频的文件，能够利用多模态目标模型分析该文件中有没有人喊救命，喊救命的人是男的还是女的，穿什么衣服等。多模态目标模型的示例包括：GPT-4V（generative pre-trained transformer-4 vision，生成式预训练转换器-4 视觉）等。在将目标问题输入给多模态目标模型之后，由多模态目标模型对该目标问题进行处理，得到与该目标问题对应的候选答案，并输出候选答案。

[82]综上所述，可以基于多模态目标模型得到每个目标问题对应的候选答案，如目标问题 d1 对应的候选答案 e1、目标问题 d2 对应的候选答案 e2、目标问题 d3 对应的候选答

案 e3、目标问题 d4 对应的候选答案 e4，以此类推。

[83]示例性的，针对每个候选答案，该候选答案可以是图像维度的候选答案，可以是文本维度的候选答案，可以是声音维度的候选答案，可以是视频维度的候选答案，还可以是多种维度的候选答案，如文本和视频组合维度的候选答案。

[84]步骤 206、基于每个目标问题对应的候选答案确定原始输入问题对应的目标答案，该目标答案可以是针对原始输入问题的最终输出结果。

[85]示例性的，基于每个目标问题对应的候选答案，服务器可以从所有候选答案中过滤内容重复的候选答案，得到至少一个剩余候选答案，并确定每个剩余候选答案在所有候选答案中的出现次数。然后，基于每个剩余候选答案对应的出现次数，服务器从所有剩余候选答案中选取目标答案。比如说，服务器将最大出现次数的剩余候选答案确定为原始输入问题对应的目标答案；或者，若最大出现次数对应的剩余候选答案在所有候选答案中的出现比例值大于预设阈值，则服务器将最大出现次数的剩余候选答案确定为原始输入问题对应的目标答案。

[86]比如说，假设一共存在 6 个候选答案，候选答案 e1 为 mmm，候选答案 e2 为 mmm，候选答案 e3 为 mmm，候选答案 e4 为 mmm，候选答案 e5 为 nnn，候选答案 e6 为 nnn，那么，在过滤内容重复的候选答案（即候选答案 e2、候选答案 e3、候选答案 e4、候选答案 e6 等）之后，得到的剩余候选答案为“mmm”和“nnn”。显然，剩余候选答案“mmm”在所有候选答案中的出现次数为 4，剩余候选答案“nnn”在所有候选答案中的出现次数为 2。

[87]服务器可以将最大出现次数的剩余候选答案“mmm”作为原始输入问题对应的目标答案，即目标答案为 mmm。或者，服务器可以判断最大出现次数对应的剩余候选答案在所有候选答案中的出现比例值（即 $4/6$ ）是否大于预设阈值，如果最大出现次数对应的出现比例值大于预设阈值，则将最大出现次数的剩余候选答案“mmm”作为原始输入问题对应的目标答案，如果最大出现次数对应的出现比例值小于或等于预设阈值，则未得到原始输入问题对应的目标答案。

[88]综上所述可以看出，针对最大出现次数的候选答案，该候选答案是出现最频繁的答案，是从所有候选答案中找到的一个共识，具有更高的置信度，可以认为候选答案是相对正确且可靠的，从而提高最终输出准确性。因此，将该候选答案作为原始输入问题对应的目标答案，且目标答案可以是针对原始输入问题的最终输出结果。

[89]步骤 207、输出目标答案对应的目标数据，该目标数据可以包括目标答案，或者，该目标数据可以包括目标答案和目标答案对应的附加问题特征。

[90]示例性的，目标答案可以包括但不限于图像、文本、声音和视频但至少一种维度的答案。比如说，该目标答案可以是图像维度的目标答案，可以是文本维度的目标答案，可以是声音维度的目标答案，可以是视频维度的目标答案，还可以是多种维度的目标答案，如文本和视频组合维度的目标答案。

[91]目标答案可以对应多个候选答案，且目标答案对应的附加问题特征包括该目标答案对应的每个候选答案对应的附加问题特征。比如说，目标答案为 mmm 时，该目标答案对应候选答案 e1、候选答案 e2、候选答案 e3 和候选答案 e4，因此，目标答案对应的附加问题特征可以包括：候选答案 e1 对应的附加问题特征（候选答案 e1 是基于初始问题特征和附加问题特征 c1 对应的目标问题 d1 得到的，因此，候选答案 e1 对应附加问题特征 c1）、候选答案 e2 对应的附加问题特征、候选答案 e3 对应的附加问题特征、候选答案 e4 对应的附加问题特征。

[92]至此成功输出原始输入问题的目标答案，完成基于多模态模型的数据处理。

[93]在一种可能的实施方式中，为了使模型具备自推理能力，可以将复杂问题分解为更简单的多步骤程序，并利用多种方式在每一个步骤中对方法进行验证。比如说，参见图 3 所示，图 3 为基于多模态模型的自推理方法的示意图，该方法可以包括输入问题、生成问题模板、问题逻辑分析、多模态模型（即多模态大模型）、技术验证等过程。

[94]第一，在输入问题过程中，可以获取原始输入问题，原始输入问题的一个例子可以如“帮我搜一下 2022 年 3 月 21 日下午 3 点左右在创新路十字路口出现的行人，这些行人中谁最有可能是这个时间点在该十字路口发生行为 X 的目标”。

[95]第二，在生成问题模板过程中，可以基于原始输入问题确定与问题模板匹配的初始问题特征。比如说，将原始输入问题转化为数学逻辑形式，通过使用键值映射将关键值替换为变量，得到初始问题特征，并得到修改后问题。比如说，将原始输入问题修改为如下问题：帮我搜下 A 左右在 B 出现的行人，这些行人中谁最有可能是 C。A：2022 年 3 月 21 日下午 3 点；B：创新路十字路口；C：行为 X。

[96]第三，在问题逻辑分析过程中，可以获取与初始问题特征关联的多个附加问题特征，并基于初始问题特征和多个附加问题特征生成多个目标问题。

[97]比如说，基于多重验证和交叉检查的方式，使用不同的方法生成分析解决方案，给

多模态目标模型提供提示，为问题生成额外的上下文，逻辑修改问题可以为：出现在 A 与 B，具备行为 C 的行为，这个行人可能具备 K 特性，K 是可以变化的，即 K 表示附加问题特征，这样，通过改变附加问题特征 K，就可以生成多个目标问题，每个目标问题可以包括初始问题特征和附加问题特征。

[98]第四，在多模态模型（即多模态大模型）过程中，可以将每个目标问题输入给多模态目标模型，由多模态目标模型输出每个目标问题对应的候选答案。比如说，使用多模态目标模型对逻辑修改问题（即目标问题）进行分析，以输出候选答案，这个候选答案利用了多模态目标模型的能力，即多模态目标模型分析该目标问题包括图像、视频、文本、语音等各种维度的可能性，多模态目标模型比较输出结果，在各种维度的信息中找到共识，多模态目标模型也可以提供更高置信度，从而得到目标问题对应的候选答案，即候选答案正确且可靠。

[99]比如说，多模态目标模型自动分析具备行为 X 且符合时间和地点的行人，同时基于生物特征确定行人的身份信息，利用视频结构化特征，如衣服颜色、是否背包等，加上大数据分析，从而自动推理分析，得到最有可能的目标对象，利用空间位置信息快速定位目标对象的当前可能位置。

[100]第五，在技术验证过程中，可以基于每个目标问题对应的候选答案确定原始输入问题对应的目标答案。比如说，为了确保在各种表达式（即多个目标问题）的输出中达成共识，可以确定每个目标问题对应的候选答案，并确定出现最频繁的候选答案作为目标答案，确保最终输出的准确性。最终不停循环输出一个最高置信列表，最后输出最有可能的目标对象，然后确定所输出的目标对象对应的附加问题特征，完成整个自推理过程，得到原始输入问题对应的目标答案。

[101]由以上技术方案可见，本申请实施例中，借助多模态目标模型得到目标答案，且通过多个目标问题得到准确可靠的目标答案，只需要消耗少量计算资源就能够从大量数据中查询到原始输入问题对应的目标答案，获取目标答案的时间更短，从而节省服务器的计算资源，并节省时间开销，能够快速定位原始输入问题对应的目标答案，提供异常逻辑分析推理能力，实现自动异常诊断，利用多模态目标模型的分析能力给出准确且可靠的目标答案，综合考虑图像、视频、语音等各种维度的可能性。通过多模态目标模型的能力，用户只需要输入简单描述性的文字，多模态目标模型就可以自动推理以及快速定位目标对象，可以提高多模态目标模型在逻辑推理问题上的性能，使得多模态目标模型能够自推理以完成优化需求。在本申请实施例中，能够快速定位异常问题，同时给出异

常逻辑分析推理，实现自动异常诊断。

[102]基于与上述方法同样的申请构思，本申请实施例提出一种基于多模态模型的数据处理装置，参见图 4 所示，图 4 为该装置的结构示意图，该装置可以包括：

[103]确定模块 41，用于基于原始输入问题确定与问题模板匹配的初始问题特征；获取模块 42，用于获取与所述初始问题特征关联的多个附加问题特征；生成模块 43，用于基于所述初始问题特征和所述多个附加问题特征生成多个目标问题，每个目标问题包括所述初始问题特征和至少一个附加问题特征；所述获取模块 42，还用于针对每个目标问题，将所述目标问题输入给已训练的多模态目标模型，由所述多模态目标模型输出所述目标问题对应的候选答案；所述确定模块 41，还用于基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答案，所述目标答案是所述原始输入问题的输出结果。

[104]示例性的，所述确定模块 41 基于原始输入问题确定与问题模板匹配的初始问题特征时具体用于：若所述问题模板包括特征类型，则确定所述原始输入问题中是否存在与所述特征类型匹配的关键值；若存在与所述特征类型匹配的关键值，则基于所述关键值确定所述初始问题特征；若不存在与所述特征类型匹配的关键值，则对所述原始输入问题进行分析，得到与所述特征类型匹配的初始问题特征；或，向用户发出针对所述特征类型的提示信息，由用户基于所述提示信息提供新的原始输入问题，若所述新的原始输入问题中存在与所述特征类型匹配的关键值，则基于所述关键值确定所述初始问题特征。

[105]示例性的，所述获取模块 42 获取与所述初始问题特征关联的多个附加问题特征时具体用于：将所述初始问题特征输入给已训练的特征分析目标模型，由所述特征分析目标模型输出与所述初始问题特征关联的多个附加问题特征；或者，将所述初始问题特征输入给所述多模态目标模型，由所述多模态目标模型输出与所述初始问题特征关联的多个附加问题特征；或者，基于所述初始问题特征分析与所述初始问题特征关联的多个附加问题特征。

[106]示例性的，所述特征分析目标模型是与目标场景类型匹配的特征分析目标模型；所述问题模板是与所述目标场景类型匹配的问题模板；其中，所述目标场景类型包括行为异常类型、车辆异常类型、或工业异常类型。

[107]示例性的，所述生成模块 43 基于所述初始问题特征和所述多个附加问题特征生成多个目标问题时具体用于：从多个附加问题特征中选取候选附加问题特征；在将初始问题特征和候选附加问题特征组合后，针对每一个组合，为所述初始问题特征和/或候选附

加问题特征添加上下文信息，得到候选附加问题特征对应的目标问题。

[108]示例性的，所述确定模块 41 基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答案时具体用于：基于每个目标问题对应的候选答案，从所有候选答案中过滤内容重复的候选答案，得到至少一个剩余候选答案，确定每个剩余候选答案在所有候选答案中的出现次数；将最大出现次数的剩余候选答案确定为原始输入问题对应的目标答案；或，若最大出现次数对应的剩余候选答案在所有候选答案中的出现比例值大于预设阈值，则将最大出现次数的剩余候选答案确定为所述原始输入问题对应的目标答案。

[109]示例性的，所述确定模块 41 基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答案之后还用于：输出所述目标答案对应的目标数据，所述目标数据包括所述目标答案，或者，所述目标数据包括所述目标答案和所述目标答案对应的附加问题特征；所述目标答案包括图像、文本、声音和视频至少一种维度的答案；所述目标答案对应多个候选答案，所述目标答案对应的附加问题特征包括所述目标答案对应的每个候选答案对应的附加问题特征。

[110]基于与上述方法同样的申请构思，本申请实施例提出一种电子设备，参见图 5 所示，电子设备包括：处理器 51 和机器可读存储介质 52，机器可读存储介质 52 存储有能够被处理器 51 执行的机器可执行指令（即基于多模态模型的数据处理指令）；处理器 51 用于执行机器可执行指令，以实现本申请上述示例公开的基于多模态模型的数据处理方法。

[111]基于与上述方法同样的申请构思，本申请实施例还提供一种机器可读存储介质，所述机器可读存储介质上存储有若干计算机指令，所述计算机指令被处理器执行时，能够实现本申请上述示例公开的基于多模态模型的数据处理方法。

[112]其中，上述机器可读存储介质可以是任何电子、磁性、光学或其它物理存储装置，可以包含或存储信息，如可执行指令、数据，等等。例如，机器可读存储介质可以是：RAM（Random Access Memory，随机存取存储器）、易失性存储器、非易失性存储器、闪存、存储驱动器（如硬盘驱动器）、固态硬盘、任何类型的存储盘（如光盘、dvd 等），或者类似的存储介质，或者它们的组合。

[113]上述实施例阐明的系统、装置、模块或单元，具体可以由计算机实体实现，或者由具有某种功能的产品来实现。一种典型的实现设备为计算机，计算机的具体形式可以是个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播

放器、导航设备、电子邮件收发设备、游戏控制台、平板电脑、可穿戴设备或者这些设备中的任意几种设备的组合。

[114]为了描述的方便，描述以上装置时，以功能将装置分为各种单元分别描述。当然，在实施本申请时可以把各单元的功能在同一个或多个软件和/或硬件中实现。

[115]本领域技术人员应明白，本申请的实施例可提供为方法、系统、或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本申请实施例可采用在包含有计算机可用程序代码的一个或多个计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

[116]本申请是参照根据本申请实施例的方法、设备（系统）、和计算机程序产品的流程图和/或方框图来描述的。应理解可以由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其它可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其它可编程数据处理设备的处理器执行的指令产生用于实现在流程图中的一个流程或多个流程和/或方框图中的一个方框或多个方框中指定的功能的装置。

[117]而且，这些计算机程序指令也可以存储在能引导计算机或其它可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图中的一个流程或者多个流程和/或方框图中的一个方框或者多个方框中指定的功能。

[118]这些计算机程序指令也可装载到计算机或其它可编程数据处理设备上，使得在计算机或者其它可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其它可编程设备上执行的指令提供用于实现在流程图中的一个流程或多个流程和/或方框图中的一个方框或多个方框中指定的功能的步骤。

[119]以上所述内容仅为本申请的实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本申请的权利要求的范围之内。

权利要求书

- 1、一种基于多模态模型的数据处理方法，其特征在于，所述方法包括：
基于原始输入问题确定与已配置的问题模板匹配的初始问题特征；
基于所述初始问题特征获取与所述初始问题特征关联的多个附加问题特征；
基于所述初始问题特征和所述多个附加问题特征生成多个目标问题，每个目标问题包括所述初始问题特征和至少一个附加问题特征；
针对每个目标问题，将所述目标问题输入给已训练的多模态目标模型，由所述多模态目标模型输出所述目标问题对应的候选答案；
基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答案。
- 2、根据权利要求 1 所述的方法，其特征在于，所述基于原始输入问题确定与已配置的问题模板匹配的初始问题特征，包括：
若所述问题模板包括特征类型，则确定所述原始输入问题中是否存在与所述特征类型匹配的关键值；
若存在与所述特征类型匹配的关键值，则基于所述关键值确定所述初始问题特征；
若不存在与所述特征类型匹配的关键值，则对所述原始输入问题进行分析，得到与所述特征类型匹配的初始问题特征；或，向用户发出针对所述特征类型的提示信息，由所述用户基于所述提示信息提供新的原始输入问题，若所述新的原始输入问题中存在与所述特征类型匹配的关键值，则基于所述关键值确定所述初始问题特征。
- 3、根据权利要求 1 所述的方法，其特征在于，所述基于所述初始问题特征获取与所述初始问题特征关联的多个附加问题特征，包括：
将所述初始问题特征输入给已训练的特征分析目标模型，由所述特征分析目标模型输出与所述初始问题特征关联的多个附加问题特征；或者，
将所述初始问题特征输入给所述多模态目标模型，由所述多模态目标模型输出与所述初始问题特征关联的多个附加问题特征；或者，
基于所述初始问题特征分析与所述初始问题特征关联的多个附加问题特征。
- 4、根据权利要求 3 所述的方法，其特征在于，
所述特征分析目标模型是与目标场景类型匹配的特征分析目标模型；
所述问题模板是与所述目标场景类型匹配的问题模板；
其中，所述目标场景类型包括行为异常类型、车辆异常类型、或工业异常类型。
- 5、根据权利要求 1 所述的方法，其特征在于，所述基于所述初始问题特征和所述多个附加问题特征生成多个目标问题，包括：

从所述多个附加问题特征中选取候选附加问题特征;

在将所述初始问题特征和所述候选附加问题特征组合后,针对每一个组合,为所述初始问题特征和/或所述候选附加问题特征添加上下文信息,得到所述候选附加问题特征对应的目标问题。

6、根据权利要求1所述的方法,其特征在于,所述基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答案,包括:

基于每个目标问题对应的候选答案,从所有候选答案中过滤内容重复的候选答案,得到至少一个剩余候选答案,并确定每个剩余候选答案在所有候选答案中的出现次数;

将最大出现次数的剩余候选答案确定为所述原始输入问题对应的目标答案;或者,若最大出现次数对应的剩余候选答案在所有候选答案中的出现比例值大于预设阈值,则将最大出现次数的剩余候选答案确定为所述原始输入问题对应的目标答案。

7、根据权利要求6所述的方法,其特征在于,所述基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答案之后,所述方法还包括:

输出所述目标答案对应的目标数据,所述目标数据包括所述目标答案,或者,所述目标数据包括所述目标答案和所述目标答案对应的附加问题特征;

其中,所述目标答案包括图像、文本、声音和视频中至少一种维度的答案;

其中,所述目标答案对应多个候选答案,所述目标答案对应的附加问题特征包括所述目标答案对应的每个候选答案对应的附加问题特征。

8、一种基于多模态模型的数据处理装置,其特征在于,所述装置包括:

确定模块,用于基于原始输入问题确定与问题模板匹配的初始问题特征;

获取模块,用于获取与所述初始问题特征关联的多个附加问题特征;

生成模块,用于基于所述初始问题特征和所述多个附加问题特征生成多个目标问题,每个目标问题包括所述初始问题特征和至少一个附加问题特征;

所述获取模块,还用于针对每个目标问题,将所述目标问题输入给已训练的多模态目标模型,由所述多模态目标模型输出所述目标问题对应的候选答案;

所述确定模块,还用于基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答案,所述目标答案是所述原始输入问题的输出结果。

9、根据权利要求8所述的装置,其特征在于,

其中,所述确定模块基于原始输入问题确定与问题模板匹配的初始问题特征时具体用于:若所述问题模板包括特征类型,则确定所述原始输入问题中是否存在与所述特征类型匹配的关键值;若存在与所述特征类型匹配的关键值,则基于所述关键值确定所述

初始问题特征；若不存在与所述特征类型匹配的关键值，则对所述原始输入问题进行解析，得到与所述特征类型匹配的初始问题特征；或，向用户发出针对所述特征类型的提示信息，由所述用户基于所述提示信息提供新的原始输入问题，若所述新的原始输入问题中存在与所述特征类型匹配的关键值，则基于所述关键值确定所述初始问题特征；

其中，所述获取模块获取与所述初始问题特征关联的多个附加问题特征时具体用于：将所述初始问题特征输入给已训练的特征分析目标模型，由所述特征分析目标模型输出与所述初始问题特征关联的多个附加问题特征；或者，将所述初始问题特征输入给所述多模态目标模型，由所述多模态目标模型输出与所述初始问题特征关联的多个附加问题特征；或者，基于所述初始问题特征分析与所述初始问题特征关联的多个附加问题特征；

其中，所述特征分析目标模型是与目标场景类型匹配的特征分析目标模型；所述问题模板是与所述目标场景类型匹配的问题模板；其中，所述目标场景类型包括行为异常类型、车辆异常类型、或工业异常类型；

其中，所述生成模块基于所述初始问题特征和所述多个附加问题特征生成多个目标问题时具体用于：从多个附加问题特征中选取候选附加问题特征；在将所述初始问题特征和候选附加问题特征组合后，针对每一个组合，为所述初始问题特征和/或所述候选附加问题特征添加上下文信息，得到候选附加问题特征对应的目标问题；

其中，所述确定模块基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答案时具体用于：基于每个目标问题对应的候选答案，从所有候选答案中过滤内容重复的候选答案，得到至少一个剩余候选答案，并确定每个剩余候选答案在所有候选答案中的出现次数；将最大出现次数的剩余候选答案确定为所述原始输入问题对应的目标答案；或，若最大出现次数对应的剩余候选答案在所有候选答案中的出现比例值大于预设阈值，则将最大出现次数的剩余候选答案确定为所述原始输入问题对应的目标答案；

其中，所述确定模块基于每个目标问题对应的候选答案确定所述原始输入问题对应的目标答案之后还用于：输出所述目标答案对应的目标数据，所述目标数据包括所述目标答案，或者，所述目标数据包括所述目标答案和所述目标答案对应的附加问题特征；其中，所述目标答案包括图像、文本、声音和视频中至少一种维度的答案；所述目标答案对应多个候选答案，所述目标答案对应的附加问题特征包括所述目标答案对应的每个候选答案对应的附加问题特征。

10、一种电子设备，其特征在于，包括：处理器和机器可读存储介质，所述机器可读存储介质存储有能够被所述处理器执行的机器可执行指令；所述处理器用于执行所述机器可执行指令，以实现根据权利要求 1-7 中任一项所述的方法。

11、一种机器可读存储介质，其特征在于，所述机器可读存储介质上存储有若干计算机指令，所述计算机指令被处理器执行时，能够实现根据权利要求 1-7 中任一项所述的方法。