(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 115470775 A
(43)申请公布日 2022.12.13

(21)申请号 202211250816.0

(22)申请日 2022.10.12

(71)申请人 京东科技信息技术有限公司
地址 100176 北京市大兴区北京经济技术开发区科创十一街18号院2号楼6层601

(72)发明人 赵宇明 刘瑞雪 陈蒙

(74)专利代理机构 北京路浩知识产权代理有限公司 11002
专利代理师 王治东

(51)Int.Cl.
G06F 40/242(2020.01)
G06F 40/117(2020.01)
G06F 40/279(2020.01)
G06F 16/36(2019.01)

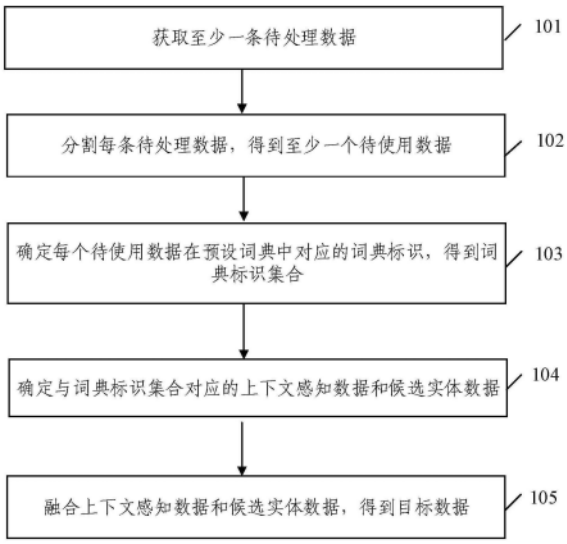
权利要求书3页 说明书13页 附图3页

(54)发明名称

数据处理方法、装置、设备、介质及产品

(57)摘要

本公开提供一种数据处理方法、装置、设备、介质及产品，方法包括：获取至少一条待处理数据；分割每条所述待处理数据，得到至少一个待使用数据；确定每个所述待使用数据在预设词典中对应的词典标识，得到词典标识集合；确定与所述词典标识集合对应的上下文感知数据和候选实体数据；融合所述上下文感知数据和所述候选实体数据，得到目标数据。本公开用以解决现有技术中的意图识别准确度低的缺陷，实现精准的识别当前问题的意图。



CN 115470775 A

1.一种数据处理方法,其特征在于,包括:

获取至少一条待处理数据;

分割每条所述待处理数据,得到至少一个待使用数据;

确定每个所述待使用数据在预设词典中对应的词典标识,得到词典标识集合;

确定与所述词典标识集合对应的上下文感知数据和候选实体数据;

融合所述上下文感知数据和所述候选实体数据,得到目标数据。

2.根据权利要求1所述的数据处理方法,其特征在于,所述分割每条所述待处理数据,得到至少一个待使用数据,包括:

基于字粒度分割策略,分割所述每条待处理数据,得到至少一个第一分割数据;

基于词粒度分割策略,分割所述每条待处理数据,得到至少一个第二分割数据;

将所述第一分割数据和所述第二分割数据的并集作为所述待使用数据。

3.根据权利要求2所述的数据处理方法,其特征在于,所述确定每个所述待使用数据在预设词典中对应的词典标识,得到词典标识集合,包括:

将每个所述第一分割数据在所述预设词典中进行匹配,得到第一匹配结果;将所述第一匹配结果在所述预设词典中对应的第一词典标识作为第一词典标识子集合;

将每个所述第二分割数据在所述预设词典中进行匹配,得到第二匹配结果;将所述第二匹配结果在所述预设词典中对应的第二词典标识作为第二词典标识子集合;

组合所述第一词典标识子集合和所述第二词典标识子集合,得到所述词典标识集合。

4.根据权利要求3所述的数据处理方法,其特征在于,所述确定与所述词典标识集合对应的上下文感知数据,包括:

确定每个所述第一词典标识对应的第一向量,得到第一文本表征矩阵;

确定每个所述第二词典标识对应的第二向量,得到第二文本表征矩阵;

将所述第一文本表征矩阵输入上下文感知模型,得到所述上下文感知模型输出的第一感知数据,以及将所述第二文本表征矩阵输入所述上下文感知模型,得到所述上下文感知模型输出的第二感知数据;

组合所述第一感知数据和所述第二感知数据,得到所述上下文感知数据;

其中,所述上下文感知模型基于文本表征矩阵样本和上下文感知数据样本训练得到。

5.根据权利要求3所述的数据处理方法,其特征在于,所述确定与所述词典标识集合对应的候选实体数据,包括:

将所述词典标识集合输入实体候选模型,得到所述实体候选模型输出的所述候选实体数据;

其中,所述实体候选模型基于词典标识集合样本和候选实体数据样本训练得到。

6.根据权利要求5所述的数据处理方法,其特征在于,所述融合所述上下文感知数据和所述候选实体数据,得到目标数据,包括:

基于所述候选实体数据,从所述第一感知数据中提取第一目标感知数据,以及从第二感知数据中提取第二目标感知数据;

将所述第一目标感知数据、第二目标感知数据和所述候选实体数据输入预设计算公式,得到所述预设计算公式输出的第一表征感知数据和第二表征感知数据;

分别对所述第一表征感知数据和所述第二表征感知数据进行池化操作,得到第一池化

感知数据和第二池化感知数据；

拼接所述第一池化感知数据和所述第二池化感知数据，得到所述目标数据。

7.根据权利要求6所述的数据处理方法，其特征在于，所述将所述第一目标感知数据、第二目标感知数据和所述候选实体数据输入预设计算公式，得到所述预设计算公式输出的第一表征感知数据和第二表征感知数据，包括：

将所述第一目标感知数据、第二目标感知数据和所述候选实体数据输入所述第一预设计算子公式，得到所述第一预设计算子公式输出的注意力权重矩阵；

将所述第一目标感知数据、第二目标感知数据、所述候选实体数据和所述注意力权重矩阵输入第二预设计算子公式，得到所述第二预设计算子公式输出的新的所述第一目标感知数据、新的所述第二目标感知数据和新的所述候选实体数据；

将所述新的第一目标感知数据、所述新的第二目标感知数据和所述新的候选实体数据输入第三预设计算子公式，得到所述第三预设计算子公式输出的所述第一表征感知数据和所述第二表征感知数据。

8.根据权利要求1-7任一项所述的数据处理方法，其特征在于，所述确定每个所述待使用数据在预设词典中对应的词典标识，得到词典标识集合之前，还包括：

获取所述待使用数据的数据长度信息；

在确定所述数据长度信息大于预设值的情况下，从所述待使用数据中提取所述数据长度信息为所述预设值的目标待使用数据，将所述目标待使用数据作为所述待使用数据；

在确定所述数据长度信息小于或等于所述预设值的情况下，执行所述确定每个所述待使用数据在预设词典中对应的词典标识，得到词典标识集合的步骤。

9.根据权利要求1-7任一项所述的数据处理方法，其特征在于，所述获取至少一条待处理数据之后，还包括：

获取与所述待处理数据对应的真实意图，其中，所述真实意图通过用户设置得到；

所述融合所述上下文感知数据和所述候选实体数据，得到目标数据之后，还包括：

预测所述目标数据对应的预测意图；

在所述真实意图和所述预测意图的一致率大于预设值的情况下，确定所述目标数据合格。

10.一种数据处理装置，其特征在于，包括：

获取模块，用于获取至少一条待处理数据；

分割模块，用于分割每条所述待处理数据，得到至少一个待使用数据；

第一确定模块，用于确定每个所述待使用数据在预设词典中对应的词典标识，得到词典标识集合；

第二确定模块，用于确定与所述词典标识集合对应的上下文感知数据和候选实体数据；

融合模块，用于融合所述上下文感知数据和所述候选实体数据，得到目标数据。

11.一种电子设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述程序时实现如权利要求1至9任一项所述的数据处理方法。

12.一种非暂态计算机可读存储介质，其上存储有计算机程序，其特征在于，该计算机

程序被处理器执行时实现如权利要求1至9任一项所述的数据处理方法。

13.一种计算机程序产品，包括计算机程序，其特征在于，所述计算机程序被处理器执行时实现如权利要求1至9任一项所述的数据处理方法。

数据处理方法、装置、设备、介质及产品

技术领域

[0001] 本公开涉及意图识别技术领域，尤其涉及一种数据处理方法、装置、设备、介质及产品。

背景技术

[0002] 意图识别任务的输入是对话上下文，包括了当前问题，输出则是当前问题的意图。

[0003] 为了识别当前问题的意图，现有技术采用确定上下文的自注意力加权来表示该上下文，再进一步确定加权对应的实体候选，以预测当前问题的意图。但是采用自注意力加权的方式，可能会导致信息的丢失，导致预测的意图出现偏差，预设准确率较低。

发明内容

[0004] 本公开提供一种数据处理方法、装置、设备、介质及产品，用以解决现有技术中的意图识别准确度低的缺陷，实现精准的识别当前问题的意图。

[0005] 本公开提供一种数据处理方法，包括：

[0006] 获取至少一条待处理数据；

[0007] 分割每条所述待处理数据，得到至少一个待使用数据；

[0008] 确定每个所述待使用数据在预设词典中对应的词典标识，得到词典标识集合；

[0009] 确定与所述词典标识集合对应的上下文感知数据和候选实体数据；

[0010] 融合所述上下文感知数据和所述候选实体数据，得到目标数据。

[0011] 根据本公开提供的一种数据处理方法，所述分割每条所述待处理数据，得到至少一个待使用数据，包括：

[0012] 基于字粒度分割策略，分割所述每条待处理数据，得到至少一个第一分割数据；

[0013] 基于词粒度分割策略，分割所述每条待处理数据，得到至少一个第二分割数据；

[0014] 将所述第一分割数据和所述第二分割数据的并集作为所述待使用数据。

[0015] 根据本公开提供的一种数据处理方法，所述确定每个所述待使用数据在预设词典中对应的词典标识，得到词典标识集合，包括：

[0016] 将每个所述第一分割数据在所述预设词典中进行匹配，得到第一匹配结果；将所述第一匹配结果在所述预设词典中对应的第一词典标识作为第一词典标识子集合；

[0017] 将每个所述第二分割数据在所述预设词典中进行匹配，得到第二匹配结果；将所述第二匹配结果在所述预设词典中对应的第二词典标识作为第二词典标识子集合；

[0018] 组合所述第一词典标识子集合和所述第二词典标识子集合，得到所述词典标识集合。

[0019] 根据本公开提供的一种数据处理方法，所述确定与所述词典标识集合对应的上下文感知数据，包括：

[0020] 确定每个所述第一词典标识对应的第一向量，得到第一文本表征矩阵；

[0021] 确定每个所述第二词典标识对应的第二向量，得到第二文本表征矩阵；

[0022] 将所述第一文本表征矩阵输入上下文感知模型，得到所述上下文感知模型输出的第一感知数据，以及将所述第二文本表征矩阵输入所述上下文感知模型，得到所述上下文感知模型输出的第二感知数据；

[0023] 组合所述第一感知数据和所述第二感知数据，得到所述上下文感知数据；

[0024] 其中，所述上下文感知模型基于文本表征矩阵样本和上下文感知数据样本训练得到。

[0025] 根据本公开提供的一种数据处理方法，所述确定与所述词典标识集合对应的候选实体数据，包括：

[0026] 将所述词典标识集合输入实体候选模型，得到所述实体候选模型输出的所述候选实体数据；

[0027] 其中，所述实体候选模型基于词典标识集合样本和候选实体数据样本训练得到。

[0028] 根据本公开提供的一种数据处理方法，所述融合所述上下文感知数据和所述候选实体数据，得到目标数据，包括：

[0029] 基于所述候选实体数据，从所述第一感知数据中提取第一目标感知数据，以及从第二感知数据中提取第二目标感知数据；

[0030] 将所述第一目标感知数据、第二目标感知数据和所述候选实体数据输入预设计算公式，得到所述预设计算公式输出的第一表征感知数据和第二表征感知数据；

[0031] 分别对所述第一表征感知数据和所述第二表征感知数据进行池化操作，得到第一池化感知数据和第二池化感知数据；

[0032] 拼接所述第一池化感知数据和所述第二池化感知数据，得到所述目标数据。

[0033] 根据本公开提供的一种数据处理方法，所述将所述第一目标感知数据、第二目标感知数据和所述候选实体数据输入预设计算公式，得到所述预设计算公式输出的第一表征感知数据和第二表征感知数据，包括：

[0034] 将所述第一目标感知数据、第二目标感知数据和所述候选实体数据输入所述第一预设计算子公式，得到所述第一预设计算子公式输出的注意力权重矩阵；

[0035] 将所述第一目标感知数据、第二目标感知数据、所述候选实体数据和所述注意力权重矩阵输入第二预设计算子公式，得到所述第二预设计算子公式输出的新的所述第一目标感知数据、新的所述第二目标感知数据和新的所述候选实体数据；

[0036] 将所述新的第一目标感知数据、所述新的第二目标感知数据和所述新的候选实体数据输入第三预设计算子公式，得到所述第三预设计算子公式输出的所述第一表征感知数据和所述第二表征感知数据。

[0037] 根据本公开提供的一种数据处理方法，所述确定每个所述待使用数据在预设词典中对应的词典标识，得到词典标识集合之前，还包括：

[0038] 获取所述待使用数据的数据长度信息；

[0039] 在确定所述数据长度信息大于预设值的情况下，从所述待使用数据中提取所述数据长度信息为所述预设值的目标待使用数据，将所述目标待使用数据作为所述待使用数据；

[0040] 在确定所述数据长度信息小于或等于所述预设值的情况下，执行所述确定每个所述待使用数据在预设词典中对应的词典标识，得到词典标识集合的步骤。

[0041] 根据本公开提供的一种数据处理方法，所述获取至少一条待处理数据之后，还包括：

[0042] 获取与所述待处理数据对应的真实意图，其中，所述真实意图通过用户设置得到；

[0043] 所述融合所述上下文感知数据和所述候选实体数据，得到目标数据之后，还包括：

[0044] 预测所述目标数据对应的预测意图；

[0045] 在所述真实意图和所述预测意图的一致率大于预设值的情况下，确定所述目标数据合格。

[0046] 本公开还提供一种数据处理装置，包括：

[0047] 获取模块，用于获取至少一条待处理数据；

[0048] 分割模块，用于分割每条所述待处理数据，得到至少一个待使用数据；

[0049] 第一确定模块，用于确定每个所述待使用数据在预设词典中对应的词典标识，得到词典标识集合；

[0050] 第二确定模块，用于确定与所述词典标识集合对应的上下文感知数据和候选实体数据；

[0051] 融合模块，用于融合所述上下文感知数据和所述候选实体数据，得到目标数据。

[0052] 本公开还提供一种电子设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述程序时实现如上述任一种所述的数据处理方法。

[0053] 本公开还提供一种非暂态计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时实现如上述任一种所述数据处理方法。

[0054] 本公开还提供一种计算机程序产品，包括计算机程序，所述计算机程序被处理器执行时实现如上任一种所述的数据处理方法。

[0055] 本公开提供的数据处理方法、装置、设备、介质及产品，通过获取至少一条待处理数据；分割每条待处理数据，得到至少一个待使用数据；确定每个待使用数据在预设词典中对应的词典标识，得到词典标识集合，可见，本公开采用对待处理数据进行分割的方式，完整的保留了待处理数据的所有信息，避免了可能出现的信息缺失的问题；确定与词典标识集合对应的上下文感知数据和候选实体数据；融合上下文感知数据和候选实体数据，得到目标数据，可见，本公开基于完整的待处理数据确定与之对应的上下文感知数据和候选实体数据，为意图识别提供了准确的数据基础，解决了现有技术意图识别准确率低的问题。

附图说明

[0056] 为了更清楚地说明本公开或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图是本公开的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

[0057] 图1是本公开提供的数据处理方法的流程示意图之一；

[0058] 图2是本公开提供的数据处理方法的流程示意图之二；

[0059] 图3是本公开提供的数据处理装置的结构示意图；

[0060] 图4是本公开提供的电子设备的结构示意图。

具体实施方式

[0061] 为使本公开实施例的目的、技术方案和优点更加清楚，下面将结合本公开实施例中的附图，对本公开实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本公开实施例一部分实施例，而不是全部的实施例。基于本公开实施例中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本公开实施例保护的范围。

[0062] 现有技术的意图识别，通过训练语言模型，来实现意图识别，该语言模型基于深度神经网络的分类模型。基础流程为：首先将输入由高维稀疏的独热码空间映射到对应的低维稠密的嵌入向量空间，然后再通过神经网络，将输入再度压缩和映射，最后将神经网络的输出表征向量送入分类器中，常用的是线性分类器，最后以交叉熵作为损失函数训练语言模型。这个过程中，有把知识融入到最终表征向量中，也有通过层次结构提取不同粒度的输入信息。

[0063] 目前的现有技术大多仅关注融入知识或者单粒度的输入，未层考虑不同粒度的输入，以及与知识之间的融合，融合方式间接、粒度单一的问题。

[0064] 本公开提供的数据处理方法，应用在意图识别技术领域，该方法可以是意图识别之前的数据处理过程，即对语料数据进行预处理，进而基于目标数据进行意图识别；也可以是意图识别的中间过程，即，对语料数据进行预处理，基于目标数据进行意图识别。

[0065] 本公开将三路表征进行融合，将字粒度、词粒度和外部知识三者进行融合，提高了意图识别的准确性。

[0066] 下面结合图1-图2描述本公开实施例的数据处理方法。

[0067] 本公开实施例提供了一种数据处理方法，该方法可以应用在服务器中，也可以应用在智能终端中。下面，以该方法应用在服务器中为例进行说明。但需要声明的是此处仅为举例说明，并不用于对保护范围进行限定。在后续的实施例中的举例也不用于对保护范围进行限定，之后不再一一说明。该方法的具体实现，如图1所示：

[0068] 步骤101，获取至少一条待处理数据。

[0069] 具体的，本公开可以应用在多个领域，例如，电商领域，通过分析用户的问题，来确定用户的购买意图；又例如，智能机器人领域，通过分析用户的语言，来确定用户的控制意图；等。下面，以本公开应用在电商领域为例进行说明。

[0070] 具体的，获取用户与客服的对话数据，其中，对话数据是以会话维度进行的，在一个会话中，用户和客服交替发出消息，对话数据通过序列$\{q_1,a_1,\ldots,q_n,a_n\}$表示。其中，其中$q_i$表示用户消息，$a_i$表示客服消息，i表示当前轮次，n表示总的对话轮次数。

[0071] 下面，以待处理数据为用户消息为例进行说明：

[0072] 提取对话数据中的用户消息，得到至少一条待处理数据，通过序列$X=\{q_1,\ldots,q_n\}$表示。

[0073] 此时，通过人工设置该用户消息对应的真实意图，通过字母y表示真实意图。

[0074] 步骤102，分割每条待处理数据，得到至少一个待使用数据。

[0075] 一个具体实施例中，获得待使用数据的具体实现方式如下所示：

[0076] 基于字粒度分割策略，分割每条待处理数据，得到至少一个第一分割数据；基于词粒度分割策略，分割每条待处理数据，得到至少一个第二分割数据；将第一分割数据和第二

分割数据的并集作为待使用数据。

[0077] 其中，通过X^{char}表示第一分割数据；通过X^{word}表示第二分割数据。

[0078] 具体的，下面以一条待处理数据为例进行说明：

[0079] 为了能够获取待处理数据的字粒度表示，将X按照字粒度分割策略进行分割，得到X^{char}，其中，$X^{char}=\{x_1^c,\dots,x_l^c\}$，其中，$X^{char}$的文本序列的长度为l，$x_i^c$表示其中第i个字。

[0080] 为了能够获取待处理数据的词粒度表示，将X按照词粒度分割策略进行分割，得到X^{word}，其中，$X^{word}=\{x_1^w,\dots,x_m^w\}$，其中，$X^{word}$的文本序列的长度为m，$x_i^w$表示其中第i个词。

[0081] 例如，一条待处理数据为我喜欢吃苹果，按照字粒度分割策略，得到的X^{char}＝{我，喜，欢，吃，苹，果}；按照词粒度分割策略，得到的X^{word}＝{我，喜欢，吃，苹果}；第一分割数据和第二分割数据的并集为$X^{char}UX^{word}$＝{我，喜，欢，吃，苹，果，喜欢，苹果}。

[0082] 本公开采用字粒度分割策略和词粒度分割策略分割待处理数据，得到的待使用数据还是完整的待处理数据，相较于现有技术利用加权去表征实体短语，不会有任何数据的丢失，为后期意图识别提供了有效的数据基础。

[0083] 并且，本公开采用字粒度分割策略和词粒度分割策略分割待处理数据，得到多粒度和细粒度的待使用数据，为后期意图识别提供了有效的数据基础。

[0084] 步骤103，确定每个待使用数据在预设词典中对应的词典标识，得到词典标识集合。

[0085] 具体的，预设词典为预先创建的存储有任意待处理数据的存储库，以及每个待处理数据对应的词典标识。其中，词典标识可以为待处理数据在预设词典中的位置信息。可以将预设词典作为是一个数组。

[0086] 具体的，预设词典的创建过程如下所示：

[0087] 以一条待处理数据样本为例进行说明：将待处理数据样本分别按照字粒度分割策略和词粒度分割策略进行分割，得到X^{char}和X^{word}。

[0088] 集合每条待处理数据样本得到的X^{char}，得到词典V^{char}；集合每条待处理数据样本得到的X^{word}，得到词典V^{word}；合并V^{char}和V^{word}，得到预设词典，其中，通过V表示预设词典。

[0089] 一个具体实施例中，获得词典标识集合的具体实现方式如下所示：

[0090] 将每个第一分割数据在预设词典中进行匹配，得到第一匹配结果；将第一匹配结果在预设词典中对应的第一词典标识作为第一词典标识子集合；将每个第二分割数据在预设词典中进行匹配，得到第二匹配结果；将第二匹配结果在预设词典中对应的第二词典标识作为第二词典标识子集合；组合第一词典标识子集合和第二词典标识子集合，得到词典标识集合。

[0091] 其中，通过S^{char}表示第一词典标识子集合；通过S^{word}表示第二词典标识子集合。

[0092] 具体的，基于字粒度分割策略得到第一分割数据和基于词粒度分割策略得到第二分割数据之后，将第一分割数据中的每个参量转换为预设词典中对应的第一词典标识，将第二分割数据中的每个参量转换为预设词典中对应的第二词典标识。

[0093] 本公开将文字信息转化为向量表示，提高了数据处理的计算速度。

[0094] 一个具体实施例中，确定每个待使用数据在预设词典中对应的词典标识，得到词

典标识集合之前，获取待使用数据的数据长度信息；比对数据长度信息和预设值的大小关系；在确定数据长度信息大于预设值的情况下，从待使用数据中提取数据长度信息为预设值的目标待使用数据，将目标待使用数据作为待使用数据；在确定数据长度信息小于或等于预设值的情况下，执行确定每个待使用数据在预设词典中对应的词典标识，得到词典标识集合的步骤。

[0095] 具体的，提取策略可以从待处理数据的最后位置开始计算，直至数据长度信息等于预设值时，将该部分待使用数据作为目标待使用数据。

[0096] 本公开在数据长度信息大于预设值的情况下，对待使用数据进行截取，以得到符合要求的待使用数据，避免数据量太大，导致的数据处理速率变慢的问题。

[0097] 具体的，数据长度信息为：按照字粒度分割策略进行分割，得到的所有X^{char}的文本序列的长度的和，定义为第一长度，以及按照词粒度分割策略进行分割，得到的所有X^{word}的文本序列的长度的和，定义为第二长度。

[0098] 具体的，比对第一长度和预设值的大小，以及第二长度和预设值的大小。

[0099] 在确定第一长度大于预设值，且，第二长度大于预设值的情况下，分别从第一分割数据和第二分割数据中提取数据长度信息为预设值的第一分割目标数据和第二分割目标数据；将第一分割目标数据和第二分割目标数据作为目标待使用数据。

[0100] 在确定第一长度大于预设值，且，第二长度小于或等于预设值的情况下，从第一分割数据中提取数据长度信息为预设值的第一分割目标数据，将第一分割目标数据和第二分割数据作为目标待使用数据。

[0101] 在确定第一长度小于或等于预设值，且，第二长度大于预设值的情况下，从第二分割数据中提取数据长度信息为预设值的第二分割目标数据，将第一分割数据和第二分割目标数据作为目标待使用数据。

[0102] 在确定第一长度小于或等于预设值，且，第二长度小于或等于预设值的情况下，执行确定每个待使用数据在预设词典中对应的词典标识，得到词典标识集合的步骤。

[0103] 具体的，$S^{char}=\{s^{c}_{1},\dots,s^{c}_{L_S}\}$，$S^{word}=\{s^{w}_{1},\dots,s^{w}_{L_S}\}$，其中。Ls为预设值。

[0104] 步骤104，确定与词典标识集合对应的上下文感知数据和候选实体数据。

[0105] 一个具体实施例中，获得上下文感知数据的具体实现方式如下所示：

[0106] 确定每个第一词典标识对应的第一向量，得到第一文本表征矩阵；确定每个第二词典标识对应的第二向量，得到第二文本表征矩阵；将第一文本表征矩阵输入上下文感知模型，得到上下文感知模型输出的第一感知数据，以及将第二文本表征矩阵输入上下文感知模型，得到上下文感知模型输出的第二感知数据；组合第一感知数据和第二感知数据，得到上下文感知数据；其中，上下文感知模型基于文本表征矩阵样本和上下文感知数据样本训练得到。

[0107] 具体的，预设词典还包括：每个待处理数据的向量表示。

[0108] 在得到词典标识集合之后，从预设词典中确定每个第一词典标识对应的第一向量和每个第二词典标识对应的第二向量，得到字粒度对应的第一文本表征矩阵和词粒度对应的第二文本表征矩阵。

[0109] 其中，第一文本表征矩阵用E^{char}表示，$E^{char}=\{e^{c}_{1},\ldots,e^{c}_{L_S}\}$；第二文本表征矩阵用$E^{word}$表示，$E^{word}=\{e^{w}_{1},\ldots,e^{w}_{L_S}\}$。

[0110] 或者，在得到词典标识集合之后，随机确定每个第一词典标识对应的第一向量和每个第二词典标识对应的第二向量，得到字粒度对应的第一文本表征矩阵和词粒度对应的第二文本表征矩阵。

[0111] 其中，上下文感知模型为神经网络模型，例如，TextCNN、RNN、Transformer等。上下文感知模型用于确定第一文本表征矩阵中的每个第一向量的上下文信息，以得到第一感知数据；确定和第二文本表征矩阵中的每个第二向量的上下文信息，以得到第二感知数据。

[0112] 其中，第一感知数据通过E'^{char}表示，$E'^{char}=\{e'^{c}_{1},\ldots,e'^{c}_{L_S}\}$；第二感知数据通过$E'^{word}$表示，$E'^{word}=\{e'^{w}_{1},\ldots,e'^{w}_{L_S}\}$，其中，$e'_i\in\mathbb{R}^{d_x}$。

[0113] 本公开在得到第一文本表征矩阵和第二文本表征矩阵之后，利用预先训练好的上下文感知模型，得到上下文感知数据，以确定待处理数据中每个字和词的上下文信息，提高了意图识别的准确性。

[0114] 一个具体实施例中，本公开在确定待处理数据的上下文感知数据的基础上，还要确定外部知识表征。通过将词典标识集合输入实体候选模型，得到实体候选模型输出的候选实体数据；其中，实体候选模型基于词典标识集合样本和候选实体数据样本训练得到。

[0115] 其中，实体候选模型为基于词典标识集合样本、候选实体数据样本、知识图谱样本和候选实体的关联关系训练得到。

[0116] 其中，候选实体数据包括：连接到M个知识图谱中的C个候选实体、以及每个候选实体在知识图谱中的起始位置和终止位置。

[0117] 具体的，将S^{char}输入实体候选模型，得到实体候选模型输出的F个第一预测候选实体数据，其中，第一预测候选实体数据包括：连接到Q个知识图谱中的第一候选实体、以及第一候选实体的在知识图谱中的起始位置和终止位置，其中，F为大于或等于1的整数，Q为大于或等于1的整数。

[0118] 将S^{word}输入实体候选模型，得到实体候选模型输出的D个第二预测候选实体数据，其中，第二预测候选实体数据包括：连接到R个知识图谱中的第二候选实体、以及第二候选实体的在知识图谱中的起始位置和终止位置，其中，D为大于或等于1的整数，R为大于或等于1的整数。

[0119] 其中，F和R的总和为C，Q和R的总和为M，

$\mathcal{C}=\{<\mathrm{start}_i,\mathrm{end}_i>,(e^{e}_{1},\ldots,e^{e}_{M})|\mathrm{i}\in 1\ldots\mathrm{C}\},e^{e}_{1}\in\mathbb{R}^{d_e}$，通过$E^{entity}$表示候选实体数据。

[0120] 当然，此处的输入实体候选模型的词典标识集合可以仅包括第一词典子集合。

[0121] 本公开基于预先设置的实体候选模型，得到与待处理数据对应的外部知识，为后期上下文感知数据和候选实体数据的融合，提供了有效的数据基础。

[0122] 步骤105，融合上下文感知数据和候选实体数据，得到目标数据。

[0123] 一个具体实施例中，融合上下文感知数据和候选实体数据，得到目标数据的具体实现方式如下所示：

[0124] 基于候选实体数据，从第一感知数据中提取第一目标感知数据，以及从第二感知数据中提取第二目标感知数据；将第一目标感知数据、第二目标感知数据和候选实体数据输入预设计算公式，得到预设计算公式输出的第一表征感知数据和第二表征感知数据；分别对第一表征感知数据和第二表征感知数据进行池化操作，得到第一池化感知数据和第二池化感知数据；拼接第一池化感知数据和第二池化感知数据，得到目标数据。

[0125] 具体的，匹配候选实体数据中的候选实体与第一感知数据中的第一分割数据，得到第三匹配结果，从第一感知数据中，提取与第三匹配结果对应第一向量，得到第一目标感知数据。以及匹配候选实体数据与第二感知数据中的第二分割数据，得到第四匹配结果，从第二感知数据中，提取与第四匹配结果对应的第二向量，得到第二目标感知数据。

[0126] 其中，通过T^{char}表示第一目标感知数据，通过T^{word}表示第二目标感知数据。

[0127] 具体的，第一目标感知数据为预设矩阵大小的矩阵，第二目标感知数据亦为预设矩阵大的矩阵，例如，第一目标感知数据和第二目标感知数据对应的矩阵均为$\mathbb{R}^{K\times d_a}$。当第一目标感知数据和/或第二目标感知数据对应的矩阵小于$\mathbb{R}^{K\times d_a}$时，用零补齐即可。最终可以得到$T^{char}\in\mathbb{R}^{C\times K\times d_a}$、$T^{word}\in\mathbb{R}^{C\times K\times d_a}$。

[0128] 其中，通过Z^{char}表示第一表征感知数据，通过Z^{word}表示第二表征感知数据。

[0129] 具体的，采用平均池化操作，分别对Z^{char}和Z^{word}进行压缩，得到Z'^{char}和Z'^{word}。拼接Z'^{char}和Z'^{word}，得到目标数据$Z^o=Z^{char}\oplus Z^{word}, Z^o\in\mathbb{R}^{2d_a}$，其中，$\oplus$表示维度的拼接。其中。$Z'^{char}$表示第一池化感知数据，$Z'^{word}$表示第二池化感知数据，$Z^o$表示目标数据。

[0130] 其中，预设计算公式为一系列计算公式，下面通过具体实施例进行具体说明：

[0131] 一个具体实施例中，将第一目标感知数据、第二目标感知数据和候选实体数据输入第一预设计算子公式，得到第一预设计算子公式输出的注意力权重矩阵；将第一目标感知数据、第二目标感知数据、候选实体数据和注意力权重矩阵输入第二预设计算子公式，得到第二预设计算子公式输出的新的第一目标感知数据、新的第二目标感知数据和新的候选实体数据；将新的第一目标感知数据、新的第二目标感知数据和新的候选实体数据输入第三预设计算子公式，得到第三预设计算子公式输出的第一表征感知数据和第二表征感知数据。

[0132] 其中，将第一感知数据、第二感知数据和候选实体数据定义为三路表征数据。

[0133] 具体的，将第一目标感知数据、第二目标感知数据和候选实体数据输入第一预设计算子公式，得到第一预设计算子公式输出的注意力权重矩阵之前，通过公式(1)、公式(2)和公式(3)对三路表征数据进行线性变换，将三路表征数据统一至相同维度：

[0134] $E'^{charproj}=E'^{char}W^{charproj}+b^{charproj}$ (1)

[0135] $E'^{wordproj}=E'^{word}W^{wordproj}+b^{wordproj}$ (2)

[0136] $E^{entityroj}=E^{entity}W^{entityproj}+b^{entityproj}$ (3)

[0137] 其中，公式(1)、公式(2)和公式(3)同时运算。

[0138] 其中，$W^{charproj}\in\mathbb{R}^{d_x\times d_a}$为常量，$b^{charproj}\in\mathbb{R}^{d_a}$为常量，$W^{wordproj}\in\mathbb{R}^{d_x\times d_a}$为常量，$b^{wordproj}\in\mathbb{R}^{d_a}$为常量，$W^{entityproj}\in\mathbb{R}^{d_e\times d_a}$为常量，$b^{entityproj}\in\mathbb{R}^{d_a}$为常量，$E^{entityroj}\in\mathbb{R}^{C\times M\times d_a}$为变量。

[0139] 此时，通过公式(1)得到了变换维度后的候选实体数据$E^{entityroj}$。

[0140] 具体的，第一预设计算子公式见公式(4)和公式(5)：

[0141] $$A=\frac{\sum_{d_a}\sum_{d_a}\sum_{d_a}T^{char}\otimes T^{word}\otimes E^{entityroj}}{\sqrt{d_a}} \quad (4)$$

[0142] $$A_{ijk}=p^{T}\left(T^{char}{}_{i}\circ T^{word}{}_{j}\circ E^{entityroj}{}_{k}\right)/\sqrt{d_a} \quad (5)$$

[0143] 其中，$A\in\mathbb{R}^{CK\times CK\times CM}$，$\otimes$代表矩阵乘操作，$p\in\mathbb{R}^{d_a}$，$\circ$表示对应元素相乘。

[0144] 通过公式(5)对注意力权重矩阵A进行逐元素归一化(softmax)操作，得到最终的注意力权重矩阵A，其中，A_{ijk}表示对任意一个元素进行softmax操作的结果。

[0145] 具体的，第二预设计算子公式见公式(6)、公式(7)、公式(8)、公式(9)、公式(10)和公式(11)：

[0146] $F^{char}=\sum_{CK}\sum_{CM}AT^{char}$ (6)

[0147] $F^{word}=\sum_{CK}\sum_{CM}AT^{word}$ (7)

[0148] $F^{entity}=\sum_{CK}\sum_{CK}AE^{entity}$ (8)

[0149] 其中，公式(6)、公式(7)和公式(8)同时运算，进而，将通过公式(6)、公式(7)和公式(8)得到的结果，带入公式(9)、公式(10)和公式(11)：

[0150] $H^{char}=T^{char}+F^{char}+F^{entity}$ (9)

[0151] $H^{word}=T^{word}+F^{word}+F^{entity}$ (10)

[0152] $H^{entity}=E^{entity}+F^{entity}$ (11)

[0153] 其中，$H^{char}\in\mathbb{R}^{C\times K\times d_a}$、$H^{word}\in\mathbb{R}^{C\times K\times d_a}$、$H^{entity}\in\mathbb{R}^{C\times M\times d_a}$。

[0154] 其中，公式(9)、公式(10)和公式(11)同时进行运算，得到新的第一目标感知数据H^{char}、新的第二目标感知数据H^{word}和新的候选实体数据H^{entity}。

[0155] 具体的，第三预设计算子公式见公式(12)和公式(13)：

[0156] $Z^{char}=MLP(MultiHeadAttn(H^{entity},H^{char},H^{char}))$ (12)

[0157] $Z^{word}=MLP(MultiHeadAttn(H^{entity},H^{word},H^{word}))$ (13)

[0158] 其中，MLP为预先创建的多层感知机模型(MultilayerPerceptron，简称MLP)，MultiHeadAttn为预先创建的神经网络模型。

[0159] 其中，Z^{char}表示第一表征感知数据，Z^{word}表示第二表征感知数据。

[0160] 本公开通过直接融合上下文感知数据和候选实体数据，为提高意图识别提供了有效的数据基础。

[0161] 一个具体实施例中，在获取至少一条待处理数据之后，获取与待处理数据对应的真实意图，其中，真实意图通过用户设置得到。并在得到目标数据之后，将目标数据输入线性分类器，得到线性分类器输出的预测意图；比对真实意图和预测意图的一致率；在真实意图和预测意图的一致率大于预设值的情况下，确定目标数据合格；在真实意图和预测意图的一致率小于或等于预设值的情况下，优化上下文感知模型、实体候选模型、和上下文感知数据与候选实体数据的融合方式，并在优化后重复执行步骤101-步骤105。

[0162] 具体的，将目标数据输入公式(14)：

[0163] $p=softmax(WZ^{o})$ (14)

[0164] 其中，p为预测意图，$W\in\mathbb{R}^{2d_a\times N}$为常量。

[0165] 具体的，比对真实意图和预测意图的一致率，计算损失情况，见公式(15)：

[0166] L＝CE(p,y) (15)

[0167] 其中,CE表示交叉熵函数,p表示预测意图,y表示真实意图。

[0168] 本公开通过确定预测意图的准确性,来进行目标数据的合格性的判定,为准确的意图识别提供了有效的数据基础。

[0169] 下面,通过图2对本公开进行具体说明:

[0170] 步骤201,获取用户消息,以及用户消息对应的真实意图。

[0171] 例如,用户消息包括用户与客户的交流时用户的问题信息,真实意图包括用户的问题诉求。下面,以用户消息为“该订单如何申请退款”为例进行说明。

[0172] 步骤202,基于字粒度分割策略,得到第一分割数据,以及基于词粒度分割策略,得到第二分割数据。

[0173] 例如,基于字粒度分割策略,得到的第一分割数据为{该,订,单,如,何,申,请,退,款},基于词粒度分割策略,得到的第二分割数据为{该,订单,如何,申请,退款}。

[0174] 步骤203,将第一分割数据在预设词典中进行匹配,得到第一词典标识子集合,以及将第二分割数据在预设词典中进行匹配,得到第二词典标识子集合。

[0175] 其中,例如,预设词典为数组[该,订,单,如,何,申,请,退,款,订单,如何,申请,退款],第一词典标识子集合为{0、1、2、3、4、5、6、7、8},第二词典标识子集合为{0、9、10、11、12}。

[0176] 步骤204,确定第一词典子集合中第一词典标识对应的第一向量,得到第一文本表征矩阵,以及确定第二词典子集合中的第二词典标识对应的第二向量,得到第二文本表征矩阵。

[0177] 具体的,每个词典标识对应一个向量表示,词典标识包括第一词典标识和第二词典标识,向量表示包括第一向量和第二向量。步骤205,将第一文本表征矩阵输入上下文感知模型,得到第一感知数据,以及将第二文本表征矩阵输入上下文感知模型,得到第二感知数据。

[0178] 步骤206,将第一词典子集合输入实体候选模型,得到候选实体数据。

[0179] 其中,候选实体数据包括:连接到M个知识图谱中的C个候选实体、以及每个候选实体在知识图谱中的起始位置和终止位置。

[0180] 步骤207,融合第一感知数据、第二感知数据和候选实体数据,得到目标数据。

[0181] 其中,按照公式(1)至公式(13)的具体方式,实现第一感知数据、第二感知数据和候选实体数据的融合。

[0182] 步骤208,通过线性分类器,预测目标数据对应的预测意图。

[0183] 步骤209,基于真实意图和预测意图的一致率,判断目标数据是否合格。

[0184] 本公开提供的数据处理方法,通过获取至少一条待处理数据;分割每条待处理数据,得到至少一个待使用数据;确定每个待使用数据在预设词典中对应的词典标识,得到词典标识集合,可见,本公开采用对待处理数据进行分割的方式,完整的保留了待处理数据的所有信息,避免了可能出现的信息缺失的问题;确定与词典标识集合对应的上下文感知数据和候选实体数据;融合上下文感知数据和候选实体数据,得到目标数据,可见,本公开基于完整的待处理数据确定与之对应的上下文感知数据和候选实体数据,为意图识别提供了准确的数据基础,解决了现有技术意图识别准确率低的问题。

[0185] 下面对本公开实施例提供的数据处理装置进行描述，下文描述的数据处理装置与上文描述的数据处理方法可相互对应参照，重复之处，不再赘述，如图3所示，该装置包括：

[0186] 获取模块301，用于获取至少一条待处理数据；

[0187] 分割模块302，用于分割每条待处理数据，得到至少一个待使用数据；

[0188] 第一确定模块303，用于确定每个待使用数据在预设词典中对应的词典标识，得到词典标识集合；

[0189] 第二确定模块304，用于确定与词典标识集合对应的上下文感知数据和候选实体数据；

[0190] 融合模块305，用于融合上下文感知数据和候选实体数据，得到目标数据。

[0191] 一个具体实施例中，分割模块302，具体用于基于字粒度分割策略，分割每条待处理数据，得到至少一个第一分割数据；基于词粒度分割策略，分割每条待处理数据，得到至少一个第二分割数据；将第一分割数据和第二分割数据的并集作为待使用数据。

[0192] 一个具体实施例中，第一确定模块303，具体用于将每个第一分割数据在预设词典中进行匹配，得到第一匹配结果；将第一匹配结果在预设词典中对应的第一词典标识作为第一词典标识子集合；将每个第二分割数据在预设词典中进行匹配，得到第二匹配结果；将第二匹配结果在预设词典中对应的第二词典标识作为第二词典标识子集合；组合第一词典标识子集合和第二词典标识子集合，得到词典标识集合。

[0193] 一个具体实施例中，第二确定模块304，具体用于确定每个第一词典标识对应的第一向量，得到第一文本表征矩阵；确定每个第二词典标识对应的第二向量，得到第二文本表征矩阵；将第一文本表征矩阵输入上下文感知模型，得到上下文感知模型输出的第一感知数据，以及将第二文本表征矩阵输入上下文感知模型，得到上下文感知模型输出的第二感知数据；组合第一感知数据和第二感知数据，得到上下文感知数据；其中，上下文感知模型基于文本表征矩阵样本和上下文感知数据样本训练得到。

[0194] 一个具体实施例中，第二确定模块304，具体用于将词典标识集合输入实体候选模型，得到实体候选模型输出的候选实体数据；其中，实体候选模型基于词典标识集合样本和候选实体数据样本训练得到。

[0195] 一个具体实施例中，融合模块305，具体用于基于候选实体数据，从第一感知数据中提取第一目标感知数据，以及从第二感知数据中提取第二目标感知数据；将第一目标感知数据、第二目标感知数据和候选实体数据输入预设计算公式，得到预设计算公式输出的第一表征感知数据和第二表征感知数据；分别对第一表征感知数据和第二表征感知数据进行池化操作，得到第一池化感知数据和第二池化感知数据；拼接第一池化感知数据和第二池化感知数据，得到目标数据。

[0196] 一个具体实施例中，融合模块305，具体用于将第一目标感知数据、第二目标感知数据和候选实体数据输入第一预设计算子公式，得到第一预设计算子公式输出的注意力权重矩阵；将第一目标感知数据、第二目标感知数据、候选实体数据和注意力权重矩阵输入第二预设计算子公式，得到第二预设计算子公式输出的新的第一目标感知数据、新的第二目标感知数据和新的候选实体数据；将新的第一目标感知数据、新的第二目标感知数据和新的候选实体数据输入第三预设计算子公式，得到第三预设计算子公式输出的第一表征感知数据和第二表征感知数据。

[0197] 一个具体实施例中，第一确定模块303，还用于获取待使用数据的数据长度信息；在确定数据长度信息大于预设值的情况下，从待使用数据中提取数据长度信息为预设值的目标待使用数据，将目标待使用数据作为待使用数据；在确定数据长度信息小于或等于预设值的情况下，执行确定每个待使用数据在预设词典中对应的词典标识，得到词典标识集合的步骤。

[0198] 一个具体实施例中，获取模块301，还用于获取与待处理数据对应的真实意图，其中，真实意图通过用户设置得到；融合模块305，还用于预测目标数据对应的预测意图；在真实意图和预测意图的一致率大于预设值的情况下，确定目标数据合格。

[0199] 图4示例了一种电子设备的实体结构示意图，如图4所示，该电子设备可以包括：处理器(processor)401、通信接口(Communications Interface)402、存储器(memory)403和通信总线404，其中，处理器401，通信接口402，存储器403通过通信总线404完成相互间的通信。处理器401可以调用存储器403中的逻辑指令，以执行数据处理方法，该方法包括：获取至少一条待处理数据；分割每条待处理数据，得到至少一个待使用数据；确定每个待使用数据在预设词典中对应的词典标识，得到词典标识集合；确定与词典标识集合对应的上下文感知数据和候选实体数据；融合上下文感知数据和候选实体数据，得到目标数据。

[0200] 此外，上述的存储器403中的逻辑指令可以通过软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本公开实施例的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备(可以是个人计算机，服务器，或者网络设备等)执行本公开各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U盘、移动硬盘、只读存储器(ROM，Read-Only Memory)、随机存取存储器(RAM，Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0201] 另一方面，本公开还提供一种计算机程序产品，所述计算机程序产品包括存储在非暂态计算机可读存储介质上的计算机程序，所述计算机程序包括程序指令，当所述程序指令被计算机执行时，计算机能够执行上述各方法所提供的数据处理方法，该方法包括：获取至少一条待处理数据；分割每条待处理数据，得到至少一个待使用数据；确定每个待使用数据在预设词典中对应的词典标识，得到词典标识集合；确定与词典标识集合对应的上下文感知数据和候选实体数据；融合上下文感知数据和候选实体数据，得到目标数据。

[0202] 又一方面，本公开还提供一种非暂态计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时实现上述各方法所提供的数据处理方法，该方法包括：获取至少一条待处理数据；分割每条待处理数据，得到至少一个待使用数据；确定每个待使用数据在预设词典中对应的词典标识，得到词典标识集合；确定与词典标识集合对应的上下文感知数据和候选实体数据；融合上下文感知数据和候选实体数据，得到目标数据。

[0203] 以上所描述的装置实施例仅仅是示意性的，其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性的劳动的情况下，即可以理解并实施。

[0204] 通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到各实施方式可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件。基于这样的理解，上述技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品可以存储在计算机可读存储介质中，如ROM/RAM、磁碟、光盘等，包括若干指令用以使得一台计算机设备(可以是个人计算机，服务器，或者网络设备等)执行各个实施例或者实施例的某些部分所述的方法。

[0205] 最后应说明的是：以上实施例仅用以说明本公开的技术方案，而非对其限制；尽管参照前述实施例对本公开进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本公开各实施例技术方案的精神和范围。

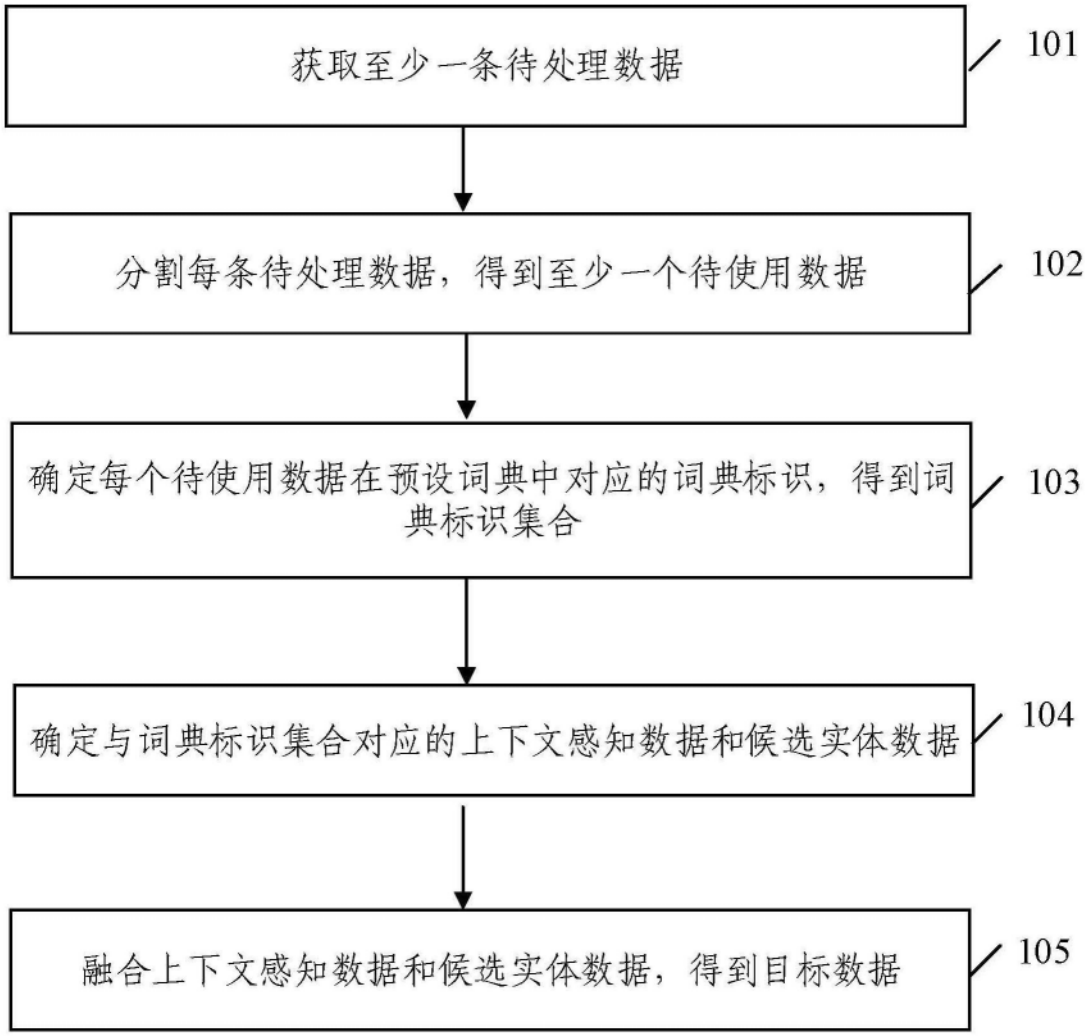

获取至少一条待处理数据
101
分割每条待处理数据，得到至少一个待使用数据
102
确定每个待使用数据在预设词典中对应的词典标识，得到词典标识集合
103
确定与词典标识集合对应的上下文感知数据和候选实体数据
104
融合上下文感知数据和候选实体数据，得到目标数据
105

图1

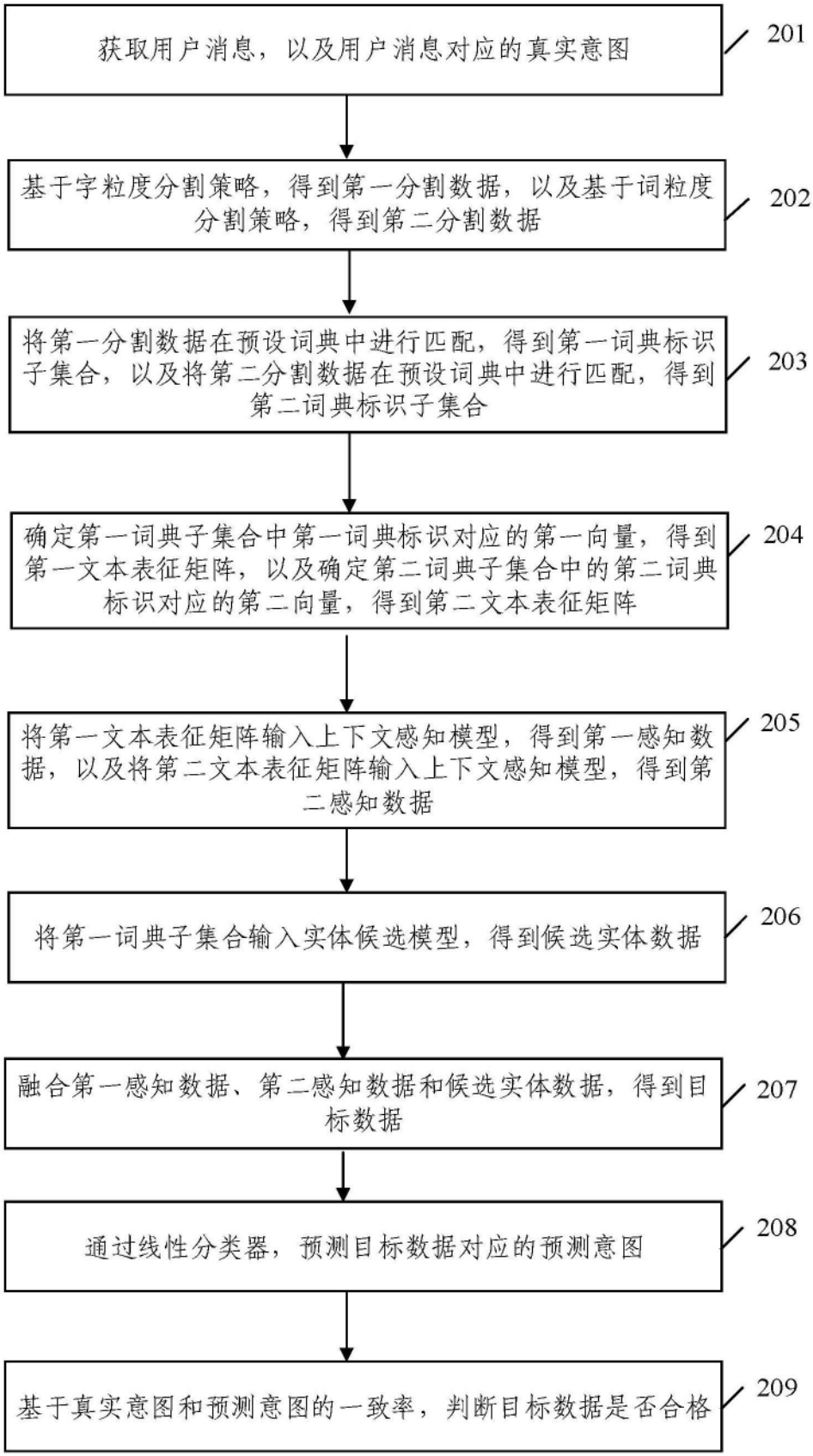

获取用户消息，以及用户消息对应的真实意图
201
基于字粒度分割策略，得到第一分割数据，以及基于词粒度分割策略，得到第二分割数据
202
将第一分割数据在预设词典中进行匹配，得到第一词典标识子集合，以及将第二分割数据在预设词典中进行匹配，得到第二词典标识子集合
203
确定第一词典子集合中第一词典标识对应的第一向量，得到第一文本表征矩阵，以及确定第二词典子集合中的第二词典标识对应的第二向量，得到第二文本表征矩阵
204
将第一文本表征矩阵输入上下文感知模型，得到第一感知数据，以及将第二文本表征矩阵输入上下文感知模型，得到第二感知数据
205
将第一词典子集合输入实体候选模型，得到候选实体数据
206
融合第一感知数据、第二感知数据和候选实体数据，得到目标数据
207
通过线性分类器，预测目标数据对应的预测意图
208
基于真实意图和预测意图的一致率，判断目标数据是否合格
209

图2

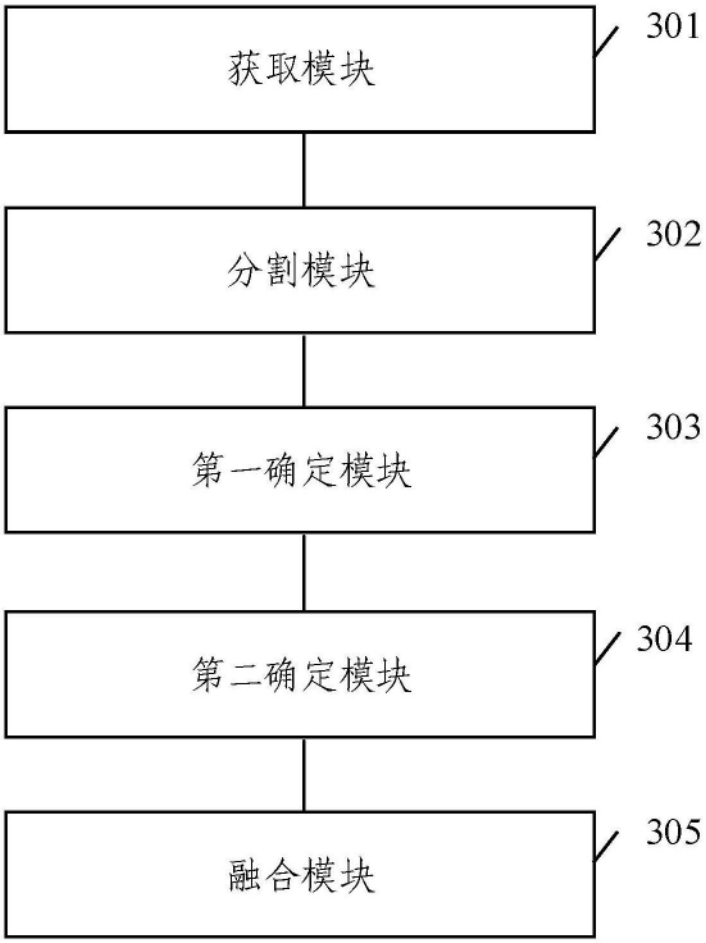

获取模块
301
分割模块
302
第一确定模块
303
第二确定模块
304
融合模块
305

图3

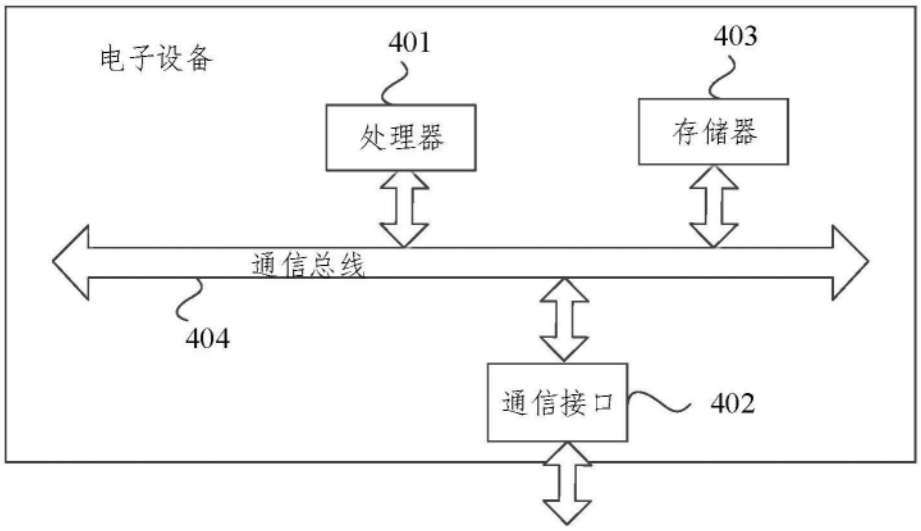

电子设备
401
处理器
403
存储器
通信总线
404
通信接口
402

图4