



(12)发明专利申请

(10)申请公布号 CN 110413788 A

(43)申请公布日 2019.11.05

(21)申请号 201910695856.8

(22)申请日 2019.07.30

(71)申请人 携程计算机技术(上海)有限公司  
地址 200335 上海市长宁区福泉路99号携程网络技术大楼

(72)发明人 江小林 郝竹林 罗超

(74)专利代理机构 上海弼兴律师事务所 31283  
代理人 薛琦 张冉

(51)Int.Cl.  
G06F 16/35(2019.01)  
G06F 17/27(2006.01)

权利要求书5页 说明书14页 附图6页

(54)发明名称

会话文本的场景类别的预测方法、系统、设备和存储介质

(57)摘要

本发明公开了一种会话文本的场景类别的预测方法、系统、设备和存储介质,所述预测方法包括:获取关键词;根据每个关键词获取数据库中的候选句子以及历史会话文本;从历史会话文本中选取目标句子;对目标句子进行分词处理并获取每个目标句子中每个分词的词向量;依次对目标句子中的每个词向量进行编码处理获取句向量;依次对历史会话文本中的每个句向量进行编码处理获取会话向量;建立用于预测会话文本属于每种场景类别的概率模型;根据概率模型获取目标会话文本属于每种场景类别的概率值;选取概率值最大时对应的场景类别作为目标会话文本对应的目标场景类别。本发明提高了现有的会话文本的场景识别的准确率以及召回率,同时降低了人工成本。

CN 110413788 A



1. 一种会话文本的场景类别的预测方法,其特征在于,所述预测方法包括:
  - 获取不同的关键词;
  - 其中,不同的所述关键词用于表征不同的场景类别;
  - 根据每个所述关键词获取数据库中包含所述关键词的候选句子;
  - 获取所述候选句子所在的不同的历史会话文本;
  - 根据所述候选句子从所述历史会话文本中选取设定数量的句子作为目标句子;
  - 对所述目标句子进行分词处理,获取分词结果;
  - 根据所述分词结果获取每个所述目标句子中每个分词对应的词向量;
  - 依次对所述目标句子中的每个所述词向量进行编码处理,获取所述目标句子对应的句向量;
  - 其中,当按照前向顺序进行编码时,所述句向量中的每个所述词向量与前向的所有词向量之间均建立语义关系;
  - 当按照后向顺序进行编码时,所述句向量中每个所述词向量与后向的所有词向量之间均建立语义关系;
  - 依次对所述历史会话文本中的每个所述句向量进行编码处理,获取所述历史会话文本对应的会话向量;
  - 其中,当按照前向顺序进行编码时,所述会话向量中的每个所述句向量与前向的所有句向量之间均建立语义关系;
  - 当按照后向顺序进行编码时,所述会话向量中每个所述句向量与后向的所有句向量之间均建立语义关系;
  - 将所述历史会话文本对应的所述会话向量作为输入,将所述历史会话文本对应的场景类别作为输出,建立用于预测会话文本属于每种场景类别的概率模型;
  - 获取目标会话文本;
  - 将所述目标会话文本输入至所述概率模型,获取所述目标会话文本属于每种场景类别的概率值;
  - 选取概率值最大时对应的场景类别作为所述目标会话文本对应的目标场景类别。
2. 如权利要求1所述的会话文本的场景类别的预测方法,其特征在于,所述对所述目标句子进行分词处理,获取分词结果的步骤包括:
  - 采用HanLP分词工具对所述目标句子进行分词处理,获取分词结果;
  - 所述根据所述分词结果获取每个所述目标句子中每个分词对应的词向量的步骤包括:
    - 采用Word2Vector模型或GloVe模型根据所述分词结果获取每个所述目标句子中每个分词对应的词向量。
3. 如权利要求1所述的会话文本的场景类别的预测方法,其特征在于,所述依次对所述目标句子中的每个所述词向量进行编码处理,获取所述目标句子对应的句向量的步骤包括:
  - 采用LSTM算法依次对所述目标句子中的每个所述词向量进行编码处理,获取每个所述词向量对应的第一中间向量;
  - 其中,排序在编码方向首位的所述词向量对应的所述第一中间向量为采用LSTM算法的隐含层对所述词向量处理后得到的向量,其他位置处的每个所述词向量对应的第一中间向

量包括当前位置处的所述词向量和前一个所述词向量对应的第一中间向量；

将在编码方向末位的所述词向量对应的所述第一中间向量作为所述目标句子对应的所述句向量；

所述依次对所述历史会话文本中的每个所述句向量进行编码处理，获取所述历史会话文本对应的会话向量的步骤包括：

采用LSTM算法依次对所述历史会话文本中的每个所述句向量进行编码处理，获取每个所述句向量对应的第二中间向量；

其中，排序在编码方向首位的所述句向量对应的所述第二中间向量为采用LSTM算法的隐含层对所述句向量处理后得到的向量，其他位置处的每个所述句向量对应的第二中间向量包括当前位置处的所述句向量和前一个所述句向量对应的第二中间向量；

将在编码方向末位的所述句向量对应的所述第二中间向量作为所述历史会话文本对应的所述会话向量。

4. 如权利要求1所述的会话文本的场景类别的预测方法，其特征在于，所述依次对所述目标句子中的每个所述词向量进行编码处理，获取所述目标句子对应的句向量的步骤包括：

采用LSTM算法依次对所述目标句子中的每个所述词向量按照前向顺序和后向顺序分别进行编码处理，获取两个第一句向量；

将两个所述第一句向量进行融合处理获得所述句向量；

所述依次对所述历史会话文本中的每个所述句向量进行编码处理，获取所述历史会话文本对应的会话向量的步骤包括：

采用LSTM算法依次对所述目标句子中的每个所述词向量按照前向顺序和后向顺序分别进行编码处理，获取两个第一会话向量；

将两个所述第一会话向量进行融合处理获得所述会话向量。

5. 如权利要求1所述的会话文本的场景类别的预测方法，其特征在于，所述依次对所述目标句子中的每个所述词向量进行编码处理，获取所述目标句子对应的句向量的步骤之后还包括：

采用Softmax算法获取所述目标句子中的每个所述词向量对应的第一权重；

根据所述第一权重对所述目标句子中的每个所述词向量进行加权处理，得到加权处理后的所述句向量；

所述依次对所述历史会话文本中的每个所述句向量进行编码处理，获取所述历史会话文本对应的会话向量的步骤之后还包括：

采用Softmax算法获取所述历史会话文本中的每个所述句向量对应的第二权重；

根据所述第二权重对所述历史会话文本中的每个所述句向量进行加权处理，得到加权处理后的所述会话向量。

6. 如权利要求1所述的会话文本的场景类别的预测方法，其特征在于，所述将所述历史会话文本对应的所述会话向量作为输入，将所述历史会话文本对应的场景类别作为输出，建立用于预测会话文本属于每种场景类别的概率模型的步骤包括：

采用Softmax算法将所述历史会话文本对应的所述会话向量作为输入，将所述历史会话文本对应的场景类别作为输出，建立用于预测会话文本属于每种场景类别的概率模型；

和/或，

所述根据所述候选句子从所述历史会话文本中选取设定数量的句子作为目标句子的步骤之后、所述对所述目标句子进行分词处理，获取分词结果的步骤之前还包括：

对所述目标句子进行预处理；和/或，

所述根据所述候选句子从所述历史会话文本中选取设定数量的句子作为目标句子的步骤包括：

从所述历史会话文本中选取所述候选句子和所述候选句子的上下文中各N个句子作为所述目标句子；其中， $N \geq 1$ ，N取值为整数。

7. 一种会话文本的场景类别的预测系统，其特征在于，所述预测系统包括关键词获取模块、候选句子获取模块、历史文本获取模块、目标句子获取模块、分词处理模块、词向量获取模块、句向量获取模块、会话向量获取模块、模型建立模块、目标文本获取模块、概率值获取模块和场景类别确定模块；

所述关键词获取模块用于获取不同的关键词；

其中，不同的所述关键词用于表征不同的场景类别；

所述候选句子获取模块用于根据每个所述关键词获取数据库中包含所述关键词的候选句子；

所述历史文本获取模块用于获取所述候选句子所在的不同的历史会话文本；

所述目标句子获取模块用于根据所述候选句子从所述历史会话文本中选取设定数量的句子作为目标句子；

所述分词处理模块用于对所述目标句子进行分词处理，获取分词结果；

所述词向量获取模块用于根据所述分词结果获取每个所述目标句子中每个分词对应的词向量；

所述句向量获取模块用于依次对所述目标句子中的每个所述词向量进行编码处理，获取所述目标句子对应的句向量；

其中，当按照前向顺序进行编码时，所述句向量中的每个所述词向量与前向的所有词向量之间均建立语义关系；

当按照后向顺序进行编码时，所述句向量中每个所述词向量与后向的所有词向量之间均建立语义关系；

所述会话向量获取模块用于依次对所述历史会话文本中的每个所述句向量进行编码处理，获取所述历史会话文本对应的会话向量；

其中，当按照前向顺序进行编码时，所述会话向量中的每个所述句向量与前向的所有句向量之间均建立语义关系；

当按照后向顺序进行编码时，所述会话向量中每个所述句向量与后向的所有句向量之间均建立语义关系；

所述模型建立模块用于将所述历史会话文本对应的所述会话向量作为输入，将所述历史会话文本对应的场景类别作为输出，建立用于预测会话文本属于每种场景类别的概率模型；

所述目标文本获取模块用于获取目标会话文本；

所述概率值获取模块用于将所述目标会话文本输入至所述概率模型，获取所述目标会

话文本属于每种场景类别的概率值；

所述场景类别确定模块用于选取概率值最大时对应的场景类别作为所述目标会话文本对应的目标场景类别。

8. 如权利要求7所述的会话文本的场景类别的预测系统,其特征在于,所述分词处理模块用于采用HanLP分词工具对所述目标句子进行分词处理,获取分词结果;

所述词向量获取模块用于采用Word2Vector模型或GloVe模型根据所述分词结果获取每个所述目标句子中每个分词对应的词向量。

9. 如权利要求7所述的会话文本的场景类别的预测系统,其特征在于,所述句向量获取模块包括第一中间向量获取单元和句向量获取单元;

所述第一中间向量获取单元用于采用LSTM算法依次对所述目标句子中的每个所述词向量进行编码处理,获取每个所述词向量对应的第一中间向量;

其中,排序在编码方向首位的所述词向量对应的所述第一中间向量为采用LSTM算法的隐含层对所述词向量处理后得到的向量,其他位置处的每个所述词向量对应的第一中间向量包括当前位置处的所述词向量和前一个所述词向量对应的第一中间向量;

所述句向量获取单元用于将在编码方向末位的所述词向量对应的所述第一中间向量作为所述目标句子对应的所述句向量;

所述会话向量获取模块包括第二中间向量获取单元和会话向量获取单元;

所述第二中间向量获取单元用于采用LSTM算法依次对所述历史会话文本中的每个所述句向量进行编码处理,获取每个所述句向量对应的第二中间向量;

其中,排序在编码方向首位的所述句向量对应的所述第二中间向量为采用LSTM算法的隐含层对所述句向量处理后得到的向量,其他位置处的每个所述句向量对应的第二中间向量包括当前位置处的所述句向量和前一个所述句向量对应的第二中间向量;

所述会话向量获取单元用于将在编码方向末位的所述句向量对应的所述第二中间向量作为所述历史会话文本对应的所述会话向量。

10. 如权利要求7所述的会话文本的场景类别的预测系统,其特征在于,所述句向量获取模块用于采用LSTM算法依次对所述目标句子中的每个所述词向量按照前向顺序和后向顺序分别进行编码处理,获取两个第一句向量,并将两个所述第一句向量进行融合处理获得所述句向量;

所述会话向量获取模块用于采用LSTM算法依次对所述目标句子中的每个所述词向量按照前向顺序和后向顺序分别进行编码处理,获取两个第一会话向量,并将两个所述第一会话向量进行融合处理获得所述会话向量。

11. 如权利要求7所述的会话文本的场景类别的预测系统,其特征在于,所述预测系统还包括权重获取模块和加权处理模块;

所述权重获取模块用于采用Softmax算法获取所述目标句子中的每个所述词向量对应的第一权重;

所述加权处理模块用于根据所述第一权重对所述目标句子中的每个所述词向量进行加权处理,得到加权处理后的所述句向量;

所述权重获取模块还用于采用Softmax算法获取所述历史会话文本中的每个所述句向量对应的第二权重;

所述加权处理模块还用于根据所述第二权重对所述历史会话文本中的每个所述句向量进行加权处理,得到加权处理后的所述会话向量。

12. 如权利要求7所述的会话文本的场景类别的预测系统,其特征在于,所述模型建立模块用于采用Softmax算法将所述历史会话文本对应的所述会话向量作为输入,将所述历史会话文本对应的场景类别作为输出,建立用于预测会话文本属于每种场景类别的概率模型;和/或,

所述预测系统还包括预处理模块;

所述预处理模块用于对所述目标句子进行预处理;和/或,

所述目标句子获取模块用于从所述历史会话文本中选取所述候选句子和所述候选句子的上下文中各N个句子作为所述目标句子;其中, $N \geq 1$ ,N取值为整数。

13. 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,所述处理器执行计算机程序时实现权利要求1-6中任一项所述的会话文本的场景类别的预测方法。

14. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求1-6中任一项所述的会话文本的场景类别的预测方法的步骤。

## 会话文本的场景类别的预测方法、系统、设备和存储介质

### 技术领域

[0001] 本发明涉及数据处理技术领域,特别涉及一种会话文本的场景类别的预测方法、系统、设备和存储介质。

### 背景技术

[0002] 对于服务型企业,用户的咨询或者反馈对企业至关重要,很多服务型企业通过设置客服系统来存储用户与客服人员的不同会话文本数据,进而分析出不同的对话型文本所属场景类别,以便于对后续的数据处理过程。

[0003] 目前,主要通过设定关键词并结合人工判定,但是该方式容易受到噪音干扰,因此准确率比较低,同时,对于海量的对话交互数据,人工审查的成本太大;另外,还可以通过CNN(卷积神经网络)算法和RNN(循环神经网络)算法的深度学习方法对对话型文本进行分类;但是,对于CNN算法,在进行对话型文本所属的场景识别时不是单纯的文本分类问题,需要上下文关联才能实现对对话的准确判别,所以在构建模型时需要考虑长文本对话的时序关系进行学习;对于RNN算法,目前大多都是使用单一网络或网络的简单变体,对于采用复杂网络结构进行对话型场景识别不适用。

### 发明内容

[0004] 本发明要解决的技术问题是为了克服现有技术中会话文本所属场景类别的确定存在准确率低等缺陷,提供一种会话文本的场景类别的预测方法、系统、设备和存储介质。

[0005] 本发明是通过下述技术方案来解决上述技术问题:

[0006] 本发明提供一种会话文本的场景类别的预测方法,所述预测方法包括:

[0007] 获取不同的关键词;

[0008] 其中,不同的所述关键词用于表征不同的场景类别;

[0009] 根据每个所述关键词获取数据库中包含所述关键词的候选句子;

[0010] 获取所述候选句子所在的不同的历史会话文本;

[0011] 根据所述候选句子从所述历史会话文本中选取设定数量的句子作为目标句子;

[0012] 对所述目标句子进行分词处理,获取分词结果;

[0013] 根据所述分词结果获取每个所述目标句子中每个分词对应的词向量;

[0014] 依次对所述目标句子中的每个所述词向量进行编码处理,获取所述目标句子对应的句向量;

[0015] 其中,当按照前向顺序进行编码时,所述句向量中的每个所述词向量与前向的所有词向量之间均建立语义关系;

[0016] 当按照后向顺序进行编码时,所述句向量中每个所述词向量与后向的所有词向量之间均建立语义关系;

[0017] 依次对所述历史会话文本中的每个所述句向量进行编码处理,获取所述历史会话文本对应的会话向量;

[0018] 其中,当按照前向顺序进行编码时,所述会话向量中的每个所述句向量与前向的所有句向量之间均建立语义关系;

[0019] 当按照后向顺序进行编码时,所述会话向量中每个所述句向量与后向的所有句向量之间均建立语义关系;

[0020] 将所述历史会话文本对应的所述会话向量作为输入,将所述历史会话文本对应的场景类别作为输出,建立用于预测会话文本属于每种场景类别的概率模型;

[0021] 获取目标会话文本;

[0022] 将所述目标会话文本输入至所述概率模型,获取所述目标会话文本属于每种场景类别的概率值;

[0023] 选取概率值最大时对应的场景类别作为所述目标会话文本对应的目标场景类别。

[0024] 较佳地,所述对所述目标句子进行分词处理,获取分词结果的步骤包括:

[0025] 采用HanLP分词工具(一种分词工具)对所述目标句子进行分词处理,获取分词结果;

[0026] 所述根据所述分词结果获取每个所述目标句子中每个分词对应的词向量的步骤包括:

[0027] 采用Word2Vector模型或GloVe模型(Word2Vector模型、GloVe模型均为一种词向量转换模型)根据所述分词结果获取每个所述目标句子中每个分词对应的词向量。

[0028] 较佳地,所述依次对所述目标句子中的每个所述词向量进行编码处理,获取所述目标句子对应的句向量的步骤包括:

[0029] 采用LSTM算法(长短时记忆网络算法)依次对所述目标句子中的每个所述词向量进行编码处理,获取每个所述词向量对应的第一中间向量;

[0030] 其中,排序在编码方向首位的所述词向量对应的所述第一中间向量为采用LSTM算法的隐含层对所述词向量处理后得到的向量,其他位置处的每个所述词向量对应的第一中间向量包括当前位置处的所述词向量和前一个所述词向量对应的第一中间向量;

[0031] 将在编码方向末位的所述词向量对应的所述第一中间向量作为所述目标句子对应的所述句向量;

[0032] 所述依次对所述历史会话文本中的每个所述句向量进行编码处理,获取所述历史会话文本对应的会话向量的步骤包括:

[0033] 采用LSTM算法依次对所述历史会话文本中的每个所述句向量进行编码处理,获取每个所述句向量对应的第二中间向量;

[0034] 其中,排序在编码方向首位的所述句向量对应的所述第二中间向量为采用LSTM算法的隐含层对所述句向量处理后得到的向量,其他位置处的每个所述句向量对应的第二中间向量包括当前位置处的所述句向量和前一个所述句向量对应的第二中间向量;

[0035] 将在编码方向末位的所述句向量对应的所述第二中间向量作为所述历史会话文本对应的所述会话向量。

[0036] 较佳地,所述依次对所述目标句子中的每个所述词向量进行编码处理,获取所述目标句子对应的句向量的步骤包括:

[0037] 采用LSTM算法依次对所述目标句子中的每个所述词向量按照前向顺序和后向顺序分别进行编码处理,获取两个第一句向量;

- [0038] 将两个所述第一句向量进行融合处理获得所述句向量；
- [0039] 所述依次对所述历史会话文本中的每个所述句向量进行编码处理,获取所述历史会话文本对应的会话向量的步骤包括:
- [0040] 采用LSTM算法依次对所述目标句子中的每个所述词向量按照前向顺序和后向顺序分别进行编码处理,获取两个第一会话向量;
- [0041] 将两个所述第一会话向量进行融合处理获得所述会话向量。
- [0042] 较佳地,所述依次对所述目标句子中的每个所述词向量进行编码处理,获取所述目标句子对应的句向量的步骤之后还包括:
- [0043] 采用Softmax算法(一种回归算法)获取所述目标句子中的每个所述词向量对应的第一权重;
- [0044] 根据所述第一权重对所述目标句子中的每个所述词向量进行加权处理,得到加权处理后的所述句向量;
- [0045] 所述依次对所述历史会话文本中的每个所述句向量进行编码处理,获取所述历史会话文本对应的会话向量的步骤之后还包括:
- [0046] 采用Softmax算法获取所述历史会话文本中的每个所述句向量对应的第二权重;
- [0047] 根据所述第二权重对所述历史会话文本中的每个所述句向量进行加权处理,得到加权处理后的所述会话向量。
- [0048] 较佳地,所述将所述历史会话文本对应的所述会话向量作为输入,将所述历史会话文本对应的场景类别作为输出,建立用于预测会话文本属于每种场景类别的概率模型的步骤包括:
- [0049] 采用Softmax算法将所述历史会话文本对应的所述会话向量作为输入,将所述历史会话文本对应的场景类别作为输出,建立用于预测会话文本属于每种场景类别的概率模型;和/或,
- [0050] 所述根据所述候选句子从所述历史会话文本中选取设定数量的句子作为目标句子的步骤之后、所述对所述目标句子进行分词处理,获取分词结果的步骤之前还包括:
- [0051] 对所述目标句子进行预处理;和/或,
- [0052] 所述根据所述候选句子从所述历史会话文本中选取设定数量的句子作为目标句子的步骤包括:
- [0053] 从所述历史会话文本中选取所述候选句子和所述候选句子的上下文中各N个句子作为所述目标句子;其中, $N \geq 1$ ,N取值为整数。
- [0054] 本发明还提供一种会话文本的场景类别的预测系统,所述预测系统包括关键词获取模块、候选句子获取模块、历史文本获取模块、目标句子获取模块、分词处理模块、词向量获取模块、句向量获取模块、会话向量获取模块、模型建立模块、目标文本获取模块、概率值获取模块和场景类别确定模块;
- [0055] 所述关键词获取模块用于获取不同的关键词;
- [0056] 其中,不同的所述关键词用于表征不同的场景类别;
- [0057] 所述候选句子获取模块用于根据每个所述关键词获取数据库中包含所述关键词的候选句子;
- [0058] 所述历史文本获取模块用于获取所述候选句子所在的不同的历史会话文本;

- [0059] 所述目标句子获取模块用于根据所述候选句子从所述历史会话文本中选取设定数量的句子作为目标句子；
- [0060] 所述分词处理模块用于对所述目标句子进行分词处理,获取分词结果；
- [0061] 所述词向量获取模块用于根据所述分词结果获取每个所述目标句子中每个分词对应的词向量；
- [0062] 所述句向量获取模块用于依次对所述目标句子中的每个所述词向量进行编码处理,获取所述目标句子对应的句向量；
- [0063] 其中,当按照前向顺序进行编码时,所述句向量中的每个所述词向量与前向的所有词向量之间均建立语义关系；
- [0064] 当按照后向顺序进行编码时,所述句向量中每个所述词向量与后向的所有词向量之间均建立语义关系；
- [0065] 所述会话向量获取模块用于依次对所述历史会话文本中的每个所述句向量进行编码处理,获取所述历史会话文本对应的会话向量；
- [0066] 其中,当按照前向顺序进行编码时,所述会话向量中的每个所述句向量与前向的所有句向量之间均建立语义关系；
- [0067] 当按照后向顺序进行编码时,所述会话向量中每个所述句向量与后向的所有句向量之间均建立语义关系；
- [0068] 所述模型建立模块用于将所述历史会话文本对应的所述会话向量作为输入,将所述历史会话文本对应的场景类别作为输出,建立用于预测会话文本属于每种场景类别的概率模型；
- [0069] 所述目标文本获取模块用于获取目标会话文本；
- [0070] 所述概率值获取模块用于将所述目标会话文本输入至所述概率模型,获取所述目标会话文本属于每种场景类别的概率值；
- [0071] 所述场景类别确定模块用于选取概率值最大时对应的场景类别作为所述目标会话文本对应的目标场景类别。
- [0072] 较佳地,所述分词处理模块用于采用HanLP分词工具对所述目标句子进行分词处理,获取分词结果；
- [0073] 所述词向量获取模块用于采用Word2Vector模型或GloVe模型根据所述分词结果获取每个所述目标句子中每个分词对应的词向量。
- [0074] 较佳地,所述句向量获取模块包括第一中间向量获取单元和句向量获取单元；
- [0075] 所述第一中间向量获取单元用于采用LSTM算法依次对所述目标句子中的每个所述词向量进行编码处理,获取每个所述词向量对应的第一中间向量；
- [0076] 其中,排序在编码方向首位的所述词向量对应的所述第一中间向量为采用LSTM算法的隐含层对所述词向量处理后得到的向量,其他位置处的每个所述词向量对应的第一中间向量包括当前位置处的所述词向量和前一个所述词向量对应的第一中间向量；
- [0077] 所述句向量获取单元用于将在编码方向末位的所述词向量对应的所述第一中间向量作为所述目标句子对应的所述句向量；
- [0078] 所述会话向量获取模块包括第二中间向量获取单元和会话向量获取单元；
- [0079] 所述第二中间向量获取单元用于采用LSTM算法依次对所述历史会话文本中的每

个所述句向量进行编码处理,获取每个所述句向量对应的第二中间向量;

[0080] 其中,排序在编码方向首位的所述句向量对应的所述第二中间向量为采用LSTM算法的隐含层对所述句向量处理后得到的向量,其他位置处的每个所述句向量对应的第二中间向量包括当前位置处的所述句向量和前一个所述句向量对应的第二中间向量;

[0081] 所述会话向量获取单元用于将在编码方向末位的所述句向量对应的所述第二中间向量作为所述历史会话文本对应的所述会话向量。

[0082] 较佳地,所述句向量获取模块用于采用LSTM算法依次对所述目标句子中的每个所述词向量按照前向顺序和后向顺序分别进行编码处理,获取两个第一句向量,并将两个所述第一句向量进行融合处理获得所述句向量;

[0083] 所述会话向量获取模块用于采用LSTM算法依次对所述目标句子中的每个所述词向量按照前向顺序和后向顺序分别进行编码处理,获取两个第一会话向量,并将两个所述第一会话向量进行融合处理获得所述会话向量。

[0084] 较佳地,所述预测系统还包括权重获取模块和加权处理模块;

[0085] 所述权重获取模块用于采用Softmax算法获取所述目标句子中的每个所述词向量对应的第一权重;

[0086] 所述加权处理模块用于根据所述第一权重对所述目标句子中的每个所述词向量进行加权处理,得到加权处理后的所述句向量;

[0087] 所述权重获取模块还用于采用Softmax算法获取所述历史会话文本中的每个所述句向量对应的第二权重;

[0088] 所述加权处理模块还用于根据所述第二权重对所述历史会话文本中的每个所述句向量进行加权处理,得到加权处理后的所述会话向量。

[0089] 较佳地,所述模型建立模块用于采用Softmax算法将所述历史会话文本对应的所述会话向量作为输入,将所述历史会话文本对应的场景类别作为输出,建立用于预测会话文本属于每种场景类别的概率模型;和/或,

[0090] 所述预测系统还包括预处理模块;

[0091] 所述预处理模块用于对所述目标句子进行预处理;和/或,

[0092] 所述目标句子获取模块用于从所述历史会话文本中选取所述候选句子和所述候选句子的上下文中各N个句子作为所述目标句子;其中, $N \geq 1$ ,N取值为整数。

[0093] 本发明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,所述处理器执行计算机程序时实现上述的会话文本的场景类别的预测方法。

[0094] 本发明还提供一种计算机可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现上述的会话文本的场景类别的预测方法的步骤。

[0095] 本发明的积极进步效果在于:

[0096] 本发明中,通过关键词获取数据库中的候选句子,进而得到对应的会话文本以及上下文句子,然后对从会话文本中挑选出的句子进行分词处理并转换为对应的词向量,再利用LSTM算法获取包含所有单词之间的语义关系的句向量,以及包含所有句子之间的语义关系的会话向量,进而建立用于预测会话文本属于每种场景类别的概率模型,来确定任意一会话文本对应的目标场景类别,从而提高了现有的会话文本的场景识别的准确率以及召

回率,同时降低了人工成本。

### 附图说明

[0097] 图1为本发明实施例1的会话文本的场景类别的预测方法的流程图。

[0098] 图2为本发明实施例1的会话文本的场景类别的预测方法中对会话文本处理的流程示意图。

[0099] 图3为本发明实施例2的会话文本的场景类别的预测方法的流程图。

[0100] 图4为本发明实施例3的会话文本的场景类别的预测系统的模块示意图。

[0101] 图5为本发明实施例4的会话文本的场景类别的预测系统的模块示意图。

[0102] 图6为本发明实施例5中的实现会话文本的场景类别的预测方法的电子设备的结构示意图。

### 具体实施方式

[0103] 下面通过实施例的方式进一步说明本发明,但并不因此将本发明限制在所述的实施例范围之中。

[0104] 实施例1

[0105] 如图1所示,本实施例的会话文本的场景类别的预测方法包括:

[0106] S101、获取不同的关键词;

[0107] 其中,不同的关键词用于表征不同的场景类别,例如,关键词包括酒店噪声、床单等。

[0108] S102、根据每个关键词获取数据库中包含关键词的候选句子;

[0109] 其中,客服系统的数据库中存储有客服人员与用户之间的所有会话记录。

[0110] S103、获取候选句子所在的不同的历史会话文本;

[0111] S104、根据候选句子从历史会话文本中选取设定数量的句子作为目标句子;

[0112] 具体地,从历史会话文本中选取候选句子和候选句子的上下文中各N个句子作为目标句子;其中, $N \geq 1$ ,N取值为整数。

[0113] S105、对目标句子进行分词处理,获取分词结果;

[0114] S106、根据分词结果获取每个目标句子中每个分词对应的词向量;

[0115] S107、依次对目标句子中的每个词向量进行编码处理,获取目标句子对应的句向量;

[0116] 其中,当按照前向顺序进行编码时,句向量中的每个词向量与前向的所有词向量之间均建立语义关系;

[0117] 当按照后向顺序进行编码时,句向量中每个词向量与后向的所有词向量之间均建立语义关系;

[0118] S108、依次对历史会话文本中的每个句向量进行编码处理,获取历史会话文本对应的会话向量;

[0119] 其中,当按照前向顺序进行编码时,会话向量中的每个句向量与前向的所有句向量之间均建立语义关系;

[0120] 当按照后向顺序进行编码时,会话向量中每个句向量与后向的所有句向量之间均

建立语义关系；

[0121] 如图2所示,以选取会话文本中包括两个目标句子为例,依次获取目标句子中每个分词的词向量,每个目标句子对应的句向量,进而得到会话文本对应的会话向量。

[0122] S109、将历史会话文本对应的会话向量作为输入,将历史会话文本对应的场景类别作为输出,建立用于预测会话文本属于每种场景类别的概率模型；

[0123] S1010、获取目标会话文本；

[0124] S1011、将目标会话文本输入至概率模型,获取目标会话文本属于每种场景类别的概率值；

[0125] S1012、选取概率值最大时对应的场景类别作为目标会话文本对应的目标场景类别。

[0126] 本实施例中,通过关键词获取数据库中的候选句子,进而得到对应的会话文本以及上下文句子,然后对从会话文本中挑选出的句子进行分词处理并转换为对应的词向量,再利用LSTM算法获取包含所有单词之间的语义关系的句向量,以及包含所有句子之间的语义关系的会话向量,进而建立用于预测会话文本属于每种场景类别的概率模型,来确定任意一会话文本对应的目标场景类别,从而提高了现有的会话文本的场景识别的准确率以及召回率,同时降低了人工成本。

[0127] 实施例2

[0128] 如图3所示,本实施例的会话文本的场景类别的预测方法是对实施例1的进一步改进,具体地:

[0129] 步骤S104之后、步骤S105之前还包括:

[0130] 对目标句子进行预处理。

[0131] 具体地,预处理的过程主要过滤掉目标句子中一些分类无效的内容,如过滤身份为员工及系统产生的语句;通过正则表达式过滤常见系统语句:如“用户离开”、“用户回来”等;替换特殊字符:如“[表情]”等;过滤纯数字;过滤不包含中文字符的语句;过滤总结部分闲聊语句,通过编辑距离计算相似度进行过滤,比如“你好”;另外,还包括对目标句子的标准化处理,比如将全角转换为半角、繁体转换为简体、大小写之间的相互转化等。

[0132] 步骤S105包括:

[0133] S1051、采用HanLP分词工具对目标句子进行分词处理,获取分词结果。

[0134] 另外,在分词处理过程中,还可以加入与场景对应的一些专业词汇以提高分词准确度。例如:在OTA(在线旅游)行业的酒店场景下,在分词处理时加入该场景对应的预授权、信用住、扣押金、返现券、大床房、到账、住二送一、住三送一、住四送一、住五送一、住六送一、住七送一、住八送一、住九送一、住十送一、满二送一、满三送一、满四送一、满五送一、满六送一、满七送一、满八送一、满九送一、满十送一、到店无房、无房涨价、坐地起价、公寓房、接送机等专业词汇。

[0135] 步骤S106包括:

[0136] S1061、采用Word2Vector模型或GloVe模型根据分词结果获取每个目标句子中每个分词对应的词向量;另外,还可以采用其他能够将每个目标句子中每个分词对应的词向量的模型。

[0137] 步骤S107包括:

[0138] 采用LSTM算法依次对目标句子中的每个词向量进行编码处理,获取每个词向量对应的第一中间向量;

[0139] 其中,排序在编码方向首位的词向量对应的第一中间向量为采用LSTM算法的隐含层对词向量处理后得到的向量,其他位置处的每个词向量对应的第一中间向量包括当前位置处的词向量和前一个词向量对应的第一中间向量;

[0140] 将在编码方向末位的词向量对应的第一中间向量作为目标句子对应的句向量,这样使得每个目标句子对应的句向量能够表征每个词向量之间的语义关系,能够保证更加准确地识别出目标句子的正确含义。

[0141] 步骤S108包括:

[0142] 采用LSTM算法依次对历史会话文本中的每个句向量进行编码处理,获取每个句向量对应的第二中间向量;

[0143] 其中,排序在编码方向首位的句向量对应的第二中间向量为采用LSTM算法的隐含层对句向量处理后得到的向量,其他位置处的每个句向量对应的第二中间向量包括当前位置处的句向量和前一个句向量对应的第二中间向量;

[0144] 将在编码方向末位的句向量对应的第二中间向量作为历史会话文本对应的会话向量。

[0145] 另外,步骤S107还包括:

[0146] 采用LSTM算法依次对目标句子中的每个词向量按照前向顺序和后向顺序分别进行编码处理,获取两个第一句向量;

[0147] 将两个第一句向量进行融合处理获得句向量;

[0148] 通过将前向顺序和后向顺序分别进行编码处理得到的两个向量进行融合,进一步提高句向量表征目标句子的正确性。

[0149] 步骤S108还包括:

[0150] 采用LSTM算法依次对目标句子中的每个词向量按照前向顺序和后向顺序分别进行编码处理,获取两个第一会话向量;

[0151] 将两个第一会话向量进行融合处理获得会话向量。

[0152] 通过将前向顺序和后向顺序分别进行编码处理得到的两个向量进行融合,进一步提高会话向量表征历史会话文本的正确性。

[0153] 例如,历史会话文本中包括的目标句子为“今天天气怎么样”,则对其进行分词处理后得到“今天”、“天气”、“怎么样”,然后由这三个分词获取对应的词向量,得到该目标句子对应的向量序列。

[0154] 采用LSTM算法依次对三个词向量按照前向顺序进行编码处理:采用LSTM算法的隐含层对“今天”对应的词向量进行处理得到对应的第一中间向量,进而得到“天气”对应的第一中间向量包括“今天”对应的第一中间向量和“天气”对应的词向量,“怎么样”对应的第一中间向量包括“天气”对应的第一中间向量和“怎么样”对应的词向量,此时将“怎么样”对应的第一中间向量作为该目标句子对应的一个第一句向量,其中第一中间向量用于表征首位词向量到当前词向量的语义。

[0155] 采用LSTM算法依次对三个词向量按照后向顺序进行编码处理:采用LSTM算法的隐含层对“怎么样”对应的词向量进行处理得到对应的第一中间向量,进而得到“天气”对应的

第一中间向量包括“怎么样”对应的第一中间向量和“天气”对应的词向量，“今天”对应的第一中间向量包括“天气”对应的第一中间向量和“今天”对应的词向量，此时将“今天”对应的第一中间向量作为该目标句子对应的一个第一句向量。

[0156] 然后将上述的两个第一句向量进行融合，这样就可以得到一个能够表征目标句子的准确度更高的句向量。具体地，两个向量融合的过程包括：如一个向量是[101]，另一个向量是[001]，融合后的结果可以为[101001]，这样通过获取的丰富信息来更准确地表征目标句子对应的句向量；或者，通过比较相同位置处的词向量，然后取两个向量的交集作为最终的融合结果。

[0157] 对于得到同一历史会话文本中的多个目标句子对应的会话向量的过程与上述过程类似，因此此处就不再赘述。

[0158] 步骤S107之后、步骤S108之前还包括：

[0159] 采用Softmax算法获取目标句子中的每个词向量对应的第一权重；

[0160] 根据第一权重对目标句子中的每个词向量进行加权处理，得到加权处理后的句向量，从而进一步提高句向量表征目标句子的准确度。

[0161] 步骤S108之后、步骤S109之前还包括：

[0162] 采用Softmax算法获取历史会话文本中的每个句向量对应的第二权重；

[0163] 根据第二权重对历史会话文本中的每个句向量进行加权处理，得到加权处理后的会话向量，从而进一步提高会话向量表征历史会话文本的准确度。

[0164] 步骤S109包括：

[0165] 采用Softmax算法将历史会话文本对应的会话向量作为输入，将历史会话文本对应的场景类别作为输出，建立用于预测会话文本属于每种场景类别的概率模型。

[0166] 本实施例中，通过关键词获取数据库中的候选句子，进而得到对应的会话文本以及上下文句子，然后对从会话文本中挑选出的句子进行分词处理并转换为对应的词向量，再利用LSTM算法获取包含所有单词之间的语义关系的句向量，以及包含所有句子之间的语义关系的会话向量，进而建立用于预测会话文本属于每种场景类别的概率模型，来确定任意一会话文本对应的目标场景类别，从而提高了现有的会话文本的场景识别的准确率以及召回率，同时降低了人工成本。

[0167] 实施例3

[0168] 如图4所示，本实施例的会话文本的场景类别的预测系统包括关键词获取模块1、候选句子获取模块2、历史文本获取模块3、目标句子获取模块4、分词处理模块5、词向量获取模块6、句向量获取模块7、会话向量获取模块8、模型建立模块9、目标文本获取模块10、概率值获取模块11和场景类别确定模块12。

[0169] 关键词获取模块1用于获取不同的关键词；

[0170] 其中，不同的关键词用于表征不同的场景类别，例如，关键词包括酒店噪声、床单等。；

[0171] 候选句子获取模块2用于根据每个关键词获取数据库中包含关键词的候选句子；

[0172] 其中，客服系统的数据库中存储有客服人员与用户之间的所有会话记录。

[0173] 历史文本获取模块3用于获取候选句子所在的不同的历史会话文本；

[0174] 目标句子获取模块4用于根据候选句子从历史会话文本中选取设定数量的句子作

为目标句子；

[0175] 具体地,从历史会话文本中选取候选句子和候选句子的上下文中各N个句子作为目标句子;其中, $N \geq 1$ ,N取值为整数。

[0176] 分词处理模块5用于对目标句子进行分词处理,获取分词结果;

[0177] 词向量获取模块6用于根据分词结果获取每个目标句子中每个分词对应的词向量;

[0178] 句向量获取模块7用于依次对目标句子中的每个词向量进行编码处理,获取目标句子对应的句向量;

[0179] 其中,当按照前向顺序进行编码时,句向量中的每个词向量与前向的所有词向量之间均建立语义关系;

[0180] 当按照后向顺序进行编码时,句向量中每个词向量与后向的所有词向量之间均建立语义关系;

[0181] 会话向量获取模块8用于依次对历史会话文本中的每个句向量进行编码处理,获取历史会话文本对应的会话向量;

[0182] 其中,当按照前向顺序进行编码时,会话向量中的每个句向量与前向的所有句向量之间均建立语义关系;

[0183] 当按照后向顺序进行编码时,会话向量中每个句向量与后向的所有句向量之间均建立语义关系;

[0184] 如图3所示,以选取会话文本中包括两个目标句子为例,依次获取目标句子中每个分词的词向量,每个目标句子对应的句向量,进而得到会话文本对应的会话向量。

[0185] 模型建立模块9用于将历史会话文本对应的会话向量作为输入,将历史会话文本对应的场景类别作为输出,建立用于预测会话文本属于每种场景类别的概率模型;

[0186] 目标文本获取模块10用于获取目标会话文本;

[0187] 概率值获取模块11用于将目标会话文本输入至概率模型,获取目标会话文本属于每种场景类别的概率值;

[0188] 场景类别确定模块12用于选取概率值最大时对应的场景类别作为目标会话。

[0189] 本实施例中,通过关键词获取数据库中的候选句子,进而得到对应的会话文本以及上下文句子,然后对从会话文本中挑选出的句子进行分词处理并转换为对应的词向量,再利用LSTM算法获取包含所有单词之间的语义关系的句向量,以及包含所有句子之间的语义关系的会话向量,进而建立用于预测会话文本属于每种场景类别的概率模型,来确定任意一会话文本对应的目标场景类别,从而提高了现有的会话文本的场景识别的准确率以及召回率,同时降低了人工成本。

[0190] 实施例4

[0191] 如图5所示,本实施例的会话文本的场景类别的预测方法是对实施例3的进一步改进,具体地:

[0192] 预测系统还包括预处理模块13;

[0193] 预处理模块13用于对目标句子进行预处理。

[0194] 具体地,预处理的过程主要过滤掉目标句子中一些分类无效的内容,如过滤身份为员工及系统产生的语句;通过正则表达式过滤常见系统语句:如“用户离开”、“用户回来”

等;替换特殊字符:如“[表情]”等;过滤纯数字;过滤不包含中文字符的语句;过滤总结部分闲聊语句,通过编辑距离计算相似度进行过滤,比如“你好”;另外,还包括对目标句子的标准化处理,比如将全角转换为半角、繁体转换为简体、大小写之间的相互转化等。

[0195] 分词处理模块5用于采用HanLP分词工具对目标句子进行分词处理,获取分词结果。

[0196] 另外,在分词处理过程中,还可以加入与场景对应的一些专业词汇以提高分词准确度。例如:在OTA行业的酒店场景下,在分词处理时加入该场景对应的预授权、信用住、扣押金、返现券、大床房、到账、住二送一、住三送一、住四送一、住五送一、住六送一、住七送一、住八送一、住九送一、住十送一、满二送一、满三送一、满四送一、满五送一、满六送一、满七送一、满八送一、满九送一、满十送一、到店无房、无房涨价、坐地起价、公寓房、接送机等专业词汇。

[0197] 词向量获取模块6用于采用Word2Vector模型或GloVe模型根据分词结果获取每个目标句子中每个分词对应的词向量;另外,还可以采用其他能够将每个目标句子中每个分词对应的词向量的模型。

[0198] 具体地,句向量获取模块7包括第一中间向量获取单元和句向量获取单元;

[0199] 第一中间向量获取单元用于采用LSTM算法依次对目标句子中的每个词向量进行编码处理,获取每个词向量对应的第一中间向量;

[0200] 其中,排序在编码方向首位的词向量对应的第一中间向量为采用LSTM算法的隐含层对词向量处理后得到的向量,其他位置处的每个词向量对应的第一中间向量包括当前位置处的词向量和前一个词向量对应的第一中间向量;

[0201] 句向量获取单元用于将在编码方向末位的词向量对应的第一中间向量作为目标句子对应的句向量,这样使得每个目标句子对应的句向量能够表征每个词向量之间的语义关系,能够保证更加准确地识别出目标句子的正确含义。

[0202] 会话向量获取模块8包括第二中间向量获取单元和会话向量获取单元;

[0203] 第二中间向量获取单元用于采用LSTM算法依次对历史会话文本中的每个句向量进行编码处理,获取每个句向量对应的第二中间向量;

[0204] 其中,排序在编码方向首位的句向量对应的第二中间向量为采用LSTM算法的隐含层对句向量处理后得到的向量,其他位置处的每个句向量对应的第二中间向量包括当前位置处的句向量和前一个句向量对应的第二中间向量;

[0205] 会话向量获取单元用于将在编码方向末位的句向量对应的第二中间向量作为历史会话文本对应的会话向量。

[0206] 另外,句向量获取模块7用于采用LSTM算法依次对目标句子中的每个词向量按照前向顺序和后向顺序分别进行编码处理,获取两个第一句向量,并将两个第一句向量进行融合处理获得句向量。

[0207] 通过将前向顺序和后向顺序分别进行编码处理得到的两个向量进行融合,进一步提高句向量表征目标句子的正确性。

[0208] 会话向量获取模块8用于采用LSTM算法依次对目标句子中的每个词向量按照前向顺序和后向顺序分别进行编码处理,获取两个第一会话向量,并将两个第一会话向量进行融合处理获得会话向量。

[0209] 通过将前向顺序和后向顺序分别进行编码处理得到的两个向量进行融合,进一步提高会话向量表征历史会话文本的正确性。

[0210] 例如,历史会话文本中包括的目标句子为“今天天气怎么样”,则对其进行分词处理后得到“今天”、“天气”、“怎么样”,然后由这三个分词获取对应的词向量,得到该目标句子对应的向量序列。

[0211] 采用LSTM算法依次对三个词向量按照前向顺序进行编码处理:采用LSTM算法的隐含层对“今天”对应的词向量进行处理得到对应的第一中间向量,进而得到“天气”对应的第一中间向量包括“今天”对应的第一中间向量和“天气”对应的词向量,“怎么样”对应的第一中间向量包括“天气”对应的第一中间向量和“怎么样”对应的词向量,此时将“怎么样”对应的第一中间向量作为该目标句子对应的一个第一句向量,其中第一中间向量用于建立首位词向量到当前词向量的语义。

[0212] 采用LSTM算法依次对三个词向量按照后向顺序进行编码处理:采用LSTM算法的隐含层对“怎么样”对应的词向量进行处理得到对应的第一中间向量,进而得到“天气”对应的第一中间向量包括“怎么样”对应的第一中间向量和“天气”对应的词向量,“今天”对应的第一中间向量包括“天气”对应的第一中间向量和“今天”对应的词向量,此时将“今天”对应的第一中间向量作为该目标句子对应的一个第一句向量。

[0213] 然后将上述的两个第一句向量进行融合,这样就可以得到一个能够表征目标句子的准确度更高的句向量。具体地,两个向量融合的过程包括:如一个向量是[101],另一个向量是[001],融合后的结果可以为[101001],这样通过获取的丰富信息来更准确地表征目标句子对应的句向量;或者,通过比较相同位置处的词向量,然后取两个向量的交集作为最终的融合结果。

[0214] 对于得到同一历史会话文本中的多个目标句子对应的会话向量的过程与上述过程类似,因此此处就不再赘述。

[0215] 预测系统还包括权重获取模块和加权处理模块;

[0216] 权重获取模块用于采用Softmax算法获取目标句子中的每个词向量对应的第一权重;

[0217] 加权处理模块用于根据第一权重对目标句子中的每个词向量进行加权处理,得到加权处理后的句向量,从而进一步提高句向量表征目标句子的准确度。

[0218] 权重获取模块还用于采用Softmax算法获取历史会话文本中的每个句向量对应的第二权重;

[0219] 加权处理模块还用于根据第二权重对历史会话文本中的每个句向量进行加权处理,得到加权处理后的会话向量,从而进一步提高会话向量表征历史会话文本的准确度。

[0220] 模型建立模块9用于采用Softmax算法将历史会话文本对应的会话向量作为输入,将历史会话文本对应的场景类别作为输出,建立用于预测会话文本属于每种场景类别的概率模型。

[0221] 本实施例中,通过关键词获取数据库中的候选句子,进而得到对应的会话文本以及上下文句子,然后对从会话文本中挑选出的句子进行分词处理并转换为对应的词向量,再利用LSTM算法获取包含所有单词之间的语义关系的句向量,以及包含所有句子之间的语义关系的会话向量,进而建立用于预测会话文本属于每种场景类别的概率模型,来确定任

意一会话文本对应的目标场景类别,从而提高了现有的会话文本的场景识别的准确率以及召回率,同时降低了人工成本。

#### [0222] 实施例5

[0223] 图6为本发明实施例5提供的一种电子设备的结构示意图。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,处理器执行程序时实现实施例1或2中任意一实施例中的会话文本的场景类别的预测方法。图6显示的电子设备30仅仅是一个示例,不应对本发明实施例的功能和使用范围带来任何限制。

[0224] 如图6所示,电子设备30可以以通用计算设备的形式表现,例如其可以为服务器设备。电子设备30的组件可以包括但不限于:上述至少一个处理器31、上述至少一个存储器32、连接不同系统组件(包括存储器32和处理器31)的总线33。

[0225] 总线33包括数据总线、地址总线和控制总线。

[0226] 存储器32可以包括易失性存储器,例如随机存取存储器(RAM) 321和/或高速缓存存储器322,还可以进一步包括只读存储器(ROM) 323。

[0227] 存储器32还可以包括具有一组(至少一个)程序模块324的程序/实用工具325,这样的程序模块324包括但不限于:操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。

[0228] 处理器31通过运行存储在存储器32中的计算机程序,从而执行各种功能应用以及数据处理,例如本发明实施例1或2中任意一实施例中的会话文本的场景类别的预测方法。

[0229] 电子设备30也可以与一个或多个外部设备34(例如键盘、指向设备等)通信。这种通信可以通过输入/输出(I/O)接口35进行。并且,模型生成的设备30还可以通过网络适配器36与一个或者多个网络(例如局域网(LAN),广域网(WAN)和/或公共网络,如因特网)通信。如图6所示,网络适配器36通过总线33与模型生成的设备30的其它模块通信。应当明白,尽管图中未示出,可以结合模型生成的设备30使用其它硬件和/或软件模块,包括但不限于:微代码、设备驱动器、冗余处理器、外部磁盘驱动阵列、RAID(磁盘阵列)系统、磁带驱动器以及数据备份存储系统等。

[0230] 应当注意,尽管在上文详细描述中提及了电子设备的若干单元/模块或子单元/模块,但是这种划分仅仅是示例性的并非强制性的。实际上,根据本发明的实施方式,上文描述的两个或更多单元/模块的特征和功能可以在一个单元/模块中具体化。反之,上文描述的一个单元/模块的特征和功能可以进一步划分为由多个单元/模块来具体化。

#### [0231] 实施例12

[0232] 本实施例提供了一种计算机可读存储介质,其上存储有计算机程序,程序被处理器执行时实现实施例1或2中任意一实施例中的会话文本的场景类别的预测方法中的步骤。

[0233] 其中,可读存储介质可以采用的更具体可以包括但不限于:便携式盘、硬盘、随机存取存储器、只读存储器、可擦拭可编程只读存储器、光存储器件、磁存储器件或上述的任意合适的组合。

[0234] 在可能的实施方式中,本发明还可以实现为一种程序产品的形式,其包括程序代码,当程序产品在终端设备上运行时,程序代码用于使终端设备执行实现实施例1或2中任意一实施例中的会话文本的场景类别的预测方法中的步骤。

[0235] 其中,可以以一种或多种程序设计语言的任意组合来编写用于执行本发明的程序

代码,程序代码可以完全地在用户设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户设备上部分在远程设备上执行或完全在远程设备上执行。

[0236] 虽然以上描述了本发明的具体实施方式,但是本领域的技术人员应当理解,这仅是举例说明,本发明的保护范围是由所附权利要求书限定的。本领域的技术人员在不背离本发明的原理和实质的前提下,可以对这些实施方式做出多种变更或修改,但这些变更和修改均落入本发明的保护范围。



图1

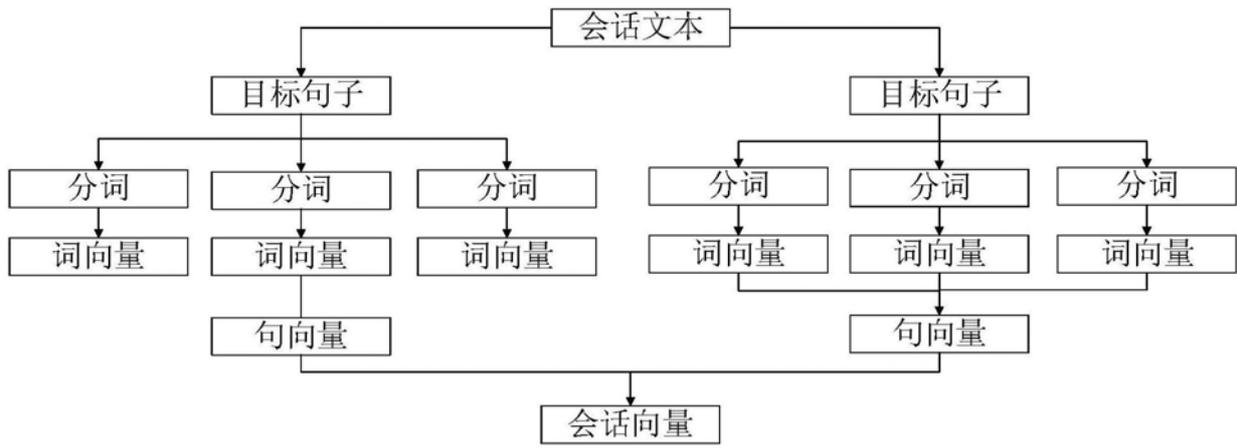


图2

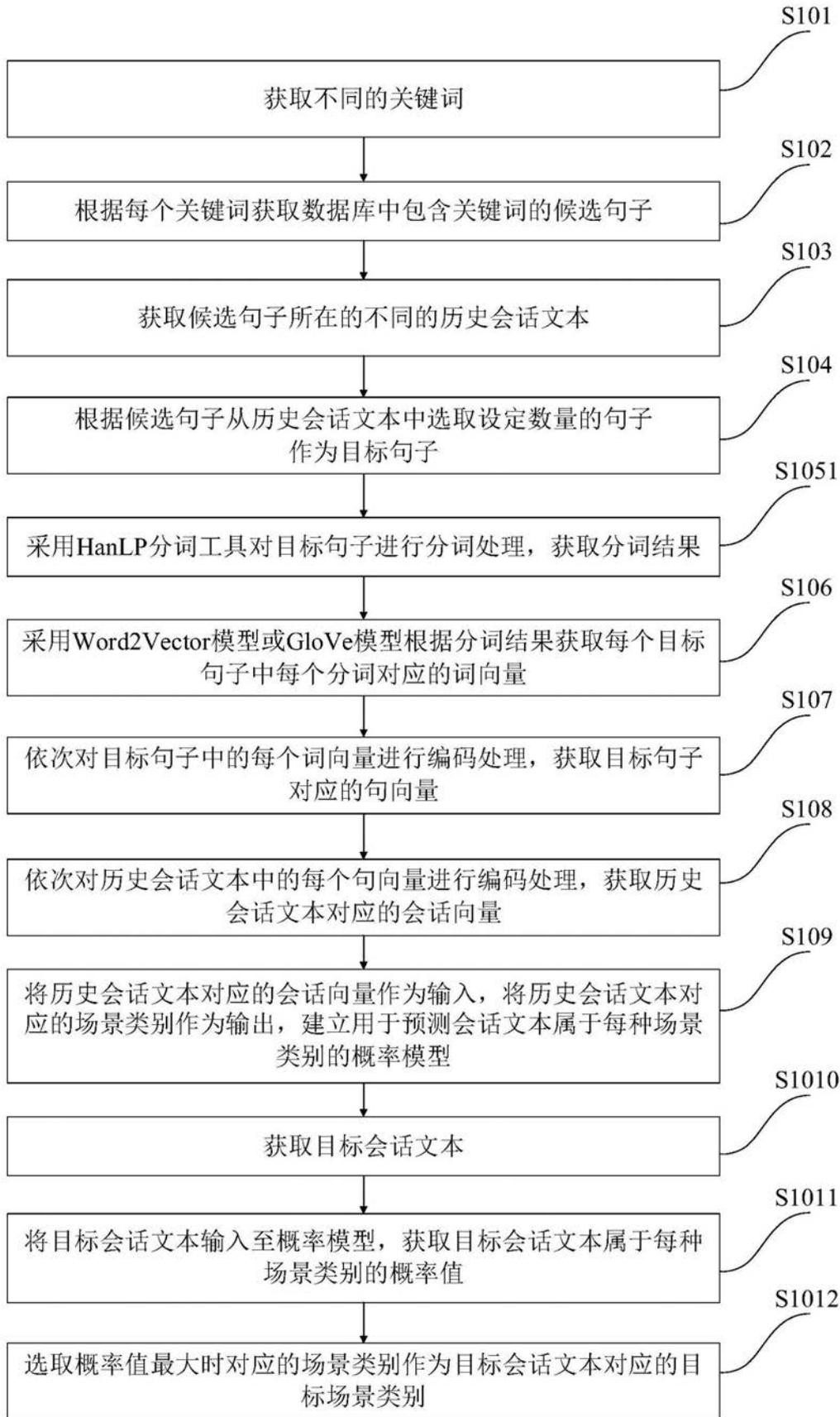


图3

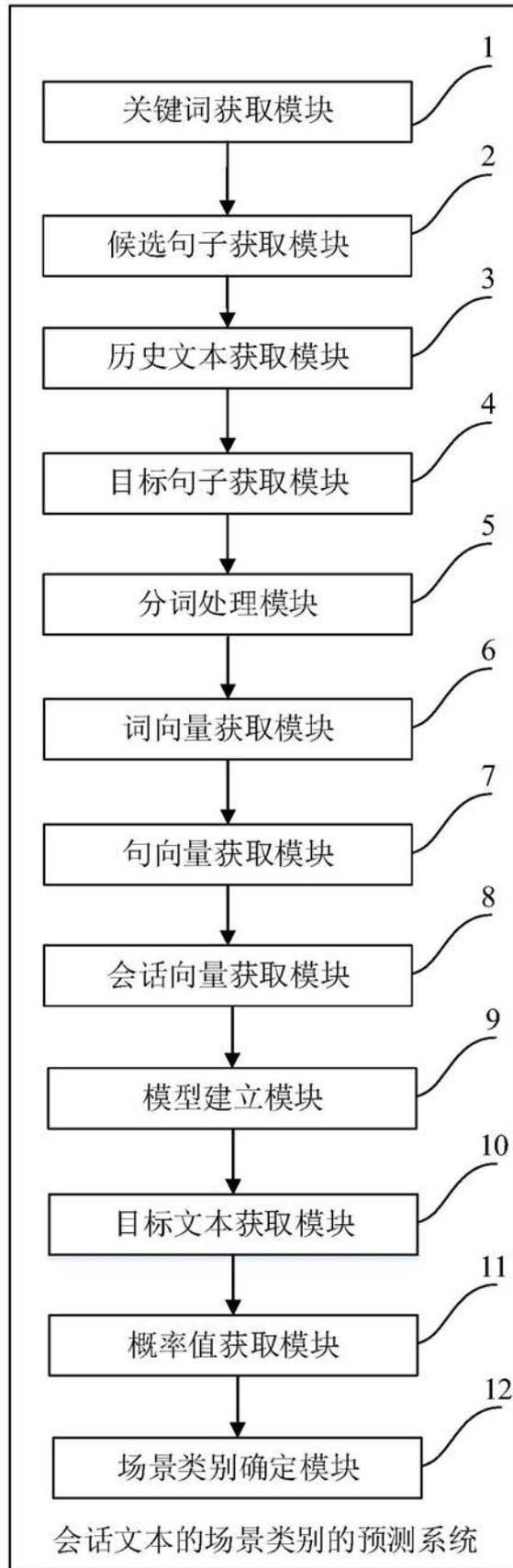


图4

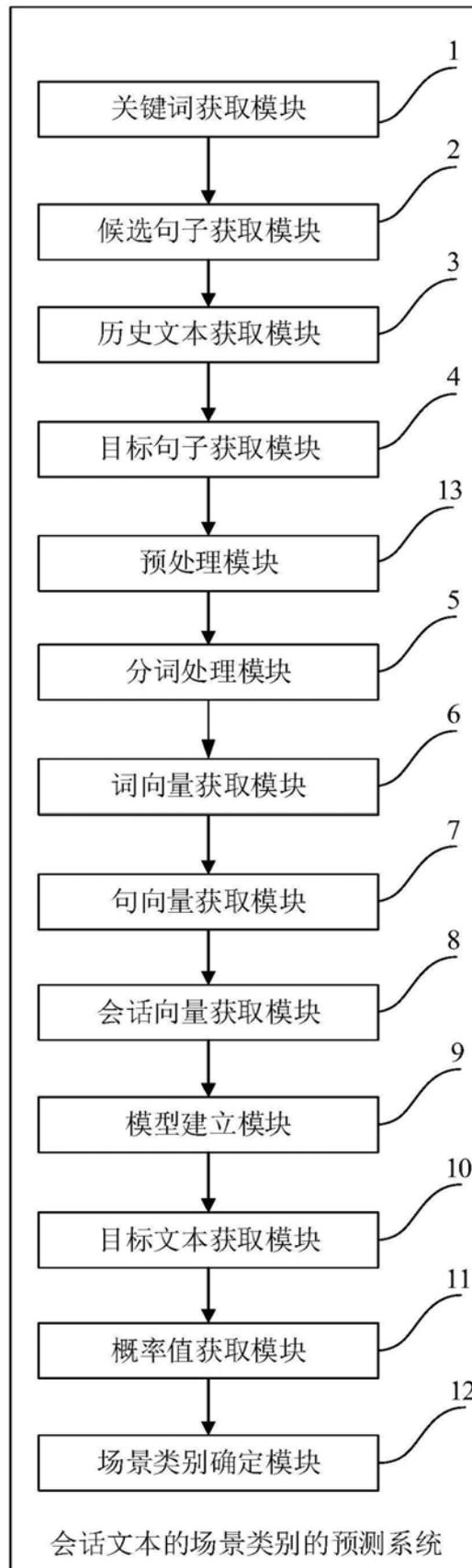


图5

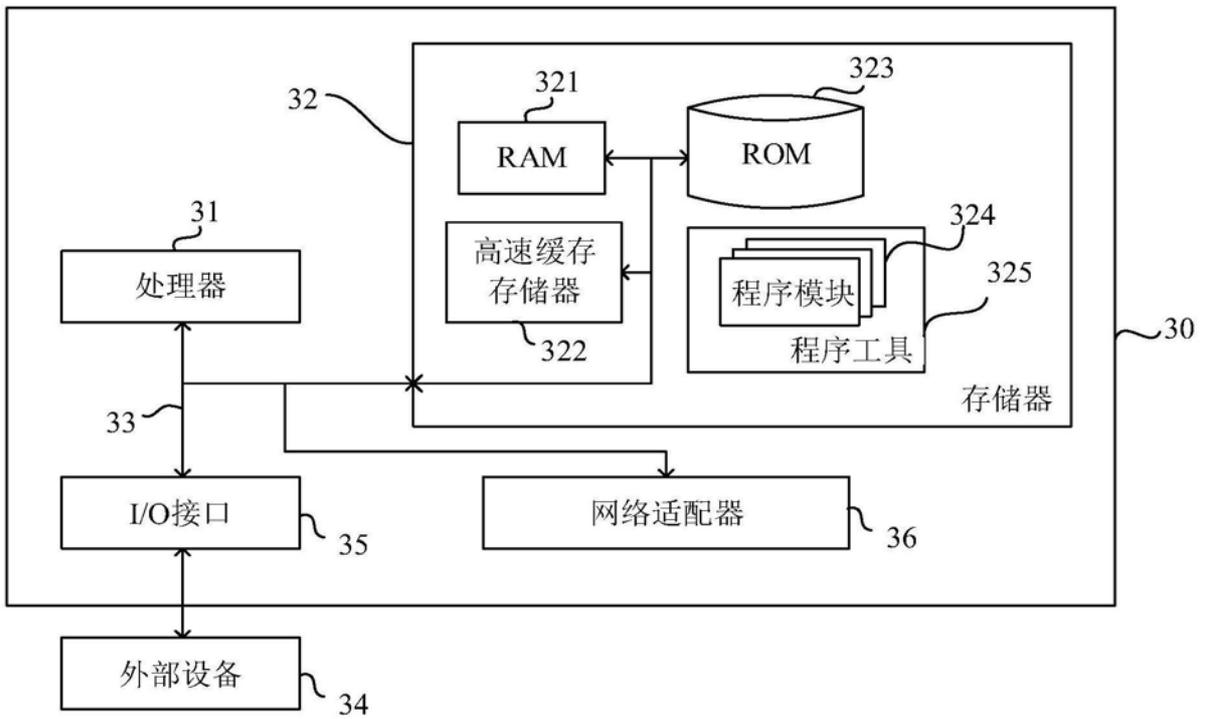


图6