



(12) 发明专利

(10) 授权公告号 CN 110489552 B

(45) 授权公告日 2021.09.21

(21) 申请号 201910645815.8

(22) 申请日 2019.07.17

(65) 同一申请的已公布的文献号  
申请公布号 CN 110489552 A

(43) 申请公布日 2019.11.22

(73) 专利权人 清华大学  
地址 100084 北京市海淀区清华园北京  
100084-82信箱

(72) 发明人 冯铃 曹楠 张慧君

(74) 专利代理机构 北京路浩知识产权代理有限公司 11002

代理人 苗晓静

(51) Int. Cl.

G06F 16/35 (2019.01)

G06F 16/958 (2019.01)

G06F 40/30 (2020.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

(56) 对比文件

CN 109918649 A, 2019.06.21

CN 107145524 A, 2017.09.08

CN 109918556 A, 2019.06.21

CN 105183876 A, 2015.12.23

KR 20140026951 A, 2014.03.06

CN 109524106 A, 2019.03.26

丁楠. “基于机器学习的大学生自杀风险预测与分析”.《现代电子技术》.2017,第40卷(第21期),

Samah Fodeh等. “Leveraging Twitter to better identify suicide risk”.《Proceedings of The First Workshop Medical Informatics and Healthcare held with the 23rd SIGKDD Conference on Knowledge Discovery and Data Mining》.2017,

审查员 姚晓斌

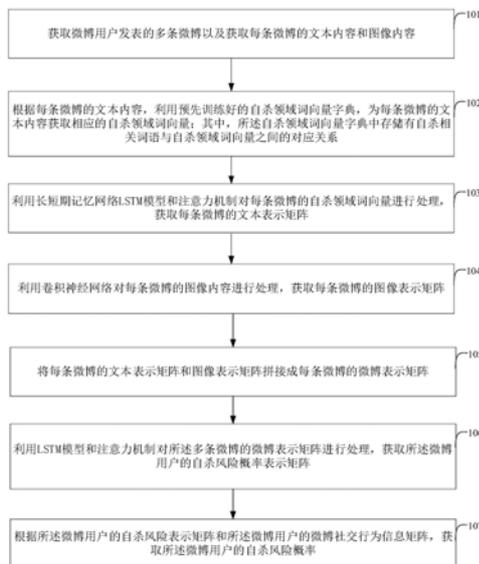
权利要求书3页 说明书14页 附图5页

(54) 发明名称

一种微博用户自杀风险检测方法及装置

(57) 摘要

本发明实施例提供一种微博用户自杀风险检测方法及装置,方法包括:获取每条微博的自杀领域词向量;利用LSTM模型和注意力机制对每条微博的自杀领域词向量进行处理,获取每条微博的文本表示矩阵;将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵;利用LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取微博用户的自杀风险概率表示矩阵;根据微博用户的自杀风险表示矩阵和微博用户的微博社交行为信息矩阵,获取微博用户的自杀风险概率。本发明实施例基于自杀领域词向量、神经网络模型和注意力机制能够较好地理解微博中的隐式自杀倾向表达,从而检便于从日常微博中检测潜在的微博自杀风险。



1. 一种微博用户自杀风险检测方法,其特征在于,包括:

获取微博用户发表的多条微博以及获取每条微博的文本内容和图像内容;

根据每条微博的文本内容,利用预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量;其中,所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系;

利用长短期记忆网络LSTM模型和注意力机制对每条微博的自杀领域词向量进行处理,获取每条微博的文本表示矩阵;

利用卷积神经网络对每条微博的图像内容进行处理,获取每条微博的图像表示矩阵;

将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵;

利用LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵;

根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵,获取所述微博用户的自杀风险概率。

2. 根据权利要求1所述的微博用户自杀风险检测方法,其特征在于,所述利用长短期记忆网络LSTM模型和注意力机制对每条微博的自杀领域词向量进行处理,获取每条微博的文本表示矩阵,具体包括:

将每条微博的自杀领域词向量  $\hat{X} = \{x_1, x_2, \dots, x_z\} \in \mathbb{R}^{z \times d_e}$  输入到LSTM模型中:

$$h_t = \text{LSTM}(x_t, h_{t-1})$$

其中,  $\hat{X}$  表示每条微博文本内容  $s_i$  对应的自杀领域词向量,  $z$  是每条微博文本内容  $s_i$  的长度,  $d_e$  是每个自杀领域词向量的长度;  $h_t, h_{t-1}$  表示在  $t$  时刻和  $t-1$  时刻的隐含层状态,根据上面模型处理得到每条微博文本内容  $s_i$  的初步文本表示矩阵  $H_i^s = \{h_1, h_2, \dots, h_z\} \in \mathbb{R}^{z \times d_e}$ ,  $\mathbb{R}$  表示实数空间;

为进一步得到初步文本表示矩阵中关于自杀的关键信息,采用注意力机制进行处理:

$$\text{Att\_I} = \text{softmax}(H_i^s W_3 + b_3)$$

其中,  $\text{Att\_I}$  是注意力向量,它表示在初步文本表示矩阵中的每个单词的权重分布;

将所述注意力向量  $\text{Att\_I}$  和所述初步文本表示矩阵  $H_i^s$  进行矩阵乘法,得到每条微博文本内容  $s_i$  的最终文本表示  $\hat{H}_i$ :

$$\hat{H}_i = (H_i^s)^T \text{Att\_I};$$

其中,  $\text{softmax}$  表示归一化指数函数,  $W_3$  表示第一类训练参数中的第三预设训练参数,  $b_3$  表示第二类训练参数中的第三预设训练参数。

3. 根据权利要求2所述的微博用户自杀风险检测方法,其特征在于,所述利用卷积神经网络对每条微博的图像内容进行处理,获取每条微博的图像表示矩阵,具体包括:

利用预先训练好的卷积神经网络对每条微博的图像内容进行如下处理,获取每条微博的图像表示矩阵  $I_i$ :

$$I_i = \tanh(O W_4 + b_4)$$

其中,0是最后一层全连接层的图像内容输入,图像表示矩阵 $I_i$ 是图像内容 $p_i$ 的图像表示, $\tanh$ 表示双曲正切函数, $W_4$ 表示第一类训练参数中的第四预设训练参数, $b_4$ 表示第二类训练参数中的第四预设训练参数。

4.根据权利要求3所述的微博用户自杀风险检测方法,其特征在于,所述将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵,具体包括:

对于每条微博 $(s_i, p_i)$ ,将与每条微博文本内容 $s_i$ 对应的最终文本表示 $\hat{H}_i$ 和每条微博的图像内容 $p_i$ 对应的图像表示矩阵 $I_i$ 拼接成微博表示 $E_i$ :

$$E_i = \hat{H}_i \oplus I_i^T$$

其中, $I_i^T$ 为 $I_i$ 的转置矩阵。

5.根据权利要求4所述的微博用户自杀风险检测方法,其特征在于,所述利用长短期记忆网络LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵,具体包括:

按照下面模型获取所述微博用户的自杀风险概率表示矩阵G:

$$h_t = \text{LSTM}(E_i, h_{t-1})$$

$$\text{Att\_II} = \text{softmax}(H^e W_5 + b_5)$$

$$G = \tanh((\text{Att\_II})^T \times H^e W_6 + b_6)$$

其中, $h_t, h_{t-1}$ 表示在t时刻和t-1时刻的隐含层状态,根据m条微博中每条微博的微博表示 $E_i$ 以及模型 $h_t = \text{LSTM}(E_i, h_{t-1})$ 得到微博用户的初步微博表示矩阵 $H^e = \{h_1, h_2, \dots, h_m\}$ ,m为从微博用户中获取的微博条数,由于并非每条微博都表达了自杀的想法,因此使用另一个注意力机制来获取微博用户的高自杀风险信息,这里计算了注意力向量Att\_II,表现出不同的风险权重,然后基于注意力向量Att\_II和初步微博表示矩阵 $H^e$ ,计算出所述微博用户的自杀风险表示矩阵G,其中, $W_5 \sim W_6$ 表示第一类训练参数中的第五和第六预设训练参数, $b_5 \sim b_6$ 表示第二类训练参数中的第五和第六预设训练参数。

6.根据权利要求5所述的微博用户自杀风险检测方法,其特征在于,根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵,获取所述微博用户的自杀风险概率,具体包括:

使用拼接操作来共同考虑所述微博用户的自杀风险表示矩阵G和所述微博用户的微博社交行为信息矩阵F,并通过一个全连接层来计算所述微博用户的自杀风险概率:

$$[y_1, y_0] = \text{softmax}(W_7 (G \oplus F) + b_7)$$

其中, $y_1, y_0$ 表示所述微博用户是否有自杀风险的可能性, $W_7$ 表示第一类训练参数中的第七预设训练参数, $b_7$ 表示第二类训练参数中的第七预设训练参数, $\oplus$ 表示矩阵拼接操作。

7.根据权利要求1所述的微博用户自杀风险检测方法,其特征在于,在所述根据每条微博的文本内容,利用预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量之前,所述方法还包括:获取预先训练好的自杀领域词向量字典的步骤,具体包括:

获取微博树洞中包含自杀词语的评论作为训练样本语句;其中,所述自杀词语属于自杀词典中的自杀词语;

对所述训练样本语句采用遮挡方式进行处理,得到带有非自杀标记的第一训练样本语

句和带有自杀标记的第二训练样本语句,其中,所述第一训练样本语句中的所有自杀词语均被替换成预设标记词语,所述第二训练样本语句中的随机插入两个所述预设标记词语;所述预设标记词语为与自杀领域无关的词语;

利用带有非自杀标记的所述第一训练样本语句、带有自杀标记的所述第二训练样本语句对预设自杀表达分类模型进行训练,获取预先训练好的自杀领域词向量字典;其中,所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系。

8. 一种微博用户自杀风险检测装置,其特征在于,包括:

第一获取模块,用于获取微博用户发表的多条微博以及获取每条微博的文本内容和图像内容;

第二获取模块,用于根据每条微博的文本内容,利用预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量;其中,所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系;

第三获取模块,用于利用长短期记忆网络LSTM模型和注意力机制对每条微博的自杀领域词向量进行处理,获取每条微博的文本表示矩阵;

第四获取模块,用于利用卷积神经网络对每条微博的图像内容进行处理,获取每条微博的图像表示矩阵;

第五获取模块,用于将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵;

第六获取模块,用于利用LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵;

第七获取模块,用于根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵,获取所述微博用户的自杀风险概率。

9. 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7任一项所述微博用户自杀风险检测方法的步骤。

10. 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7任一项所述微博用户自杀风险检测方法的步骤。

## 一种微博用户自杀风险检测方法及装置

### 技术领域

[0001] 本发明涉及计算机技术领域,尤其涉及一种微博用户自杀风险检测方法及装置。

### 背景技术

[0002] 据世界卫生组织发布的数据表明,每年世界上有超过80万人死于自杀,平均每40秒就有一个。在心理学领域,快速精准的检测并预防自杀成为了一个重要的任务。

[0003] 传统的自杀风险检测方法一般需要受试者填写一个问卷或者接受一次面谈。但是,这些方法只适用于一小群人,特别是对于那些受到心理创伤之后往往隐藏自己的感受并拒绝寻求他人帮助的人来说,这些方法无法发挥作用。

[0004] 最近,由于社交媒体(如论坛和微博)具有大规模,低成本和开放的优势,其使得研究人员能够克服之前的限制并及时发现个人的自杀意念。尽管已经有了一些相关工作利用社交媒体数据来进行用户自杀风险检测,但是由于用户的隐式和非真实的表达,这些工作的效果也一定程度上受到限制。

[0005] 统计数据表示,有自杀倾向的人倾向于在已自杀者的微博评论区中透露他们真实的内心感受。这样的微博评论区称之为“树洞”,现在已经有数百个树洞存在于新浪微博之中,其中最大的树洞已经包含了160000人发的1700000条评论。

[0006] 目前希望通过利用社交媒体上的相关数据,来检测微博用户的自杀风险。

### 发明内容

[0007] 针对现有技术中的问题,本发明实施例提供一种微博用户自杀风险检测方法及装置。

[0008] 第一方面,本发明实施例提供了一种微博用户自杀风险检测方法,包括:

[0009] 获取微博用户发表的多条微博以及获取每条微博的文本内容和图像内容;

[0010] 根据每条微博的文本内容,利用预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量;其中,所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系;

[0011] 利用长短期记忆网络LSTM模型和注意力机制对每条微博的自杀领域词向量进行处理,获取每条微博的文本表示矩阵;

[0012] 利用卷积神经网络对每条微博的图像内容进行处理,获取每条微博的图像表示矩阵;

[0013] 将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵;

[0014] 利用LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵;

[0015] 根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵,获取所述微博用户的自杀风险概率。

[0016] 进一步地,所述利用长短期记忆网络LSTM模型和注意力机制对每条微博的自杀领

域词向量进行处理,获取每条微博的文本表示矩阵,具体包括:

[0017] 将每条微博的自杀领域词向量  $\hat{X} = \{x_1, x_2, \dots, x_z\} \in \mathbb{R}^{z \times d_e}$  输入到LSTM模型中:

[0018]  $h_t = \text{LSTM}(x_t, h_{t-1})$

[0019] 其中,  $\hat{X}$  表示每条微博文本内容  $s_i$  对应的自杀领域词向量,  $z$  是每条微博文本内容  $s_i$  的长度,  $d_e$  是每个自杀领域词向量的长度;  $h_t, h_{t-1}$  表示在  $t$  时刻和  $t-1$  时刻的隐含层状态, 根据上面模型处理得到每条微博文本内容  $s_i$  的初步文本表示矩阵  $H_i^s = \{h_1, h_2, \dots, h_z\} \in \mathbb{R}^{z \times d_e}$ ,  $\mathbb{R}$  表示实数空间;

[0020] 为进一步得到初步文本表示矩阵中关于自杀的关键信息,采用注意力机制进行处理:

[0021]  $Att\_I = \text{softmax}(H_i^s W_3 + b_3)$

[0022] 其中,  $Att\_I$  是注意力向量,它表示在初步文本表示矩阵中的每个单词的权重分布;

[0023] 将所述注意力向量  $Att\_I$  和所述初步文本表示矩阵  $H_i^s$  进行矩阵乘法,得到每条微博文本内容  $s_i$  的最终文本表示  $\hat{H}_i$  :

[0024]  $\hat{H}_i = (H_i^s)^T Att\_I$ ;

[0025] 其中,  $\text{softmax}$  表示归一化指数函数,  $W_3$  表示第一类训练参数中的第三预设训练参数,  $b_3$  表示第二类训练参数中的第三预设训练参数。

[0026] 进一步地,所述利用卷积神经网络对每条微博的图像内容进行处理,获取每条微博的图像表示矩阵,具体包括:

[0027] 利用预先训练好的卷积神经网络对每条微博的图像内容进行如下处理,获取每条微博的图像表示矩阵  $I_i$  :

[0028]  $I_i = \tanh(OW_4 + b_4)$

[0029] 其中,  $O$  是最后一层全连接层的图像内容输入,图像表示矩阵  $I_i$  是图像内容  $p_i$  的图像表示,  $\tanh$  表示双曲正切函数,  $W_4$  表示第一类训练参数中的第四预设训练参数,  $b_4$  表示第二类训练参数中的第四预设训练参数。

[0030] 进一步地,所述将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵,具体包括:

[0031] 对于每条微博  $(s_i, p_i)$ ,将与每条微博文本内容  $s_i$  对应的最终文本表示  $\hat{H}_i$  和每条微博的图像内容  $p_i$  对应的图像表示矩阵  $I_i$  拼接成微博表示  $E_i$  :

[0032]  $E_i = \hat{H}_i \oplus I_i^T$

[0033] 其中,  $I_i^T$  为  $I_i$  的转置矩阵。

[0034] 进一步地,所述利用长短期记忆网络LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵,具体包括:

[0035] 按照下面模型获取所述微博用户的自杀风险概率表示矩阵  $G$ :

[0036]  $h_t = \text{LSTM}(E_i, h_{t-1})$

[0037]  $\text{Att\_II} = \text{softmax}(H^e W_5 + b_5)$

[0038]  $G = \tanh(((\text{Att\_II})^T \times H^e) W_6 + b_6)$

[0039] 其中,  $h_t, h_{t-1}$  表示在  $t$  时刻和  $t-1$  时刻的隐含层状态, 根据  $m$  条微博中每条微博的微博表示  $E_i$  以及模型  $h_t = \text{LSTM}(E_i, h_{t-1})$  得到微博用户的初步微博表示矩阵  $H^e = \{h_1, h_2, \dots, h_m\}$ ,  $m$  为从微博用户中获取的微博条数, 由于并非每条微博都表达了自杀的想法, 因此使用另一个注意力机制来获取微博用户的高自杀风险信息, 这里计算了注意力向量  $\text{Att\_II}$ , 表现出不同的风险权重, 然后基于注意力向量  $\text{Att\_II}$  和初步微博表示矩阵  $H^e$ , 计算出所述微博用户的自杀风险表示矩阵  $G$ , 其中,  $W_5 \sim W_6$  表示第一类训练参数中的第五和第六预设训练参数,  $b_5 \sim b_6$  表示第二类训练参数中的第五和第六预设训练参数。

[0040] 进一步地, 根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵, 获取所述微博用户的自杀风险概率, 具体包括:

[0041] 使用拼接操作来共同考虑所述微博用户的自杀风险表示矩阵  $G$  和所述微博用户的微博社交行为信息矩阵  $F$ , 并通过一个全连接层来计算所述微博用户的自杀风险概率:

[0042]  $[y_1, y_0] = \text{softmax}(W_7(G \oplus F) + b_7)$

[0043] 其中,  $y_1, y_0$  表示所述微博用户是否有自杀风险的可能性,  $W_7$  表示第一类训练参数中的第七预设训练参数,  $b_7$  表示第二类训练参数中的第七预设训练参数,  $\oplus$  表示矩阵拼接操作。

[0044] 进一步地, 在所述根据每条微博的文本内容, 利用预先训练好的自杀领域词向量字典, 为每条微博的文本内容获取相应的自杀领域词向量之前, 所述方法还包括: 获取预先训练好的自杀领域词向量字典的步骤, 具体包括:

[0045] 获取微博树洞中包含自杀词语的评论作为训练样本语句; 其中, 所述自杀词语属于自杀词典中的自杀词语;

[0046] 对所述训练样本语句采用遮挡方式进行处理, 得到带有非自杀标记的第一训练样本语句和带有自杀标记的第二训练样本语句, 其中, 所述第一训练样本语句中的所有自杀词语均被替换成预设标记词语, 所述第二训练样本语句中的随机插入两个所述预设标记词语; 所述预设标记词语为与自杀领域无关的词语;

[0047] 利用带有非自杀标记的所述第一训练样本语句、带有自杀标记的所述第二训练样本语句对预设自杀表达分类模型进行训练, 获取预先训练好的自杀领域词向量字典; 其中, 所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系。

[0048] 第二方面, 本发明实施例还提供了一种微博用户自杀风险检测装置, 包括:

[0049] 第一获取模块, 用于获取微博用户发表的多条微博以及获取每条微博的文本内容和图像内容;

[0050] 第二获取模块, 用于根据每条微博的文本内容, 利用预先训练好的自杀领域词向量字典, 为每条微博的文本内容获取相应的自杀领域词向量; 其中, 所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系;

[0051] 第三获取模块, 用于利用长短期记忆网络 LSTM 模型和注意力机制对每条微博的自杀领域词向量进行处理, 获取每条微博的文本表示矩阵;

[0052] 第四获取模块, 用于利用卷积神经网络对每条微博的图像内容进行处理, 获取每

条微博的图像表示矩阵；

[0053] 第五获取模块,用于将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵；

[0054] 第六获取模块,用于利用LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵；

[0055] 第七获取模块,用于根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵,获取所述微博用户的自杀风险概率。

[0056] 第三方面,本发明实施例还提供了一种电子设备,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第一方面所述微博用户自杀风险检测方法的步骤。

[0057] 第四方面,本发明实施例还提供了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第一方面所述微博用户自杀风险检测方法的步骤。

[0058] 由上面技术方案可知,本发明实施例提供的微博用户自杀风险检测方法及装置,先从微博用户中获取多条微博并获取每条微博的文本内容和图像内容,然后根据每条微博的文本内容,利用预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量,接着利用长短期记忆网络LSTM模型和注意力机制对每条微博的自杀领域词向量进行处理,获取每条微博的文本表示矩阵,并利用卷积神经网络对每条微博的图像内容进行处理,获取每条微博的图像表示矩阵;然后将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵,并继续利用LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵,最后根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵,获取所述微博用户的自杀风险概率。本发明实施例基于自杀领域词向量、神经网络模型和注意力机制能够较好地理解微博中的隐式自杀倾向表达,从而便于从日常微博中检测潜在的微博自杀风险。

## 附图说明

[0059] 为了更清楚地说明本发明实施例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发明的一些实施例,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。

[0060] 图1为本发明一实施例提供的微博用户自杀风险检测方法的流程图；

[0061] 图2为本发明一实施例提供的训练自杀领域词向量的过程示意图；

[0062] 图3为本发明一实施例提供的预测微博用户自杀风险的模型结构图；

[0063] 图4为本发明一实施例提供的微博用户自杀风险检测装置的结构示意图；

[0064] 图5为本发明一实施例提供的电子设备的结构示意图。

## 具体实施方式

[0065] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例

中的附图,对本发明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发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,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,都属于本发明保护的范围。

[0066] 本发明实施例提供的微博用户自杀风险检测方法,利用微博用户的微博信息,构建了一种基于自杀领域的词向量和增强深度学习模型对文本自杀表达的理解与风险感知的微博用户自杀风险检测方法。下面将通过具体实施例对本发明实施例提供的微博用户自杀风险检测方法及装置进行详细说明。

[0067] 图1示出了本发明实施例提供的微博用户自杀风险检测方法的流程图。如图1所示,本发明实施例提供的微博用户自杀风险检测方法包括如下步骤:

[0068] 步骤101:获取微博用户发表的多条微博以及获取每条微博的文本内容和图像内容。

[0069] 步骤102:根据每条微博的文本内容,利用预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量;其中,所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系。

[0070] 在本步骤中,可以根据预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量。其中,所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系。

[0071] 在介绍本步骤具体内容之前,需要先解释一下获取预先训练好的自杀领域词向量字典的过程:

[0072] 首先,获取微博树洞中包含自杀词语的评论作为训练样本语句;其中,所述自杀词语属于自杀词典中的自杀词语;

[0073] 其次,对所述训练样本语句采用遮挡方式进行处理,得到带有非自杀标记的第一训练样本语句和带有自杀标记的第二训练样本语句,其中,所述第一训练样本语句中的所有自杀词语均被替换成预设标记词语,所述第二训练样本语句中的随机插入两个所述预设标记词语;所述预设标记词语为与自杀领域无关的词语;

[0074] 最后,利用带有非自杀标记的所述第一训练样本语句、带有自杀标记的所述第二训练样本语句对预设自杀表达分类模型进行训练,获取预先训练好的自杀领域词向量字典;其中,所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系。

[0075] 对于上述训练过程,具体可参见图2所示的内容,先从微博树洞中获取包含自杀词语的训练样本语句,如图2中所示的:I don't want to suicide.I really hope the end of the world comes soon.Everyone is going to die,接着对所述训练样本语句采用遮挡方式进行处理,得到带有非自杀标记的第一训练样本语句和带有自杀标记的第二训练样本语句,其中,所述第一训练样本语句中的所有自杀词语均被替换成预设标记词语,所述第二训练样本语句中的随机插入两个所述预设标记词语;所述预设标记词语为与自杀领域无关的词语;参见图2所示,对图2中所示的训练样本语句I don't want to suicide.I really hope the end of the world comes soon.Everyone is going to die采用遮挡方式进行处理,得到带有非自杀标记的第一训练样本语句和带有自杀标记的第二训练样本语句。其中,本实施例选择50%的句子用“[mask]”替换所有与自杀有关的单词/短语。对于

其余50%的句子,本实施例在每个句子中随机插入两个“[mask]”,以避免仅根据是否包含单词“[mask]”来对句子进行分类。如图2所示。第一训练样本语句Masked sentence 1是本实施例用“[mask]”替换所有与自杀相关的单词/短语的句子,而第二训练样本语句Masked sentence 2是本实施例随机插入两个“[mask]”的句子。本实施例将Masked sentence 1标记为0(非自杀),Masked sentence 2标记为1(自杀)。在本实施例中,需要说明的是,虽然已经有一些关于词向量的优秀工作,但由于缺乏领域信息限制了它们在自杀检测方面的表现。基于一系列预训练好的词向量和自杀相关词典,本实施例的目标是生成与自杀相关的词向量,这可以加强与自杀相关的词汇的敏感性。在这项研究中,本实施例称之为面向自杀的词向量。由于社交媒体中的自杀用户经常在他们的帖子中使用与自杀相关的单词/短语,本实施例使用中国自杀词典来生成与自杀领域相关的词向量。中国自杀词典分析了106万个活跃微博用户的帖子,列出了与自杀意念有关的2168个单词/短语。这些单词/短语属于13个类别,每个单词/短语的自杀风险权重从1到3分配,表明自杀的相关性。由于预训练的词向量已经包含丰富的语义信息和语境信息,本实施例只需要用自杀相关的信息来丰富现有的词向量。本实施例采用遮挡分类任务来解决这个问题。一般来说,如果一个句子表达了自杀意念,那么它应该包含与自杀相关的单词/短语。因此,本实施例从树洞数据集中选择10000个句子来进行训练,并确保每个句子都包含多个在中文自杀词典中出现的单词/短语。此外,本实施例利用所选择的句子进行自杀表达分类,只有在包含至少一个与自杀相关的单词/短语时,才会将句子视为自杀表达。这样,本实施例会进行句子级别的分类,以调整预先训练过的词向量,并让他们理解哪个单词/短语与自杀表达相关。在训练过程中,对于每次迭代,本实施例选择50%的句子用“[mask]”替换所有与自杀有关的单词/短语。对于其余50%的句子,本实施例在每个句子中随机插入两个“[mask]”,以避免仅根据是否包含单词“[mask]”来对句子进行分类。如图2所示。Masked sentence 1是本实施例用“[mask]”替换所有与自杀相关的单词/短语的句子,而Masked sentence 2是本实施例随机插入两个“[mask]”的句子。本实施例将Masked sentence 1标记为0(非自杀),Masked sentence 2标记为1(自杀)。由于自杀相关单词/短语与预训练词向量中的其他单词/短语之间没有明确的界限,通过这种自杀表达分类,在这里本实施例强制自杀相关的单词/短语包含丰富的自杀领域信息。在自杀表达分类模型收敛于树洞数据集后,本实施例获得了自杀领域的词向量,它包含了来自预训练词向量的语义信息和来自自杀字典的自杀信息。

[0076] 如图2所示,假设一个微博用户在微博树洞中写的句子 $A = \{w_1, w_2, \dots, w_n\}$ ,自杀表达分类的目的是分类这个句子是否包含关于自杀意念的表达。在这种情况下,本实施例令 $\hat{X} = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^{n \times d_e}$ 表示A的词向量, $\mathbb{R}$ 表示实数空间,其中 $d_e$ 是每个词向量的长度。图2显示了自杀表达分类模型的结构。本实施例使用LSTM层和全连接层提取文本特征以进行分类。本实施例将词向量X以如下方式输入到LSTM中: $h_t = \text{LSTM}(x_t, h_{t-1})$ , $[k_1, k_2] = \text{softmax}((HW_1 + b_1)^T W_2 + b_2)$ ,其中, $h_t, h_{t-1}$ 表示在t时刻和t-1时刻的隐含层状态, $H = \{h_1, h_2, \dots, h_n\}$ 是句子A的文本表示, $k_1, k_2$ 代表句子中包含自杀表达的可能性。

[0077] 在本步骤中,目的是训练出自杀领域词向量,在训练过程中,利用模型的反向传播优化算法不断的调整输入的词向量,当观察到自杀表达分类模型已经拟合后,可以取出词向量,进而获取到训练好的自杀领域词向量字典。

[0078] 在本步骤102中,在获取到训练好的自杀领域词向量后,根据每条微博的文本内容,利用预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量;例如图3所示的,利用预先训练好的自杀领域词向量字典,获取与微博的文本内容 $w_1, w_2$ 和 $w_3$ (也即“why I live”,这是将为什么活着翻译后的文本)对应的自杀领域词向量 $x_1, x_2$ 和 $x_3$ 。

[0079] 步骤103:利用长短期记忆网络LSTM模型和注意力机制对每条微博的自杀领域词向量进行处理,获取每条微博的文本表示矩阵。

[0080] 在本步骤中,假设一个微博用户发表的多个微博为 $\hat{T}$ ,其中, $\hat{T}=\{(s_1, p_1), (s_2, p_2), \dots, (s_m, p_m)\}$ ,这里面 $m$ 表示此用户的微博总数, $(s_i, p_i)$ 表示此用户的第 $i$ 条微博的文本和图片。令 $\hat{X}=\{x_1, x_2, \dots, x_z\} \in \mathbb{R}^{z \times d_e}$ 表示 $s_i$ 的词向量,其中 $z$ 是文本的长度, $d_e$ 是每个词向量的长度,图3展示了自杀检测模型的具体结构。

[0081] 在本步骤中,参见图3所示,将每条微博的自杀领域词向量 $\hat{X}=\{x_1, x_2, \dots, x_z\} \in \mathbb{R}^{z \times d_e}$ 输入到LSTM模型中:

$$[0082] \quad h_t = \text{LSTM}(x_t, h_{t-1})$$

[0083] 其中, $\hat{X}$ 表示每条微博文本内容 $s_i$ 对应的自杀领域词向量, $z$ 是每条微博文本内容 $s_i$ 的长度, $d_e$ 是每个自杀领域词向量的长度; $h_t, h_{t-1}$ 表示在 $t$ 时刻和 $t-1$ 时刻的隐含层状态,根据上面模型处理得到每条微博文本内容 $s_i$ 的初步文本表示矩阵 $H_i^s = \{h_1, h_2, \dots, h_z\} \in \mathbb{R}^{z \times d_e}$ ;

[0084] 为进一步得到初步文本表示矩阵中关于自杀的关键信息,采用注意力机制进行处理:

$$[0085] \quad \text{Att\_I} = \text{softmax}(H_i^s W_3 + b_3)$$

[0086] 其中,Att\_I是注意力向量,它表示在初步文本表示矩阵中的每个单词的权重分布;

[0087] 将所述注意力向量Att\_I和所述初步文本表示矩阵 $H_i^s$ 进行矩阵乘法,得到每条微博文本内容 $s_i$ 的最终文本表示 $\hat{H}_i$ :

$$[0088] \quad \hat{H}_i = (H_i^s)^T \text{Att\_I};$$

[0089] 其中,softmax表示归一化指数函数, $W_3$ 表示第一类训练参数中的第三预设训练参数, $b_3$ 表示第二类训练参数中的第三预设训练参数。

[0090] 步骤104:利用卷积神经网络对每条微博的图像内容进行处理,获取每条微博的图像表示矩阵。

[0091] 在本步骤中,参见图3所示,利用预先训练好的卷积神经网络对每条微博的图像内容进行如下处理,获取每条微博的图像表示矩阵 $I_i$ :

$$[0092] \quad I_i = \tanh(0W_4 + b_4)$$

[0093] 其中, $0$ 是最后一层全连接层的图像内容输入,图像表示矩阵 $I_i$ 是图像内容 $p_i$ 的图像表示,tanh表示双曲正切函数, $W_4$ 表示第一类训练参数中的第四预设训练参数, $b_4$ 表示第二类训练参数中的第四预设训练参数。

[0094] 需要说明的是,在获取图像表示矩阵时,本实施例使用了预训练好的34层卷积神经网络。此外,为了方便本次实验,本实施例将其最后一层全连接层从 $512 \times 1000$ 改为了 $512 \times d_e$ 。

[0095] 此外,需要说明的是,当某条微博包含内容时,将采用默认的图像内容进行相应的处理,获取与所述默认的图像内容对应的图像表示矩阵。其中,所述默认的图像内容应该为不包含任何自杀相关信息的图像内容,优选地,所述默认的图像内容应该为不带有明显感情色彩的图像内容。

[0096] 步骤105:将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵。

[0097] 在本步骤中,参见图3所示,对于每条微博 $(s_i, p_i)$ ,将与每条微博文本内容 $s_i$ 对应的最终文本表示 $\hat{H}_i$ 和每条微博的图像内容 $p_i$ 对应的图像表示矩阵 $I_i$ 拼接成微博表示 $E_i$ :

$$[0098] \quad E_i = \hat{H}_i \oplus I_i^T$$

[0099] 其中, $I_i^T$ 为 $I_i$ 的转置矩阵。

[0100] 步骤106:利用LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵。

[0101] 在本步骤中,参见图3所示,按照下面模型获取所述微博用户的自杀风险概率表示矩阵G:

$$[0102] \quad h_t = \text{LSTM}(E_i, h_{t-1})$$

$$[0103] \quad \text{Att\_II} = \text{softmax}(H^e W_5 + b_5)$$

$$[0104] \quad G = \tanh(((\text{Att\_II})^T \times H^e) W_6 + b_6)$$

[0105] 其中, $h_t, h_{t-1}$ 表示在t时刻和t-1时刻的隐含层状态,根据m条微博中每条微博的微博表示 $E_i$ 以及模型 $h_t = \text{LSTM}(E_i, h_{t-1})$ 得到微博用户的初步微博表示矩阵 $H^e = \{h_1, h_2, \dots, h_m\}$ ,m为从微博用户中获取的微博条数,由于并非每条微博都表达了自杀的想法,因此使用另一个注意力机制来获取微博用户的高自杀风险信息,这里计算了注意力向量Att\_II,表现出不同的风险权重,然后基于注意力向量Att\_II和初步微博表示矩阵 $H^e$ ,计算出所述微博用户的自杀风险表示矩阵G,其中, $W_5 \sim W_6$ 表示第一类训练参数中的第五和第六预设训练参数, $b_5 \sim b_6$ 表示第二类训练参数中的第五和第六预设训练参数。

[0106] 步骤107:根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵,获取所述微博用户的自杀风险概率。

[0107] 在本步骤中,所述微博用户的微博社交行为信息矩阵是指包含了微博用户的微博社交行为信息的矩阵,这里的微博社交行为信息可以指昵称、出生年月、爱好、在微博上的活跃度、粉丝数、粉丝背景、一段时间内总共发表的微博数目等等。需要说明的是,除了采用微博用户的自杀风险表示矩阵以外,进一步结合所述微博用户的微博社交行为信息矩阵,可以得到更为准确的自杀风险预测结果。

[0108] 在本步骤中,从用户的个人资料中提取了大约12个左右的特征组成了所述微博用户的微博社交行为信息矩阵。

[0109] 在本步骤中,参见图3所示,使用拼接操作来共同考虑所述微博用户的自杀风险表示矩阵G和所述微博用户的微博社交行为信息矩阵F,并通过一个全连接层来计算所述微博

用户的自杀风险概率：

$$[0110] \quad [y_1, y_0] = \text{softmax}(W_7(G \oplus F) + b_7)$$

[0111] 其中,  $y_1, y_0$  表示所述微博用户是否有自杀风险的可能性,  $W_7$  表示第一类训练参数中的第七预设训练参数,  $b_7$  表示第二类训练参数中的第七预设训练参数,  $\oplus$  表示矩阵拼接操作。

[0112] 在本步骤中, 需要说明的是,  $y_1$  和  $y_0$  都是 0-1 之间的数, 且  $y_1$  和  $y_0$  之和为 1。其中,  $y_1$  表示有自杀风险,  $y_0$  表示无自杀风险, 假设  $y_1$  大于  $y_0$ , 则表示该用户有自杀风险, 若  $y_1$  小于  $y_0$ , 则表示该用户无自杀风险。

[0113] 需要说明的是, 本发明实施例研究的目的是理解微博中的隐式表达, 以便从日常微博中检测潜在的微博自杀风险。在本实施例中, 基于树洞内容, 构建基于树洞内容的自杀领域词向量, 基于树洞评论的表达, 增强自杀相关词汇在词向量中的敏感性。此外, 在本实施例中, 提出分层的注意机制, 能够较好地理解微博中的隐式自杀倾向表达, 从而检便于从日常微博中检测潜在的微博自杀风险。

[0114] 由上面技术方案可知, 本发明实施例提供的微博用户自杀风险检测方法, 先将树洞评论输入自杀表达分类模型, 获取自杀相关词语对应的词向量 (也即自杀领域词向量字典), 之后从微博用户中获取多条微博并获取每条微博的文本内容和图像内容, 然后根据自杀领域词向量字典获取每条微博的文本内容对应的自杀领域词向量, 接着利用长短期记忆网络 LSTM 模型和注意力机制对每条微博的自杀领域词向量进行处理, 获取每条微博的文本表示矩阵, 并利用卷积神经网络对每条微博的图像内容进行处理, 获取每条微博的图像表示矩阵; 然后将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵, 并继续利用 LSTM 模型和注意力机制对所述多条微博的微博表示矩阵进行处理, 获取所述微博用户的自杀风险概率表示矩阵, 最后根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵, 获取所述微博用户的自杀风险概率。本发明实施例基于自杀领域词向量、神经网络模型和注意力机制能够较好地理解微博中的隐式自杀倾向表达, 从而检便于从日常微博中检测潜在的微博自杀风险。

[0115] 基于上述实施例的内容, 在本实施例中, 上述步骤 103 可通过如下方式实现:

[0116] 将每条微博的自杀领域词向量  $\hat{X} = \{x_1, x_2, \dots, x_z\} \in \mathbb{R}^{z \times d_e}$  输入到 LSTM 模型中:

$$[0117] \quad h_t = \text{LSTM}(x_t, h_{t-1})$$

[0118] 其中,  $\hat{X}$  表示每条微博文本内容  $s_i$  对应的自杀领域词向量,  $z$  是每条微博文本内容  $s_i$  的长度,  $d_e$  是每个自杀领域词向量的长度;  $h_t, h_{t-1}$  表示在  $t$  时刻和  $t-1$  时刻的隐含层状态, 根据上面模型处理得到每条微博文本内容  $s_i$  的初步文本表示矩阵  $H_i^s = \{h_1, h_2, \dots, h_z\} \in \mathbb{R}^{z \times d_e}$ ;

[0119] 为进一步得到初步文本表示矩阵中关于自杀的关键信息, 采用注意力机制进行处理:

$$[0120] \quad \text{Att\_I} = \text{softmax}(H_i^s W_3 + b_3)$$

[0121] 其中,  $\text{Att\_I}$  是注意力向量, 它表示在初步文本表示矩阵中的每个单词的权重分布;

[0122] 将所述注意力向量  $\text{Att\_I}$  和所述初步文本表示矩阵  $H_i^s$  进行矩阵乘法, 得到每条微

博文本内容 $s_i$ 的最终文本表示 $\hat{H}_i$ ：

$$[0123] \quad \hat{H}_i = (H_i^s)^T \text{Att\_I};$$

[0124] 其中,softmax表示归一化指数函数, $W_3$ 表示第一类训练参数中的第三预设训练参数, $b_3$ 表示第二类训练参数中的第三预设训练参数。

[0125] 基于上述实施例的内容,在本实施例中,上述步骤104可通过如下方式实现:

[0126] 利用预先训练好的卷积神经网络对每条微博的图像内容进行如下处理,获取每条微博的图像表示矩阵 $I_i$ ：

$$[0127] \quad I_i = \tanh(OW_4 + b_4)$$

[0128] 其中,0是最后一层全连接层的图像内容输入,图像表示矩阵 $I_i$ 是图像内容 $p_i$ 的图像表示,tanh表示双曲正切函数, $W_4$ 表示第一类训练参数中的第四预设训练参数, $b_4$ 表示第二类训练参数中的第四预设训练参数。

[0129] 基于上述实施例的内容,在本实施例中,上述步骤105可通过如下方式实现:

[0130] 对于每条微博 $(s_i, p_i)$ ,将与每条微博文本内容 $s_i$ 对应的最终文本表示 $\hat{H}_i$ 和每条微博的图像内容 $p_i$ 对应的图像表示矩阵 $I_i$ 拼接成微博表示 $E_i$ ：

$$[0131] \quad E_i = \hat{H}_i \oplus I_i^T$$

[0132] 其中, $I_i^T$ 为 $I_i$ 的转置矩阵。

[0133] 基于上述实施例的内容,在本实施例中,上述步骤106可通过如下方式实现:

[0134] 按照下面模型获取所述微博用户的自杀风险概率表示矩阵G:

$$[0135] \quad h_t = \text{LSTM}(E_i, h_{t-1})$$

$$[0136] \quad \text{Att\_II} = \text{softmax}(H^e W_5 + b_5)$$

$$[0137] \quad G = \tanh(((\text{Att\_II})^T \times H^e) W_6 + b_6)$$

[0138] 其中, $h_t, h_{t-1}$ 表示在t时刻和t-1时刻的隐含层状态,根据m条微博中每条微博的微博表示 $E_i$ 以及模型 $h_t = \text{LSTM}(E_i, h_{t-1})$ 得到微博用户的初步微博表示矩阵 $H^e = \{h_1, h_2, \dots, h_m\}$ ,m为从微博用户中获取的微博条数,由于并非每条微博都表达了自杀的想法,因此使用另一个注意力机制来获取微博用户的高自杀风险信息,这里计算了注意力向量Att\_II,表现出不同的风险权重,然后基于注意力向量Att\_II和初步微博表示矩阵 $H^e$ ,计算出所述微博用户的自杀风险表示矩阵G,其中, $W_5 \sim W_6$ 表示第一类训练参数中的第五和第六预设训练参数, $b_5 \sim b_6$ 表示第二类训练参数中的第五和第六预设训练参数。

[0139] 基于上述实施例的内容,在本实施例中,上述步骤107可通过如下方式实现:

[0140] 使用拼接操作来共同考虑所述微博用户的自杀风险表示矩阵G和所述微博用户的微博社交行为信息矩阵F,并通过一个全连接层来计算所述微博用户的自杀风险概率:

$$[0141] \quad [y_1, y_0] = \text{softmax}(W_7(G \oplus F) + b_7)$$

[0142] 其中, $y_1, y_0$ 表示所述微博用户是否有自杀风险的可能性, $W_7$ 表示第一类训练参数中的第七预设训练参数, $b_7$ 表示第二类训练参数中的第七预设训练参数, $\oplus$ 表示矩阵拼接操作。其中,所述第一类训练参数和所述第二类训练参数均遵从于正态分布 $U(-0.001, 0.001)$ ,第一类训练参数中的各预设训练参数和第二类训练参数中的各预设训练参数可以

根据实际需要进行设置。

[0143] 图4示出了本发明实施例提供的微博用户自杀风险检测装置的结构示意图。如图4所示,本发明实施例提供的微博用户自杀风险检测装置包括:第一获取模块21、第二获取模块22、第三获取模块23、第四获取模块24、第五获取模块25、第六获取模块26和第七获取模块27,其中:

[0144] 第一获取模块21,用于获取微博用户发表的多条微博以及获取每条微博的文本内容和图像内容;

[0145] 第二获取模块22,用于根据每条微博的文本内容,利用预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量;其中,所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系;

[0146] 第三获取模块23,用于利用长短期记忆网络LSTM模型和注意力机制对每条微博的自杀领域词向量进行处理,获取每条微博的文本表示矩阵;

[0147] 第四获取模块24,用于利用卷积神经网络对每条微博的图像内容进行处理,获取每条微博的图像表示矩阵;

[0148] 第五获取模块25,用于将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵;

[0149] 第六获取模块26,用于利用LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵;

[0150] 第七获取模块27,用于根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵,获取所述微博用户的自杀风险概率。

[0151] 在本实施例中,所述第三获取模块23具体用于:

[0152] 将每条微博的自杀领域词向量  $\hat{X} = \{x_1, x_2, \dots, x_z\} \in \mathbb{R}^{z \times d_e}$  输入到LSTM模型中:

[0153]  $h_t = \text{LSTM}(x_t, h_{t-1})$

[0154] 其中,  $\hat{X}$  表示每条微博文本内容  $s_i$  对应的自杀领域词向量,  $z$  是每条微博文本内容  $s_i$  的长度,  $d_e$  是每个自杀领域词向量的长度;  $h_t, h_{t-1}$  表示在  $t$  时刻和  $t-1$  时刻的隐含层状态,根据上面模型处理得到每条微博文本内容  $s_i$  的初步文本表示矩阵  $H_i^s = \{h_1, h_2, \dots, h_z\} \in \mathbb{R}^{z \times d_e}$ ;

[0155] 为进一步得到初步文本表示矩阵中关于自杀的关键信息,采用注意力机制进行处理:

[0156]  $Att\_I = \text{softmax}(H_i^s W_3 + b_3)$

[0157] 其中,  $Att\_I$  是注意力向量,它表示在初步文本表示矩阵中的每个单词的权重分布;

[0158] 将所述注意力向量  $Att\_I$  和所述初步文本表示矩阵  $H_i^s$  进行矩阵乘法,得到每条微博文本内容  $s_i$  的最终文本表示  $\hat{H}_i$  :

[0159]  $\hat{H}_i = (H_i^s)^T Att\_I$ ;

[0160] 其中,  $\text{softmax}$  表示归一化指数函数,  $W_3$  表示第一类训练参数中的第三预设训练参数,  $b_3$  表示第二类训练参数中的第三预设训练参数。

[0161] 在本实施例中,所述第四获取模块24,具体用于:

[0162] 利用预先训练好的卷积神经网络对每条微博的图像内容进行如下处理,获取每条微博的图像表示矩阵 $I_i$ :

$$[0163] \quad I_i = \tanh(0W_4 + b_4)$$

[0164] 其中,0是最后一层全连接层的图像内容输入,图像表示矩阵 $I_i$ 是图像内容 $p_i$ 的图像表示, $\tanh$ 表示双曲正切函数, $W_4$ 表示第一类训练参数中的第四预设训练参数, $b_4$ 表示第二类训练参数中的第四预设训练参数。

[0165] 在本实施例中,所述第五获取模块25,具体用于:

[0166] 对于每条微博 $(s_i, p_i)$ ,将与每条微博文本内容 $s_i$ 对应的最终文本表示 $\hat{H}_i$ 和每条微博的图像内容 $p_i$ 对应的图像表示矩阵 $I_i$ 拼接成微博表示 $E_i$ :

$$[0167] \quad E_i = \hat{H}_i \oplus I_i^T$$

[0168] 其中, $I_i^T$ 为 $I_i$ 的转置矩阵。

[0169] 在本实施例中,所述第六获取模块26,具体用于:

[0170] 按照下面模型获取所述微博用户的自杀风险概率表示矩阵G:

$$[0171] \quad h_t = \text{LSTM}(E_i, h_{t-1})$$

$$[0172] \quad \text{Att\_II} = \text{softmax}(H^e W_5 + b_5)$$

$$[0173] \quad G = \tanh(((\text{Att\_II})^T \times H^e) W_6 + b_6)$$

[0174] 其中, $h_t, h_{t-1}$ 表示在t时刻和t-1时刻的隐含层状态,根据m条微博中每条微博的微博表示 $E_i$ 以及模型 $h_t = \text{LSTM}(E_i, h_{t-1})$ 得到微博用户的初步微博表示矩阵 $H^e = \{h_1, h_2, \dots, h_m\}$ ,m为从微博用户中获取的微博条数,由于并非每条微博都表达了自杀的想法,因此使用另一个注意力机制来获取微博用户的高自杀风险信息,这里计算了注意力向量Att\_II,表现出不同的风险权重,然后基于注意力向量Att\_II和初步微博表示矩阵 $H^e$ ,计算出所述微博用户的自杀风险表示矩阵G,其中, $W_5 \sim W_6$ 表示第一类训练参数中的第五和第六预设训练参数, $b_5 \sim b_6$ 表示第二类训练参数中的第五和第六预设训练参数。

[0175] 在本实施例中,所述第七获取模块27,具体用于:

[0176] 使用拼接操作来共同考虑所述微博用户的自杀风险表示矩阵G和所述微博用户的微博社交行为信息矩阵F,并通过一个全连接层来计算所述微博用户的自杀风险概率:

$$[0177] \quad [y_1, y_0] = \text{softmax}(W_7 (G \oplus F) + b_7)$$

[0178] 其中, $y_1, y_0$ 表示所述微博用户是否有自杀风险的可能性, $W_7$ 表示第一类训练参数中的第七预设训练参数, $b_7$ 表示第二类训练参数中的第七预设训练参数, $\oplus$ 表示矩阵拼接操作。

[0179] 由于本发明实施例提供的微博用户自杀风险检测装置,可以用于执行上述实施例所述的的微博用户自杀风险检测方法,其工作原理和有益效果类似,故此处不再详述,具体内容可参见上述实施例的介绍。

[0180] 基于相同的发明构思,本发明又一实施例提供了一种电子设备,参见图5,所述电子设备具体包括如下内容:处理器301、存储器302、通信接口303和通信总线304;

[0181] 其中,所述处理器301、存储器302、通信接口303通过所述通信总线304完成相互间的通信;

[0182] 所述处理器301用于调用所述存储器302中的计算机程序,所述处理器执行所述计算机程序时实现上述微博用户自杀风险检测方法的全部步骤,例如,所述处理器执行所述计算机程序时实现下述过程:获取微博用户发表的多条微博以及获取每条微博的文本内容和图像内容;根据每条微博的文本内容,利用预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量;其中,所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系;利用长短期记忆网络LSTM模型和注意力机制对每条微博的自杀领域词向量进行处理,获取每条微博的文本表示矩阵;利用卷积神经网络对每条微博的图像内容进行处理,获取每条微博的图像表示矩阵;将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵;利用LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵;根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵,获取所述微博用户的自杀风险概率。

[0183] 基于相同的发明构思,本发明又一实施例提供了一种非暂态计算机可读存储介质,该非暂态计算机可读存储介质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微博用户自杀风险检测方法的全部步骤,例如,所述处理器执行所述计算机程序时实现下述过程:获取微博用户发表的多条微博以及获取每条微博的文本内容和图像内容;根据每条微博的文本内容,利用预先训练好的自杀领域词向量字典,为每条微博的文本内容获取相应的自杀领域词向量;其中,所述自杀领域词向量字典中存储有自杀相关词语与自杀领域词向量之间的对应关系;利用长短期记忆网络LSTM模型和注意力机制对每条微博的自杀领域词向量进行处理,获取每条微博的文本表示矩阵;利用卷积神经网络对每条微博的图像内容进行处理,获取每条微博的图像表示矩阵;将每条微博的文本表示矩阵和图像表示矩阵拼接成每条微博的微博表示矩阵;利用LSTM模型和注意力机制对所述多条微博的微博表示矩阵进行处理,获取所述微博用户的自杀风险概率表示矩阵;根据所述微博用户的自杀风险表示矩阵和所述微博用户的微博社交行为信息矩阵,获取所述微博用户的自杀风险概率。

[0184] 此外,上述的存储器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,可以存储在一个计算机可读存储介质中。基于这样的理解,本发明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)执行本发明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。

[0185] 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发明实施例方案的目的。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。

[0186] 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)执行各个实施例或者实施例的某些部分所述的用于心理压力检测的多模态融合方法。

[0187] 最后应说明的是:以上实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,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发明各实施例技术方案的精神和范围。

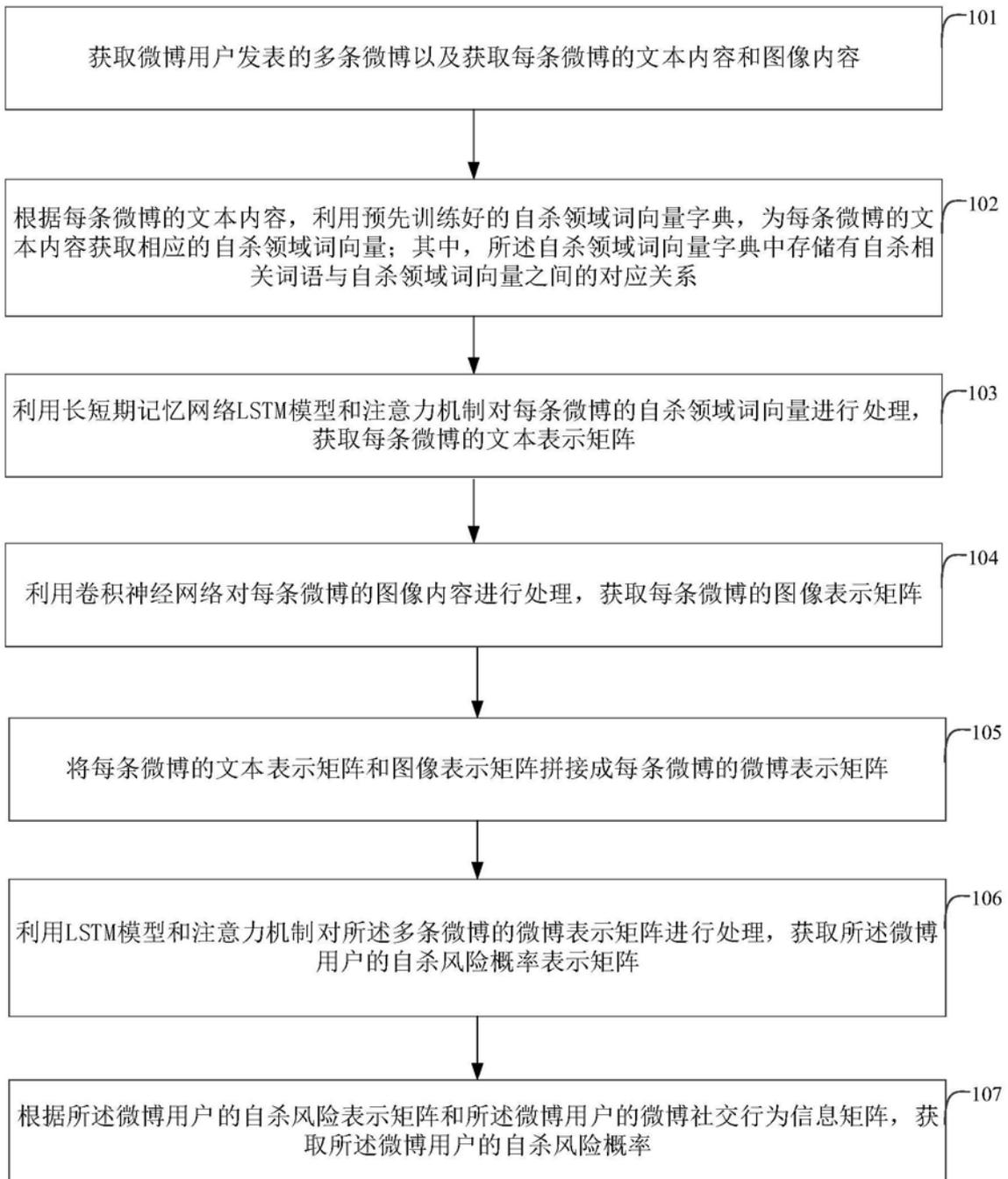


图1

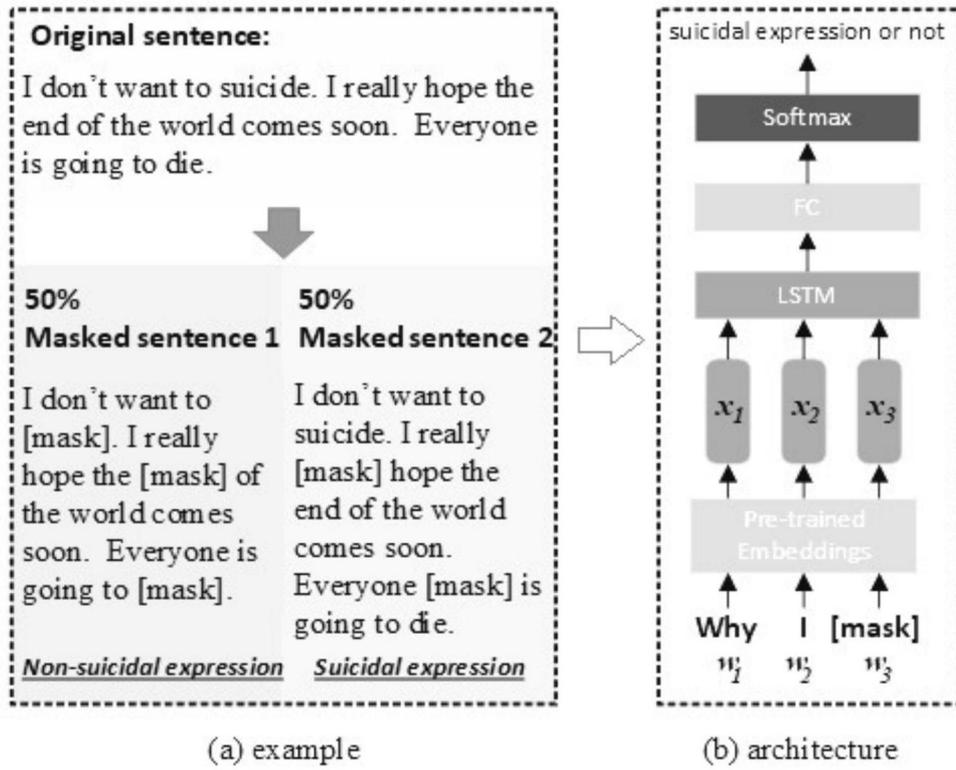


图2

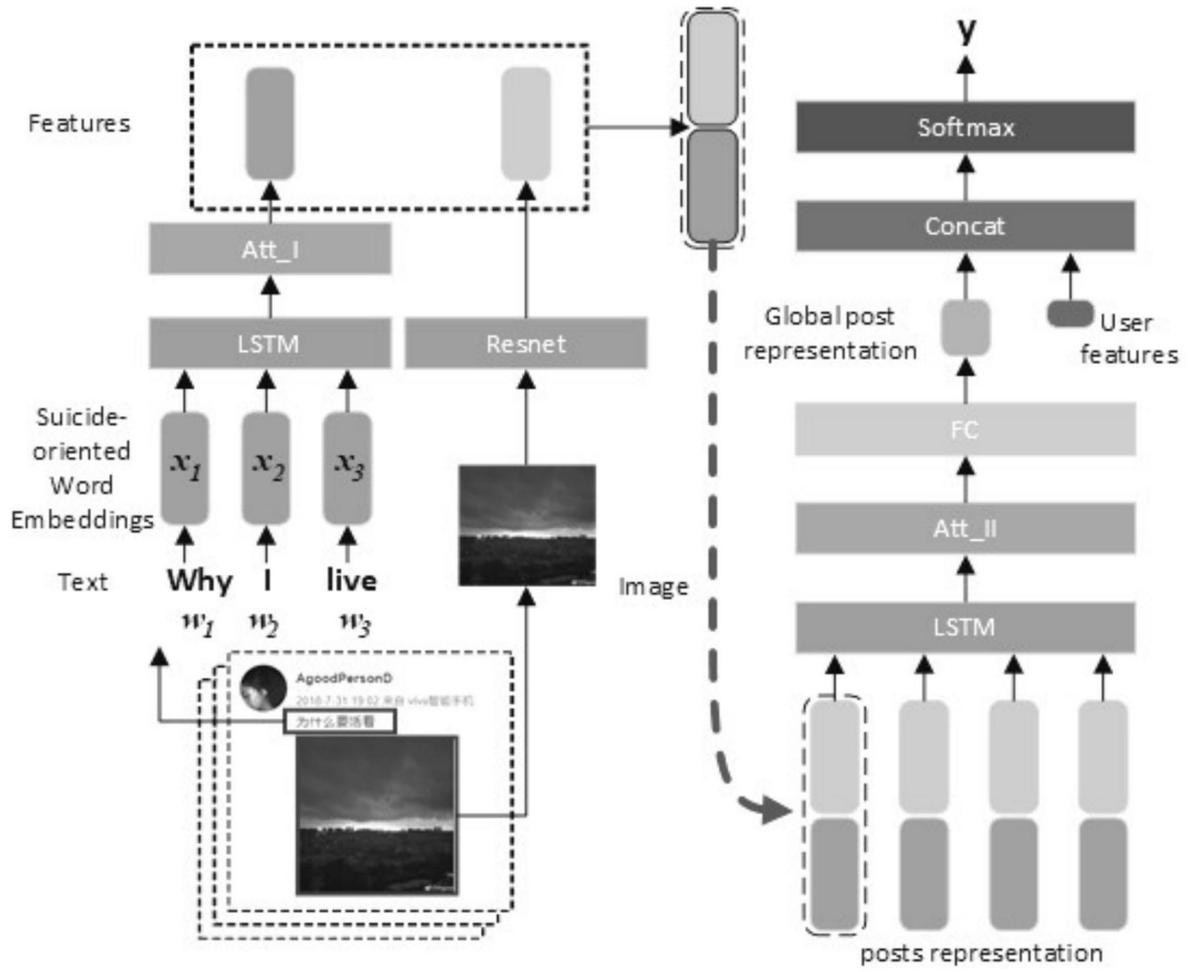


图3

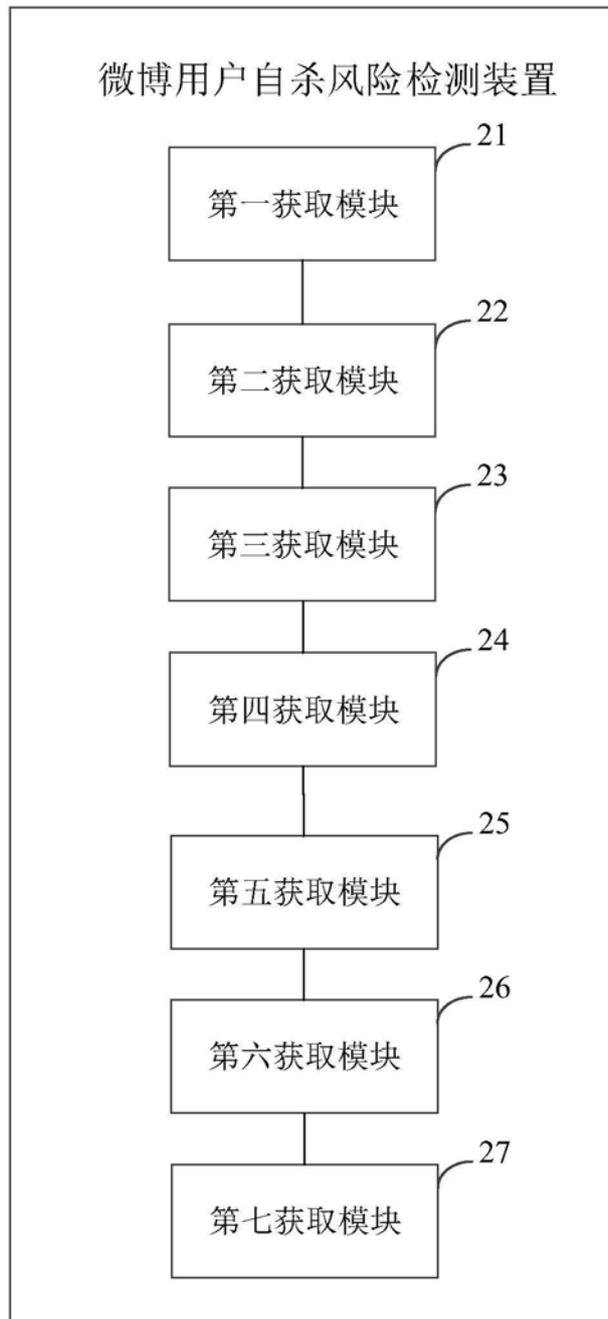


图4

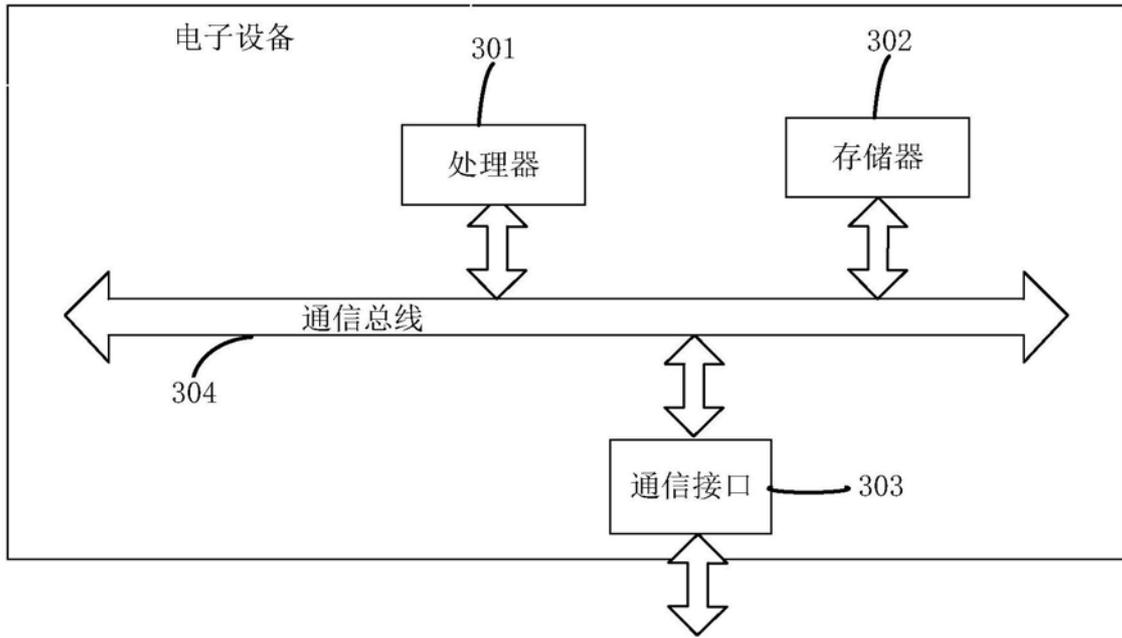


图5