



(12) 发明专利申请

(10) 申请公布号 CN 115422945 A

(43) 申请公布日 2022. 12. 02

(21) 申请号 202211139407.3

G06N 3/08 (2006.01)

(22) 申请日 2022.09.19

(71) 申请人 福州大学

地址 350108 福建省福州市闽侯县福州大学城乌龙江北大道2号福州大学

(72) 发明人 陈羽中 朱文龙 饶孟宇 万宇杰

(74) 专利代理机构 福州元创专利商标代理有限公司 35100

专利代理师 郭东亮 蔡学俊

(51) Int.Cl.

G06F 40/30 (2020.01)

G06F 40/211 (2020.01)

G06F 16/36 (2019.01)

G06K 9/62 (2022.01)

G06N 3/04 (2006.01)

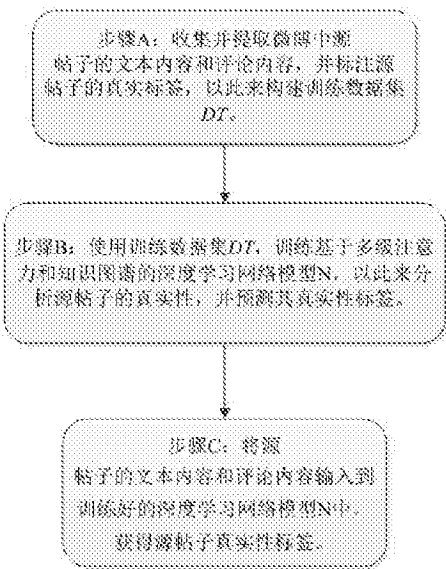
权利要求书8页 说明书16页 附图3页

(54) 发明名称

一种融合情感挖掘的谣言检测方法及系统

(57) 摘要

本发明提出一种融合情感挖掘的谣言检测方法,所述方法包括以下步骤:步骤A:收集并提取社交网络媒体中源帖子的文本内容和评论内容,并人工标注源帖子的真实标签,形成训练数据集DT;步骤B:使用训练数据集DT,训练基于多级注意力和知识图谱的深度学习网络模型N,训练内容包括分析源帖子的真实性和预测源帖子的真实性标签;步骤C:将源帖子的文本内容和评论内容输入到训练好的深度学习网络模型N中,获得源帖子的真实性标签;本发明可以提升对微博进行谣言检测的准确性。



1. 一种融合情感挖掘的谣言检测方法,其特征在于:所述方法包括以下步骤;

步骤A:收集并提取社交网络媒体中源帖子的文本内容和评论内容,并人工标注源帖子的真实标签,形成训练数据集DT;

步骤B:使用训练数据集DT,训练基于多级注意力和知识图谱的深度学习网络模型N,训练内容包括分析源帖子的真实性和预测源帖子的真实性标签;

步骤C:将源帖子的文本内容和评论内容输入到训练好的深度学习网络模型N中,获得源帖子的真实性标签。

2. 根据权利要求1所述的一种融合情感挖掘的谣言检测方法,其特征在于:所述步骤B包括以下步骤;

步骤B1:对训练数据集DT中的每个训练样本进行编码,以得到文本内容的初始表征向量 T^{st} 、评论内容的初始表征向量 T^{rt} 以及句法邻接矩阵 A^{sk} ;

步骤B2:根据句法知识子图构造算法从知识图谱和句法依赖图中生成文本内容的对应的句法知识子图SK,并得到其邻接矩阵 A^{sk} ,然后对其节点进行编码,得到句法知识子图SK的节点知识表示向量 H^{sk} ;

步骤B3:将步骤B1得到的文本内容初始表征向量 T^{st} 输入到双向长短期记忆网络Bi-LSTM中,得到上下文增强的文本内容表征向量 H^{st} ,令 $U^{st}=H^{st}$;然后,将表征向量 T^{st} 和评论内容初始表征向量 T^{rt} 一起输入到多头交叉注意力机制中,得到基于文本内容的评论表征向量 P^{sr} ,同时将表征向量 T^{st} 输入到多头自注意力机制中,得到文本内容增强表征向量 P^s ;然后通过把基于文本内容的评论表征向量 P^{sr} 和文本内容增强表征向量 P^s 分别输入到池化层中进行平均池化操作,得到平均池化评论内容句子表征向量 \bar{P}^{sr} 和平均池化文本内容增强表征向量 \bar{P}^s ;

步骤B4:将子图SK的节点知识表示向量 H^{sk} 和将步骤B3得到的表征向量 U^{st} 分别输入到两个具有K层的图卷积网络中,记为文本知识图卷积网络SKGCN和文本内容图卷积网络SCGCN,用于学习外部知识信息和提取句法信息;同时利用知识引导机制将文本内容图卷积网络SCGCN的每层节点与文本知识图卷积网络SKGCN进行知识引导,得到源帖子的图知识表征向量 V^{sks} ;

步骤B5:通过使用交叉注意力机制将B4步骤得到的图知识表征向量 V^{sks} 和句子表征向量 U^{st} 融合,得到知识增强型句子级表征向量 E^{sd} ,以进一步提高模型提取信息的能力;然后通过多头自注意力机制将 E^{sd} 进一步加强,得到聚合词级信息的句子表征 E^{mt} ;再通过门控机制减少来自不规范句子的噪声,得到源帖子情感表征向量 E^{sf} ;

步骤B6:将源帖子对应的平均池化评论内容句子表征向量 \bar{P}^{sr} 和平均池化文本内容增强表征向量 \bar{P}^s 一起输入到多头交叉注意力机制中,并通过平均池化得到评论内容的综合语义表征 C^{sr} ;然后将平均池化文本内容增强表征向量 \bar{P}^s 和评论内容的综合语义表征 C^{sr} 输入到融合门控机制中,得到源帖子的细粒度语义表征向量 V^t ;

步骤B7:将步骤B5得到的情感表征向量 E^{sf} 与步骤B6得到的源帖子的细粒度语义表征向量 V^t 结合,得到最终表征向量 E^f ;然后将 E^f 输入全连接层和softmax函数,得到预测结果;再根据目标损失函数loss,利用反向传播方法计算深度学习网络模型中各参数的梯度,并利用随机梯度下降方法更新各参数;

步骤B8:当深度学习网络模型N产生的损失值迭代变化小于所给定阈值、或者达到最大迭代次数,则终止深度学习网络模型N的训练过程。

3.根据权利要求2所述的一种融合情感挖掘的谣言检测方法,其特征在于:所述步骤B1包括以下步骤:

步骤B11:对训练集DT进行遍历,对其中源帖子的文本内容和评论内容进行分词处理并去除停用词之后,DT中的每个训练样本表示为 $dt = (st, rt, l)$;其中 st 为源帖子的文本内容, rt 为源帖子对应的评论内容, l 为源帖子对应的真实性标签, $l \in \{\text{一般事实, 谣言, 未经证实的传闻, 被辟谣的谣言}\}$;

源帖子的文本内容 st 表示为:

$$st = (w_1^{st} \dots w_i^{st} \dots w_n^{st}) \quad \text{公式一:}$$

其中, w_i^{st} 为文本内容 st 中的第 i 个词, $i = 1, 2, \dots, n$, n 为源帖子文本内容 st 的词语数量;

源帖子的评论内容 rt 表示为:

$$rt = (w_1^{rt} \dots w_j^{rt} \dots w_m^{rt}) \quad \text{公式二:}$$

其中, w_j^{rt} 为评论内容 rt 中的第 j 个词, $j = 1, 2, \dots, m$, m 为评论内容 rt 的词语数量;

步骤B12:对步骤B11得到文本内容 $st = (w_1^{st} \dots w_i^{st} \dots w_n^{st})$ 进行编码,得到文本内容 st 的初始表征向量 T^{st} ; T^{st} 表示为:

$$T^{st} = \{x_1^{st}, x_2^{st}, \dots, x_i^{st}, \dots, x_n^{st}\} \quad \text{公式三:}$$

其中,在预训练的词向量矩阵 $L \in \mathbb{R}^d \times |V|$ 中能查找得到 x_i^{st} , $x_i^{st} \in \mathbb{R}^d$ 为第 i 个词 w_i^{st} 所对应的词向量, d 表示词向量的维度, $|V|$ 是词典 V 中的词语数;

步骤B13:对步骤B11得到评论内容 $rt = (w_1^{rt} \dots w_j^{rt} \dots w_m^{rt})$ 进行编码,得到评论内容 rt 的初始表征向量 T^{rt} ; T^{rt} 表示为:

$$T^{rt} = \{x_1^{rt}, x_2^{rt}, \dots, x_j^{rt}, \dots, x_m^{rt}\} \quad \text{公式四:}$$

其中,在预训练的词向量矩阵 $L \in \mathbb{R}^d \times |V|$ 中能查找得到 x_j^{rt} , $x_j^{rt} \in \mathbb{R}^d$ 表示第 j 个词 w_j^{rt} 所对应的词向量, d 表示词向量的维度, $|V|$ 是词典 V 中的词语数;

步骤B14:对文本内容 st 进行句法依赖解析得到对应的句法依赖树DTD,以及 n 阶句法邻接矩阵 A^{st} ;句法依赖树DTD表示为,

$$DTD = \{(w_i^{st}, w_j^{st}) | i \in [1, n], j \in [1, n], i \neq j\} \quad \text{公式五:}$$

其中, (w_i^{st}, w_j^{st}) 表示文本内容词 w_i^{st} 和文本内容词 w_j^{st} 之间存在句法依赖关系。

4.根据权利要求3所述的一种融合情感挖掘的谣言检测方法,其特征在于:所述步骤B2包括以下步骤:

步骤B21:将句法依赖树DTD中的每个原始单词节点作为根节点,从知识图谱中拓展hop层来生成子节点,每层选取 u 个与上一层节点在知识图谱中有边连接的节点作为该层节点,即每个种子节点有 $q = \frac{u^{hop+1} - 1}{u - 1}$ 个拓展子节点,最终得到所有节点总数为 $z = n + n * q$ 的句法

知识子图SK,以及z阶邻接矩阵 A^{SK} ;句法知识子图SK表示为,

$$SK = \begin{cases} (w_i^{st}, w_j^{st}) | i \in [1, n], j \in [1, n], i \neq j \\ (w_i^{st}, w_j^{kg}) | i \in [1, n], j \in [1, u] \\ (w_i^{kg}, w_j^{kg}) | i \in [1, u^{tr-1}], j \in [1, u^{tr}], i \neq j, tr \in [2, hop], hop \geq 2 \end{cases}$$

公式六;

其中, (w_i^{st}, w_j^{kg}) 表示知识节点词 w_j^{kg} 是文本内容词 w_i^{st} 的拓展节点, (w_i^{kg}, w_j^{kg}) 表示知识节点词 w_j^{kg} 是知识节点词 w_i^{kg} 的知识拓展子节点, (w_i^{st}, w_j^{st}) 表示文本内容词 w_i^{st} 和文本内容词 w_j^{st} 之间有句法依赖关系, u是在知识图谱中选取的节点数, hop是所拓扑的层数;

步骤B22:利用知识图谱嵌入对句法知识子图SK的节点进行编码,得到其节点知识表示向量为 $H^{SK} = \{h_1^{sk}, h_2^{sk}, \dots, h_n^{sk}\}$, 令 $G^{SK,0} = \{g_{d,1}^0, g_{d,2}^0, \dots, g_{d,s}^0\} = H^{SK}$ 作为文本知识图卷积网络SKGCN的初始输入;在预训练的知识词向量矩阵 $L^{kg} \in \mathbb{R}^{d \times |V|}$ 中能查找得到 h_i^{sk} , $h_i^{sk} \in \mathbb{R}^d$ 为第i个词 w_i^{kg} 所对应的知识词向量,其中d表示知识词向量的维度, |V|是知识词嵌入V中的词语数。

5. 根据权利要求4所述的一种融合情感挖掘的谣言检测方法,其特征在于:所述步骤B3包括以下步骤;

步骤B31:将文本内容的初始表征向量 $T^{st} = \{x_1^{st}, x_2^{st}, \dots, x_n^{st}\}$ 依次分别输入第一个双向长短期记忆网络的前向层和反向层,以得到前向隐层的状态向量序列和反向隐层的状态向量序列,即 $\vec{H}^{st} = \{\vec{h}_1^{st}, \vec{h}_2^{st}, \dots, \vec{h}_i^{st}, \dots, \vec{h}_n^{st}\}$ 和 $\overleftarrow{H}^{st} = \{\overleftarrow{h}_1^{st}, \overleftarrow{h}_2^{st}, \dots, \overleftarrow{h}_i^{st}, \dots, \overleftarrow{h}_n^{st}\}$,

其中 $\vec{h}_i^{st} = f(x_i^{st}, \overrightarrow{h_{i-1}^{st}})$, $\overleftarrow{h}_i^{st} = f(x_i^{st}, \overleftarrow{h_{i+1}^{st}})$, $\vec{h}_i^{st}, \overleftarrow{h}_i^{st} \in \mathbb{R}^d$, $i = 1, 2, \dots, n$, 为激活函数;经过连接得到上下文增强的文本内容表征向量 $H^{st} = \{h_1^{st}, h_2^{st}, \dots, h_i^{st}, \dots, h_n^{st}\}$, 其中, $h_i^{st} = [\vec{h}_i^{st}, \overleftarrow{h}_i^{st}]$, $H^{st} \in \mathbb{R}^{n \times d}$, $h_i^{st} \in \mathbb{R}^{2d}$, $i = 1, 2, \dots, n$, “:”表示向量连接操作; H^{st} 即为 U^{st} ;

步骤B32:将文本内容st的初始表征向量 T^{st} 和评论内容rt的初始表征向量 T^{rt} 一起输入到一个多头交叉注意力机制中,得到基于文本内容的评论表征向量 P^{sr} ,其计算公式如下:

$$P^{sr} = \text{MultiHead}(T^{st}, T^{rt}, T^{rt}) \quad \text{公式七};$$

$$\text{MultiHead}(Q', K', V') = \text{Concat}(\text{head}^1, \text{head}^2, \dots, \text{head}_h) W_o \quad \text{公式八};$$

$$\text{head}_i = \text{Attention}(Q' W_i^Q, K' W_i^K, V' W_i^V) \quad \text{公式九};$$

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{l}}\right)V \quad \text{公式十};$$

其中, MultiHead表示多头注意力机制, Q' 、 K' 、 V' 表示多头注意力机制的输入向量, 文本内容的初始表征向量 T^{st} 作为矩阵 Q' , 相应的评论内容rt的初始表征向量 T^{rt} 作为 K' 和 V' ; head_i 为对 Q' 、 K' 、 V' 的第i个子向量使用注意力机制 $\text{Attention}(\cdot)$ 计算得到的输出向量, h是多头注意力机制的头数, W_o 为多头注意力机制的训练参数, $W_i^Q, W_i^K, W_i^V \in \mathbb{R}^{d \times d/h}$ 是线性

投影的权重矩阵, \sqrt{l} 是比例因子;

步骤B33: 将文本内容初始表征向量 T^{st} 输入到多头自注意力机制中, 得到文本内容增强表征向量 P^s , 其计算公式如下:

$$P^s = \text{MultiHead}(T^{st}, T^{st}, T^{st}) \quad \text{公式十一};$$

$$\text{MultiHead}(Q', K', V') = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h) W_1 \quad \text{公式十二};$$

$$\text{head}_i = \text{Attention}(Q' W_i^Q, K' W_i^K, V' W_i^V) \quad \text{公式十三};$$

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{l}}\right)V \quad \text{公式十四};$$

其中, MultiHead 表示多头注意力机制, Q', K', V' 表示多头注意力机制的输入向量, 文本内容的初始表征向量 T^{st} 作为矩阵 Q', K' 和 V' ; head_i 为对 Q', K', V' 的第 i 个子向量使用注意力机制 $\text{Attention}(\cdot)$ 计算得到的输出向量, h 是多头注意力机制的头数, W_1 为多头注意力机制的训练参数, $W_i^Q, W_i^K, W_i^V \in \mathbb{R}^{d \times d/k}$ 是线性投影的权重矩阵, \sqrt{l} 是比例因子;

步骤B34: 将基于文本内容的评论表征向量 P^{sr} 和文本内容增强表征向量 P^s 分别输入到池化层中进行平均池化操作, 得到平均池化评论内容句子表征向量 \hat{P}^{sr} 和平均池化文本内容增强表征向量 \hat{P}^s , 其计算公式如下:

$$\hat{P}^s = \text{MeanPool}(P^s) \quad \text{公式十五};$$

$$\hat{P}^{sr} = \text{MeanPool}(P^{sr}) \quad \text{公式十六};$$

其中, $\hat{P}^{sr}, \hat{P}^s \in \mathbb{R}^d$, MeanPool 是平均池化函数。

6. 根据权利要求5所述的一种融合情感挖掘的谣言检测方法, 其特征在于: 所述步骤B4包括以下步骤:

步骤B41: 将步骤B22得到的子图节点知识表征向量 $G^{SK,0}$ 输入文本知识图卷积网络SKGCN第一层图卷积网络, 使用邻接矩阵 A^{SK} 对每个子图节点的向量表示进行更新, 输出 $G^{SK,1}$, 并作为下一层图卷积网络的输入;

$$\text{其中, } G^{SK,1} \text{ 表示为: } G^{SK,1} = \{g_{d,1}^1, g_{d,2}^1, \dots, g_{d,n}^1\} \quad \text{公式十七};$$

其中, $g_{d,i}^1 \in \mathbb{R}^{2d}$ 是第一层图卷积网络中节点 i 的输出, $g_{d,i}^1$ 的计算公式如下:

$$h_{d,i}^1 = \sum_{j=1}^z A_{ij}^{SK} W^{SK} g_{d,j}^0 \quad \text{公式十八};$$

$$g_{d,i}^1 = \text{relu}\left(\frac{h_{d,i}^1}{d_i + 1} + b^{SK}\right) \quad \text{公式十九};$$

其中, $b^{SK} \in \mathbb{R}^{2d}$ 为偏置项; W^{SK}, b^{SK} 均为可学习的参数, $W^{SK} \in \mathbb{R}^{2d \times 2d}$ 为权重矩阵, relu 是激活函数; SKGCN中的节点 i 与评论内容中的第 i 个词 w_i^{sr} 相对应, 节点间的边代表词间存在知识连接关系, d_i 表示节点 i 的度, 为了防止节点 i 的度为0导致运算出错, 选择 $d_i + 1$ 作为除数;

步骤B42:对于文本内容图卷积网络SCGCN,将步骤B31得到的上下文增强的文本内容表征向量 U^{st} 输入SCGCN第一层图卷积网络,利用邻接矩阵 A^{sk} 对每个词语的向量表示进行更新,输出 $U^{st,1}$,

其中, $U^{st,1}$ 表示为: $U^{st,1} = \{u_{st,1}^1, u_{st,2}^1, \dots, u_{st,n}^1\}$ 公式二十;

其中, $u_{st,i}^1 \in \mathbb{R}^{2d}$ 是第一层图卷积网络中节点i的输出, $u_{st,i}^1$ 的计算公式如下:

$$u_{st,i}^1 = \text{relu} \left(\frac{\sum_j A_{ij}^{st} W_{st}^{st} u_{st,j}^0}{d_i + 1} + b_{st}^{st} \right) \quad \text{公式二十一};$$

其中, W_{st}^{st} 、 b_{st}^{st} 均为可学习的参数, $W_{st}^{st} \in \mathbb{R}^{2d \times 2d}$ 为权重矩阵, $b_{st}^{st} \in \mathbb{R}^{2d}$ 为偏置项;relu是激活函数;图卷积网络中的节点i与评论内容中的第i个词 w_i^{st} 相对应,图卷积网络中节点间的边代表评论内容中词语间的句法依赖关系, d_i 表示节点i的度,为了防止节点i的度为0导致运算出错,选择 d_i+1 作为除数;

对于知识引导机制,对SKGCN的第一层输出 $G^{sk,1}$ 舍弃除当前评论内容句子中的词以外的内容,得到关于文本内容的第一层知识表示 $G^{qc,1} = \{g_{qc,1}^1, g_{qc,2}^1, \dots, g_{qc,n}^1\}$,接下来使用交叉注意力机制将其与SCGCN第一层输出 $U^{st,1}$ 结合,以获得具有知识的评论内容句子表征 $G^{sd,1}$,并作为SCGCN下一层的输入,

其中, $G^{sd,1}$ 表示为: $G^{sd,1} = \{g_{sd,1}^1, g_{sd,2}^1, \dots, g_{sd,n}^1\}$ 公式二十二;

其中,SCGCN第一层图卷积网络中节点i的通过知识引导机制的输出为 $g_{sd,i}^1 \in \mathbb{R}^{2d}$, $g_{sd,i}^1$ 的计算公式如下:

$$g_{sd,i}^1 = \sum_{j=1}^n \alpha_j u_{st,j}^1 \quad \text{公式二十三};$$

$$\alpha_i = \frac{\exp(f(g_{qc}^1, u_{st,i}^1))}{\sum_{j=1}^n \exp(f(g_{qc}^1, u_{st,j}^1))} \quad \text{公式二十四};$$

$$f(g_{qc}^1, u_{st,i}^1) = \sum_{j=1}^n (u_{st,i}^1)^T g_{qc,j}^1 \quad \text{公式二十五};$$

其中, $(\cdot)^T$ 表示转置操作, α_i 是评论内容s中第i个词的有关知识的注意力权重;

步骤B43:SKGCN和SCGCN的下一层图卷积网络的输入为 $G^{sk,1}$ 和 $G^{sd,1}$,重复步骤B41、B42;

其中,对于SKGCN, $G^{sk,k} \in \mathbb{R}^{x \times 2d}$ 为第k层图卷积网络的输出,将其作为第k+1层图卷积网络的输入,迭代结束后得到图卷积表征向量 $G^{sk,K} \in \mathbb{R}^{x \times 2d}$;对于SCGCN, $U^{st,k} \in \mathbb{R}^{n \times 2d}$ 为第k层图卷积网络的输出,通过知识交互机制将 $U^{st,k}$ 与 $G^{sd,k}$ 作为第k+1层图卷积网络的输入,经过不断迭代,最后结束后得到图卷积表征向量 $V^{eks} \in \mathbb{R}^{n \times 2d}$;其中 $1 \leq k \leq K$,K为图卷积网络的层数。

7. 根据权利要求6所述的一种融合情感挖掘的谣言检测方法,其特征在于:所述步骤B5

包括以下步骤:

步骤B51:将步骤B31得到的上下文增强的文本内容表征向量 U^{st} 和步骤B43得到的 V^{sks} 输入注意力网络,通过注意力网络选择重要的知识信息,得到知识增强型句子级表征向量 E^{sd} ,计算公式如下:

$$E^{sd} = \sum_{i=1}^n \varepsilon_i u_i^{st} \quad \text{公式二十六;}$$

$$\varepsilon_i = \frac{\exp(f(v_{sks}, u_i^{st}))}{\sum_{j=1}^n \exp(f(v_{sks}, u_j^{st}))} \quad \text{公式二十七;}$$

$$f(v_{sks}, u_i^{st}) = \sum_{j=1}^n (u_i^{st})^T v_{sks,j} \quad \text{公式二十八;}$$

其中, $(\cdot)^T$ 表示转置操作, ε_i 是评论内容s中第i个词的注意力权重;

步骤B52:将步骤51得到的知识增强型句子级表征向量 E^{sd} 输入多头自注意力机制中,得到聚合词级信息的句子表征向量 E^{mt} ,

$$E^{mt} = \text{MultiHead}(E^{sd}, E^{sd}, E^{sd}) \quad \text{公式二十九;}$$

步骤B53:针对不规范句子对模型带来的噪声,将聚合词级信息的句子表征向量 E^{mt} 输入门控函数以过滤无关信息,得到向量 E^{sda} ;然后将其输入多层感知机(MLP),得出源帖子的情感表征向量 E^{sf} ;具体的计算过程如下:

$$E^{sda} = \text{relu}(W_1^{sda} E^{mt} + b_1^{sda}) \quad \text{公式三十;}$$

$$E^{sf} = \text{mlp}(W_2^{sda} E^{sda} + b_2^{sda}) \quad \text{公式三十一;}$$

其中, W_1^{sda} 、 W_2^{sda} 、 b_1^{sda} 和 b_2^{sda} 均为可学习的参数, $W_1^{sda} \in \mathbb{R}^{n \times 2d}$ 和 $W_2^{sda} \in \mathbb{R}^{n \times d}$ 为权重矩阵, $b_1^{sda} \in \mathbb{R}^{2d}$ 和 $b_2^{sda} \in \mathbb{R}^d$ 为偏置项。

8. 根据权利要求7所述的一种融合情感挖掘的谣言检测方法,其特征在于:所述步骤B6包括以下步骤:

步骤B61:将源帖子对应的所有平均池化评论内容句子表征向量 \bar{P}^{sr} 和平均池化文本内容增强表征向量 \bar{P}^s 一起输入到多头交叉注意力机制中,并通过平均池化得到评论内容的综合语义表征 C^{sr} ,其计算过程如下:

$$C' = \text{MultiHead}(\bar{P}^s, \bar{P}^{sr}, \bar{P}^{sr}) \quad \text{公式三十二;}$$

$$C^{sr} = \text{MeanPool}(C') \quad \text{公式三十三;}$$

其中, $C^{sr} \in \mathbb{R}^d$, MeanPool是平均池化函数;

步骤B62:将平均池化文本内容增强表征向量 \bar{P}^s 和评论内容的综合语义表征 C^{sr} 共同输入到融合门控机制中,得到源帖子的细粒度语义表征向量 V^t ,其计算过程如下:

$$\alpha = \sigma(w_1 \bar{P}^s + w_2 C^{sr} + b) \quad \text{公式三十四;}$$

$$V^t = \bar{P}^s \odot \alpha + C^{\sigma^*} \odot (1 - \alpha) \quad \text{公式三十五};$$

其中, $\sigma(\cdot) = \frac{1}{1 + \exp(\cdot)}$ 是sigmoid激活函数, $w_1, w_2 \in \mathbb{R}^{d \times 1}$ 和 $b \in \mathbb{R}$ 是融合门控机制中可学习参数, \odot 为点积操作。

9. 根据权利要求8所述的一种融合情感挖掘的谣言检测方法, 其特征在于: 所述步骤B7包括以下步骤:

步骤B71: 将步骤B53得到的源帖子情感表征向量 E^{sf} 和步骤B62得到的 V^t 进行连接, 得到最终表征向量 E^f , 计算公式如下:

$$E^f = \text{Concat}(E^{sf}, V^t) \quad \text{公式三十六};$$

其中, $E^f \in \mathbb{R}^{2d}$, Concat是向量连接操作。

步骤B72: 最终表征向量 E^f 输入到全连接层, 并使用softmax归一化, 计算文本内容对应属于各类别的概率, 计算公式如下:

$$y = W_3 E^f + b \quad \text{公式三十七};$$

$$p^c(y) = \text{softmax}(y) \quad \text{公式三十八};$$

其中, y 是全连接层的输出向量, $W_3 \in \mathbb{R}^{3 \times 2d}$ 为全连接层权重矩阵, $b \in \mathbb{R}^3$ 为全连接层的偏置项, $p^c(y)$ 是预测该文本内容对应类别为 c 的概率, $0 \leq p^c(y) \leq 1$, $c \in \{\text{一般事实, 谣言, 未经证实的传闻, 被辟谣的谣言}\}$;

步骤B73: 用交叉熵作为损失函数计算损失值, 通过梯度优化算法Adam进行学习率更新, 利用反向传播迭代更新模型参数, 以最小化损失函数来训练模型; 最小化损失函数loss的计算公式如下:

$$\text{loss} = - \sum_{c \in C} c \log(p^c(y)) + (1 - c) \log(1 - p^c(y)) + \lambda \|\theta\|_2^2 \quad \text{公式三十九};$$

其中, $\lambda \|\theta\|_2^2$ 是L2正则化项, λ 是学习率, θ 包含所有参数, c 为该文本内容对应的真实性标签。

10. 一种融合情感挖掘的谣言检测系统, 采用权利要求1至9任一项中所述的谣言检测方法, 其特征在于: 所述社交网络媒体为微博, 谣言检测系统包括以下模块:

数据收集模块: 用于提取微博中源帖子的文本内容和评论内容, 并对源帖子的真实性进行标注, 构建训练集;

预处理模块: 用于对训练集中的训练样本进行预处理, 包括分词处理、去除停用词;

编码模块: 用于在预训练的词向量矩阵中查找经过预处理的文本内容和评论内容中词的词向量, 得到文本内容的初始表征向量和评论内容的初始表征向量, 在预训练好的知识图谱词向量矩阵中查找句法知识子图中节点的词向量, 得到评论内容有关的句法知识子图初始表征向量;

网络训练模块: 用于将文本内容的初始表征向量和评论内容的初始表征向量和句法知识子图初始表征向量输入到深度学习网络中, 得到最终表征向量并以此训练深度学习网络, 利用该表征向量属于某一类别的概率以及训练集中的标注作为损失, 以最小化损失为目标来对整个深度学习网络进行训练, 得到基于多级注意力和知识图谱的深度学习网络模

型;

谣言检测模块:利用NLP工具,提取输入的源帖子文本内容和评论内容中的语义和情感信息,之后利用训练好的基于多级注意力和知识图谱的深度学习网络模型对输入的源帖子文本内容和评论内容进行分析处理,输出预测的源帖子真实性标签。

一种融合情感挖掘的谣言检测方法及系统

技术领域

[0001] 本发明涉及自然语言处理技术领域,尤其是一种融合情感挖掘的谣言检测方法及系统。

背景技术

[0002] 谣言检测(Rumor Detection),也称为虚假新闻检测,是自然语言处理(NLP)领域的一项重要任务。谣言检测可以看作是一个有监督学习的文本分类问题,一般可以分为是谣言和不是谣言两种类型。随着互联网技术的发展,诸如微博、推特等社交网络平台在大众生活中迅速流行起来。在社交网络平台上,人们不仅仅是信息的接收者也是内容的创造者。社交网络平台大大加速了人与人之间信息交流的速度和深度。社交网络平台能够及时全面地提供关于世界各地发生的事件,因此越来越多的人热衷于在社交网络平台上参与热点话题的讨论和交流。这种讨论和交流一方面促进了新闻的传播和扩散,使人们能够更加方便快捷地了解正在发生的事情。然而在这样的便利环境下,社交网络平台也降低了不实信息的传播成本。虚假谣言通常使用虚假的或者伪造的图像以及煽动性的语言,误导读者并迅速传播。虚假谣言的传播会对社会造成大规模的负面效应,引起社会动荡。

[0003] 近年来,随着深度学习技术的兴起,该技术也被谣言检测任务广泛应用。其中最常用的神经网络是卷积神经网络(CNN)和递归神经网络(RNN)。由于CNN在从文本中捕获语义信息方面表现良好,因此一些研究人员将它们应用于基于内容的谣言检测中。然而,该网络不能充分利用句子中的上下文信息,而这对于建模一个方面与其上下文之间的语义关系至关重要。因此,基于CNN的神经网络模型的性能在谣言检测任务中受到限制。针对这个问题,许多研究人员采用了RNN,尤其是长短期记忆(LSTM)和门控循环单元(GRU),来提取谣言的上下文语义信息。与CNN不同,RNN将一个句子视为一个词序列,按时间顺序取每个词,将隐藏层的输出作为下一个隐藏层的输入,不断学习序列数据中的上下文信息。Ma等人利用循环神经网络来捕获每个源帖子及其转发评论之间的语义变化,并根据语义变化进行预测。基于RNN的神经网络模型在谣言检测中明显优于基于CNN的神经网络模型。

[0004] 研究人员指出,给定帖子的谣言特征往往是由几个关键词决定的,而不是由上下文中的所有词决定的。而RNN无法精确估计不同上下文词对整体语义的贡献。相比之下,注意力机制可以通过计算每个上下文词对给定帖子语义的注意力权重并利用这个注意力权重来计算帖子的语义表示来捕捉每个上下文词的重要性。

[0005] 然而,这些神经网络模型大多都忽略了帖子中的情感信息,它代表了发布者对于帖子内容的情绪,这对于正确判断帖子的真实性标签尤为重要。最近有学者专注于发现假谣言和真谣言之间的独特情感特征。Ajao等人验证了新闻的真实性(真假)与情感词的使用之间存在关系,并设计一个情感特征(消极词和积极词的数量之比)来帮助检测假新闻。此外,Giachanou等人基于情感词典从新闻内容中提取情感特征用于谣言检测。然而,现有的相关研究忽略了在情感方面所需要的句法依赖信息和外部知识信息,使得情感信息没有被充分提取。

发明内容

[0006] 本发明提出一种融合情感挖掘的谣言检测方法及系统,可以提升对微博进行谣言检测的准确性。

[0007] 本发明采用以下技术方案。

[0008] 一种融合情感挖掘的谣言检测方法,所述方法包括以下步骤:

[0009] 步骤A:收集并提取社交媒体中源帖子的文本内容和评论内容,并人工标注源帖子的真实标签,形成训练数据集DT;

[0010] 步骤B:使用训练数据集DT,训练基于多级注意力和知识图谱的深度学习网络模型N,训练内容包括分析源帖子的真实性和预测源帖子的真实性标签;

[0011] 步骤C:将源帖子的文本内容和评论内容输入到训练好的深度学习网络模型N中,获得源帖子的真实性标签。

[0012] 所述步骤B包括以下步骤:

[0013] 步骤B1:对训练数据集DT中的每个训练样本进行编码,以得到文本内容的初始表征向量 T^{st} 、评论内容的初始表征向量 T^{rt} 以及句法邻接矩阵 A^{st} ;

[0014] 步骤B2:根据句法知识子图构造算法从知识图谱和句法依赖图中生成文本内容的对应的句法知识子图SK,并得到其邻接矩阵 A^{SK} ,然后对其节点进行编码,得到句法知识子图SK的节点知识表示向量 H^{SK} ;

[0015] 步骤B3:将步骤B1得到的文本内容初始表征向量 T^{st} 输入到双向长短期记忆网络Bi-LSTM中,得到上下文增强的文本内容表征向量 H^{st} ,令 $U^{st}=H^{st}$;然后,将表征向量 T^{sk} 和评论内容初始表征向量 T^{rt} 一起输入到多头交叉注意力机制中,得到基于文本内容的评论表征向量 P^{sr} ,同时将表征向量 T^{st} 输入到多头自注意力机制中,得到文本内容增强表征向量 P_s ;然后通过把基于文本内容的评论表征向量 P^{sr} 和文本内容增强表征向量 P_s 分别输入到池化层中进行平均池化操作,得到平均池化评论内容句子表征向量 \hat{P}^{sr} 和平均池化文本内容增强表征向量 \hat{P}_s ;

[0016] 步骤B4:将子图SK的节点知识表示向量 H^{SK} 和将步骤B3得到的表征向量 U^{st} 分别输入到两个具有K层的图卷积网络中,记为文本知识图卷积网络SKGCN和文本内容图卷积网络SCGCN,用于学习外部知识信息和提取句法信息;同时利用知识引导机制将文本内容图卷积网络SCGCN的每层节点与文本知识图卷积网络SKGCN进行知识引导,得到源帖子的图知识表征向量 V^{sks} ;

[0017] 步骤B5:通过使用交叉注意力机制将B4步骤得到的图知识表征向量 V^{sks} 和句子表征向量 U^{st} 融合,得到知识增强型句子级表征向量 E^{sd} ,以进一步提高模型提取信息的能力;然后通过多头自注意力机制将 E^{sd} 进一步加强,得到聚合词级信息的句子表征 E^{mt} ;再通过门控机制减少来自不规范句子的噪声,得到源帖子情感表征向量 E^{sf} ;

[0018] 步骤B6:将源帖子对应的平均池化评论内容句子表征向量 \hat{P}^{sr} 和平均池化文本内容增强表征向量 \hat{P}_s 一起输入到多头交叉注意力机制中,并通过平均池化得到评论内容的综合语义表征 C^{sr} ;然后将平均池化文本内容增强表征向量 \hat{P}_s 和评论内容的综合语义表征 C^{sr} 输入到融合门控机制中,得到源帖子的细粒度语义表征向量 V^t ;

[0019] 步骤B7:将步骤B5得到的情感表征向量 E^{sf} 与步骤B6得到的源帖子的细粒度语义表

征向量 V^t 结合,得到最终表征向量 E^f ;然后将 E^f 输入全连接层和softmax函数,得到预测结果;再根据目标损失函数loss,利用反向传播方法计算深度学习网络模型中各参数的梯度,并利用随机梯度下降方法更新各参数;

[0020] 步骤B8:当深度学习网络模型N产生的损失值迭代变化小于所给定阈值、或者达到最大迭代次数,则终止深度学习网络模型N的训练过程。

[0021] 所述步骤B1包括以下步骤:

[0022] 步骤B11:对训练集DT进行遍历,对其中源帖子的文本内容和评论内容进行分词处理并去除停用词之后,DT中的每个训练样本表示为 $dt = (st, rt, l)$;其中 st 为源帖子的文本内容, rt 为源帖子对应的评论内容, l 为源帖子对应的真实性标签, $l \in \{\text{一般事实, 谣言, 未经证实的传闻, 被辟谣的谣言}\}$;

[0023] 源帖子的文本内容 st 表示为:

[0024] $st = (w_1^{st} \cdots w_i^{st} \cdots w_n^{st})$ 公式一;

[0025] 其中, w_i^{st} 为文本内容 st 中的第 i 个词, $i = 1, 2, \cdots, n$, n 为源帖子文本内容 st 的词语数量;

[0026] 源帖子的评论内容 rt 表示为:

[0027] $rt = (w_1^{rt} \cdots w_j^{rt} \cdots w_m^{rt})$ 公式二;

[0028] 其中, w_j^{rt} 为评论内容 rt 中的第 j 个词, $j = 1, 2, \cdots, m$, m 为评论内容 rt 的词语数量;

[0029] 步骤B12:对步骤B11得到文本内容 $st = (w_1^{st} \cdots w_i^{st} \cdots w_n^{st})$ 进行编码,得到文本内容 st 的初始表征向量 T^{st} ; T^{st} 表示为:

[0030] $T^{st} = \{x_1^{st}, x_2^{st}, \cdots, x_i^{st}, \cdots, x_n^{st}\}$ 公式三;

[0031] 其中,在预训练的词向量矩阵 $L \in \mathbb{R}^d \times |V|$ 中能查找得到 x_i^{st} , $x_i^{st} \in \mathbb{R}^d$ 为第 i 个词 w_i^{st} 所对应的词向量, d 表示词向量的维度, $|V|$ 是词典 V 中的词语数;

[0032] 步骤B13:对步骤B11得到评论内容 $rt = (w_1^{rt} \cdots w_j^{rt} \cdots w_m^{rt})$ 进行编码,得到评论内容 rt 的初始表征向量 T^{rt} ; T^{rt} 表示为:

[0033] $T^{rt} = \{x_1^{rt}, x_2^{rt}, \cdots, x_j^{rt}, \cdots, x_m^{rt}\}$ 公式四;

[0034] 其中,在预训练的词向量矩阵 $L \in \mathbb{R}^d \times |V|$ 中能查找得到 x_j^{rt} , $x_j^{rt} \in \mathbb{R}^d$ 表示第 j 个词 w_j^{rt} 所对应的词向量, d 表示词向量的维度, $|V|$ 是词典 V 中的词语数;

[0035] 步骤B14:对文本内容 st 进行句法依赖解析得到对应的句法依赖树DTD,以及 n 阶句法邻接矩阵 A^{st} ;句法依赖树DTD表示为,

[0036] $DTD = \{(w_i^{st}, w_j^{st}) | i \in [1, n], j \in [1, n], i \neq j\}$ 公式五;

[0037] 其中, (w_i^{st}, w_j^{st}) 表示文本内容词 w_i^{st} 和文本内容词 w_j^{st} 之间存在句法依赖关系。

[0038] 所述步骤B2包括以下步骤:

[0039] 步骤B21:将句法依赖树DTD中的每个原始单词节点作为根节点,从知识图谱中拓展hop层来生成子节点,每层选取 u 个与上一层节点在知识图谱中有边连接的节点作为该层节点,即每个种子节点有 $q = \frac{u^{hop+1}-1}{u-1}$ 个拓展子节点,最终得到所有节点总数为 $z = n + n * q$ 的句法知识子图SK,以及 z 阶邻接矩阵 A^{SK} ;句法知识子图SK表示为,

$$[0040] \quad SK = \begin{cases} (w_i^{st}, w_j^{st}) | i \in [1, n], j \in [1, n], i \neq j \\ (w_i^{st}, w_j^{kg}) | i \in [1, n], j \in [1, u] \\ (w_i^{kg}, w_j^{kg}) | i \in [1, u^{tr}-1], j \in [1, u^{tr}], i \neq j, tr \in [2, hop], hop \geq 2 \end{cases}$$

公式六:

[0041] 其中, (w_i^{st}, w_j^{kg}) 表示知识节点词 w_j^{kg} 是文本内容词 w_i^{st} 的拓展节点, (w_i^{kg}, w_j^{kg}) 表示知识节点词 w_j^{kg} 是知识节点词 w_i^{kg} 的知识拓展子节点, (w_i^{st}, w_j^{st}) 表示文本内容词 w_i^{st} 和文本内容词 w_j^{st} 之间有句法依赖关系, u 是在知识图谱中选取的节点数, hop 是所拓扑的层数;

[0042] 步骤B22: 利用知识图谱嵌入对句法知识子图SK的节点进行编码, 得到其节点知识表示向量为 $H^{SK} = \{h_1^{sk}, h_2^{sk}, \dots, h_n^{sk}\}$, 令 $G^{SK,0} = \{g_{d,1}^0, g_{d,2}^0, \dots, g_{d,s}^0\} = H^{SK}$ 作为文本知识图卷积网络SKGCN的初始输入; 在预训练的知识词向量矩阵 $L^{kg} \in \mathbb{R}^{d \times |V|}$ 中能查找得到 h_i^{sk} , $h_i^{sk} \in \mathbb{R}^d$ 为第 i 个词 w_i^{kg} 所对应的知识词向量, 其中 d 表示知识词向量的维度, $|V|$ 是知识词嵌入 V 中的词语数。

[0043] 所述步骤B3包括以下步骤:

[0044] 步骤B31: 将文本内容的初始表征向量 $T^{st} = \{x_1^{st}, x_2^{st}, \dots, x_n^{st}\}$ 依次分别输入第一个双向长短期记忆网络的前向层和反向层, 以得到前向隐层的状态向量序列和反向隐层的状态向量序列, 即 $\vec{H}^{st} = \{\vec{h}_1^{st}, \vec{h}_2^{st}, \dots, \vec{h}_n^{st}\}$ 和 $\overleftarrow{H}^{st} = \{\overleftarrow{h}_1^{st}, \overleftarrow{h}_2^{st}, \dots, \overleftarrow{h}_n^{st}\}$, 其中 $\vec{h}_i^{st} = f(x_i^{st}, \vec{h}_{i-1}^{st})$, $\overleftarrow{h}_i^{st} = f(x_i^{st}, \overleftarrow{h}_{i+1}^{st})$, $\vec{h}_i^{st}, \overleftarrow{h}_i^{st} \in \mathbb{R}^d, i = 1, 2, \dots, n$, f 为激活函数; 经过连接得到上下文增强的文本内容表征向量 $H^{st} = \{h_1^{st}, h_2^{st}, \dots, h_i^{st}, \dots, h_n^{st}\}$, 其中, $h_i^{st} = [\vec{h}_i^{st}, \overleftarrow{h}_i^{st}]$, $H^{st} \in \mathbb{R}^{n \times d}, h_i^{st} \in \mathbb{R}^{2d}, i = 1, 2, \dots, n$, “:” 表示向量连接操作; H^{st} 即为 U^{st} ;

[0045] 步骤B32: 将文本内容 st 的初始表征向量 T^{st} 和评论内容 rt 的初始表征向量 T^{rt} 一起输入到一个多头交叉注意力机制中, 得到基于文本内容的评论表征向量 P^{sr} , 其计算公式如下:

$$[0046] \quad P^{sr} = \text{MultiHead}(T^{st}, T^{rt}, T^{rt}) \quad \text{公式七};$$

$$[0047] \quad \text{MultiHead}(Q', K', V') = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h) W_o \quad \text{公式八};$$

$$[0048] \quad \text{head}_i = \text{Attention}(Q' W_i^Q, K' W_i^K, V' W_i^V) \quad \text{公式九};$$

$$[0049] \quad \text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad \text{公式十};$$

[0050] 其中, MultiHead 表示多头注意力机制, Q', K', V' 表示多头注意力机制的输入向量, 文本内容的初始表征向量 T^{st} 作为矩阵 Q' , 相应的评论内容 rt 的初始表征向量 T^{rt} 作为 K' 和 V' ; head_i 为对 Q', K', V' 的第 i 个子向量使用注意力机制 $\text{Attention}(\cdot)$ 计算得到的输出向量, h 是多头注意力机制的头数, W_o 为多头注意力机制的训练参数, $W_i^Q, W_i^K, W_i^V \in \mathbb{R}^{d \times d/h}$ 是线性投影的权重矩阵, \sqrt{d} 是比例因子;

[0051] 步骤B33: 将文本内容初始表征向量 T^{st} 输入到多头自注意力机制中, 得到文本内容增强表征向量 P^s , 其计算公式如下:

$$[0052] \quad P^s = \text{MultiHead}(T^{st}, T^{st}, T^{st}) \quad \text{公式十一};$$

[0053] $\text{MultiHead}(Q', K', V') = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h) W_1$ 公式十二;

[0054] $\text{head}_i = \text{Attention}(Q' W_i^Q, K' W_i^K, V' W_i^V)$ 公式十三;

[0055] $\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V$ 公式十四;

[0056] 其中, MultiHead表示多头注意力机制, Q', K', V' 表示多头注意力机制的输入向量, 文本内容的初始表征向量 T^{st} 作为矩阵 Q', K' 和 V' ; head_i 为对 Q', K', V' 的第 i 个子向量使用注意力机制 $\text{Attention}(\cdot)$ 计算得到的输出向量, h 是多头注意力机制的头数, W_1 为多头注意力机制的训练参数, $W_i^Q, W_i^K, W_i^V \in \mathbb{R}^{d \times d/h}$ 是线性投影的权重矩阵, \sqrt{d} 是比例因子;

[0057] 步骤B34: 将基于文本内容的评论表征向量 P^{sr} 和文本内容增强表征向量 P^{s} 分别输入到池化层中进行平均池化操作, 得到平均池化评论内容句子表征向量 \bar{P}^{sr} 和平均池化文本内容增强表征向量 \bar{P}^{s} , 其计算公式如下:

[0058] $\bar{P}^{\text{s}} = \text{MeanPool}(P^{\text{s}})$ 公式十五;

[0059] $\bar{P}^{\text{sr}} = \text{MeanPool}(P^{\text{sr}})$ 公式十六;

[0060] 其中, $\bar{P}^{\text{sr}}, \bar{P}^{\text{s}} \in \mathbb{R}^d$, MeanPool是平均池化函数。

[0061] 所述步骤B4包括以下步骤:

[0062] 步骤B41: 将步骤B22得到的子图节点知识表征向量 $G^{\text{SK}, 0}$ 输入文本知识图卷积网络SKGCN第一层图卷积网络, 使用邻接矩阵 A^{SK} 对每个子图节点的向量表示进行更新, 输出 $G^{\text{SK}, 1}$, 并作为下一层图卷积网络的输入;

[0063] 其中, $G^{\text{SK}, 1}$ 表示为: $G^{\text{SK}, 1} = \{g_{d,1}^1, g_{d,2}^1, \dots, g_{d,n}^1\}$ 公式十七;

[0064] 其中, $g_{d,i}^1 \in \mathbb{R}^{2d}$ 是第一层图卷积网络中节点 i 的输出, $g_{d,i}^1$ 的计算公式如下:

[0065]
$$h_{d,i}^1 = \sum_{j=1}^S A_{ij}^{\text{SK}} W^{\text{SK}} g_{d,j}^0$$
 公式十八;

[0066]
$$g_{d,i}^1 = \text{relu}\left(\frac{h_{d,i}^1}{d_i + 1} + b^{\text{SK}}\right)$$
 公式十九;

[0067] 其中, $b^{\text{SK}} \in \mathbb{R}^{2d}$ 为偏置项; $W^{\text{SK}}, b^{\text{SK}}$ 均为可学习的参数, $W^{\text{SK}} \in \mathbb{R}^{2d \times 2d}$ 为权重矩阵, relu是激活函数; SKGCN中的节点 i 与评论内容中的第 i 个词 w_i^{ss} 相对应, 节点间的边代表词语间存在知识连接关系, d_i 表示节点 i 的度, 为了防止节点 i 的度为0导致运算出错, 选择 $d_i + 1$ 作为除数;

[0068] 步骤B42: 对于文本内容图卷积网络SCGCN, 将步骤B31得到的上下文增强的文本内容表征向量 U^{st} 输入SCGCN第一层图卷积网络, 利用邻接矩阵 A^{SK} 对每个词语的向量表示进行更新, 输出 $U^{\text{st}, 1}$,

[0069] 其中, $U^{\text{st}, 1}$ 表示为: $U^{\text{st}, 1} = \{u_{\text{st},1}^1, u_{\text{st},2}^1, \dots, u_{\text{st},n}^1\}$ 公式二十;

[0070] 其中, $u_{\text{st},i}^1 \in \mathbb{R}^{3d}$ 是第一层图卷积网络中节点 i 的输出, $u_{\text{st},i}^1$ 的计算公式如下:

$$[0071] \quad u_{st,i}^1 = \text{relu} \left(\frac{\sum_j^n A_{ij}^{st} W_{st}^{st} u_{st,j}^0}{d_i + 1} + b_3^{st} \right) \quad \text{公式二十一};$$

[0072] 其中, W_{st}^{st} 、 b_3^{st} 均为可学习的参数, $W_{st}^{st} \in \mathbb{R}^{2d \times 2d}$ 为权重矩阵, $b_3^{st} \in \mathbb{R}^{2d}$ 为偏置项; relu是激活函数;图卷积网络中的节点i与评论内容中的第i个词 w_i^{st} 相对应,图卷积网络中节点间的边代表评论内容中词语间的句法依赖关系, d_i 表示节点i的度,为了防止节点i的度为0导致运算出错,选择 d_i+1 作为除数;

[0073] 对于知识引导机制,对SKGCN的第一层输出 $G^{SK,1}$ 舍弃除当前评论内容句子中的词以外的内容,得到关于文本内容的第一层知识表示 $G^{qc,1} = \{g_{qc,1}^1, g_{qc,2}^1, \dots, g_{qc,n}^1\}$,接下来使用交叉注意力机制将其与SCGCN第一层输出 $U^{st,1}$ 结合,以获得具有知识的评论内容句子表征 $G^{SD,1}$,并作为SCGCN下一层的输入,

$$[0074] \quad \text{其中, } G^{SD,1} \text{表示为: } G^{SD,1} = \{g_{sd,1}^1, g_{sd,2}^1, \dots, g_{sd,n}^1\} \quad \text{公式二十二};$$

[0075] 其中, SCGCN第一层图卷积网络中节点i的通过知识引导机制的输出为 $g_{sd,i}^1 \in \mathbb{R}^{2d}$, $g_{sd,i}^1$ 的计算公式如下:

$$[0076] \quad g_{sd,i}^1 = \sum_{i=1}^n \alpha_i u_{st,i}^1 \quad \text{公式二十三};$$

$$[0077] \quad \alpha_i = \frac{\exp(f(g_{qc}^1, u_{st,i}^1))}{\sum_{j=1}^n \exp(f(g_{qc}^1, u_{st,j}^1))} \quad \text{公式二十四};$$

$$[0078] \quad f(g_{qc}^1, u_{st,i}^1) = \sum_{j=1}^n (u_{st,i}^1)^T g_{qc,j}^1 \quad \text{公式二十五};$$

[0079] 其中, $(\cdot)^T$ 表示转置操作, α_i 是评论内容s中第i个词的有关知识的注意力权重;

[0080] 步骤B43: SKGCN和SCGCN的下一层图卷积网络的输入为 $G^{SK,1}$ 和 $G^{SD,1}$,重复步骤B41、B42;

[0081] 其中,对于SKGCN, $G^{SK,k} \in \mathbb{R}^{n \times 2d}$ 为第k层图卷积网络的输出,将其作为第k+1层图卷积网络的输入,迭代结束后得到图卷积表征向量 $G^{SK,K} \in \mathbb{R}^{n \times 2d}$;对于SCGCN, $U^{st,k} \in \mathbb{R}^{n \times 2d}$ 为第k层图卷积网络的输出,通过知识交互机制将 $U^{st,k}$ 与 $G^{SD,k}$ 作为第k+1层图卷积网络的输入,经过不断迭代,最后结束后得到图卷积表征向量 $V^{sk,K} \in \mathbb{R}^{n \times 2d}$,其中 $1 \leq k \leq K$, K为图卷积网络的层数。

[0082] 所述步骤B5包括以下步骤;

[0083] 步骤B51:将步骤B31得到的上下文增强的文本内容表征向量 U^{st} 和步骤B43得到的 V^{sk} 输入注意力网络,通过注意力网络选择重要的知识信息,得到知识增强型句子级表征向量 E^{sd} ,计算公式如下:

$$[0084] \quad E^{sd} = \sum_{i=1}^n \varepsilon_i u_i^{st} \quad \text{公式二十六};$$

$$[0085] \quad \varepsilon_i = \frac{\exp(f(v_{sks}, u_i^{st}))}{\sum_{j=1}^n \exp(f(v_{sks}, u_j^{st}))} \quad \text{公式二十七;}$$

$$[0086] \quad f(v_{sks}, u_i^{st}) = \sum_{j=1}^n (u_i^{st})^T v_{sks,j} \quad \text{公式二十八;}$$

[0087] 其中, $(\cdot)^T$ 表示转置操作, ε_i 是评论内容 s 中第 i 个词的注意力权重;

[0088] 步骤B52: 将步骤51得到的知识增强型句子级表征向量 E^{sd} 输入多头自注意力机制中, 得到聚合词级信息的句子表征向量 E^{mt} ,

$$[0089] \quad E^{mt} = \text{MultiHead}(E^{sd}, E^{sd}, E^{sd}) \quad \text{公式二十九;}$$

[0090] 步骤B53: 针对不规范句子对模型带来的噪声, 将聚合词级信息的句子表征向量 E^{mt} 输入门控函数以过滤无关信息, 得到向量 E^{sda} ; 然后将其输入多层感知机 (MLP), 得出源帖子的情感表征向量 E^{sf} ; 具体的计算过程如下:

$$[0091] \quad E^{sda} = \text{relu}(W_1^{sda} E^{mt} + b_1^{sda}) \quad \text{公式三十;}$$

$$[0092] \quad E^{sf} = \text{mlp}(W_2^{sda} E^{sda} + b_2^{sda}) \quad \text{公式三十一;}$$

[0093] 其中, W_1^{sda} 、 W_2^{sda} 、 b_1^{sda} 和 b_2^{sda} 均为可学习的参数, $W_1^{sda} \in \mathbb{R}^{n \times 2d}$ 和 $W_2^{sda} \in \mathbb{R}^{n \times d}$ 为权重矩阵, $b_1^{sda} \in \mathbb{R}^{2d}$ 和 $b_2^{sda} \in \mathbb{R}^d$ 为偏置项。

[0094] 所述步骤B6包括以下步骤:

[0095] 步骤B61: 将源帖子对应的所有平均池化评论内容句子表征向量 \tilde{P}^{sr} 和平均池化文本内容增强表征向量 \tilde{P}^s 一起输入到多头交叉注意力机制中, 并通过平均池化得到评论内容的综合语义表征 C^{sr} , 其计算过程如下:

$$[0096] \quad C' = \text{MultiHead}(\tilde{P}^s, \tilde{P}^{sr}, \tilde{P}^{sr}) \quad \text{公式三十二;}$$

$$[0097] \quad C^{sr} = \text{MeanPool}(C') \quad \text{公式三十三;}$$

[0098] 其中, $C^{sr} \in \mathbb{R}^d$, MeanPool 是平均池化函数;

[0099] 步骤B62: 将平均池化文本内容增强表征向量 \tilde{P}^s 和评论内容的综合语义表征 C^{sr} 共同输入到融合门控机制中, 得到源帖子的细粒度语义表征向量 V^t , 其计算过程如下:

$$[0100] \quad \alpha = \sigma(w_1 \tilde{P}^s + w_2 C^{sr} + b) \quad \text{公式三十四;}$$

$$[0101] \quad V^t = \tilde{P}^s \odot \alpha + C^{sr} \odot (1 - \alpha) \quad \text{公式三十五;}$$

[0102] 其中, $\sigma(\cdot) = \frac{1}{1 + \exp(-\cdot)}$ 是sigmoid激活函数, $w_1, w_2 \in \mathbb{R}^{d \times 1}$ 和 $b \in \mathbb{R}$ 是融合门控机制中可学习参数, \odot 为点积操作。

[0103] 所述步骤B7包括以下步骤:

[0104] 步骤B71: 将步骤B53得到的源帖子情感表征向量 E^{sf} 和步骤B62得到的 V^t 进行连接, 得到最终表征向量 E^f , 计算公式如下:

$$[0105] \quad E^f = \text{Concat}(E^{sf}, V^t) \quad \text{公式三十六;}$$

[0106] 其中, $E^f \in \mathbb{R}^{2d}$, Concat 是向量连接操作。

[0107] 步骤B72: 最终表征向量 E^f 输入到全连接层, 并使用softmax归一化, 计算文本内容

对应属于各类别的概率,计算公式如下:

[0108] $y = W_3 E^f + b$ 公式三十七;

[0109] $p^c(y) = \text{softmax}(y)$ 公式三十八;

[0110] 其中, y 是全连接层的输出向量, $W_3 \in \mathbb{R}^{3 \times 2d}$ 为全连接层权重矩阵, $b \in \mathbb{R}^3$ 为全连接层的偏置项, $p^c(y)$ 是预测该文本内容对应类别为 c 的概率, $0 \leq p^c(y) \leq 1$, $c \in \{\text{一般事实, 谣言, 未经证实的传闻, 被辟谣的谣言}\}$;

[0111] 步骤B73:用交叉熵作为损失函数计算损失值,通过梯度优化算法Adam进行学习率更新,利用反向传播迭代更新模型参数,以最小化损失函数来训练模型;最小化损失函数loss的计算公式如下:

$$[0112] \quad \text{loss} = - \sum_{c \in \mathcal{C}} c \log(p^c(y)) + (1 - c) \log(1 - p^c(y)) + \lambda \|\theta\|_2^2$$

公式三十九;

[0113] 其中, $\lambda \|\theta\|_2^2$ 是L2正则化项, λ 是学习率, θ 包含所有参数, c 为该文本内容对应的真实性标签。

[0114] 一种融合情感挖掘的谣言检测系统,采用以上所述的谣言检测方法,所述社交网络媒体为微博,谣言检测系统包括以下模块:

[0115] 数据收集模块:用于提取微博中源帖子的文本内容和评论内容,并对源帖子的真实性进行标注,构建训练集;

[0116] 预处理模块:用于对训练集中的训练样本进行预处理,包括分词处理、去除停用词;

[0117] 编码模块:用于在预训练的词向量矩阵中查找经过预处理的文本内容和评论内容中词的词向量,得到文本内容的初始表征向量和评论内容的初始表征向量,在预训练好的知识图谱词向量矩阵中查找句法知识子图中节点的词向量,得到评论内容有关的句法知识子图初始表征向量;

[0118] 网络训练模块:用于将文本内容的初始表征向量和评论内容的初始表征向量和句法知识子图初始表征向量输入到深度学习网络中,得到最终表征向量并以此训练深度学习网络,利用该表征向量属于某一类别的概率以及训练集中的标注作为损失,以最小化损失为目标来对整个深度学习网络进行训练,得到基于多级注意力和知识图谱的深度学习网络模型;

[0119] 谣言检测模块:利用NLP工具,提取输入的源帖子文本内容和评论内容中的语义和情感信息,之后利用训练好的基于多级注意力和知识图谱的深度学习网络模型对输入的源帖子文本内容和评论内容进行分析处理,输出预测的源帖子真实性标签。

[0120] 与现有技术相比,本发明具有以下有益效果:本发明利用知识图谱和子图生成策略得到对应评论句子的句法知识子图,之后对评论内容和文本内容分别编码,同时通过两个图卷积网络以及知识引导机制学习评论内容中句法依存关系和外部知识,并利用门控机制过滤句子噪声来增强评论句子的表示。本发明还利用多级注意力机制,学习文本内容与评论内容之间的细粒度语义信息。相比于以往,本发明能够利用细粒度的语义信息和丰富的情感信息来增强谣言的特征表示,从而进一步提高谣言检测的精度并增强其鲁棒性。

附图说明

[0121] 下面结合附图和具体实施方式对本发明进一步详细的说明：

[0122] 附图1是本发明实施例的方法实现流程示意图；

[0123] 附图2是本发明实施例中模型架构示意图；

[0124] 附图3是本发明实施例的系统结构示意图。

具体实施方式

[0125] 下面结合附图及实施例对本发明做进一步说明。

[0126] 应该指出，以下详细说明都是示例性的，旨在对本申请提供进一步的说明。除非另有指明，本文使用的所有技术和科学术语具有与本申请所属技术领域的普通技术人员通常理解的相同含义。

[0127] 需要注意的是，这里所使用的术语仅是为了描述具体实施方式，而非意图限制根据本申请的示例性实施方式。如在这里所使用的，除非上下文另外明确指出，否则单数形式也意图包括复数形式，此外，还应当理解的是，当在本说明书中使用术语“包含”和/或“包括”时，其指明存在特征、步骤、操作、器件、组件和/或它们的组合。

[0128] 如图所示，一种融合情感挖掘的谣言检测方法，所述方法包括以下步骤：

[0129] 步骤A：收集并提取社交网络媒体中源帖子的文本内容和评论内容，并人工标注源帖子的真实标签，形成训练数据集DT；

[0130] 步骤B：使用训练数据集DT，训练基于多级注意力和知识图谱的深度学习网络模型N，训练内容包括分析源帖子的真实性和预测源帖子的真实性标签；

[0131] 步骤C：将源帖子的文本内容和评论内容输入到训练好的深度学习网络模型N中，获得源帖子的真实性标签。

[0132] 所述步骤B包括以下步骤：

[0133] 步骤B1：对训练数据集DT中的每个训练样本进行编码，以得到文本内容的初始表征向量 T^{st} 、评论内容的初始表征向量 T^{rt} 以及句法邻接矩阵 A^{st} ；

[0134] 步骤B2：根据句法知识子图构造算法从知识图谱和句法依赖图中生成文本内容的对应的句法知识子图SK，并得到其邻接矩阵 A^{SK} ，然后对其节点进行编码，得到句法知识子图SK的节点知识表示向量 H^{SK} ；

[0135] 步骤B3：将步骤B1得到的文本内容初始表征向量 T^{st} 输入到双向长短期记忆网络Bi-LSTM中，得到上下文增强的文本内容表征向量 H^{st} ，令 $U^{st}=H^{st}$ ；然后，将表征向量 T^{st} 和评论内容初始表征向量 T^{rt} 一起输入到多头交叉注意力机制中，得到基于文本内容的评论表征向量 P^{sr} ，同时将表征向量 T^{st} 输入到多头自注意力机制中，得到文本内容增强表征向量 P^s ；然后通过把基于文本内容的评论表征向量 P^{sr} 和文本内容增强表征向量 P^s 分别输入到池化层中进行平均池化操作，得到平均池化评论内容句子表征向量 \tilde{P}^{sr} 和平均池化文本内容增强表征向量 \tilde{P}^s ；

[0136] 步骤B4：将子图SK的节点知识表示向量 H^{SK} 和将步骤B3得到的表征向量 U^{st} 分别输入到两个具有K层的图卷积网络中，记为文本知识图卷积网络SKGCN和文本内容图卷积网络SCGCN，用于学习外部知识信息和提取句法信息；同时利用知识引导机制将文本内容图卷积网络SCGCN的每层节点与文本知识图卷积网络SKGCN进行知识引导，得到源帖子的图知识表

征向量 V^{sk_s} ;

[0137] 步骤B5:通过使用交叉注意力机制将B4步骤得到的图知识表征向量 V^{sk_s} 和句子表征向量 U^{st} 融合,得到知识增强型句子级表征向量 E^{sd} ,以进一步提高模型提取信息的能力;然后通过多头自注意力机制将 E^{sd} 进一步加强,得到聚合词级信息的句子表征 E^{st} ;再通过门控机制减少来自不规范句子的噪声,得到源帖子情感表征向量 E^{sf} ;

[0138] 步骤B6:将源帖子对应的平均池化评论内容句子表征向量 \bar{P}^{sr} 和平均池化文本内容增强表征向量 \bar{P}^{st} 一起输入到多头交叉注意力机制中,并通过平均池化得到评论内容的综合语义表征 C^{sr} ;然后将平均池化文本内容增强表征向量 \bar{P}^{st} 和评论内容的综合语义表征 C^{sr} 输入到融合门控机制中,得到源帖子的细粒度语义表征向量 V^t ;

[0139] 步骤B7:将步骤B5得到的情感表征向量 E^{sf} 与步骤B6得到的源帖子的细粒度语义表征向量 V^t 结合,得到最终表征向量 E^f ;然后将 E^f 输入全连接层和softmax函数,得到预测结果;再根据目标损失函数loss,利用反向传播方法计算深度学习网络模型中各参数的梯度,并利用随机梯度下降方法更新各参数;

[0140] 步骤B8:当深度学习网络模型N产生的损失值迭代变化小于所给定阈值、或者达到最大迭代次数,则终止深度学习网络模型N的训练过程。

[0141] 所述步骤B1包括以下步骤:

[0142] 步骤B11:对训练集DT进行遍历,对其中源帖子的文本内容和评论内容进行分词处理并去除停用词之后,DT中的每个训练样本表示为 $dt = (st, rt, l)$;其中 st 为源帖子的文本内容, rt 为源帖子对应的评论内容, l 为源帖子对应的真实性标签, $l \in \{\text{一般事实, 谣言, 未经证实的传闻, 被辟谣的谣言}\}$;

[0143] 源帖子的文本内容 st 表示为:

[0144] $st = (w_1^{st} \cdots w_j^{st} \cdots w_n^{st})$ 公式一;

[0145] 其中, w_i^{st} 为文本内容 st 中的第 i 个词, $i = 1, 2, \dots, n$, n 为源帖子文本内容 st 的词语数量;

[0146] 源帖子的评论内容 rt 表示为:

[0147] $rt = (w_1^{rt} \cdots w_j^{rt} \cdots w_m^{rt})$ 公式二;

[0148] 其中, w_j^{rt} 为评论内容 rt 中的第 j 个词, $i = 1, 2, \dots, m$, m 为评论内容 rt 的词语数量;

[0149] 步骤B12:对步骤B11得到文本内容 $st = (w_1^{st} \cdots w_j^{st} \cdots w_n^{st})$ 进行编码,得到文本内容 st 的初始表征向量 T^{st} ; T^{st} 表示为:

[0150] $T^{st} = \{x_1^{st}, x_2^{st}, \dots, x_j^{st}, \dots, x_n^{st}\}$ 公式三;

[0151] 其中,在预训练的词向量矩阵 $L \in \mathbb{R}^d \times |V|$ 中能查找得到 x_i^{st} , $x_i^{st} \in \mathbb{R}^d$ 为第 i 个词 w_i^{st} 所对应的词向量, d 表示词向量的维度, $|V|$ 是词典 V 中的词语数;

[0152] 步骤B13:对步骤B11得到评论内容 $rt = (w_1^{rt} \cdots w_j^{rt} \cdots w_m^{rt})$ 进行编码,得到评论内容 rt 的初始表征向量 T^{rt} ; T^{rt} 表示为:

[0153] $T^{rt} = \{x_1^{rt}, x_2^{rt}, \dots, x_j^{rt}, \dots, x_m^{rt}\}$ 公式四;

[0154] 其中,在预训练的词向量矩阵 $L \in \mathbb{R}^d \times |V|$ 中能查找得到 x_j^{rt} , $x_j^{rt} \in \mathbb{R}^d$ 表示第 j 个词 w_j^{rt} 所对应的词向量, d 表示词向量的维度, $|V|$ 是词典 V 中的词语数;

[0155] 步骤B14:对文本内容st进行句法依赖解析得到对应的句法依赖树DTD,以及n阶句法邻接矩阵 A^{st} ;句法依赖树DTD表示为,

$$[0156] \quad DTD = \{(w_i^{st}, w_j^{st}) | i \in [1, n], j \in [1, n], i \neq j\} \quad \text{公式五;}$$

[0157] 其中, (w_i^{st}, w_j^{st}) 表示文本内容词 w_i^{st} 和文本内容词 w_j^{st} 之间存在句法依赖关系。

[0158] 所述步骤B2包括以下步骤:

[0159] 步骤B21:将句法依赖树DTD中的每个原始单词节点作为根节点,从知识图谱中拓展hop层来生成子节点,每层选取u个与上一层节点在知识图谱中有边连接的节点作为该层节点,即每个种子节点有 $q = \frac{u^{hop+1}-1}{u-1}$ 个拓展子节点,最终得到所有节点总数为 $z = n+n*q$ 的句法知识子图SK,以及z阶邻接矩阵 A^{SK} ;句法知识子图SK表示为,

$$[0160] \quad SK = \begin{cases} (w_i^{st}, w_j^{st}) | i \in [1, n], j \in [1, n], i \neq j \\ (w_i^{st}, w_j^{kg}) | i \in [1, n], j \in [1, u] \\ (w_i^{kg}, w_j^{kg}) | i \in [1, u^{tr-1}], j \in [1, u^{tr}], i \neq j, tr \in [2, hop], hop \geq 2 \end{cases}$$

公式六;

[0161] 其中, (w_i^{st}, w_j^{kg}) 表示知识节点词 w_j^{kg} 是文本内容词 w_i^{st} 的拓展节点, (w_i^{kg}, w_j^{kg}) 表示知识节点词 w_j^{kg} 是知识节点词 w_i^{kg} 的知识拓展子节点, (w_i^{st}, w_j^{st}) 表示文本内容词 w_i^{st} 和文本内容词 w_j^{st} 之间有句法依赖关系,u是在知识图谱中选取的节点数,hop是所拓扑的层数;

[0162] 步骤B22:利用知识图谱嵌入对句法知识子图SK的节点进行编码,得到其节点知识表示向量为 $H^{SK} = \{h_1^{sk}, h_2^{sk}, \dots, h_n^{sk}\}$,令 $G^{SK,0} = \{g_{d,1}^0, g_{d,2}^0, \dots, g_{d,s}^0\} = H^{SK}$ 作为文本知识图卷积网络SKGCN的初始输入;在预训练的知识词向量矩阵 $L^{kg} \in \mathbb{R}^{d \times |V|}$ 中能查找得到 $h_i^{sk}, h_i^{sk} \in \mathbb{R}^d$ 为第i个词 w_i^{kg} 所对应的知识词向量,其中d表示知识词向量的维度,|V|是知识词嵌入V中的词语数。

[0163] 所述步骤B3包括以下步骤:

[0164] 步骤B31:将文本内容的初始表征向量 $T^{st} = \{x_1^{st}, x_2^{st}, \dots, x_n^{st}\}$ 依次分别输入第一个双向长短期记忆网络的前向层和反向层,以得到前向隐层的状态向量序列和反向隐层的状态向量序列,即 $\vec{H}^{st} = \{\vec{h}_1^{st}, \vec{h}_2^{st}, \dots, \vec{h}_i^{st}, \dots, \vec{h}_n^{st}\}$ 和 $\overleftarrow{H}^{st} = \{\overleftarrow{h}_1^{st}, \overleftarrow{h}_2^{st}, \dots, \overleftarrow{h}_i^{st}, \dots, \overleftarrow{h}_n^{st}\}$,

[0165] 其中 $\vec{h}_i^{st} = f(x_i^{st}, \overleftarrow{h}_{i-1}^{st})$, $\overleftarrow{h}_i^{st} = f(x_i^{st}, \overleftarrow{h}_{i+1}^{st})$, $\vec{h}_i^{st}, \overleftarrow{h}_i^{st} \in \mathbb{R}^d, i = 1, 2, \dots, n$, f为激活函数;经过连接得到上下文增强的文本内容表征向量 $H^{st} = \{h_1^{st}, h_2^{st}, \dots, h_i^{st}, \dots, h_n^{st}\}$,其中, $h_i^{st} = [\vec{h}_i^{st}, \overleftarrow{h}_i^{st}]$, $H^{st} \in \mathbb{R}^{n \times d}, h_i^{st} \in \mathbb{R}^{2d}, i = 1, 2, \dots, n$,“:”表示向量连接操作; H^{st} 即为 U^{st} ;

[0166] 步骤B32:将文本内容st的初始表征向量 T^{st} 和评论内容rt的初始表征向量 T^{rt} 一起输入到一个多头交叉注意力机制中,得到基于文本内容的评论表征向量 P^{sr} ,其计算公式如下:

$$[0167] \quad P^{sr} = \text{MultiHead}(T^{st}, T^{rt}, T^{rt}) \quad \text{公式七;}$$

$$[0168] \quad \text{MultiHead}(Q', K', V') = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h) W_o \quad \text{公式八;}$$

[0169] $\text{head}_i = \text{Attention}(Q'W_i^Q, K'W_i^K, V'W_i^V)$ 公式九;

[0170] $\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V$ 公式十;

[0171] 其中, MultiHead表示多头注意力机制, Q', K', V' 表示多头注意力机制的输入向量, 文本内容的初始表征向量 T^{st} 作为矩阵 Q' , 相应的评论内容 r_t 的初始表征向量 T^{rt} 作为 K' 和 V' ; head_i 为对 Q', K', V' 的第 i 个子向量使用注意力机制 $\text{Attention}(\cdot)$ 计算得到的输出向量, h 是多头注意力机制的头数, W_0 为多头注意力机制的训练参数, $W_i^Q, W_i^K, W_i^V \in \mathbb{R}^{d \times d/h}$ 是线性投影的权重矩阵, \sqrt{d} 是比例因子;

[0172] 步骤B33: 将文本内容初始表征向量 T^{st} 输入到多头自注意力机制中, 得到文本内容增强表征向量 P^s , 其计算公式如下:

[0173] $P^s = \text{MultiHead}(T^{\text{st}}, T^{\text{st}}, T^{\text{st}})$ 公式十一;

[0174] $\text{MultiHead}(Q', K', V') = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W_1$ 公式十二;

[0175] $\text{head}_i = \text{Attention}(Q'W_i^Q, K'W_i^K, V'W_i^V)$ 公式十三;

[0176] $\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V$ 公式十四;

[0177] 其中, MultiHead表示多头注意力机制, Q', K', V' 表示多头注意力机制的输入向量, 文本内容的初始表征向量 T^{st} 作为矩阵 Q', K' 和 V' ; head_i 为对 Q', K', V' 的第 i 个子向量使用注意力机制 $\text{Attention}(\cdot)$ 计算得到的输出向量, h 是多头注意力机制的头数, W_1 为多头注意力机制的训练参数, $W_i^Q, W_i^K, W_i^V \in \mathbb{R}^{d \times d/h}$ 是线性投影的权重矩阵, \sqrt{d} 是比例因子;

[0178] 步骤B34: 将基于文本内容的评论表征向量 P^{sr} 和文本内容增强表征向量 P^s 分别输入到池化层中进行平均池化操作, 得到平均池化评论内容句子表征向量 \hat{P}^{sr} 和平均池化文本内容增强表征向量 \hat{P}^s , 其计算公式如下:

[0179] $\hat{P}^s = \text{MeanPool}(P^s)$ 公式十五;

[0180] $\hat{P}^{\text{sr}} = \text{MeanPool}(P^{\text{sr}})$ 公式十六;

[0181] 其中, $\hat{P}^{\text{sr}}, \hat{P}^s \in \mathbb{R}^d$, MeanPool是平均池化函数。

[0182] 所述步骤B4包括以下步骤;

[0183] 步骤B41: 将步骤B22得到的子图节点知识表征向量 $G^{\text{SK}, 0}$ 输入文本知识图卷积网络SKGCN第一层图卷积网络, 使用邻接矩阵 A^{SK} 对每个子图节点的向量表示进行更新, 输出 $G^{\text{SK}, 1}$, 并作为下一层图卷积网络的输入;

[0184] 其中, $G^{\text{SK}, 1}$ 表示为: $G^{\text{SK}, 1} = \{g_{d,1}^1, g_{d,2}^1, \dots, g_{d,n}^1\}$ 公式十七;

[0185] 其中, $g_{d,i}^1 \in \mathbb{R}^{2d}$ 是第一层图卷积网络中节点 i 的输出, $g_{d,i}^1$ 的计算公式如下:

[0186] $h_{d,i}^1 = \sum_{j=1}^n A_{ij}^{\text{SK}} W^{\text{SK}} g_{d,j}^0$ 公式十八;

[0187] $g_{d,i}^1 = \text{relu}\left(\frac{h_{d,i}^1}{d_i + 1} + b^{\text{SK}}\right)$ 公式十九;

[0188] 其中, $b^{SK} \in \mathbb{R}^{2d}$ 为偏置项; W^{SK} 、 b^{SK} 均为可学习的参数, $W^{SK} \in \mathbb{R}^{2d \times 2d}$ 为权重矩阵, relu 是激活函数; SKGCN 中的节点 i 与评论内容中的第 i 个词 w_i^{st} 相对应, 节点间的边代表词语间存在知识连接关系, d_i 表示节点 i 的度, 为了防止节点 i 的度为 0 导致运算出错, 选择 $d_i + 1$ 作为除数;

[0189] 步骤 B42: 对于文本内容图卷积网络 SCGCN, 将步骤 B31 得到的上下文增强的文本内容表征向量 U^{st} 输入 SCGCN 第一层图卷积网络, 利用邻接矩阵 A^{SK} 对每个词语的向量表示进行更新, 输出 $U^{st,1}$,

[0190] 其中, $U^{st,1}$ 表示为: $U^{st,1} = \{u_{st,1}^1, u_{st,2}^1, \dots, u_{st,n}^1\}$ 公式二十;

[0191] 其中, $u_{st,i}^1 \in \mathbb{R}^{2d}$ 是第一层图卷积网络中节点 i 的输出, $u_{st,i}^1$ 的计算公式如下:

$$[0192] \quad u_{st,i}^1 = \text{relu} \left(\frac{\sum_j A_{ij}^{st} W^{st} u_{st,j}^0}{d_i + 1} + b_3^{st} \right) \quad \text{公式二十一};$$

[0193] 其中, W^{st} 、 b_3^{st} 均为可学习的参数, $W^{st} \in \mathbb{R}^{2d \times 2d}$ 为权重矩阵, $b_3^{st} \in \mathbb{R}^{2d}$ 为偏置项; relu 是激活函数; 图卷积网络中的节点 i 与评论内容中的第 i 个词 w_i^{st} 相对应, 图卷积网络中节点间的边代表评论内容中词语间的句法依赖关系, d_i 表示节点 i 的度, 为了防止节点 i 的度为 0 导致运算出错, 选择 $d_i + 1$ 作为除数;

[0194] 对于知识引导机制, 对 SKGCN 的第一层输出 $G^{SK,1}$ 舍弃除当前评论内容句子中的词以外的内容, 得到关于文本内容的第一层知识表示 $G^{QC,1} = \{g_{qc,1}^1, g_{qc,2}^1, \dots, g_{qc,n}^1\}$, 接下来使用交叉注意力机制将其与 SCGCN 第一层输出 $U^{st,1}$ 结合, 以获得具有知识的评论内容句子表征 $G^{SD,1}$, 并作为 SCGCN 下一层的输入,

[0195] 其中, $G^{SD,1}$ 表示为: $G^{SD,1} = \{g_{sd,1}^1, g_{sd,2}^1, \dots, g_{sd,n}^1\}$ 公式二十二;

[0196] 其中, SCGCN 第一层图卷积网络中节点 i 的通过知识引导机制的输出为 $g_{sd,i}^1 \in \mathbb{R}^{2d}$, $g_{sd,i}^1$ 的计算公式如下:

$$[0197] \quad g_{sd,i}^1 = \sum_{j=1}^n \alpha_j u_{st,j}^1 \quad \text{公式二十三};$$

$$[0198] \quad \alpha_i = \frac{\exp(f(g_{qc}^1, u_{st,i}^1))}{\sum_{j=1}^n \exp(f(g_{qc}^1, u_{st,j}^1))} \quad \text{公式二十四};$$

$$[0199] \quad f(g_{qc}^1, u_{st,i}^1) = \sum_{j=1}^n (u_{st,i}^1)^T g_{qc,j}^1 \quad \text{公式二十五};$$

[0200] 其中, $(\cdot)^T$ 表示转置操作, α_i 是评论内容 s 中第 i 个词的有关知识的注意力权重;

[0201] 步骤 B43: SKGCN 和 SCGCN 的下一层图卷积网络的输入为 $G^{SK,1}$ 和 $G^{SD,1}$, 重复步骤 B41、B42;

[0202] 其中, 对于 SKGCN, $G^{SK,k} \in \mathbb{R}^{s \times 2d}$ 为第 k 层图卷积网络的输出, 将其作为第 $k+1$ 层图卷积网络的输入, 迭代结束后得到图卷积表征向量 $G^{SK,K} \in \mathbb{R}^{s \times 2d}$; 对于 SCGCN, $U^{st,k} \in \mathbb{R}^{n \times 2d}$ 为第 k 层图卷积网络的输出, 通过知识交互机制将 $U^{st,k}$ 与 $G^{SD,k}$ 作为第 $k+1$ 层图卷积网络的输

入,经过不断迭代,最后结束后得到图卷积表征向量 $V^{sks} \in \mathbb{R}^{n \times 2d}$;其中 $1 \leq k \leq K$, K 为图卷积网络的层数。

[0203] 所述步骤B5包括以下步骤;

[0204] 步骤B51:将步骤B31得到的上下文增强的文本内容表征向量 U^{st} 和步骤B43得到的 V^{sks} 输入注意力网络,通过注意力网络选择重要的知识信息,得到知识增强型句子级表征向量 E^{sd} ,计算公式如下:

$$[0205] \quad E^{sd} = \sum_{i=1}^n \varepsilon_i U_i^{st} \quad \text{公式二十六;}$$

$$[0206] \quad \varepsilon_i = \frac{\exp(f(v_{sks}, U_i^{st}))}{\sum_{j=1}^n \exp(f(v_{sks}, U_j^{st}))} \quad \text{公式二十七;}$$

$$[0207] \quad f(v_{sks}, U_i^{st}) = \sum_{j=1}^n (U_i^{st})^T v_{sks,j} \quad \text{公式二十八;}$$

[0208] 其中, $(\cdot)^T$ 表示转置操作, ε_i 是评论内容 s 中第 i 个词的注意力权重;

[0209] 步骤B52:将步骤51得到的知识增强型句子级表征向量 E^{sd} 输入多头自注意力机制中,得到聚合词级信息的句子表征向量 E^{mt} ,

$$[0210] \quad E^{mt} = \text{MultiHead}(E^{sd}, E^{sd}, E^{sd}) \quad \text{公式二十九;}$$

[0211] 步骤B53:针对不规范句子对模型带来的噪声,将聚合词级信息的句子表征向量 E^{mt} 输入门控函数以过滤无关信息,得到向量 E^{sda} ;然后将其输入多层感知机(MLP),得出源帖子的情感表征向量 E^{sf} ;具体的计算过程如下:

$$[0212] \quad E^{sda} = \text{relu}(W_1^{sda} E^{mt} + b_1^{sda}) \quad \text{公式三十;}$$

$$[0213] \quad E^{sf} = \text{mlp}(W_2^{sda} E^{sda} + b_2^{sda}) \quad \text{公式三十一;}$$

[0214] 其中, W_1^{sda} 、 W_2^{sda} 、 b_1^{sda} 和 b_2^{sda} 均为可学习的参数, $W_1^{sda} \in \mathbb{R}^{n \times 2d}$ 和 $W_2^{sda} \in \mathbb{R}^{n \times d}$ 为权重矩阵, $b_1^{sda} \in \mathbb{R}^{2d}$ 和 $b_2^{sda} \in \mathbb{R}^d$ 为偏置项。

[0215] 所述步骤B6包括以下步骤;

[0216] 步骤B61:将源帖子对应的所有平均池化评论内容句子表征向量 \bar{P}^{sr} 和平均池化文本内容增强表征向量 \bar{P}^s 一起输入到多头交叉注意力机制中,并通过平均池化得到评论内容的综合语义表征 C^{sr} ,其计算过程如下:

$$[0217] \quad C' = \text{MultiHead}(\bar{P}^s, \bar{P}^{sr}, \bar{P}^{sr}) \quad \text{公式三十二;}$$

$$[0218] \quad C^{sr} = \text{MeanPool}(C') \quad \text{公式三十三;}$$

[0219] 其中, $C^{sr} \in \mathbb{R}^d$,MeanPool是平均池化函数;

[0220] 步骤B62:将平均池化文本内容增强表征向量 \bar{P}^s 和评论内容的综合语义表征 C^{sr} 共同输入到融合门控机制中,得到源帖子的细粒度语义表征向量 V^t ,其计算过程如下:

$$[0221] \quad \alpha = \sigma(w_1 \bar{P}^s + w_2 C^{sr} + b) \quad \text{公式三十四;}$$

$$[0222] \quad V^t = \bar{P}^s \odot \alpha + C^{sr} \odot (1 - \alpha) \quad \text{公式三十五;}$$

[0223] 其中, $\sigma(\cdot) = \frac{1}{1+\exp(-\cdot)}$ 是sigmoid激活函数, $w_1, w_2 \in \mathbb{R}^{d \times 1}$ 和 $b \in \mathbb{R}$ 是融合门控机制中可学习参数, \odot 为点积操作。

[0224] 所述步骤B7包括以下步骤:

[0225] 步骤B71: 将步骤B53得到的源帖子情感表征向量 E^{sf} 和步骤B62得到的 V^t 进行连接, 得到最终表征向量 E^f , 计算公式如下:

[0226] $E^f = \text{Concat}(E^{sf}, V^t)$ 公式三十六;

[0227] 其中, $E^f \in \mathbb{R}^{2d}$, Concat 是向量连接操作。

[0228] 步骤B72: 最终表征向量 E^f 输入到全连接层, 并使用softmax归一化, 计算文本内容对应属于各类别的概率, 计算公式如下:

[0229] $y = W_3 E^f + b$ 公式三十七;

[0230] $p^c(y) = \text{softmax}(y)$ 公式三十八;

[0231] 其中, y 是全连接层的输出向量, $W_3 \in \mathbb{R}^{3 \times 2d}$ 为全连接层权重矩阵, $b \in \mathbb{R}^3$ 为全连接层的偏置项, $p^c(y)$ 是预测该文本内容对应类别为 c 的概率, $0 \leq p^c(y) \leq 1$, $c \in \{\text{一般事实, 谣言, 未经证实的传闻, 被辟谣的谣言}\}$;

[0232] 步骤B73: 用交叉熵作为损失函数计算损失值, 通过梯度优化算法Adam进行学习率更新, 利用反向传播迭代更新模型参数, 以最小化损失函数来训练模型; 最小化损失函数 loss 的计算公式如下:

$$[0233] \quad \text{loss} = - \sum_{c \in C} c \log(p^c(y)) + (1-c) \log(1-p^c(y)) + \lambda \|\theta\|_2^2$$

公式三十九;

[0234] 其中, $\lambda \|\theta\|_2^2$ 是L2正则化项, λ 是学习率, θ 包含所有参数, c 为该文本内容对应的真实性标签。

[0235] 一种融合情感挖掘的谣言检测系统, 采用以上所述的谣言检测方法, 所述社交网络媒体为微博, 谣言检测系统包括以下模块:

[0236] 数据收集模块: 用于提取微博中源帖子的文本内容和评论内容, 并对源帖子的真实性进行标注, 构建训练集;

[0237] 预处理模块: 用于对训练集中的训练样本进行预处理, 包括分词处理、去除停用词;

[0238] 编码模块: 用于在预训练的词向量矩阵中查找经过预处理的文本内容和评论内容中词的词向量, 得到文本内容的初始表征向量和评论内容的初始表征向量, 在预训练好的知识图谱词向量矩阵中查找句法知识子图中节点的词向量, 得到评论内容有关的句法知识子图初始表征向量;

[0239] 网络训练模块: 用于将文本内容的初始表征向量和评论内容的初始表征向量和句法知识子图初始表征向量输入到深度学习网络中, 得到最终表征向量并以此训练深度学习网络, 利用该表征向量属于某一类别的概率以及训练集中的标注作为损失, 以最小化损失为目标来对整个深度学习网络进行训练, 得到基于多级注意力和知识图谱的深度学习网络模型;

[0240] 谣言检测模块: 利用NLP工具, 提取输入的源帖子文本内容和评论内容中的语义和情感信息, 之后利用训练好的基于多级注意力和知识图谱的深度学习网络模型对输入的源

帖子文本内容和评论内容进行分析处理,输出预测的源帖子真实性标签。

[0241] 以上所述,仅是本发明的较佳实施例而已,并非是对本发明作其它形式的限制,任何熟悉本专业的技术人员可能利用上述揭示的技术内容加以变更或改型为等同变化的等效实施例。但是凡是未脱离本发明技术方案内容,依据本发明的技术实质对以上实施例所作的任何简单修改、等同变化与改型,仍属于本发明技术方案的保护范围。

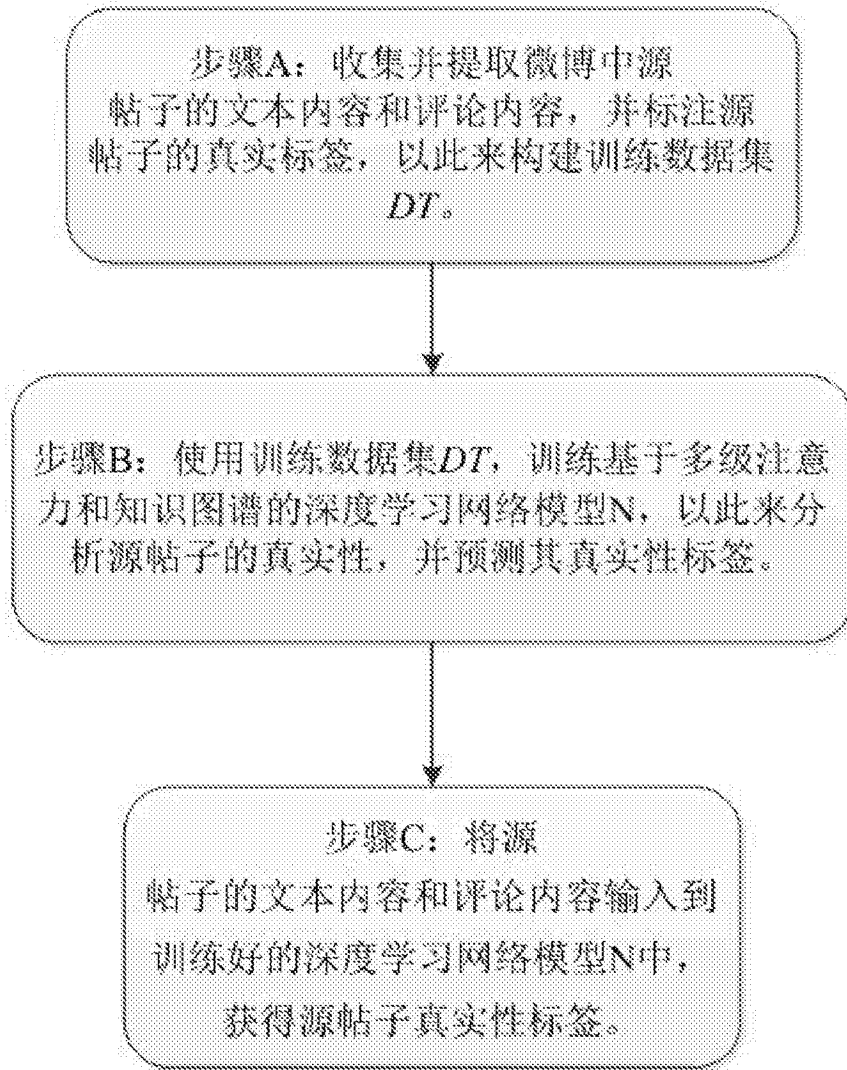


图1

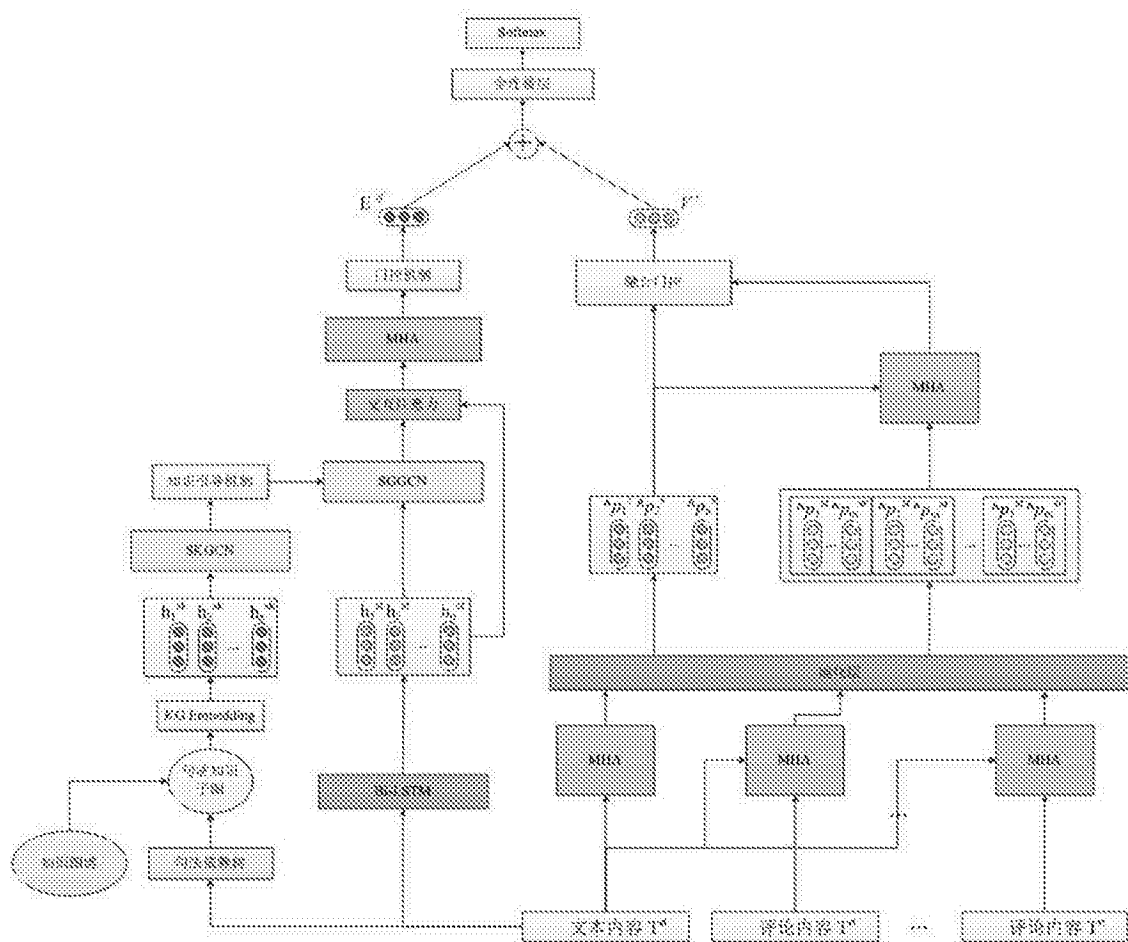


图2

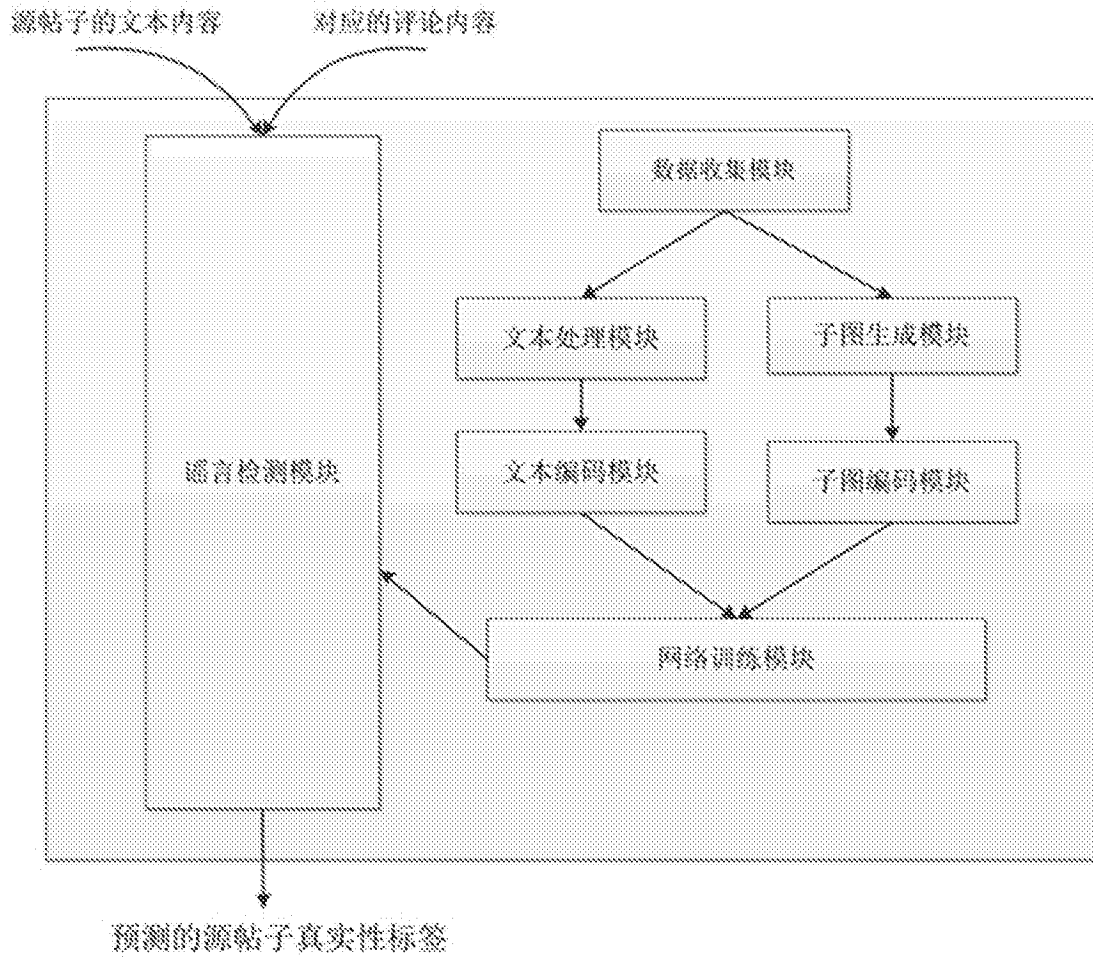


图3