(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 110188194 B
(45)授权公告日 2020.12.01

(21)申请号 201910354263.5
(22)申请日 2019.04.26
(65)同一申请的已公布的文献号
申请公布号 CN 110188194 A
(43)申请公布日 2019.08.30
(73)专利权人 哈尔滨工业大学(深圳)
地址 518055 广东省深圳市南山区桃源街道深圳大学城哈尔滨工业大学校区
(72)发明人 廖清 韩浩 丁烨 漆舒汉 蒋琳 王轩
(74)专利代理机构 广州三环专利商标代理有限公司 44202
代理人 麦小婵 郝传鑫
(51)Int.Cl.
G06F 16/35(2019.01)
G06Q 50/00(2012.01)
审查员 狄晓斐

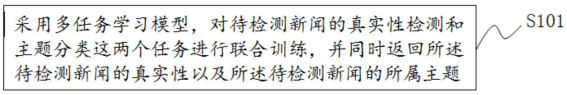
权利要求书2页 说明书9页 附图2页

(54)发明名称

一种基于多任务学习模型的假新闻检测方法及系统

(57)摘要

本发明公开了一种基于多任务学习模型的假新闻检测方法及系统。在一个实施例中：采用多任务学习模型，对待检测新闻的真实性检测和主题分类这两个任务进行联合训练，并同时返回所述待检测新闻的真实性以及所述待检测新闻的所属主题。本发明实施例的教导，能够同时检测出新闻的真实性以及新闻的所属主题，且提高了假新闻检测和主题分类的准确性。



CN 110188194 B

1.一种基于多任务学习模型的假新闻检测方法,其特征在于,包括:

采用多任务学习模型,对待检测新闻的真实性检测和主题分类这两个任务进行联合训练,并同时返回所述待检测新闻的真实性以及所述待检测新闻的所属主题;其中,所述多任务学习模型包括嵌入层、表示层和多任务层;

所述嵌入层,将所述待检测新闻的文本内容和上下文信息作为原始数据嵌入到一个低维空间,以此获得文本内容和上下文信息的数学向量表示,并输入所述表示层;所述文本内容指新闻文本,所述上下文信息是指新闻的元数据;

所述表示层,对文本嵌入词向量使用GRU(Gated Recurrent Unit,门控循环单元)层和CNN(Convolutional Neural Network,卷积神经网络)模型进行文本特征提取,对上下文嵌入向量使用另一个CNN模型和上下文特征向量层进行上下文特征提取,并将所提取出来的文本特征和上下文特征组合成完整特征;

所述多任务层,根据所述完整特征,对待检测新闻的真实性检测和主题分类进行联合训练,以同时完成真实性检测任务和主题分类任务,并输出真实性标签和主题标签。

2.如权利要求1所述的假新闻检测方法,其特征在于,所述嵌入层,将所述待检测新闻的陈述中的每个词汇嵌入到低维空间,每个词汇获得一个固定长度的词向量,将所述待检测新闻的作者姓名、作者所属政党、地点嵌入到一个低维空间以此得到一个固定长度的低维向量,即数学输入的表示。

3.如权利要求2所述的假新闻检测方法,其特征在于,所述表示层,使用GRU层作为特征提取器对文本特征进行特征提取,通过对GRU层的输出进行最大池化得到一组文本特征的表示,同时使用CNN来捕捉GRU层输出的特征,以此获得另一组文本特征;拼接来自所述嵌入层获得的上下文特征向量,以此获得一组上下文特征的表示,同时使用CNN来捕捉所述上下文信息之间的关联关系,以此得到另一组上下文特征的表示;将所提取出来的文本特征和上下文特征进行组合以得到待检测新闻的完整特征的表示。

4.一种基于多任务学习模型的假新闻检测系统,其特征在于,包括:

检测单元,用于采用多任务学习模型,对待检测新闻的真实性检测和主题分类这两个任务进行联合训练,并同时返回所述待检测新闻的真实性以及所述待检测新闻的所属主题;其中,所述多任务学习模型包括嵌入层、表示层和多任务层;

所述嵌入层,将所述待检测新闻的文本内容和上下文信息作为原始数据嵌入到一个低维空间,以此获得文本内容和上下文信息的数学向量表示,并输入所述表示层;所述文本内容指新闻文本,所述上下文信息是指新闻的元数据;

所述表示层,对文本嵌入词向量使用GRU(Gated Recurrent Unit,门控循环单元)层和CNN(Convolutional Neural Network,卷积神经网络)模型进行文本特征提取,对上下文嵌入向量使用另一个CNN模型和上下文特征向量层进行上下文特征提取,并将所提取出来的文本特征和上下文特征组合成完整特征;

所述多任务层,用于根据所述完整特征,对待检测新闻的真实性检测和主题分类进行联合训练,以同时完成真实性检测任务和主题分类任务,并输出真实性标签和主题标签。

5.如权利要求4所述的假新闻检测系统,其特征在于,所述嵌入层,具体用于将所述待检测新闻的陈述中的每个词汇嵌入到低维空间,每个词汇获得一个固定长度的词向量,将所述待检测新闻的作者姓名、作者所属政党、地点嵌入到一个低维空间以此得到一个固定

长度的低维向量,即数学输入的表示。

6.如权利要求5所述的假新闻检测系统,其特征在于,所述表示层,具体用于使用GRU层作为特征提取器对文本特征进行特征提取,通过对GRU层的输出进行最大池化得到一组文本特征的表示,同时使用CNN来捕捉GRU层输出的特征,以此获得另一组文本特征;拼接来自所述嵌入层获得的上下文特征向量,以此获得一组上下文特征的表示,同时使用CNN来捕捉所述上下文信息之间的关联关系,以此得到另一组上下文特征的表示;将所提取出来的文本特征和上下文特征进行组合以得到待检测新闻的完整特征的表示。

一种基于多任务学习模型的假新闻检测方法及系统

技术领域

[0001] 本发明涉及新闻检测技术领域，尤其涉及一种基于多任务学习模型的假新闻检测方法及系统。

背景技术

[0002] 社交媒体对新闻传播是一把双刃剑。一方面，它成本低廉，容易获取，而且通过快速的传播，它允许用户消费和分享新闻。另一方面，它可以产生有害的假新闻，即一些有意含有错误信息的低质量新闻。假新闻的快速传播对社会和个人有着巨大的潜在危害。举个例子，在2016年美国总统大选中，最受欢迎的假新闻在Facebook上的传播范围要比主流媒体的新闻还要大。

[0003] 因此，社交媒体中的假新闻检测引起了研究者和政治家们的注意。但是社交媒体的假新闻检测有着独特的特性而且呈现出新的挑战。

[0004] 首先，假新闻的内容是被有意制造用来误导读者，这使得我们很难根据其新闻的内容来判断是否是假新闻。因此，我们需要类似用户在社交媒体上的活跃度等辅助信息来将假新闻从真新闻中区分出来。

[0005] 第二，用户基于假新闻的社交活动产生了大量不完整、非结构化和充满噪声的数据，这使得利用这些数据变得非常困难。

[0006] 在对现有技术的研究过程中，本发明的发明人发现，现有的假新闻检测方法主要集中在从新闻文本内容中提取词汇特征来对其真实性进行预判，当新闻文本内容较短时，其准确性难以令人满意。

[0007] 为了全面并准确地检测假新闻，许多事实核查机构和社交媒体平台投入了大量的人力物力，推动了假新闻检测检测的相关算法改进和技术发展，但也没有能够很好地解决上述问题。

发明内容

[0008] 本发明所要解决的技术问题在于，提供一种基于多任务学习模型的假新闻检测方法及系统，能够同时检测出新闻的真实性以及新闻的所属主题，且提高了假新闻检测和主题分类的准确性。

[0009] 为解决上述问题，本发明实施例提供一种基于多任务学习模型的假新闻检测方法，包括：

[0010] 采用多任务学习模型，对待检测新闻的真实性检测和主题分类这两个任务进行联合训练，并同时返回所述待检测新闻的真实性以及所述待检测新闻的所属主题。

[0011] 优选地，所述多任务学习模型包括嵌入层、表示层和多任务层；

[0012] 所述嵌入层，将所述待检测新闻的文本内容和上下文信息作为原始数据嵌入到一个低维空间，以此获得文本内容和上下文信息的数学向量表示，并输入所述表示层；

[0013] 所述表示层，对文本嵌入词向量使用GRU(Gated Recurrent Unit，门控循环单元)

层和CNN(Convolutional Neural Network,卷积神经网络)模型进行文本特征提取,对上下文嵌入向量使用另一个CNN模型和上下文特征向量层进行上下文特征提取,并将所提取出来的文本特征和上下文特征组合成完整特征;

[0014] 所述多任务层,根据所述完整特征,对待检测新闻的真实性检测和主题分类进行联合训练,以同时完成真实性检测任务和主题分类任务,并输出真实性标签和主题标签。

[0015] 优选地,所述嵌入层,将所述待检测新闻的陈述中的每个词汇嵌入到低维空间,每个词汇获得一个固定长度的词向量,将所述待检测新闻的作者姓名、作者所属政党、地点等信息嵌入到一个低维空间以此得到一个固定长度的低维向量,即数学输入的表示。

[0016] 优选地,所述表示层,使用GRU层作为特征提取器对文本特征进行特征提取,通过对GRU层的输出进行最大池化得到一组文本特征的表示,同时使用CNN来捕捉GRU层输出的特征,以此获得另一组文本特征;拼接来自所述嵌入层获得的上下文特征向量,以此获得一组上下文特征的表示,同时使用CNN来捕捉所述上下文信息之间的关联关系,以此得到另一组上下文特征的表示;将所提取出来的文本特征和上下文特征进行组合以得到待检测新闻的完整特征的表示。

[0017] 本发明实施例还提供一种基于多任务学习模型的假新闻检测系统,包括:

[0018] 检测单元,用于采用多任务学习模型,对待检测新闻的真实性检测和主题分类这两个任务进行联合训练,并同时返回所述待检测新闻的真实性以及所述待检测新闻的所属主题。

[0019] 优选地,所述多任务学习模型包括嵌入层、表示层和多任务层;

[0020] 所述嵌入层,将所述待检测新闻的文本内容和上下文信息作为原始数据嵌入到一个低维空间,以此获得文本内容和上下文信息的数学向量表示,并输入所述表示层;

[0021] 所述表示层,对文本嵌入词向量使用GRU(Gated Recurrent Unit,门控循环单元)层和CNN(Convolutional Neural Network,卷积神经网络)模型进行文本特征提取,对上下文嵌入向量使用另一个CNN模型和上下文特征向量层进行上下文特征提取,并将所提取出来的文本特征和上下文特征组合成完整特征;

[0022] 所述多任务层,用于根据所述完整特征,对待检测新闻的真实性检测和主题分类进行联合训练,以同时完成真实性检测任务和主题分类任务,并输出真实性标签和主题标签。

[0023] 优选地,所述嵌入层,具体用于将所述待检测新闻的陈述中的每个词汇嵌入到低维空间,每个词汇获得一个固定长度的词向量,将所述待检测新闻的作者姓名、作者所属政党、地点等信息嵌入到一个低维空间以此得到一个固定长度的低维向量,即数学输入的表示。

[0024] 优选地,所述表示层,具体用于使用GRU层作为特征提取器对文本特征进行特征提取,通过对GRU层的输出进行最大池化得到一组文本特征的表示,同时使用CNN来捕捉GRU层输出的特征,以此获得另一组文本特征;拼接来自所述嵌入层获得的上下文特征向量,以此获得一组上下文特征的表示,同时使用CNN来捕捉所述上下文信息之间的关联关系,以此得到另一组上下文特征的表示;将所提取出来的文本特征和上下文特征进行组合以得到待检测新闻的完整特征的表示。

[0025] 实施本发明实施例,具有如下有益效果:

[0026] 在本发明一个实施例的教导中，通过采用多任务学习模型，对待检测新闻的真实性检测和主题分类这两个任务进行联合训练，并同时返回所述待检测新闻的真实性以及所述待检测新闻的所属主题，上述方案能够同时检测出新闻的真实性以及新闻的所属主题，且提高了假新闻检测和主题分类的准确性。

附图说明

[0027] 图1是本发明一个实施例提供的一种基于多任务学习模型的假新闻检测方法的流程示意图；

[0028] 图2是本发明一个实施例提供的不同主题和不同作者的新闻的可信度分布示意图；

[0029] 图3是本发明一个实施例提供的多任务学习模型(FDML模型)的示意图；

[0030] 图4是本发明一个实施例提供的用于提取文本特征的Text-CNN模型的示意图。

具体实施方式

[0031] 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，所描述的实施例仅仅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。

[0032] 应当理解，当在本说明书和所附权利要求书中使用时，术语“包括”和“包含”指示所描述特征、整体、步骤、操作、元素和/或组件的存在，但并不排除一个或多个其它特征、整体、步骤、操作、元素、组件和/或其集合的存在或添加。

[0033] 还应当理解，在此本发明说明书中所使用的术语仅仅是出于描述特定实施例的目的而并不意在限制本发明。如在本发明说明书和所附权利要求书中所使用的那样，除非上下文清楚地指明其它情况，否则单数形式的“一”、“一个”及“该”意在包括复数形式。

[0034] 还应当进一步理解，在本发明说明书和所附权利要求书中使用的术语“和/或”是指相关联列出的项中的一个或多个的任何组合以及所有可能组合，并且包括这些组合。

[0035] 请参阅图1。

[0036] 一种基于多任务学习模型的假新闻检测方法，包括：

[0037] S101、采用多任务学习模型，对待检测新闻的真实性检测和主题分类这两个任务进行联合训练，并同时返回所述待检测新闻的真实性以及所述待检测新闻的所属主题。

[0038] 如图2所示，图2为不同主题和不同作者的新闻的可信度分布示意图，由图2可以看出，某些特定主题的新闻，被归类为假新闻的概率较高；同样的，某些作者也可能会被归类为有很高的意图发表假新闻。

[0039] 对于新闻真实性与新闻主题之间的存在着某些不可忽略的关系，本实施例采用多任务学习模型(Integrated Multi-Task Model，即用于假新闻检测的多任务学习模型(fake news detection multi-task learning(FDML))，对待检测新闻的真实性检测和主题分类进行联合训练，以提高假新闻检测的准确性。

[0040] 需要强调的是，本实施例在FDML模型中，假新闻检测任务和新闻主题分类任务是以统一的方式，进行同步训练学习的。

[0041] 在一优选实施例中，为了便于后续技术方案的描述，先定义FDML模型中一些术语，具体如下：

[0042] 1)新闻。

[0043] 设D＝{d_1,d_2,...,$d_{|D|}$}集合为包含|D|新闻，对于每一条新闻d＝{x,c}，x表示新闻的文本内容，c表示与新闻的文本内容对应的上下文信息。

[0044] 2)新闻标签。

[0045] 每一条新闻都与之相关标签y＝{$y^{m=1}$,$y^{m=2}$}，其中$y^{m=1}$表示新闻的主题标签，$y^{m=2}$表示相应的新闻的真实性标签。为了便于后面的描述，我们将索引标记出来，即m＝1表示主题分类任务，m＝2表示假新闻检测任务。在FDML模型中，假新闻检测任务与主题分类任务是同时进行的联合训练。

[0046] 3)主题分类。

[0047] 给定一组新闻D＝{d_1,d_2,...,$d_{|D|}$}，找到$\hat{y}^{m=1}$为每一条未标记的新闻，这样$\hat{y}^{m=1}=y^{m=1}$。

[0048] 4)假新闻检测。

[0049] 给定一组新闻D＝{d_1,d_2,...,$d_{|D|}$}，找到$\hat{y}^{m=2}$为每一条未标记的新闻，这样$\hat{y}^{m=2}=y^{m=2}$。请注意，对于假新闻检测的问题，新闻的主题标签$\hat{y}^{m=1}$并不一定是与新闻d＝{x,c}必需有关的消息。

[0050] 5)多任务假新闻检测。

[0051] 给定一组D＝{d_1,d_2,...,$d_{|D|}$}，找到一个能够同时预测新闻d＝{x,c}的主题标签$\hat{y}^{m=1}$和真实性标签$\hat{y}^{m=2}$的模型F，其中，

[0052] $F_d=\{\hat{y}^{m=1},\hat{y}^{m=2}\}=\{y^{m=1},y^{m=2}\}=y$。

[0053] 请参阅图3。

[0054] 所述多任务学习模型包括嵌入层、表示层和多任务层。

[0055] 所述嵌入层，将所述待检测新闻的文本内容和上下文信息作为原始数据嵌入到一个低维空间，以此获得文本内容和上下文信息的数学向量表示，并输入所述表示层。

[0056] 具体的，所述嵌入层，将所述待检测新闻的陈述中的每个词汇嵌入到低维空间，每个词汇获得一个固定长度的词向量，将所述待检测新闻的作者姓名、作者所属政党、地点等信息嵌入到一个低维空间以此得到一个固定长度的低维向量，即数学输入的表示。

[0057] 所述表示层，对文本嵌入词向量使用GRU(Gated Recurrent Unit，门控循环单元)层和CNN(Convolutional Neural Network，卷积神经网络)模型进行文本特征提取，对上下文嵌入向量使用另一个CNN模型和上下文特征向量层进行上下文特征提取，并将所提取出来的文本特征和上下文特征组合成完整特征。

[0058] 所述表示层可以对假新闻检测任务和主题分类任务同时执行训练学习。

[0059] 对于新闻的文本特征向量，FDML模型首先使用GRU层作为特征提取器，对文本内容进行编码，获得更高阶的表示。同时对GRU层的输出进一步应用最大池化和CNN模型以捕捉高阶特征。最后，FDML模型将提取的文本特征分别表示为后续的假新闻检测器和主题分类器的输入。

[0060] 同时，所述表示层使用另一个CNN模型和上下文特征向量层来捕捉新闻的上下文信息的特征表示。最后，FDML模型将提取的上下文信息的特征表示分别表示为后续的假新闻检测器和主题分类器的输入。

[0061] 最后将所提取出来的文本特征和上下文特征，根据其表示形式组合成完整特征，并输入多任务层进行下一步处理。

[0062] 需要强调的是，在表示层的应用的两个CNN模型，一个适用于对新闻的文本内容进行文本特征提取，另一个适用于对新闻上下文信息进行上下文特征提取，两者是不同的设计。

[0063] 所述多任务层，根据所述完整特征，对待检测新闻的真实性检测和主题分类进行联合训练，以同时完成真实性检测任务和主题分类任务，并输出真实性标签和主题标签。

[0064] 可以理解的是，对于多任务层，由于对待检测新闻的真实性检测和主题分类进行联合训练，即两个任务是同时执行的，所以输出主题标签和真实性标签是相互影响的。

[0065] 请参阅图4。

[0066] 在一优选实施例中，一条新闻由文本内容和上下文信息两部分组成。文本内容，指新闻文本，例如："Virginia governor Ralph Northam defies calls to resign over racist photo"，即标题、摘要、简报或声明等文本。上下文信息，是指新闻的元数据，例如，由"Jonathan Martin"于"Feb.2,2019"在"New York Times"上发表，即作者、时间、地点或期刊等元数据。

[0067] 文本内容可以看作是一个列表，因此，现有的词汇嵌入方法，如word2vec、fastText、GloVe均可直接使用。将每个词汇表示为一个低维向量，因为它可以有效地编码语义信息。考虑包含T个词汇$\{x_1,x_2,...,x_T\}$，词汇嵌入的目的是生成一个固定长度的向量，由变长序列x表示并通过一些深度神经网络进行训练，如RNN或CNN。

[0068] 与文本内容不同，语境内容是范畴性的，在退出时无法在语义上匹配上下文信息。因此，对于上下文信息，嵌入模型必须在学习过程中接受训练。此外，由于每个类别不同，如作者、出版商、时间，以及领域和词汇分布，需要按类别分开如$\{c_1,c_2,...,c_N\}$，对嵌入特征向量进行训练，如图3所示。

[0069] 在一优选实施例中，所述表示层，使用GRU层作为特征提取器对文本特征进行特征提取，通过对GRU层的输出进行最大池化得到一组文本特征的表示，同时使用CNN来捕捉GRU层输出的特征，以此获得另一组文本特征；拼接来自所述嵌入层获得的上下文特征向量，以此获得一组上下文特征的表示，同时使用CNN来捕捉所述上下文信息之间的关联关系，以此得到另一组上下文特征的表示；将所提取出来的文本特征和上下文特征进行组合以得到待检测新闻的完整特征的表示。

[0070] 当新闻的文本内容很短的时候，对其真实性的检测难度变得很大，对于这个问题，传统的假新闻检测手段并不能给出一个较为满意的解决方案。

[0071] 为了解决这个问题，本实施例在文本内容上附加了上下文信息，以协助完成假新闻的侦查任务。因为上下文信息对主题分类也很有用，所以它也用于主题分类。

[0072] 在FDML模型中，将所提取出来的文本特征和上下文特征组合成完整特征输入到假新闻检测器和主题分类器。

[0073] 传统的RNN在无法有效的捕获长期依赖，因此，本实施例使用GRU作为文本特征提

取器。对于每一个时间步长t，计算GRU的过程描述如下：

[0074] $x'_t = Ex'_t$

[0075] $r_t = \sigma(W_r x'_t + U_r h_{t-1})$

[0076] $z_t = \sigma(W_z x'_t + U_z h_{t-1})$

[0077] $h'_t = \tanh(W_h x'_t + U_h(h_{t-1} \cdot r_t))$

[0078] $h_t = (1-z_t) \cdot h_{t-1} + z_t \cdot h'_t$；

[0079] 其中，x'_t表示词汇x_t的词汇嵌入向量，E为词汇嵌入矩阵，r_t为重置门，z_t为更新门，并且(W,U)是GRU中的参数。

[0080] 如图3所示，在GRU层计算得到对文本内容(主题信息)进行有效编码的隐状态表示$\{h_1, h_2, ..., h_T\}$，使用多个卷积核对GRU的隐状态进行特征提取。卷积运算采用卷积核$w \in R^{pk}$，其中k为隐状态的参数，使用多个卷积核对GRU的隐状态进行特征提取。

[0081] 给定长度为r的一个隐藏状态区域$h_{i:i+p-1}$，则，$u_i = f(w \cdot h_{i:i+p-1} + b)$，其中b是偏置项，f是非线性函数。根据每个区域的GRU隐状态，我们可以得到特征映射

[0082] $u = [u_1, u_2, ..., u_{T-r+1}]$。

[0083] 如图4所示，展示了Text-CNN具有2/3/4滤波区域，每个滤波区域配置2个滤波器。

[0084] 为了捕获上下文信息之间的依赖关系，本实施例使用另一个CNN执行类似的操作。但输入是上下文信息的嵌入向量，而不是GRU层的隐状态$\{h_1, h_2, ..., h_T\}$。

[0085] 考虑到作者的信用历史代表了作者的可信度，并且已经作为载体呈现，它没有参与在上述计算过程中，而是与上下文信息进行串联运算，组成上下文信息的最终表示形式用于假新闻检测。在形式上，

[0086] $s^{m=2} = CNN([c_1, c_2, ..., c_N]) \oplus Ch$，

[0087] 其中Ch为信用记录。

[0088] 对于主题分类任务，本实施例只需连接N个嵌入向量表示上下文信息。在形式上，

[0089] $s^{m=1} = c_1 \oplus c_2 \oplus ... \oplus c_N$。

[0090] 在一优选实施例中，得到完整特征，即实现了全连接层用于融合文本特征和上下文特征。多任务层的执行描述如下：

[0091] $\hat{y}_m = Soft\max(w_2^m \cdot f(w_1^m r^m + b_1^m) + b_2^m)$；

[0092] 其中，(w_1^m, w_2^m)为全连接层的权重，(b_1^m, b_2^m)为对应的偏置，f为激活函数，$r^m = \upsilon^m \oplus s^m$，$\oplus$表示串联运算。

[0093] 损失函数为假新闻检测与主题分类的线性组合：

[0094] $L = CE(\hat{y}^{m=1}, y^{m=1}) + CE(\hat{y}^{m=2}, y^{m=2}) + \lambda \|\Theta\|_2^2$；

[0095] 其中，Θ代表模型参数，λ是权衡模型参数的正则化系数。

[0096] 基于上述任一实施例的教导，根据以下现象：

[0097] 1)新闻设计的某些话题被归类为假新闻的概率很高；

[0098] 2)一些作者有很高的意图发布假新闻。

[0099] FDML模型有效地融合了文本内容和上下文信息，并利用了新闻主题、作者的可信

度分布与新闻的真实性之间的相关性，进行多任务学习，提高了假新闻检测的准确性。

[0100] 本发明实施例还提供一种基于多任务学习模型的假新闻检测系统，包括：

[0101] 检测单元，用于采用多任务学习模型，对待检测新闻的真实性检测和主题分类这两个任务进行联合训练，并同时返回所述待检测新闻的真实性以及所述待检测新闻的所属主题。

[0102] 如图2所示，图2为不同主题和不同作者的新闻的可信度分布示意图，由图2可以看出，某些特定主题的新闻，被归类为假新闻的概率较高；同样的，某些作者也可能会被归类为有很高的意图发表假新闻。

[0103] 对于新闻真实性与新闻主题之间的存在着某些不可忽略的关系，本实施例采用多任务学习模型(Integrated Multi-Task Model，即用于假新闻检测的多任务学习模型(fake news detection multi-task learning(FDML))，对待检测新闻的真实性检测和主题分类进行联合训练，以提高假新闻检测的准确性。

[0104] 需要强调的是，本实施例在FDML模型中，假新闻检测任务和新闻主题分类任务是以统一的方式，进行同步训练学习的。

[0105] 请参阅图3。

[0106] 所述多任务学习模型包括嵌入层、表示层和多任务层。

[0107] 所述嵌入层，将所述待检测新闻的文本内容和上下文信息作为原始数据嵌入到一个低维空间，以此获得文本内容和上下文信息的数学向量表示，并输入所述表示层。

[0108] 具体的，所述嵌入层，将所述待检测新闻的陈述中的每个词汇嵌入到低维空间，每个词汇获得一个固定长度的词向量，将所述待检测新闻的作者姓名、作者所属政党、地点等信息嵌入到一个低维空间以此得到一个固定长度的低维向量，即数学输入的表示。

[0109] 所述表示层，对文本嵌入词向量使用GRU(Gated Recurrent Unit，门控循环单元)层和CNN(Convolutional Neural Network，卷积神经网络)模型进行文本特征提取，对上下文嵌入向量使用另一个CNN模型和上下文特征向量层进行上下文特征提取，并将所提取出来的文本特征和上下文特征组合成完整特征。

[0110] 所述表示层可以对假新闻检测任务和主题分类任务同时执行训练学习。

[0111] 对于新闻的文本特征向量，FDML模型首先使用GRU层作为特征提取器，对文本内容进行编码，获得更高阶的表示。同时对GRU层的输出进一步应用最大池化和CNN模型以捕捉高阶特征。最后，FDML模型将提取的文本特征分别表示为后续的假新闻检测器和主题分类器的输入。

[0112] 同时，所述表示层使用另一个CNN模型和上下文特征向量层来捕捉新闻的上下文信息的特征表示。最后，FDML模型将提取的上下文信息的特征表示分别表示为后续的假新闻检测器和主题分类器的输入。

[0113] 最后将所提取出来的文本特征和上下文特征，根据其表示形式组合成完整特征，并输入多任务层进行下一步处理。

[0114] 需要强调的是，在表示层的应用的两个CNN模型，一个适用于对新闻的文本内容进行文本特征提取，另一个适用于对新闻上下文信息进行上下文特征提取，两者是不同的设计。

[0115] 所述多任务层，根据所述完整特征，对待检测新闻的真实性检测和主题分类进行

联合训练，以同时完成真实性检测任务和主题分类任务，并输出真实性标签和主题标签。

[0116] 可以理解的是，对于多任务层，由于对待检测新闻的真实性检测和主题分类进行联合训练，即两个任务是同时执行的，所以输出主题标签和真实性标签是相互影响的。

[0117] 请参阅图4。

[0118] 在一优选实施例中，一条新闻由文本内容和上下文信息两部分组成。文本内容，指新闻文本，例如："Virginia governor Ralph Northam defies calls to resign over racist photo"，即标题、摘要、简报或声明等文本。上下文信息，是指新闻的元数据，例如，由"Jonathan Martin"于"Feb.2,2019"在"New York Times"上发表，即作者、时间、地点或期刊等元数据。

[0119] 文本内容可以看作是一个列表，因此，现有的词汇嵌入方法，如word2vec、fastText、GloVe均可直接使用。将每个词汇表示为一个低维向量，因为它可以有效地编码语义信息。考虑包含T个词汇$\{x_1,x_2,...,x_T\}$，词汇嵌入的目的是生成一个固定长度的向量，由变长序列x表示并通过一些深度神经网络进行训练，如RNN或CNN。

[0120] 与文本内容不同，语境内容是范畴性的，在退出时无法在语义上匹配上下文信息。因此，对于上下文信息，嵌入模型必须在学习过程中接受训练。此外，由于每个类别不同，如作者、出版商、时间，以及领域和词汇分布，需要按类别分开如$\{c_1,c_2,...,c_N\}$，对嵌入特征向量进行训练，如图3所示。

[0121] 在一优选实施例中，所述表示层，使用GRU层作为特征提取器对文本特征进行特征提取，通过对GRU层的输出进行最大池化得到一组文本特征的表示，同时使用CNN来捕捉GRU层输出的特征，以此获得另一组文本特征；拼接来自所述嵌入层获得的上下文特征向量，以此获得一组上下文特征的表示，同时使用CNN来捕捉所述上下文信息之间的关联关系，以此得到另一组上下文特征的表示；将所提取出来的文本特征和上下文特征进行组合以得到待检测新闻的完整特征的表示。

[0122] 当新闻的文本内容很短的时候，对其真实性的检测难度变得很大，对于这个问题，传统的假新闻检测手段并不能给出一个较为满意的解决方案。

[0123] 为了解决这个问题，本实施例在文本内容上附加了上下文信息，以协助完成假新闻的侦查任务。因为上下文信息对主题分类也很有用，所以它也用于主题分类。

[0124] 在FDML模型中，将所提取出来的文本特征和上下文特征组合成完整特征输入到假新闻检测器和主题分类器。

[0125] 传统的RNN在无法有效的捕获长期依赖，因此，本实施例使用GRU作为文本特征提取器。对于每一个时间步长t，计算GRU的过程描述如下：

[0126] $x'_t=Ex'_t$

[0127] $r_t=\sigma(W_r x'_t+U_r h_{t-1})$

[0128] $z_t=\sigma(W_z x'_t+U_z h_{t-1})$

[0129] $h'_t=\tanh(W_h x'_t+U_h(h_{t-1}\cdot r_t))$

[0130] $h_t=(1-z_t)\cdot h_{t-1}+z_t\cdot h'_t$；

[0131] 其中，x'_t表示词汇x_t的词汇嵌入向量，E为词汇嵌入矩阵，r_t为重置门，z_t为更新门，并且(W,U)是GRU中的参数。

[0132] 如图3所示，在GRU层计算得到对文本内容(主题信息)进行有效编码的隐状态表示

$\{h_1,h_2,...,h_T\}$，使用多个卷积核对GRU的隐状态进行特征提取。卷积运算采用卷积核$w\in R^{pk}$，其中k为隐状态的参数，使用多个卷积核对GRU的隐状态进行特征提取。

[0133] 给定长度为r的一个隐藏状态区域$h_{i:i+p-1}$，则，$u_i=f(w\cdot h_{i:i+p-1}+b)$，其中b是偏置项，f是非线性函数。根据每个区域的GRU隐状态，我们可以得到特征映射

[0134] $u=[u_1,u_2,...,u_{T-r+1}]$。

[0135] 如图4所示，展示了Text-CNN具有2/3/4滤波区域，每个滤波区域配置2个滤波器。

[0136] 为了捕获上下文信息之间的依赖关系，本实施例使用另一个CNN执行类似的操作。但输入是上下文信息的嵌入向量，而不是GRU层的隐状态$\{h_1,h_2,...,h_T\}$。

[0137] 考虑到作者的信用历史代表了作者的可信度，并且已经作为载体呈现，它没有参与在上述计算过程中，而是与上下文信息进行串联运算，组成上下文信息的最终表示形式用于假新闻检测。在形式上，

[0138] $s^{m=2}=CNN([c_1,c_2,...,c_N])\oplus Ch$，

[0139] 其中Ch为信用记录。

[0140] 对于主题分类任务，本实施例只需连接N个嵌入向量表示上下文信息。在形式上，

[0141] $s^{m=1}=c_1\oplus c_2\oplus...\oplus c_N$。

[0142] 在一优选实施例中，得到完整特征，即实现了全连接层用于融合文本特征和上下文特征。多任务层的执行描述如下：

[0143] $\hat{y}_m=Soft\max(w_2^m\cdot f(w_1^m r^m+b_1^m)+b_2^m)$；

[0144] 其中，(w_1^m,w_2^m)为全连接层的权重，(b_1^m,b_2^m)为对应的偏置，f为激活函数，$r^m=v^m\oplus s^m$，⊕表示串联运算。

[0145] 损失函数为假新闻检测与主题分类的线性组合：

[0146] $L=CE(\hat{y}^{m=1},y^{m=1})+CE(\hat{y}^{m=2},y^{m=2})+\lambda\|\Theta\|_2^2$；

[0147] 其中，Θ代表模型参数，λ是权衡模型参数的正则化系数。

[0148] 基于上述任一实施例的教导，根据以下现象：

[0149] 1)新闻设计的某些话题被归类为假新闻的概率很高；

[0150] 2)一些作者有很高的意图发布假新闻。

[0151] FDML模型有效地融合了文本内容和上下文信息，并利用了新闻主题、作者的可信度分布与新闻的真实性之间的相关性，进行多任务学习，提高了假新闻检测的准确性。

[0152] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，所述的存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory，ROM)或随机存储记忆体(Random Access Memory，RAM)等。

[0153] 以上所述是本发明的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也视为本发明的保护范围。

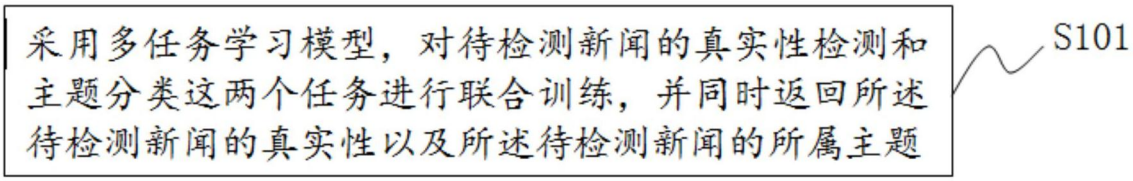

采用多任务学习模型，对待检测新闻的真实性检测和主题分类这两个任务进行联合训练，并同时返回所述待检测新闻的真实性以及所述待检测新闻的所属主题
S101

图1

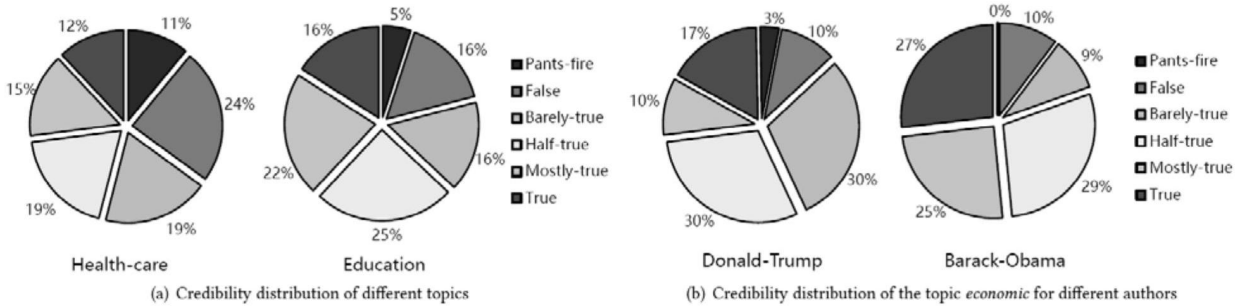

12%
11%
15%
24%
19%
19%
Health-care
16%
5%
16%
22%
16%
25%
Education
Pants-fire
False
Barely-true
Half-true
Mostly-true
True
17%
3%
10%
10%
30%
30%
Donald-Trump
27%
0%
10%
9%
25%
29%
Barack-Obama
(a) Credibility distribution of different topics
(b) Credibility distribution of the topic economic for different authors

图2

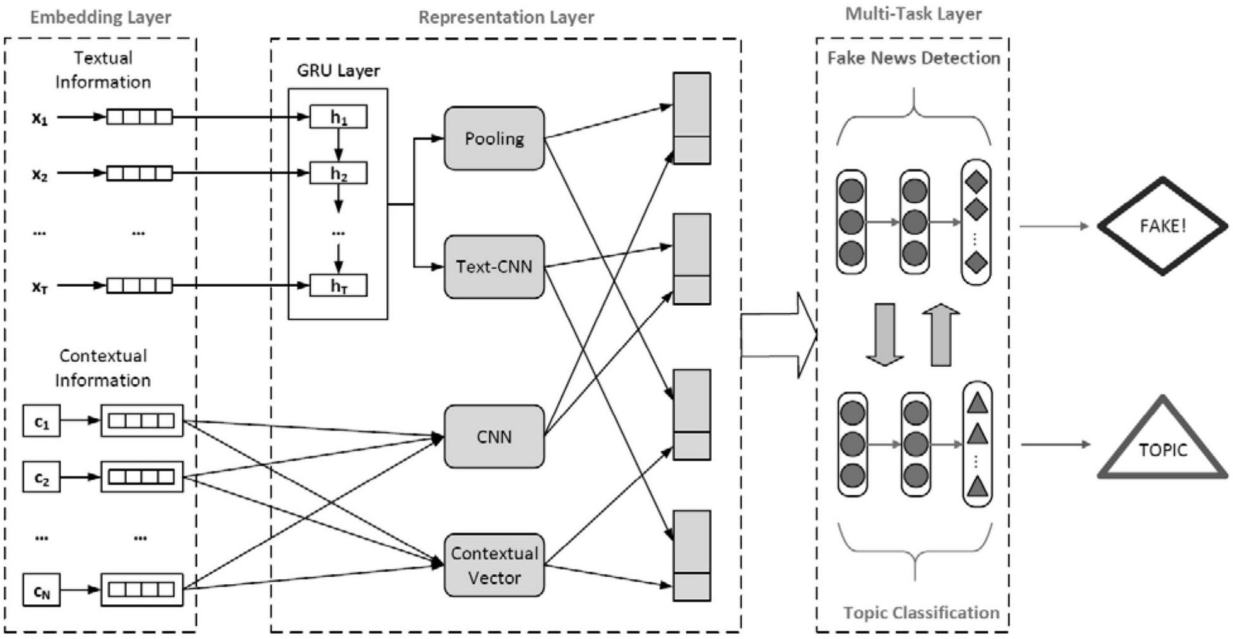

Embedding Layer
Representation Layer
Multi-Task Layer
Textual Information
GRU Layer
Pooling
Text-CNN
CNN
Contextual Vector
Contextual Information
Fake News Detection
Topic Classification
FAKE!
TOPIC

图3

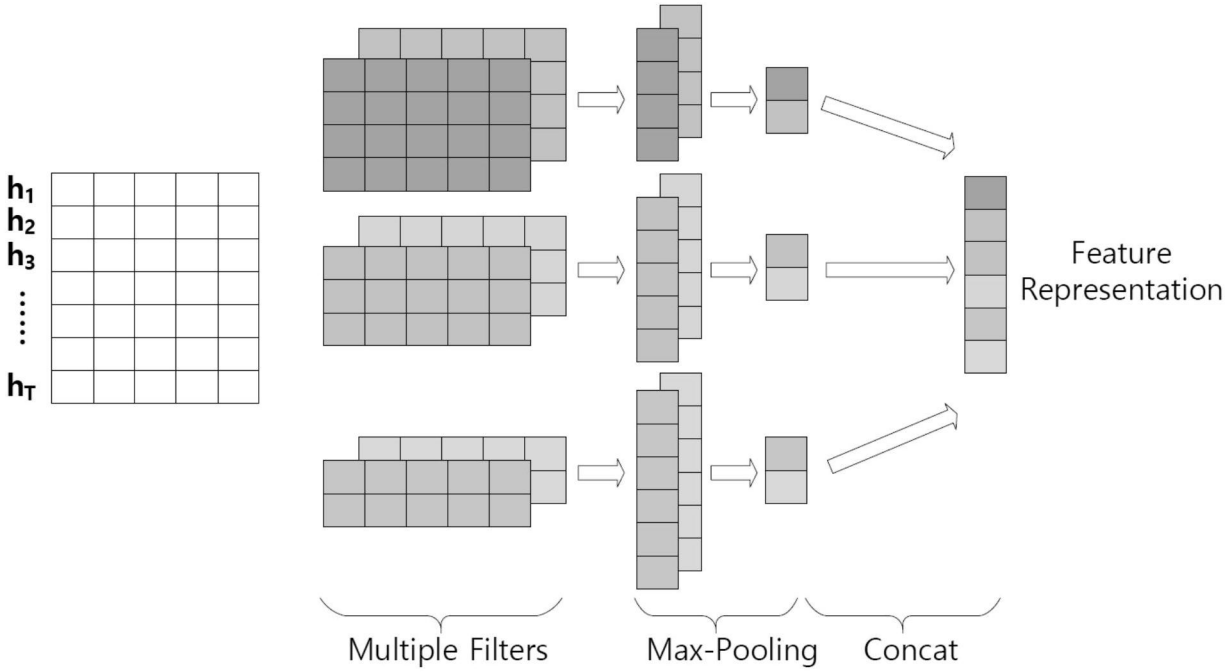

h1
h2
h3
hT
Feature
Representation
Multiple Filters
Max-Pooling
Concat

图4