

数据分析与知识发现
Data Analysis and Knowledge Discovery
ISSN 2096-3467, CN 10-1478/G2

《数据分析与知识发现》网络首发论文

题目：融合新闻传播模式和传播者情感偏好的虚假新闻检测研究
作者：蒋涛，潘云辉，崔鹏
网络首发日期：2024-06-05
引用格式：蒋涛，潘云辉，崔鹏. 融合新闻传播模式和传播者情感偏好的虚假新闻检测研究[J/OL]. 数据分析与知识发现.
<https://link.cnki.net/urlid/10.1478.G2.20240605.0919.004>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

融合新闻传播模式和传播者情感偏好的虚假新闻检测研究

蒋涛, 潘云辉, 崔鹏
(贵州财经大学信息学院 贵阳 550025)

摘要:

[目的]解决现有的基于新闻传播模式的虚假新闻检测研究未能充分挖掘并融合传播者情感偏好特征的问题,提升虚假新闻检测模型准确率。

[方法]构建了一种融合新闻传播模式和传播者情感偏好的虚假新闻检测模型,从传播者历史发文中挖掘传播者情感偏好特征,并用新闻文本的情感特征丰富新闻文本特征,再以新闻传播模式为基础,通过图卷积神经网络以及混合的池化函数,融合新闻传播者情感偏好和新闻传播模式,最后将丰富后的新闻文本特征与池化函数的结果融合,输入到分类器中得到最终的分类结果。

[结果]在公开的数据集 GossipCop、PolitiFact 上将模型与其他 5 个模型进行比较,该模型的精确率分别达到 97.39% 和 90.48%,优于基线模型,证明了该模型的有效性。

[局限]本研究暂未考虑到传播者跟风转发等特殊情况。

[结论]融合新闻传播模式和传播者情感偏好的模型能够有效提高虚假新闻检测识别的准确率。

关键词: 虚假新闻检测; 图卷积神经网络; 社会网络; 情感分析; 传播模式

分类号: TP391, G203

DOI:

Fake News Detection Incorporating News Propagation Models and Propagator Sentiment Preference

Jiang Tao, Pan Yunhui, Cui Peng

(School of Information, Guizhou University of Finance and Economics, Guiyang 550025, China)

Abstract:

[Objective] In response to solve the problem that existing research on fake news detection based on news propagation model fails to fully explore and incorporate the propagator sentiment preference features, and to improve the accuracy of the fake news detection model.

[Methods] Based on the news propagation mode, the sentiment preference of the news propagator and the news propagation mode are fused through the graph convolutional neural network and the mixed pooling function, and finally the enriched news text features are fused with the results of the pooling function and input into the classifier to obtain the final classification result.

[Results] The model is compared with five other models on the publicly available datasets GossipCop and PolitiFact. The precision of the model reaches 97.39% and 90.48% respectively, surpassing the baseline models and demonstrating the effectiveness of the proposed model.

[Limitations] One limitation of this study is that it did not account for situations where disseminators mindlessly repost news without considering their own preferences.

[Conclusions] The method of integrating news propagation models with propagators' sentiment preference can effectively improve the accuracy of fake news detection.

Keywords: Fake news detection; Graph Convolutional Neural Networks; Social network; Sentiment analysis; propagation model

1 引言

随着自媒体技术的不断发展,传播者能够突破时间、地点的限制,自由地接收和发表观点,但由于社交媒体的匿名性,促进了网络虚假新闻的滋生和传播。虚假新闻是指与时事相关的,旨在使人们相信但又缺乏具体资料以证实其实真实性的言论^[1]。目前,社交媒体平台主要以用户举报和人工审核的方式识别虚假新闻,这种方法需要花费大量的时间,较为繁琐,并不能满足当下的需求。

近年来,如何高效及时地检测出虚假新闻是需要解决的热点问题^[2-4]。大部分虚假新闻检测模型是基于新闻文本的,如利用新闻内容的语义特征或语义一致性的模型^[3,4]。但高明的虚假新闻发布者,往往会模仿真实新闻的语言表达,使得虚假新闻与真实新闻在语义特征上几乎没有差别,导致这类模型会很难识别出这类虚假新闻。现有的利用新闻传播路径的虚假新闻检测研究,大多通过计算传播者回复或转推文本的情感分数来表示传播者的情感偏好,而未能充分挖掘并融合传播者的情感偏好特征,并用于虚假新闻检测。

因此,本文从传播者的历史发文中挖掘出丰富的传播者情感偏好特征表示,并基于新闻传播模式,利用图神经网络对传播者的情感偏好和传播模式建模,构建融合新闻传播模式和传播者情感偏好的虚假新闻检测模型(User Sentiment Preference-aware Graph Convolution Network, USPGCN)。在现实世界的数据集 GossipCop 和 PolitiFact 上,精确率分别达到了 97.39%和 90.48%,并通过消融实验证明了各模块的有效性。

2 相关研究

深度学习虚假新闻检测模型主要分为基于文本的检测模型、基于社会网络的检测模型。

基于文本的虚假新闻检测模型主要依靠文本的语义特征来区分新闻的真假。Ajao 等人^[5]利用新闻文本的情感特征来检测虚假新闻。Luvembe 等人^[6]利用新闻发布者情感和社会情感作为双重情感来检测虚假新闻。Zhu 等人^[7]从新闻文本中提取实体,构建实体去偏置框架来提高虚假新闻检测模型的泛化性。张国标等人^[4]在提取文本特征、图像特征的基础上,将图像映射为语义标签,通过计算文本与图像语义的一致性来构建虚假新闻检测模型。曾江峰等人^[8]利用主题模型挖掘新闻主题,并与预训练模型 BERT^[9]挖掘出的动态上下文语义特征相融合,以实现虚假新闻检测。刘赏等人^[10]利用新闻标题与新闻正文的差异性来挖掘新闻文本以及与新闻文本相关联的文本来区别新闻的真假。

基于社会网络的模型主要是利用新闻在社交媒体上的传播模式来区分虚假新闻^[11]。新闻传播模式是指新闻在社交媒体上传播的路径以及传播的方式,包含了新闻从发布者到新闻传播者之间的传播规模、传播速度等方面的特征。虚假新闻与真实新闻在传播方式上存在一定的差异。例如虚假新闻通常先是被一个普通用户发布,然后被一些意见领袖转发和支持,最后被大量普通用户转发;相反,

真实新闻通常是由意见领袖发布的,并由许多普通用户直接转发。Vosoughi 等人分析^[12]推特上的新闻传播,发现与真实新闻相比,虚假新闻能够传播得更快、更广、更深。充分利用新闻传播模式可以提高虚假新闻检测模型的准确性^[13]。深度学习方法可以利用新闻传播的路径,将新闻传播路径用图来表示,从而得到不同新闻的传播模式。

为了对新闻的传播模式建模, Ma 等人^[14]使用 LSTM 从新闻传播的树状图中挖掘虚假新闻传播的顺序传播。Yu 等人^[15]使用 CNN 来获取新闻传播图的局部相邻的特征。然而基于 RNN 的方法未能充分利用传播的扩散结构,基于 CNN 的方法未能充分利用传播结构图中的全局结构,为了更充分的挖掘新闻的传播模式特征, Bian 等人^[16]在前人的基础上利用两个 GCN 分别捕捉自下而上和自上而下的两种传播模式,以此来充分对新闻传播模式建模。新闻传播者的思想观念各不相同,导致不同的新闻传播者偏向于接受和赞同不同的新闻,为了获得传播者赞同新闻的偏好, Dou 等人^[17]将传播者的历史发文进行编码,以此得到传播者的偏好表示,再利用 GCN 对新闻传播模式建模,实现虚假新闻检测,但该方法未能充分考虑传播者历史发文中的情感偏好,并与传播模式融合,用于虚假新闻检测。

传播者情感方面的偏好作为传播者的特征表示已经广泛用于社会网络研究^[18]、情感分析^[19]、信息行为^[20]、虚假评论检测^[21]等领域,一些学者也将其用于虚假新闻检测任务中。如 Fu 等人^[22]从归一化的回复序列中提取文本特征,并利用百度情感分析 API 计算情感分数,将文本特征与情感分数组合作为微博表示,最后通过 RvNN 模型^[14]捕获新闻传播路径上的微博表示,但该方法使用的 RvNN 模型未能充分利用新闻传播的扩散结构。Moschopoulos 等人^[23]构建了一个与传播者偏好相关的知识图谱,利用 VADER 情感分析工具^[24],计算传播者回复和转推文本的情感分数,并结合知识图谱合成传播者情感偏好特征,再将传播者情感偏好与新闻传播图相关特征融合,该方法未能充分将传播者情感偏好与新闻传播模式融合。

现有的基于新闻传播模式虚假新闻检测模型,大多未能充分挖掘并融合传播者的情感偏好特征,并用于虚假新闻检测。因此本文利用情感分析模型,从传播者的历史推文中提取出情感偏好特征作为传播者特征,根据新闻传播者的转发行为,建立新闻传播图以获得新闻的传播模式,利用图卷积神经网络构建融合新闻传播模式和传播者情感偏好的虚假新闻检测模型,以用于虚假新闻检测。

3 研究框架与模型

3.1 问题定义

虚假新闻检测任务可以看成是一个二分类问题,每个新闻都是真实($y=0$)和虚假($y=1$)的其中一种。 $S \in \{W_1, W_2, W_3, \dots, W_n\}$ 代表新闻文本序列,获取新闻传播者序列 $v_i \in \{v_1, v_2, v_3, \dots, v_n\}$, 进一步获取每一位传播者近期历史发文 U_m , 其中 $n \in \{0, 1, 2, \dots\}$ 表示传播者编号, $t \in \{0, 1, 2, \dots, 199\}$ 是传播者发文的编号, U_m 代表第 n 个传播者的第 t 条推文,学习一个函数: $W'_t \rightarrow Y, Y \in \{0, 1\}$ 是新闻序列的标签。

3.2 模型建立

新闻是否被传播与新闻传播者的偏好之间存在关系,这是区别真实新闻和虚假新闻的重要特征,因此本文提出一种融合新闻传播模式和传播者情感偏好的模

型框架。基于传播者与新闻本身的联级关系，融合传播模式和传播者情感偏好，以此来建立模型，模型结构如图 1 所示，其主要结构如下：（1）新闻内容特征以及传播者情感偏好提取模块；（2）新闻社会网络传播图构建模块；（3）池化函数模块；（4）信息融合与分类预测模块。

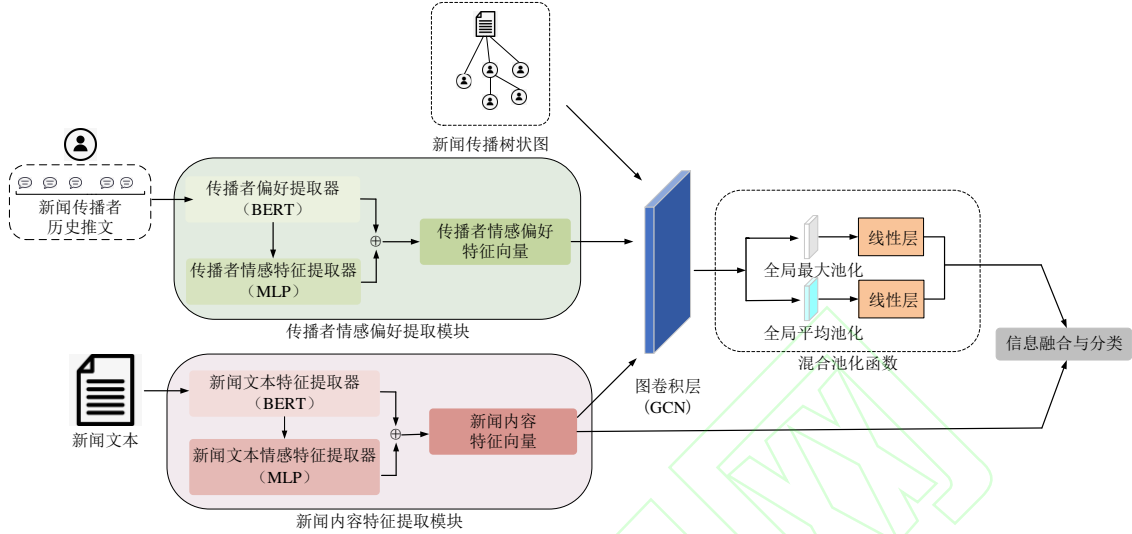


图 1 USPGCN 网络框架

Fig.1 USPGCN Network Framework

下面是对以上模块的详细介绍。

（1）新闻内容特征以及传播者情感偏好提取

新闻内容是鉴别虚假新闻的重要信息。王昊等人^[25]提出了语义增强的方法，提取了新闻内容文本中深层次的情感特征来丰富新闻文本特征，我们为了利用新闻内容的情感特征来丰富新闻文本特征，借鉴王昊等人的方法，引入一个 BERT+MLP 模型来提取新闻内容的情感特征。具体步骤为首先利用一个情感分析数据集对训练出一个基于 BERT 模型的 MLP 二分类情感分析模型。然后利用 BERT 模型提取新闻内容文本特征 X_t 。再将新闻内容文本特征 X_t 输入到所训练好的 MLP 中，得出新闻内容情感特征 X'_t 。最后将新闻内容文本特征 X_t 和新闻内容情感特征 X'_t 拼接，得到新闻内容特征 V_t 。这一过程如公式(1)~(3)所示：

$$X_t = \text{BERT}([\text{CLS}]S[\text{SEP}]) \quad (1)$$

$$X'_t = \text{MLP}(X_t) \quad (2)$$

$$V_t = X_t \oplus X'_t \quad (3)$$

其中 $S = \{w_1, w_2, w_3, \dots, w_n\}$ 是新闻文本序列。其中公式(1)用来得到新闻内容的向量表示，在此基础上，我们借鉴王昊等人^[25]提出的语义增强的方法，提出利用公式(2)和公式(3)来提取新闻内容中的情感特征。

传播者大量的历史发文蕴含了传播者丰富的情感偏好特征。针对传播者情感偏好的提取的问题，Xiao 等人^[26]在挖掘电子商务中的用户偏好时，利用细粒度情感分析模型挖掘用户情感偏好。受此启发，本文利用情感分析模型从传播者历史发文中挖掘丰富的情感特征，与传播者历史发文的文本特征拼接，作为传播者

情感偏好。具体做法与新闻内容特征向量的提取方法类似，将传播者近期所发的 T 条推文分别输入 BERT 模型得到推文文本特征 M_m ，再做平均转变成一条文本特征向量 N_n 。然后将文本特征向量 N_n 输入到所训练好的 MLP 中提取情感特征 N'_n ，最后将情感特征向量 N'_n 与文本特征向量 N_n 横向拼接，得到传播者发文情感偏好特征向量 O_n 。具体过程如公式(4)~(7)所示：

$$M_m = \text{BERT}([\text{CLS}]U_m[\text{SEP}]) \quad (4)$$

$$N_n = (\sum M_m) / T \quad (5)$$

$$N'_n = \text{MLP}(N_n) \quad (6)$$

$$O_n = N_n \oplus N'_n \quad (7)$$

其中 n 是传播者编号， t 是传播者发文的编号， U_m 代表第 n 个传播者的第 t 条推文。我们提出使用公式(4)和公式(5)用来得到单个新闻传播者 n 的文本特征，受到 Xiao 等人^[26]挖掘用户情感偏好的做法的启发；公式(6)和公式(7)的使用与公式(2)和公式(3)类似，但用来挖掘传播者历史发文中的情感偏好，同样是借鉴了王昊等人^[25]的提取文本中情感特征的做法。

(2) 新闻社会网络传播图构建模块

为了获得某篇新闻的传播模式，我们使用了和 UPFD 相同的策略^[17]构建新闻社会网络传播图。如图 1 中的树状图，根节点表示新闻，其他节点都是该新闻的传播者。所构建树状的新闻社会网络传播图，作为图卷积神经网络的邻接矩阵。将上文提取的新闻内容特征向量 V_i 作为根节点特征，传播者情感偏好特征 O_n 作为其他节点的特征。本文使用的图卷积神经网络的卷积层采用的方法是 SAGEConv^[27]，该方法是图卷积神经网络常用的方法。它的公式如下：

$$X'_i = \text{ReLU}(W_1 X_i + W_2 \sum_{j \in N(i)} \frac{1}{|N(i)|} X_j) \quad (8)$$

其中， X_i 是节点 i 的输入特征， X'_i 是节点 i 的输出特征向量， W_1 和 W_2 是可学习权重 $N(i)$ 是节点 i 的邻居节点的集合， $|N(i)|$ 是邻居节点的数量， X_j 是与节点 i 相邻的节点 j 的特征向量。通过该公式，实现了传播者情感偏好和新闻传播模式的初次融合。

(3) 池化函数模块

图卷积神经网络在处理社会网络图时，只使用全局最大池化会忽略全局节点特征，只使用全局平均池化会忽略局部节点特征，都在一定程度上造成了信息丢失。因此本文参考 Tong 等人^[28]的方法，提出了混合的池化函数，如图 2 所示，相比于 Tong 等人设置超参数作为全局平均池化和全局混合池化的权重的方法，本文将两个池化层的输出结果放入一层全连接神经网络中，训练可学习参数 w_1 和 w_2 来权衡全局平均池化和全局混合池化的比重，提高了训练效率，最

后将两个全连接神经网络的输出结果直接相加，得到输出向量。这一过程如公式(9)~(11)所示：

$$P_{max} = W_1(\max\{X'_i\}) + b_1 \quad (9)$$

$$P_{mean} = W_2(\frac{1}{N} \sum_{i=1}^N X'_i) + b_2 \quad (10)$$

$$P_{out} = P_{max} + P_{mean} \quad (11)$$

其中 $X_i \in \{X'_1, X'_2, X'_3, \dots, X'_n\}$ 是卷积层输出的各个节点， W_1 和 W_2 是可学习权重， b_1 和 b_2 是可学习偏置项， P_{out} 是池化函数的输出。

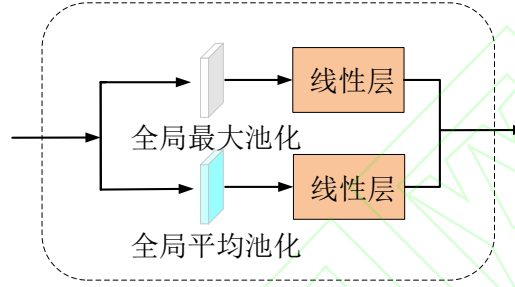


图2 池化函数结构图

Fig.2 Pooling Framework

(4) 信息融合与分类预测模块

本文使用了 Moschopoulos 等人^[23]提出的对不同特征进行融合的方法，将新闻内容特征向量 V_i 与池化函数的输出 P_{out} 拼接，作为最终新闻整体特征向量 V_{final} ，我们将 V_{final} 输入到全连接层里，并通过 softmax 函数得出分类标签。这一过程由公式得到：

$$V_{final} = V_{text} \oplus P_{out} \quad (12)$$

$$p = \text{softmax}(W_c V_{final} + b_c) \quad (13)$$

其中 W_c 和 b_c 是 USPGCN 模型的可学习参数。

4 实验与分析

4.1 数据集

本文的实验使用 Shu 等人提出的公共的虚假新闻检测数据集 GossipCop 和 PolitiFact^[29]。训练集、验证集和测试集按 2:1:7 划分，其中正向代表新闻是虚假新闻，负向代表新闻是真实新闻，数据集的统计结果如表 1 所示：

原始数据集中还包含新闻内容的文本、传播者的用户配置文件、转推以及参与的时间戳。

数据集的节点包括新闻节点和传播者节点，每条新闻根据上文“3.2 模型建立”中提到的策略，被构建成单个新闻节点连接多个传播者节点，传播者节点之间相互连接的树状结构，这种树状的图结构可以用于表示新闻的传播模式。

我们使用的传播者历史推文数据，来自 UPFD 公开的对数据集原始推文数据的补充数据。

为了得到能提取文本情感特征的 MLP 模型，我们用 BERT+MLP 在斯坦福大学 Twitter 情感分析数据集^[30]上做二分类情感分析。

斯坦福大学 Twitter 情感分析数据集中有 1578672 条数据，我们从中随机选取四万条数据用于我们的实验，正向情感数量与负向情感数量比为 1:1。

表 1 数据集统计指标

Table1 Statistics of the Dataset						
数据集	图的总数	真实新闻的数量	虚假新闻的数量	节点总数	边的总数	平均每个图节点数
PolitiFact	314	157	157	41054	40740	131
GossipCop	5464	2732	2732	314262	309798	58

4.2 实验设置

本文使用准确率(accuracy)、F1 值、召回率(recall)和精确率(precision)作为评价指标。

本文模型 USPGCN 的图嵌入维度为 128，优化算法为 Adam，将学习率设置成了 0.001，batch size 设置为 128，使用 ReLU 作为激活函数，其他超参数设置如表 2 所示。我们在 NVIDIA RTX 4060(8G)上完成训练，使用版本为 2.0.1 的 Pytorch 框架，图卷积神经网络模型是使用 PyTorch-Geometric 包完成的，使用的 Python 版本是 3.9.13。

表 2 USPGCN 模型参数设置

Table2 Parameter Setting of the USPGCN Model	
USPGCN 模型参数	参数取值
每个传播者历史推文数量	200
提取新闻的 BERT 最大输入长度	512
提取推文的 BERT 最大输入长度	16
BERT 模型	BERT-Large
图嵌入维度	128
优化算法	Adam
L2 正则化常数	0.001
批量大小	128
学习率	0.001
激活函数	ReLU
GossipCop 数据集的 epoch	80
PolitiFact 数据集的 epoch	50
图卷积层数	1
损失函数	负对数似然损失

对于斯坦福大学 Twitter 情感分析数据集，利用预训练模型 BERT+MLP 对其做二分类情感分析，MLP 的层数设置为 2，使用 ReLU 作为激活函数，其他参数设置如表 3 所示。该步骤在 NVIDIA A100(40G)上完成训练，使用版本为 1.11 的 PyTorch 框架，使用的 Python 版本是 3.8.16。最终模型准确率达到 72%。

表 3 BERT+MLP 参数设置

Table3 Parameter Setting of the BERT+MLP	
BERT+MLP 模型参数	参数取值
BERT 模型隐藏层维度	768

BERT 模型	BERT-Base-Uncased
第一层 MLP 输入维度	768
第一层 MLP 输出维度	300
第二层 MLP 输入维度	300
第二层 MLP 输出维度	2
损失函数	交叉熵损失
激活函数	ReLU

4.3 基线模型

为了证明我们模型的有效性，我们将本文提出的模型与其他先进模型进行对比，具体如下：

(1) BiGCN^[16]:利用双向图神经网络考虑新闻正向传播和反向传播的虚假新闻检测模型。

(2) UPFD^[17]:通过 BERT 对新闻传播者历史发文进行编码作为传播者发文偏好，并采用将新闻内容与新闻传播结构相结合的实现虚假新闻检测。

(3) SVM^[31]: 利用支持向量机，提取新闻文本特征，进行虚假新闻检测。

(4) RFC^[32]: 根据新闻特征，利用随机森林分类器进行虚假新闻检测。

(5) KAN^[33]: 利用 Transformer 的编码器和注意力机制将外部知识用于虚假新闻检测。将新闻内容中的实体链接到 Wikidata 知识图谱^[34]中的实体和实体上下文，把知识图谱中的实体和实体上下文作为外部知识。

表 4 算法性能对比

Table4 Algorithm Performance

数据集	指标	BiGCN	UPFD	SVM	RFC	KAN	USPGCN
GossipCop	Precision	0.9279	0.9691	0.7493	0.7015	0.7764	0.9739
	Recall	0.8481	0.9731	0.6254	0.6707	0.7696	0.9803
	F1	0.8897	0.9707	0.5955	0.6691	0.7713	0.9761
	Accuracy	0.8902	0.9709	0.6643	0.6918	0.7766	0.9767
PolitiFact	Precision	0.8039	0.8714	0.746	0.7470	0.8687	0.9048
	Recall	0.7626	0.8103	0.6826	0.7361	0.8499	0.8125
	F1	0.7871	0.8425	0.6466	0.7362	0.8539	0.8586
	Accuracy	0.7883	0.8431	0.6694	0.7406	0.8586	0.8597

表 4 为本文提出的 USPGCN 与传统机器学习模型 SVM、RFC，基于外部知识的深度学习模型 KAN，以及基于传播的深度学习模型 BiGCN、UPFD 进行比较，USPGCN 模型在 GossipCop 数据集和 PolitiFact 数据集上的精确率为 0.9739 和 0.9048。相比于对比的实验中表现最好的模型 UPFD，USPGCN 模型的精确率分别提升了 0.48% 和 3.34%，这说明传播者情感偏好和池化函数能更深层次、更有效地区分真假新闻在传播方面的不同。与基于外部知识的模型 KAN 相比，USPGCN 的精确率等指标高于 KAN 的指标。

由此可见，本文提出的模型高于其他检测模型，验证了本文所提出的模型的有效性。

4.4 消融实验

为了验证本文提出的模型的各个模块对实验结果的影响，我们设计了模型变体，对模型进行消融分析，如表 5 所示。

四种消融模型的解释如下：

(1)-USP:去除情感特征提取器,不使用情感特征提取器,去除原模型中的 MLP,直接使用 BERT 的输出。

(2)-Pooling:去除本文使用的池化函数,使用全局最大池化替代原本模型中的池化函数。

(3) - P&-U:将情感特征提取器和池化函数一起去除。

(4) -Fusion 去除信息融合模块,将图卷积神经网络的输出直接输入分类器。

表 5 消融实验

Table5 Ablation Study

数据集	指标	-USP	-Pooling	-P&-U	-Fusion	USPGCN
GossipCop	F1	0.9756	0.9745	0.9707	0.9730	0.9761
	Accuracy	0.9757	0.9746	0.9709	0.9731	0.9767
PolitiFact	F1	0.8547	0.8450	0.8425	0.8356	0.8586
	Accuracy	0.8552	0.8462	0.8431	0.8371	0.8597

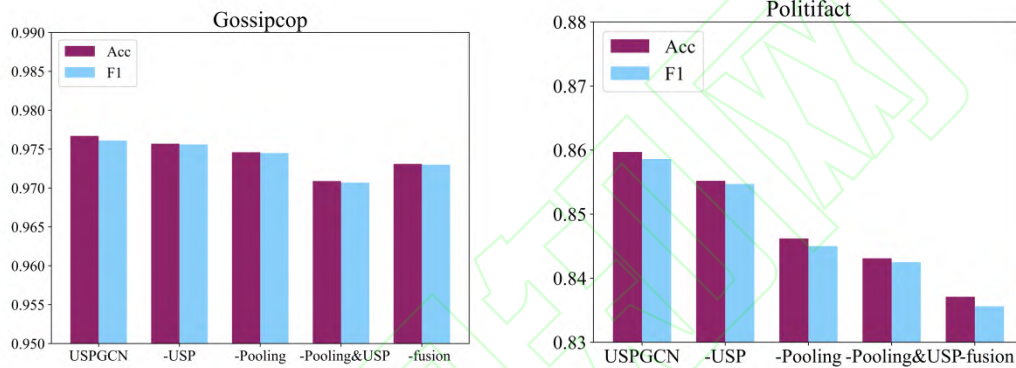


图 3 消融实验结果

Fig.3 Results of Ablation Experiments

从图 3 可以看出,去除任意模块都会一定程度上降低模型性能。具体来说,不加入情感特征提取器,也就是不提取传播者情感偏好,模型性能就会降低,这说明相比于单纯的挖掘传播者偏好,传播者情感偏好可以更深层的挖掘传播者偏好,可以让模型更准确的捕捉新闻的传播模式,从而获得更好的效果。

对于池化函数,可以看出去除本文提出的池化函数,改用全局最大池化,模型的性能也有所降低,这说明将全局最大池化和全局平均池化一起使用,相比于使用单一的池化函数,能更好的考虑图卷积神经网络的卷积层提取出的特征,从而达到提升模型性能的作用。

将新闻传播者情感偏好和池化函数模块一起去除,可以看出与单独去除传播者情感偏好或单独去除池化函数相比,模型性能大幅下降,这更加说明通过新闻传播者情感偏好和池化函数可以深入挖掘新闻传播者发文偏好,从而提高模型性能。

信息融合模块也一定程度上的提高了模型性能,说明将文本信息和传播图的信息相融合有助于丰富信息,提高检测能力。

5 结论

针对现有的基于传播者偏好的虚假新闻检测研究,大多利用传播者发文的语义信息特征,未将充分提取潜在的情感偏好,并用于虚假新闻检测的问题,我们提出了一种融合新闻传播模式和传播者情感偏好的虚假新闻检测模型

(USPGCN)。在公开的数据集 GossipCop、PolitiFact 上的实验结果表明, 我们的模型 USPGCN 与当前较好的模型作比较, 具有较高的准确率。另一方面, 我们构造了一种混合的池化函数, 将全局最大池化和全局平均池化结合在了一起。消融实验表明, 作为一种混合的池化函数, 与单一池化函数相比, 提高了检测的准确率。

实验结果证明了我们所提出的模型的有效性和合理性。然而本文仍存在一些不足, 未来将尝试考虑传播者跟风发文等异常情况下的传播者情感偏好提取。此外本文仅尝试在英文真实数据集上进行实验, 且仅涉及到政治新闻和娱乐新闻, 未来将尝试在中文语境的真实数据集或其他领域的新闻数据集上进行实验分析。

参考文献:

- [1] 闵庆飞, 刘晓丹. 谣言研究综述: 基于媒介演变的视角[J]. 情报杂志, 2015, 34(4): 104-109. (Min Qingfei, Liu Xiaodan. A Review of Rumor Study Based on Media Evolution[J]. Journal of Intelligence, 2015, 34(4): 104-109.)
- [2] 刘帅, 傅丽芳. 融合外部知识和用户交互特征的虚假新闻检测[J]. 数据分析与知识发现, 2023, 7(11): 79-87. (Liu Shuai, Fu Lifang. Identifying Fake News with External Knowledge and User Interaction Features[J]. Data Analysis and Knowledge Discovery, 2023, 7(11): 79-87.)
- [3] 李悦晨, 钱玲飞, 马静. 基于 BERT-RCNN 模型的微博谣言早期检测研究[J]. 情报理论与实践, 2021, 44(7): 173-177. (Li Yuechen, Qian Lingfei, Ma Jing. Early Detection of Micro Blog Rumors Based on BERT-RCNN Model[J]. Information Studies: Theory & Application, 2021, 44(7): 173-177.)
- [4] 张国标, 李洁. 融合多模态内容语义一致性的社交媒体虚假新闻检测[J]. 数据分析与知识发现, 2021, 5(5): 21-29. (Zhang Guobiao, Li Jie. Detecting Social Media Fake News with Semantic Consistency Between Multi-model Contents[J]. Data Analysis and Knowledge Discovery, 2021, 5(5): 21-29.)
- [5] Ajao O, Bhowmik D, Zargari S. Sentiment Aware Fake News Detection on Online Social Networks[C]. ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019: 2507-2511.
- [6] Luvembe A M, Li W, Li S, et al. Dual Emotion Based Fake News Detection: A Deep Attention-Weight Update Approach[J]. Information Processing & Management, 2023, 60(4): 103354.
- [7] Zhu Y, Sheng Q, Cao J, et al. Generalizing to the Future: Mitigating Entity Bias In Fake News Detection[C]. Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2022: 2120-2125.
- [8] 曾江峰, 程征, 黄泳潼等. 融合 BERT 和主题模型的谣言检测方法[J/OL]. 情报科学:1-27[2023-10-12]. <http://kns.cnki.net/kcms/detail/22.1264.G2.20230915.1108.009.html>. (Zeng Jiangfeng, Cheng Zheng, Huang Yoongtong, et al. A Rumor Detection Method Integrating BERT and Topic Model[J/OL]. Information Science: 1-27[2023-10-12]. <http://kns.cnki.net/kcms/detail/22.1264.G2.20230915.1108.009.html>.)
- [9] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding[J]. arXiv Preprint arXiv: 1810.04805.
- [10] 刘赏, 沈逸凡. 基于新闻标题-正文差异性的虚假新闻检测方法[J]. 数据分析与知识发现, 2023, 7(2): 97-107. (Liu Shang, Shen Yifan. A Fake News Detection Method Based on News Title-Content Difference[J]. Data Analysis and Knowledge Discovery, 2023, 7(2): 97-107)
- [11] Shu K, Mahudeswaran D, Wang S, et al. Hierarchical Propagation Networks for Fake News Detection: Investigation and Exploitation[C]. Proceedings of the International AAAI Conference on Web and Social Media, 2020, 14: 626-637.
- [12] Vosoughi, Soroush, Deb Roy, and Sinan Aral. The Spread of True and False News Online[J]. Science, 2018,

359(6380): 1146-1151.

- [13] Rosenfeld N, Szanto A, Parkes D C. A Kernel of Truth: Determining Rumor Veracity on Twitter by Diffusion Pattern Alone[C]. Proceedings of the Web Conference 2020. 2020: 1018-1028.
- [14] Ma J, Gao W, Wong K F. Rumor Detection on Twitter with Tree Structured Recursive Neural Networks[C]. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, 2018: 1980–1989.
- [15] Yu F, Liu Q, Wu S; Wang L, et al. A Convolutional Approach for Misinformation Identification[C]. Proceedings of the 26th International Joint Conference on Artificial Intelligence, 2017: 3901-3907.
- [16] Bian T, Xiao X, Xu T, et al. Rumor Detection on Social Media with Bi-directional Graph Convolutional Networks[C]. Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(1): 549-556.
- [17] Dou Y, Shu K, Xia C, et al. User Preference-aware Fake News Detection[C]. Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021: 2051-2055.
- [18] Păvăloaia V D, Teodor E M, Fotache D, et al. Opinion Mining on Social Media Data: Sentiment Analysis of User Preferences[J]. Sustainability, 2019, 11(16): 4459.
- [19] Zhu X, Wu J, Zhu L, et al. Exploring User Historical Semantic and Sentiment Preference for Microblog Sentiment Classification[J]. Neurocomputing, 2021, 464: 141-150.
- [20] Zhou Q. Detecting the Public's Information Behaviour Preferences in Multiple Emergency Events[J]. Journal of Information Science, 2023, 49(4): 871-886.
- [21] Vidanagama D U, Silva A T P, Karunananda A S. Ontology Based Sentiment Analysis For Fake Review Detection[J]. Expert Systems with Applications, 2022, 206: 117869.
- [22] Fu C, Chen K, Pan X, et al. Rumor Detection Based on the Temporal Sentiment[C]. China National Conference on Big Data and Social Computing. Singapore: Springer Nature Singapore, 2022: 275-290.
- [23] Moschopoulos V, Tsourma M, Drosou A, et al. Misinformation Detection Based on News Dispersion[C]. 2023 24th International Conference on Digital Signal Processing (DSP). IEEE, 2023: 1-5.
- [24] Hutto C, Gilbert E. Vader: A parsimonious Rule-based Model for Sentiment Analysis of Social Media Text[C]. Proceedings of the International AAAI Conference on Web and Social Media. 2014, 8(1): 216-225.
- [25] 王昊, 龚丽娟, 周泽聿等. 融合语义增强的社交媒体虚假信息检测方法研究[J]. 数据分析与知识发现, 2023, 7(2): 48-60. (Wang Hao, Gong Lijuan, Zhou Zeyu. A Study on Social Media False Information Detection Method Incorporating Semantic Enhancement[J]. Data Analysis and Knowledge Discovery, 2023, 7(2): 48-60.)
- [26] Xiao Y, Li C, Thürer M, et al. User Preference Mining Based on Fine-grained Sentiment Analysis[J]. Journal of Retailing and Consumer Services, 2022, 68: 103013.
- [27] Hamilton W, Ying Z, Leskovec J. Inductive Representation Learning on Large Graphs[J]. Advances in Neural Information Processing Systems, 2017: 30.
- [28] Tong Z, Aihara K, Tanaka G. A Hybrid Pooling Method for Convolutional Neural Networks[C]. Neural Information Processing: 23rd International Conference, ICONIP 2016, Kyoto, Japan, October 16-21, 2016, Proceedings, Part II 23. Springer International Publishing, 2016: 454-461.
- [29] Shu K, Mahudeswaran D, Wang S, et al. FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media[J]. Big Data, 2020, 8(3): 171-188.
- [30] THINKNOOK. Twitter Sentiment Analysis Training Corpus (Dataset) [EB/OL]. [2023-3-30]. <http://thinknook.com/twitter-sentiment-analysis-training-corpus-dataset-2012-09-22>
- [31] Yang F, Liu Y, Yu X, et al. Automatic Detection of Rumor on Sina Weibo[C]. Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics. 2012: 1-7.
- [32] Kwon S, Cha M, Jung K, et al. Prominent Features of Rumor Propagation in Online Social Media[C]. Proceedings of the IEEE 13th International Conference on Data Mining. 2013: 1103–1108.
- [33] Dun Y, Tu K, Chen C, et al. KAN: Knowledge-aware Attention Network for Fake News Detection[C].

Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(1): 81-89.

[34] Vrandečić D, Krötzsch M. Wikidata: A Free Collaborative Knowledgebase[J]. Communications of the ACM, 2014, 57(10): 78–85.

通讯作者 (Corresponding author) : 崔鹏 (Cui Peng) , ORCID: 0000-0001-5911-3008, E-mail: pcui@mail.gufe.edu.cn。

基金项目: 本文得到了国家自然科学基金 (项目编号: 21963004)、贵州省教育厅科技拔尖人才项目 (项目编号: 黔教技[2022]080) 以及贵州省高等学校区块链与金融科技重点实验室建设项目 (项目编号: 黔教技[2023]014) 的支持。

The work is supported by and the National Natural Science Foundation of China (No. 21963004), Guizhou Provincial Department of Education Science and Technology Top Talent Project (QJJ[2022]080), and the Key Laboratory Program of Blockchain and Fintech of Department of Education of Guizhou Province (QJJ[2023]014).

作者贡献声明:

蒋涛: 设计研究方案, 进行实验, 论文起草与修改;

潘云辉: 方法指导, 论文修改;

崔鹏: 提出研究思路, 论文修改及最终版本修订。

利益冲突声明:

所有作者声明不存在利益冲突关系。