

MMCSA: 一种跨模态的假新闻检测方法

赵越¹, 郝琨¹, 赵敬^{2,3}, 信俊昌^{2,3}

(1. 东北大学 医学与生物信息工程学院, 辽宁 沈阳 110169; 2. 东北大学 计算机科学与工程学院, 辽宁 沈阳 110169; 3. 辽宁省大数据管理与分析重点实验室, 辽宁 沈阳 110169)

摘要: 目前基于新闻内容的假新闻检测方法没有考虑到不同模态更高层的语义关联, 缺少可以依据的信息对新闻进行判断, 从而缺乏对有重要辨别特征的新闻的社交网络信息进行有效使用. 针对这个问题, 提出了基于新闻内容的假新闻检测方法, 通过提取文本、图像和视频等多模态新闻的高层语义特征, 分析不同模态高层语义信息, 设计跨模态主题一致性和跨模态情感一致性计算方法. 在此基础上, 设计了一种跨模态内容语义一致性的假新闻检测模型 MMCSA (multi-modal feature content semantic consistency). 实验证明, 相比于传统方法, 所提出的 MMCSA 有较好的检测效果.

关键词: 假新闻检测; 内容语义一致性; 跨模态主题一致性; 跨模态情感一致性

中图分类号: TP 391 文献标志码: A 文章编号: 1005-3026(2024)01-0018-08

MMCSA: A Cross-Modal Approach to Fake News Detection

ZHAO Yue¹, HAO Kun¹, ZHAO Jing^{2,3}, XIN Jun-chang^{2,3}

(1. School of Medicine & Biological Information Engineering, Northeastern University, Shenyang 110169, China; 2. School of Computer Science & Engineering, Northeastern University, Shenyang 110169, China; 3. Key Laboratory of Big Data Management and Analytics, Liaoning Province, Shenyang 110169, China. Corresponding author: XIN Jun-chang, E-mail: xinjunchang@mail.neu.edu.cn)

Abstract: Current fake news detection methods based on news content do not take into account the higher-level semantic correlation of different modalities, and lack information that can be used to judge news, thus lacking effective use of social network information for news with important distinguishing features. Address to this problem, a fake news detection method based on news content is proposed. By extracting high-level semantic features of multi-modal news such as text, images and videos, the high-level semantic information of different modalities is analyzed, and the cross-modal topic consistency and cross-modal emotional consistency are designed. On this basis, a fake news detection model MMCSA (multi-modal feature content semantic consistency) is designed with cross-modal content semantic consistency. Experiments show that the proposed MMCSA has better detection effect than the traditional method.

Key words: fake news detection; content semantic consistency; cross-modal topic consistency; cross-modal emotional consistency

随着近年来互联网的普及, 人们获得信息途径更加便利, 为假新闻传播提供了土壤^[1]. 以医疗领域为例, 假的医疗新闻可能会误导大众的医学认知, 进而传播大量“伪医学”, 甚至有些人宁肯相信网络上的虚假信息而非医生的建议^[2-3]. 比如所谓的“偏方治疑难杂症”“枸杞补肾”等, 人们并不关心信息的来源是否可靠, 仅仅相信假新

闻的一面之词, 而且听到的重复信息越多, 就对此越深信不疑^[4]. 假医学新闻的传播几乎没有任何代价, 而且容易病毒式传播, 严重损害了公众的权益和健康. 因此, 准确检测出虚假的医学新闻, 将错误的医学信息排除于大众视野之外, 具有十分重要的实际意义^[5].

现有的检测假医学新闻的方法可分为两

收稿日期: 2022-07-28

基金项目: 国家自然科学基金资助项目(62072089); 中央高校基本科研业务费专项资金资助项目(N2116016, N2104001, N2019007, N2224001-10).

作者简介: 赵越(1966-), 女, 辽宁沈阳人, 东北大学教授.

类^[6-7]:一类是根据新闻内容利用专家系统、集体智慧、事实真理等信息对新闻进行检测,但是,依赖新闻内容检测假新闻可能会出现信息闭塞的情况,即没有可以依据的信息对新闻进行判断,无法从新闻文本内容本身判断真假^[8-9];另一类是利用社交传播过程中产生的传播信息对新闻进行检测^[10-11],但是,依赖社交环境检测虚假新闻会导致时间延迟,即当虚假新闻开始出现时,第一时间无法判断真假,需要传播信息的补充才可以做出判断。

文献[12]提出运用卷积神经网络(CNN)进行文本分类,进而检测真假新闻.Rashkin等^[13]将假新闻与真实新闻进行新闻语言比较,然后发现了不可靠文本的语言特征.Castelo等^[14]发现新闻涵盖了各种各样的主题和语言风格,提出了一个新的主题不可知分类策略.文献[15-16]通过分析新闻网页的内容来确定新闻的真实性.Chen等^[17]提出利用循环神经网络(RNN)捕获相关帖子随时间变化的上下文.文献[18]使用CNN-RNN模型提取出新闻的视觉特征,并利用注意力机制融合视觉特征的频域和像素域特征.文献[19]抽取视觉模态语义特征,采用注意力机制提取文本特征并与图像特征进行拼接.Wang等^[20]利用未标注的新闻信息并用对抗神经网络学习新闻的通用特征.文献[21]对社交媒体假新闻内容深入剖析,融合了文本与图像的多种特征,构建出多模态特征融合的假新闻检测模型.文献[22]设计了一种图像与文本内容语义一致性计算方法。

本文的具体工作如下:

1) 为了充分利用不同模态之间的语义关联,提取不同模态的语义特征.设计跨模态主题一致性和跨模态情感一致性计算方法,解决了缺少考虑内容语义关联的问题,提高假新闻检测的准确度;

2) 在此基础上,设计了跨模态内容语义一致性的假新闻检测模型MMSC;

3) 使用真实数据集,与现有算法进行大量的对比实验,验证所提出方法的有效性。

1 跨模态的假新闻检测算法

本节提出跨模态主题一致性和跨模态情感一致性计算方法,实现多模态高层语义信息之间的互补,提高了基于新闻内容的假新闻检测的准确度。

1.1 跨模态主题一致性

为充分挖掘出文本信息与图像信息之间的关联性和交互性特征,首先提取文本关键词,得到图像的关键词,并利用神经网络模型训练出图像与文字的映射关系;然后,通过计算文本主题词与图像主题词之间的相似度;最后,计算跨模态主题一致性向量.本文文本模态采用Word2Vec词向量表示方法,根据主题词向量表示,利用向量之间的余弦距离,如式(1)所示,计算出主题词之间的相似度值。

$$\cos\theta = \frac{\sum_{i=1}^n (A_i \times B_i)}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} = \frac{A \times B}{|A| \times |B|}. \quad (1)$$

式中, A, B 为两个词向量。

算法1描述了跨模态主题一致性向量的计算过程^[23].首先,利用关键帧提取方法得到视频的图像信息;其次,利用TextRank算法^[24]提取出新闻文本关键词,并借助百度API工具得到图像的实体和场景作为图像主题词;然后,根据训练好的词向量模型并调用相似度方法,计算图像每个主题词与文本主题词之间的相似度,得到图像主题词 I_i^{topic} 与文本主题词之间的相似度最大值 Similarity_n .最后,将图像主题词的概率值 $\text{value}_n = \{V_1, V_2, \dots, V_5\}$ 与对应相似度值 Similarity_n 相乘计算出跨模态主题一致性向量 f_{topicSim} .

算法1 跨模态主题一致性向量

输入:新闻的文本、图像、视频信息 text, img, video.

输出:跨模态主题一致性向量 f_{topicSim} .

1. $V_Image_i \leftarrow$ 提取 i 个视频的关键帧
2. $T_n^{\text{topic}} \leftarrow \text{getTextTopic}(\text{text})$ //提取文本主题词
3. $I_m^{\text{topic}} \leftarrow \text{getImgTopic}(\text{img})$ //提取图像主题词
4. $V_k^{\text{topic}} \leftarrow \text{getImgTopic}(V_Image_i)$ //提取关键帧主题词
5. $IV_{n+k}^{\text{topic}} \leftarrow I_n^{\text{topic}} + V_k^{\text{topic}}$ //将关键帧主题词合并
6. //计算文本与图像主题词相似度
7. //初始化文本图像主题相似度 Similarity 列表
8. for 图像主题词 IV_{n+k}^{topic} do //遍历图像主题词
9. $\text{maxSimi} \leftarrow 0$ //相似度值初始值为 0
10. for 文本主题词 T_n^{topic} do //遍历图像主题词
11. $\text{topicSimilarity}_{ij} \leftarrow \text{Word2Vec.similarity}(IV_i^{\text{topic}}, T_j^{\text{topic}})$
12. $\text{maxSimi} \leftarrow \text{Max}(\text{topicSimilarity}_{ij})$
13. $\text{Similarity.append}(\text{maxSimi})$

14. topicSim \leftarrow Similarity_n \times value_{I_n}
18. rankDesc(topicSim)//从大到小进行排序
19. $f_{\text{topicSim}} = [\text{consistency}_1, \dots, \text{consistency}_5]$
20. return f_{topicSim} //返回主题一致性向量

1.2 跨模态情感一致性

如前所述,假新闻可能将真实新闻的视频重新处理,更改语音信息或者图像信息,导致新闻文本信息与语音、图像不一致.为了检测出文本和视频信息不相关的新闻,提出通过判断文本与音频的情感倾向是否一致,进而验证文本和视频信息是否一致.采用情感分析方法,提取出文本和音频特征的情感类别,计算两种模态信息的情感一致性,进而辅助判断视频是否和文本表达同一内容,提高假新闻检测的性能.下面介绍跨模态情感一致性特征具体计算过程.

首先,通过文本情感词典,对文本中的正面情感词、负面情感词、程度副词和否定词进行标记,依次设置上述4类词汇的权重,如表1所示.

表1 情感词典权重设置

Table 1 Emotional dictionary weight setting	
情感词类型	权重设置
正面情感词	2
负面情感词	5
程度副词	most/very/more/ish/insufficiently/over/inverse 4/3/2/0.5/-0.3/-0.5/1
否定词	-1

然后,提取音频的情感特征.通过音频抽取方法,将视频中的音频数据抽取出来,获得音频的梅尔倒频谱系数(MFCC)特征.采用基于MFCC和CNN的语音情感识别算法,提取出音频中的情感特征.使用汉语情感语料库(CASIA)进行语音情感模型训练.根据不同程度的音频情绪设置情感值,取值范围是 $[-1, 1]$.

最后,根据文本、音频的情感特征计算出跨模态情感一致性.本文利用情感词典提取出文本的情感特征 T^{senti} ,采用基于MFCC特征和神经网络的语音情感分析方法得到音频的情感特征 A^{senti} .根据式(2)计算出两种情感特征之间的距离,该距离大小代表文本和音频的情感一致性.

$$\text{distance} = |T^{\text{senti}} - A^{\text{senti}}|. \quad (2)$$

其中, $\text{distance} \in [0, 2]$,这个距离越小表示文本和音频在情感上越一致.

算法2描述了跨模态情感一致性计算过程.首先,本文利用音频抽取方法,将视频的音频信

息提取出来.然后,提取出音频的MFCC特征,将其输入到音频情感分析方法中,获得音频的情感得分.最后,根据文本情感分析方法提取出文本情感得分.

算法2 跨模态情感一致性

输入:新闻文本和视频信息 text, video.

输出:文本音频情感一致性特征 f_{sentiSim} .

1. audio \leftarrow 抽取 video 的音频信息
2. audioFeatures \leftarrow getMFCC(audio)
3. $T^{\text{senti}} \leftarrow$ getTextSenti(text)//得到文本情感分类
4. $A^{\text{senti}} \leftarrow$ getAudioSenti(audioFeatures)
5. $f_{\text{sentiSim}} \leftarrow |T^{\text{senti}} - A^{\text{senti}}|$
6. return f_{sentiSim}

1.3 跨模态内容语义一致性假新闻检测模型

基于新闻内容的假新闻检测任务中,新闻的文本信息发挥着重要作用.首先,利用Word2Vec向量模型对预处理后的文本词语进行向量化表示,得到特征向量 W .然后,由于文本与文本之间存在上下文关系,本文采用双向长短期记忆(Bi-LSTM)进行处理,得到文本特征向量,如式(3)所示.

$$f_{\text{text}} = \text{Bi-LSTM}(W). \quad (3)$$

为了进一步分析新闻的图像、视频信息,提取新闻内容的高层语义特征,本文提出通过跨模态的主题提取及情感分析方法,得到跨模态主题一致性特征向量和跨模态情感一致性特征.通过融合新闻的文本特征与跨模态内容语义一致性特征,设计多因素融合的假新闻检测方法.

为了有效地合并多个模态特征选择线性连接,如式(4)所示.

$$M = f_{\text{text}} \oplus f_{\text{topicSim}} \oplus f_{\text{sentiSim}} \quad (4)$$

其中: f_{topicSim} 表示跨模态主题一致性特征向量; f_{sentiSim} 表示跨模态情感一致性特征向量; M 表示新闻内容多模态特征相连接后的总特征向量;符号 \oplus 表示向量连接操作.

本文利用多层感知器作为新闻内容检测器的分类算法,将新闻内容多模态特征相连接后的总特征向量 M 输入到MLP分类器中,通过模型训练得到新闻内容检测结果.调用sklearn模块的MLP分类器进行模型训练,MLPclassifier方法是MLP分类器的核心方法,本文通过调用sklearn的GridSearchCV模块进行自动选择参数.

跨模态内容语义一致性的假新闻检测模型的总体框架图如图1所示,框架包含三部分,分别是一致性特征提取器(consistency feature

extraction)、新闻内容检测器(content detector)和视频处理(video processing)。视频处理是对新闻的视频数据信息进行处理,分别提取到视频的音频和关键帧信息,关键帧作为视频的图像信息。然后将音频和关键帧输入到一致性特征提取器(consistency feature extraction)中,一致性特征提取器对不同模态特征进行特征提取,得到文本特

征向量、文本主题词、图像主题词以及文本、音频情感特征,通过计算跨模态内容语义一致性,得到跨模态主题一致性特征和跨模态情感一致性特征,最后,利用新闻内容检测器(content detector)将不同模态特征进行合并,进行模型训练,得到新闻内容检测结果。

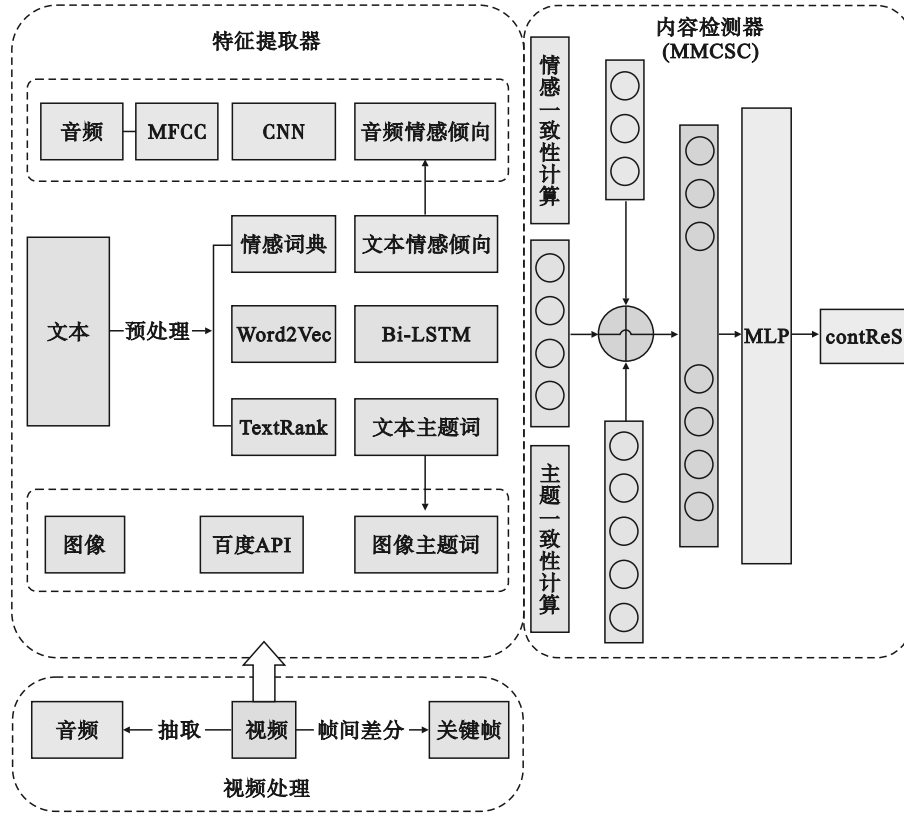


图 1 跨模态内容语义一致性的假新闻检测模型框架

Fig. 1 Framework of fake news detection model for cross modal content semantic consistency

2 实验设计及结果分析

2.1 实验设置

实验环境:本文进行模型训练及开发过程的实验环境配置如表 2 所示。

数据集:所有实验采用公开数据集 Twitter

表 2 实验环境

实验环境	环境配置
操作系统	Windows10 64 位
CPU	Intel(R) Core(TM)i7-8750H CPU
内存	8.00 GB
编程语言	Python3.6
开发工具	IntelliJ IDEA-Community

and Weibo, 本文主要使用 Weibo 数据集。Weibo 数据集被广泛应用在假新闻检测领域中,该数据集每条新闻仅保留一个代表性图片。公开数据集不包含视频数据,但本文需要验证引入音频特征后检测模型的有效性。因此,本文根据新闻中相关的视频链接手动下载对应视频数据。处理后数据集的详细信息如表 3 所示。

表 3 数据集信息

真实新闻	虚假新闻	图片数目	视频数目(补充)
4 749	4 779	3 852	1 200

评价指标:假新闻检测属于二分类任务,分类算法常用的评价指标是准确度、精准度、召回率及 F1 值等。无论是二分类还是多分类,分类正确的样本数占总样本数的比例就是准确度。二分

类问题的预测结果可以根据情况分成以下 4 类, 如表 4 所示.

表 4 二分类预测情况

二分类	阳性	阴性
正	TP(真阳)	TN(真阴)
负	FP(假阳)	FN(假阴)

准确度 Accuracy、精准度 Precision、召回率 Recall 和 F1 值如式(5)~式(8)所示.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (5)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (6)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (7)$$

$$F1 = \frac{2\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (8)$$

2.2 基线方法

为全面测试提出的 MMCSC 算法的性能, 选择不同的基线模型与其进行对比, 分别设置单模态基线模型和多模态基线模型对比实验. 目前, 单模态模型主要有基于单文本的检测模型和基于单视觉模态的检测模型. 为更准确地验证本文提出的跨模态内容语义一致性假新闻检测方法的有效性, 对模型进行消去分析, 设计去掉跨模态主题一致性的 MMCSC_text 模型和 MMCSC_sentiSim 模型以及去掉跨模态情感一致性的 MMCSC_topicSim 模型和 MMCSC_img 模型.

2.2.1 基于单文本模态的检测模型

CNN: 文献[13]提出运用 CNN 进行文本分类, 进而检测真假新闻.

CNN-LSTM: CNN 虽能够获取本地信息, 但是无法获取文本中上下文依赖关系, 长短时记忆网络(long short-term memory, LSTM)则可以解决该问题, 但 LSTM 在提取特征时会出现特征丢失的情况.

2.2.2 基于单视觉模态的检测模型

VGG: 在目前图像处理研究中, VGG 模型被广泛用作视觉特征提取器, 但是该方法难以捕获到图像与图像之间的语义共性.

MVNN: 文献[19]利用 CNN 模型自动捕获了假新闻图像在频域的复杂模式, 使用 CNN-RNN 模型提取出新闻的视觉特征, 然后利用注意力机制融合视觉特征的频域和像素域特征.

2.2.3 基于多模态的检测模型

EANN: 文献[21]提出一种利用对抗神经网络

实现端到端的假新闻检测模型, 该模型利用 VGG19 提取视觉模态特征, 同时采用 Text-CNN 抽取文本模态特征.

MMCSC_topicSim: 通过融合新闻的文本特征和图像特征, 提取出文本与图像、视频的主题一致性向量, 将文本特征向量与跨模态主题一致性特征向量进行拼接, 输入到 MLP 分类器中.

MMCSC_sentiSim: 通过融合新闻的文本特征和音频特征, 将文本特征向量与文本音频情感一致性特征向量进行拼接, 输入到 MLP 分类器中.

MMCSC_img: 对文本、图像的主题进行分析, 得到跨模态主题一致性向量, 将该向量输入到 MLP 分类器中, 不考虑文本特征向量.

2.3 结果分析

对基于多模态的检测模型分别进行消去分析, 得到单模态模型 EANN_text 和 MMCSC_text. 下面从单模态数据和多模态数据两个角度进行分析.

2.3.1 基于单模态数据的实验分析

首先, 比较基于单文本模态的假新闻检测方法, 对比实验分别是基线方法 CNN 和 CNN-LSTM, 为了保证纯文本分类模型对比的公平性, 将基于多模态的检测模型 EANN 以及本文提出的 MMCSC 进行消去分析, 使模型仅输入新闻文本特征, 得到 MMCSC_text 和 EANN_text 算法. 通过实验训练得到 CNN 和 CNN-LSTM 模型的训练损失值趋势图, 如图 2 所示. 训练损失值明显是随着迭代次数的增加而不断降低. 实验结果表明, CNN 模型最终停止训练时的损失值比 CNN-LSTM 模型高, 可以说明 CNN-LSTM 模型训练效果较好, LSTM 获取了文本中上下文依赖关系, 提高了模型的性能.

MMCSC_text 和 EANN_text 模型的训练损失值趋势图如图 3 所示. 实验结果表明, EANN_text 模型的损失值收敛速度较快, 在第 15 次迭代时模型便达到了较优的训练结果, 第 20 次迭代时模型趋于稳定. 但是最终模型停止训练时 MMCSC_text 模型的损失值低于 EANN_text. 因此, EANN_text 模型比 MMCSC_text 模型更快达到较好的训练结果, 并且 EANN_text 模型更稳定, 但是最终 MMCSC_text 模型训练效果优于 EANN_text 模型.

在此基础上, 计算以上 4 种基于单文本模态测试方法的性能指标, 通过进一步比较基于单模

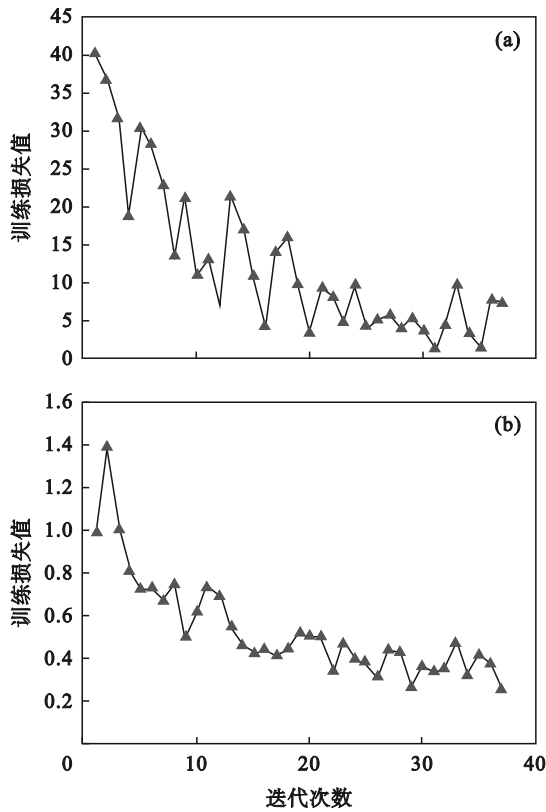


图2 基于单文本检测模型训练损失值趋势图
Fig. 2 Trend chart of training loss value based on single text detection model
(a)—CNN 模型训练损失值;
(b)—CNN-LSTM 模型训练损失值.

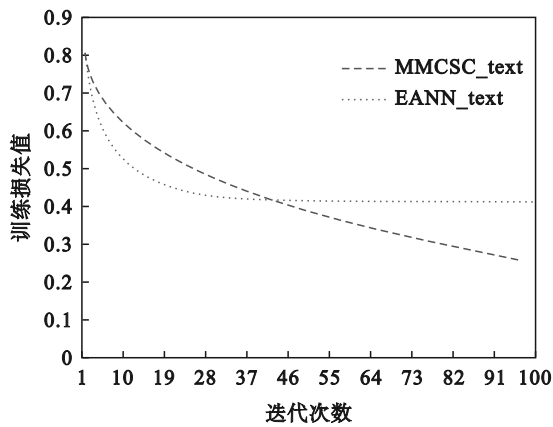


图3 EANN_text,MMCS_text模型训练损失值
Fig. 3 Training loss value of EANN_text and MMCS_text

态数据的检测模型方法,验证MMCS模型的有效性,模型性能指标计算结果如表5所示.基于单文本的检测模型分类效果均优于单视觉模型,说明新闻的文本特征在假新闻检测中起重要作用,而图像特征可以作为辅助特征,但不能脱离新闻文本特征,仅依靠新闻的图像信息进行假新闻检测.

表5 基于单模态的假新闻检测模型性能指标
Table 5 Performance index of fake news detection model based on single modal

模态	模型	Accuracy	Precision	Recall	F1
单文本	CNN	0.721	0.732	0.731	0.731
	CNN-LSTM	0.761	0.742	0.712	0.733
	EANN_text	0.764	0.747	0.740	0.764
	TestGCN	0.771	0.768	0.767	0.767
	BertGCN	0.774	0.770	0.768	0.769
	MMCS_text	0.780	0.778	0.781	0.770
单视觉	VGG	0.642	0.680	0.701	0.672
	MVNN	0.701	0.721	0.712	0.718

从以上实验可知,MMCS_text模型性能明显高于CNN-LSTM模型.虽然,EANN_text模型比MMCS_text模型更快达到较好的训练结果且EANN_text模型更稳定,但MMCS_text模型检测效果略高于EANN_text模型.

2.3.2 基于多模态数据的实验分析

为进一步验证本文提出的跨模态内容语义一致性假新闻检测方法的有效性,首先,对比MMCS,MMCS_text,MMCS_img,MMCS_topicSim和MMCS_sentiSim模型训练损失值趋势.其中MMCS_text和MMCS_img是基于单模态数据的检测方法,而MMCS,MMCS_topicSim和MMCS_sentiSim是基于多模态数据的检测方法.通过实验训练得到以上5种模型的训练损失值趋势图,如图4所示.

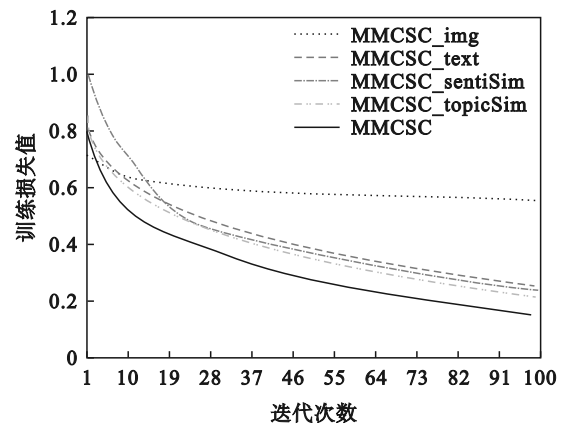


图4 MMCS模型的训练损失值趋势
Fig. 4 Trend of training loss value of MMCS model

从图4中可以看出,MMCS模型训练时损失值的收敛速度最快且最终模型停止训练时的损失值是最低的.根据移除后的模型训练时损失值趋势的收敛情况的实验结果,本文将模型组件的重要程度进行排序,排序结果如下:跨模态主

题一致性<文本<文本 \oplus 跨模态情感一致性<文本 \oplus 跨模态主题一致性。

从以上实验结果可知,对于虚假新闻检测任务,新闻的文本特征在假新闻检测中的作用至关重要,是检测的基础特征,所提取的跨模态主题一致性特征和跨模态情感一致性特征均可以一定程度上提高检测模型的性能。

接下来,对比 EANN 模型与 MMCSC 模型的性能。由于 EANN 模型仅用了新闻的文本及图像特征,未对新闻的视频信息进行分析。因此,采用 Weibo 数据集对比 MMCSC_topicSim 与 EANN 模

型,然后补充视频信息,对比未融合跨模态情感一致性的 EANN 与融合了跨模态情感一致性 MMCSC 模型的性能,得到不同检测模型的评价指标如表 6 所示。结果表明,本文提出的融合跨模态内容语义一致性的假新闻检测方法可以有效提升虚假新闻检测的性能。针对 Weibo 数据集,EANN 模型和 MMCSC_topicSim 模型在性能表现上差别不大,但加入了视频数据后,MMCSC 模型性能有了明显提升,尤其是在精度上提高了 2.1 个百分点。

表 6 基于多模态的假新闻检测模型性能指标

Table 6 Performance index of fake news detection model based on multi-modality

数据集	模型类型	Accuracy	Precision	Recall	F1
Weibo	EANN	0.803	0.835	0.682	0.786
	MMCSC_topicSim	0.808	0.820	0.701	0.793
补充视频信息	EANN	0.803	0.836	0.691	0.787
	MMCSC	0.815	0.857	0.701	0.806

3 结 语

传统假新闻检测方法没有考虑到不同模态更高层的语义关联,从而导致缺少可以依据的信息对新闻进行判断。针对以上问题,本文提出跨模态主题一致性和跨模态情感一致性计算方法,提高了基于新闻内容的假新闻检测的准确度。设计了跨模态内容语义一致性的假新闻检测模型 MMCSC;通过与现有算法进行对比实验,表明了所提方法的有效性。

参考文献:

- [1] 白红义,曹诗语,陈斌.2020年虚假新闻研究报告[J].新闻记者,2021(1):23-37.
(Bai Hong-yi, Cao Shi-yu, Chen Bin. Fake news 2020 research report[J]. *Shanghai Journalism Review*, 2021(1): 23-37.)
- [2] Przybyla P. Capturing the style of fake news[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York, 2020:490-497.
- [3] Lazer D M J, Baum M A, Benkler Y, et al. The science of fake news[J]. *Science*, 2018, 359(6380):1094-1096.
- [4] Jack C. Lexicon of lies: terms for problematic information[J]. *Data & Society*, 2017, 3(22):1094-1096.
- [5] Shu K, Sliva A, Wang S H, et al. Fake news detection on social media[J]. *ACM SIGKDD Explorations Newsletter*, 2017, 19(1):22-36.
- [6] Wang W Y. "Liar, liar pants on fire": a new benchmark dataset for fake news detection [EB/OL]. (2017-05-01) [2022-07-02]. 2017: arXiv: 1705.00648. <https://arxiv.org/abs/1705.00648.pdf>.
- [7] Wang Y Q, Yang W F, Ma F L, et al. Weak supervision for

fake news detection via reinforcement learning [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(1):516-523.

- [8] Vosoughi S, Roy D, Aral S. The spread of true and false news online[J]. *Science*, 2018, 359(6380):1146-1151.
- [9] Rozin P, Royzman E B. Negativity bias, negativity dominance, and contagion [J]. *Personality and Social Psychology Review*, 2001, 5(4):296-320.
- [10] Otto L, Glogger I, Boukes M. The softening of journalistic political communication: a comprehensive framework model of sensationalism, soft news, infotainment, and tabloidization [J]. *Communication Theory*, 2017, 27(2):136-155.
- [11] Sachsman D B. Sensationalism: murder, mayhem, mudslinging, scandals, and disasters in 19th-century reporting[R]. New York:Routledge, 2017.
- [12] Kim Y. Convolutional neural networks for sentence classification [C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, 2014:1746-1751.
- [13] Rashkin H, Choi E, Jang J Y, et al. Truth of varying shades: analyzing language in fake news and political fact-checking [C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen, 2017:2931-2937.
- [14] Castelo S, Almeida T, Elghafari A, et al. A topic-agnostic approach for identifying fake news pages [C]//Proceedings of the 2019 World Wide Web Conference. San Francisco, 2019:975-980.
- [15] Verónica P R, Bennett K B, Alexandra L, et al. Automatic detection of fake news [C]//Proceedings of the 27th International Conference on Computational Linguistics. Santa Fe, 2018:3391-3401.
- [16] Hannah R K, Eunsol C, Jin Y J, et al. Truth of varying shades: analyzing language in fake news and political fact-checking [C]//Proceeding of the Conference on Empirical Methods in NLP. Copenhagen, 2017:2931-2937.
- [17] Chen T, Li X, Yin H Z, et al. Call attention to rumors: deep

