

# 基于TextRank与BERT预训练模型的新闻评论观点句识别方法

王红斌<sup>1,2</sup>, 李伊全<sup>1,2,3</sup>, 李 辉<sup>1\*</sup>

<sup>1</sup>昆明理工大学信息工程与自动化学院, 云南 昆明

<sup>2</sup>昆明理工大学云南省人工智能重点实验室, 云南 昆明

<sup>3</sup>昆明理工大学云南省计算机技术应用重点实验室, 云南 昆明

收稿日期: 2022年4月29日; 录用日期: 2022年6月1日; 发布日期: 2022年6月8日

## 摘 要

由于用户的观点句与新闻内容高度相关,对新闻评论进行观点句识别时需要关注新闻文本这一额外信息。本文针对新闻文本通常很长, BERT并不能很好地处理长序列文本的问题。提出了将TextRank算法与BERT预训练模型相结合的方法,利用TextRank算法从新闻文本中提取出新闻摘要,在不缺失语义信息情况下将较长的新闻文本表示为较短的文本。再将新闻摘要信息与评论通过BERT模型得到语义融合表示向量,最后在全连接层将融合表示向量转换为评论是否为观点句的概率。本文与近年流行的深度学习文本分类模型进行了对比,在准确率上取得了79.80%的最佳效果,说明了模型的有效性。并在NLPC&2012微博观点句识别数据集取得了准确率为80.38%的最佳效果,验证了模型具有一定的泛化能力。

## 关键词

观点挖掘, 观点句识别, 文本分类, 深度学习, 新闻评论

# Opinion Sentences Recognition in News Comments Based on TextRank and BERT Pre-Training Model

Hongbin Wang<sup>1,2</sup>, Yitong Li<sup>1,2,3</sup>, Hui Li<sup>1\*</sup>

<sup>1</sup>Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming Yunnan

<sup>2</sup>Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming Yunnan

\*通讯作者。

文章引用: 王红斌, 李伊全, 李辉. 基于TextRank与BERT预训练模型的新闻评论观点句识别方法[J]. 计算机科学与应用, 2022, 12(6): 1489-1498. DOI: 10.12677/csa.2022.126148

<sup>3</sup>Yunnan Key Laboratory of Computer Technology Application, Kunming University of Science and Technology, Kunming Yunnan

Received: Apr. 29<sup>th</sup>, 2022; accepted: Jun. 1<sup>st</sup>, 2022; published: Jun. 8<sup>th</sup>, 2022

## Abstract

Because the user's opinion sentence is highly related to the news content, the additional information of the news text needs to be paid attention to when identifying the opinion sentence of the news comment. In this article, news texts are usually very long, and BERT cannot handle the problem of long sequence texts well. A method of combining the TextRank algorithm with the BERT pre-training model is proposed. The TextRank algorithm is used to extract news summaries from the news text, and the longer news text is expressed as a shorter text without missing semantic information. Then the news summary information and comments are obtained through the BERT model to obtain the semantic fusion representation vector, and finally the fusion representation vector is converted into the probability of whether the comment is an opinion sentence in the fully connected layer. This paper compares with the popular deep learning text classification model in recent years, and achieves the best effect of 79.80% in accuracy rate, which shows the effectiveness of the model. And in the NLPCC&2012 Weibo opinion sentence recognition dataset, the best effect was achieved with an accuracy rate of 80.38%, which verified that the model has a certain generalization ability.

## Keywords

Opinion Mining, Opinion Sentences Recognition, Text Classification, Deep Learning, News Comments

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

近年来,互联网的高速发展使得社会进入信息化时代,每天都有大量的新闻及对应的海量新闻评论诞生,网络用户通常会对热点新闻发表大量的评论。新闻评论通常由多个句子组成,但并非所有的句子都能表达用户的观点。实际上评论中有一部分句子属于观点句,另一部分并不属于。正确识别热点新闻评论中的观点句能够有效地从评论中筛选出用户对于某篇新闻的态度及看法,从而引导舆情,降低新闻舆情事件所产生的消极影响。因此,对新闻评论进行观点句识别具有重要意义。

根据 NLPCC&2012 对观点句的定义,凡是表达对特定事物或对象的评价的句子都称为观点句,只包含内心自我情感、意愿或心情的句子都不是观点句。本文将观点句识别任务看作二分类任务,即对评论中每个句子进行分类,用标签 Y 来表示观点句,标签 N 表示非观点句。传统的分类方法通常只针对评论进行分类,然而在新闻评论观点句识别中,我们发现用户表达的观点句与新闻的内容高度相关,因此新闻文本信息不能忽视。而 BERT 模型在处理新闻这样的长文本时会截断一部分,导致整体语义缺失。因此在此基础上提出一种将 TextRank [1]算法与 BERT (bidirectional encoder representations) [2]模型结合的方法。首先通过 TextRank 算法提取出新闻文本的几个关键句,将这些关键句组成一个简单摘要,再将新

闻的每条评论与新闻摘要信息一起传到 BERT 模型中获得文本融合表示, 最后送入全连接层, 利用激活函数将全连接层的输出转换为是否为观点句的概率, 通过融入新闻摘要信息来提高观点句识别效果。

## 2. 相关工作

目前, 观点句识别所使用的方法主要有基于情感词典方法、基于机器学习方法和基于深度学习的方法。基于情感词典的观点句识别, 主要是识别句子是否含有所构建的情感词典中的情感词来判断。Kim [3] 等人提出给定一个主题, 模型能自动识别与该主题及情感相近的观点句。该系统包含一个用于确定词语情感的模块和另一个用于在句子中组合情感的模块, 通过建立观点词集, 计算句子中观点词的强度来实现观点句识别。为了识别语料库中表达的情感, 并估计微博数据所表达的情感。Shen [4] 等人对情感词进行分类, 构建了多个词典, 包括否定词词典、程度词词典以及感叹词词典, 首先计算微博条目中每个句子的权重, 然后对第一句和最后一句进行特殊处理, 最后将所有权重相加来综合计算每条微博的情感指数。

基于机器学习的方法主要是将观点句识别作为传统的分类任务, 依据有监督或无监督方式, 选取合适的观点句特征并结合机器学习分类算法实现观点句识别。Pak [5] 等人使用朴素贝叶斯分类器与 n-gram 算法标注出词性, 建立了一个情感分类器, 能够确定一个新闻的积极、消极和中性情感, 从而自动对推特中的评论进行情感分析及观点识别。Turney [6] 提出了一种简单的无监督学习算法, 主要用于英文评论观点句识别, 对评论进行推荐和不推荐分类, 其中词性、标点符号和句子位置等可以作为观点句分类的指导, 并依据先前的语料研究确定句子是否属于观点句。刘培玉[7] 等人提出一种基于隐马尔可夫模型(HMM)的主观句识别方法。该方法首先从训练语料中抽取具有明显分类效果的七类主客观特征, 然后每个句子应用 HMM 进行特征角色类别标注, 并依据标注的结果计算句子的权重, 最终识别主观句。Mihalcea [8] 等人提出在双语词典或平行语料库利之间架起一座桥梁, 用主观词和客观词作为分类特征训练朴素贝叶斯分类器来实现观点句识别。Pang [9] 等人考虑的问题不是根据主题, 而是根据整体情绪, 分别比较了多种分类算法(朴素贝叶斯、最大熵分类和支持向量机), 以及各种特征和特征权值选择策略在基于监督学习的情感分类中的效果。谢丽星[10] 等人提出了一种基于 SVM 的层次结构多策略方法, 包括对于三种情感分析的方法进行了深入研究, 包括表情符号的规则方法、情感词典的规则方法、基于 SVM 的层次结构的多策略方法, 实验表明基于 SVM 的层次结构多策略方法效果最好。Hu [11] 等人通过挖掘客户评价过的产品特征识别每次评论中的意见句, 判断每个意见句是积极的还是消极的, 再总结结果。刘荣[12] 等人提出一种基于语义模式和 BootStrapping 的半监督观点句识别方法。首先对标注样本进行语义模式的抽取, 并将抽取出的观点句语义模式与非观点句语义模式转化为特征; 其次, 将得到的语义模式特征、词汇特征和语法特征运用于小规模标注样本的训练, 从而得到朴素贝叶斯分类器, 继续使用 BootStrapping 的方法对模型进行迭代训练; 最后实现最终的观点句识别。Shi [13] 等人在分析微博的基础上构建了特征集, 将连续的感叹号或问号、网络用词等视为观点句的特征, 并使用 SVM 作为分类器来识别观点句。Wang [14] 等人对中文微博情感分析中的观点句识别和元素提取进行研究。在观点句识别方面, 提出了一种基于情感词和情感影响因子的微博语义情感倾向性计算算法。在元素提取中, 应用主题术语分类和关联规则, 并结合一系列的剪枝、筛选和定界规则提取微博评价对象。通过意见句识别和要素提取的相互过滤, 实现了对观点句的识别。

深度学习方法主要将文本传递给神经网络去获得文本表示, 再用激活函数去获得分类概率。刘丹[15] 等人将总结出的 6 种词性特征表示为向量后与对应的文本词向量进行首尾拼接, 从而解决原始词向量中词性信息较弱的问题, 并送入到卷积神经网络中去对句子进行主客观分类。林思琦[16] 等人使用 LSTM 网络来识别句子的语义信息, 用注意力机制来对观点句相关的信息进行表征, 将句子的情感、主题、位置等观点句特征融入编码层和注意力机制中, 来提升观点句抽取的效果。王晓涵[17] 等人从微博原文抽取

出关键词，将其视为特征与对应的微博评论向量进行拼接，再送入卷积神经网络去进行观点句识别。

从观点句识别的研究现状可以看出，基于情感词典的方法局限性在于有考虑到词与词之间的关系和词语的上下文信息，并且比较依赖情感词典的质量。而基于机器学习的方法需要使用大量特征进行结合来强化观点句识别效果。虽然现在深度学习兴起使得研究人员无需再费心于大量的特征工程，但在观点句识别任务中依然会利用各种特征来辅助观点句分类模型。现在人人都可以利用热门的 BERT 预训练模型去获得良好的文本向量表示，再在预训练模型中对自己的特定任务进行微调。在预训练模型基础上融入外部信息理论上可以进一步提高观点句识别效果。本文的新闻评论观点句识别需要额外关注新闻文本信息，通过将新闻文本与评论一同送入 BERT 模型进行编码，可以将新闻与评论这两种文本的信息进行语义融合。但新闻文本序列过长，BERT 模型会截断一部分文本，导致新闻整体信息缺失。因此本文提出将 TextRank 算法与 BERT 模型相结合的方法，以此来提高观点句识别效果。

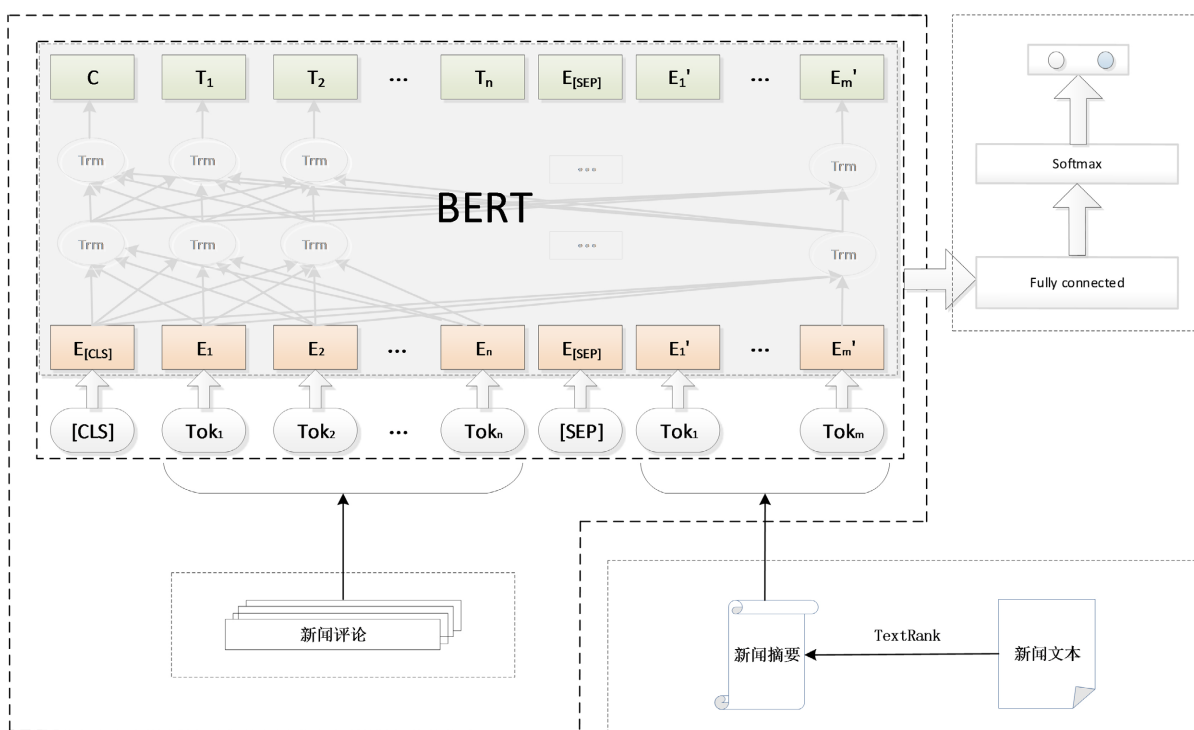


Figure 1. Illustration of model

图 1. 模型示意图

### 3. 模型介绍

新闻评论观点句识别旨在从非结构的新闻评论文本中识别出表达观点的语句。由于新闻评论中用户表达的观点与新闻内容高度相关，所以将新闻摘要信息与评论信息进行深度融合，理论上可以有效的对新闻评论进行观点句识别。该识别方法的流程如图 1 所示。

1) 新闻摘要抽取层。使用 TextRank 将新闻文本中的关键语句抽取出来作为其对应的摘要，可以通过少量语句对新闻文本进行概括性表达，且其对应的文本表示中也包含了新闻的部分语义信息。从文本中提取关键句时，将文本中的每个句子分别看作一个节点，并计算每个句子的 TextRank 分数。计算每个句子分数时，除了考虑链接句的重要性之外，还要考虑两个句子之间的相似度，用相似度来表示两个节点之间的权重。如果两个句子具有相似性，就说这两个句子对应的节点之间存在一条无向有权边。衡量

句子之间相似性如公式(1)所示。

$$\text{Similarity}(V_i, V_j) = \frac{\{w_k | w_k \in V_i \& w_k \in V_j\}}{\log(|V_i|) + \log(|V_j|)}, \quad (1)$$

其中,  $V_i$  和  $V_j$  分别表示句子  $i$  和句子  $j$  的节点, 也代表句子词的集合。  $w_k$  表示句子中的第  $k$  个词, 分子部分用于统计  $w_k$  在两个句子中同时出现的次数, 分母部分中对句子中词的数量求对数之和, 这样可以有效地限制长句子在相似度计算上的优势。根据式(1)计算得出两个节点之间的相似度后, 去除相似度较低的边连接, 从而构建出节点连接图, 每个节点 TextRank 分数计算如公式(2)所示。

$$WS(V_i) = (1-d) + d \times \sum_{V_j \in \text{In}(V_i)} \left( \frac{\text{Similarity}(V_i, V_j)}{\sum_{V_k \in \text{Out}(V_j)} \text{Similarity}(V_j, V_k)} \times WS(V_j) \right), \quad (2)$$

其中  $WS(V_i)$  代表  $V_i$  本身的 TextRank 分数,  $WS(V_j)$  代表  $V_j$  本身的 TextRank 分数,  $d$  为阻尼系数, 防止某个节点的得分为 0。  $\text{In}(V_i)$  代表所有指向节点  $V_i$  的节点集合, 即入链集合。  $\text{Out}(V_j)$  代表节点  $V_j$  指向的所有节点集合, 也就是出链集合。  $V_k$  代表出链集合的第  $k$  个节点。每个句子都要将自身的分数通过计算权重占总权重的比例来分配给其它句子。最开始时会将所有句子的分数初始化为 1, 通过公式(2)来对每个句子的分数进行多次迭代, 收敛时的最后分数就是句子的最终得分, 其中收敛条件是图中的任意一个节点的误差率小于给定的极限值。

接下来对每个句子的 TextRank 分数进行从大到小排序, 分数最高的前  $n$  个句子作为新闻文本对应的摘要  $Abs$ , 如公式(3)所示,  $WS(V_1, \dots, V_n)$  代表新闻所有句子的最终分数。

$$Abs = \text{Top}_n(WS(V_1, \dots, V_n)). \quad (3)$$

2) 文本融合表示层。将抽取得到的文本摘要  $Abs$ , 作为该新闻对应评论语句的背景知识补充, 与评论语句一起输入文本表示层, 针对评论的语义信息与其对应新闻文本的语义信息进行深度融合。BERT 是自然语言处理领域中表现稳定且优秀的预训练模型, 可以通过该模型获取高质量的文本表示。具体来说, 将评论语句  $Obj$  与新闻摘要  $Abs$  一同输入到 BERT 模型中, 如公式(4)所示。

$$R = \text{BERT}[Obj, Abs], \quad (4)$$

其中  $Obj$  为评论语句,  $Abs$  为抽取到的文本摘要,  $R$  为经由 BERT 模型训练得到的融合文本表示。

3) 全连接层和输出层。在得到了融合新闻摘要信息的文本表示后, 将其传入全连接层中。全连接层的意义在于将融合新闻摘要信息的文本表示向量  $R$  转换为评论是否为观点句的概率值。这里使用 Softmax 激活函数来转换概率值。最后将概率值通过输出层输出。如公式(5)所示。

$$y = \text{soft max}(WR + b), \quad (5)$$

其中  $W$  为权重矩阵,  $b$  为偏置项。训练过程使用负对数似然损失函数进行参数更新。

## 4. 实验结果与分析

### 4.1. 实验数据集

目前缺少能应用于新闻评论观点句识别的公共数据集, 因此本文利用爬虫手段从网易新闻爬取了约 200 篇新闻及对应的评论。从中选择了评论数在 300 条以上的 28 篇新闻, 总共 11,898 条评论作为数据集, 涵盖了数码科技、金融、名人明星、刑事案件、民事纠纷、新冠肺炎、国际新闻等领域。经过人工标注后数据集标签分布如下表所示, 其中  $Y$  表示观点句,  $N$  表示非观点句。最后将数据集随机划分为训练集、

验证集和测试集。另外我们使用了 NLPCC&2012 微博观点句识别的公共数据集来测试模型性能，以此来验证模型的泛化能力，该数据集涵盖了多个微博话题下的用户微博文本，并且已经由专业人员清洗了噪声数据。本文模型使用的实验数据集如表 1 所示，公共数据集如表 2 所示。

**Table 1.** Distribution of datasets

**表 1.** 数据集分布

类别	训练集	验证集	测试集
Y	2189	993	988
N	4116	1804	1808
总数	6305	2797	2796

**Table 2.** Distribution of public datasets

**表 2.** 公共数据集分布

类别	训练集	验证集	测试集
Y	1109	546	552
N	607	300	302
总数	1716	846	854

## 4.2. 评价指标

本文采用准确率(Accuracy)、精确率(Precision)、召回率(Recall)、F1 值等常见的指标作为模型的评价指标。四种指标的公式如下：

$$\text{Accuracy} = \frac{TP + TN}{N}, \quad (6)$$

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (7)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (8)$$

$$\text{F1} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (9)$$

其中  $TP$  表示模型正确预测为观点句的样本数， $TN$  表示模型正确预测为非观点句的样本数， $FP$  表示模型错误预测为观点句的样本数， $FN$  表示模型错误预测为非观点句的样本数， $N$  表示评论样本总数， $N = TP + TN + FP + FN$ 。

## 4.3. 参数设置

本文模型设置的文本最大长度序列为 100，Batch size 设为 24，学习率为  $5e-5$ ，Epoch 设为 30。

## 4.4. 实验结果与分析

### 4.4.1. 参数选取实验

由于 TextRank 算法的本质是对文本中的每个句子进行打分并进行排序，选取  $n$  个句子作为关键句，

再将这  $n$  个句子组成摘要。因此本文设置了参数选取实验来探索不同的关键句数量对模型效果的影响。实验结果如表 3 所示:

**Table 3.** The influence of different number of key sentences on model effect

**表 3.** 不同关键句数量对模型效果影响

关键句数量 $n$	准确率	精确率	召回率	F1 值
2	0.7880	0.6908	0.7291	0.7095
3	0.7887	0.7039	0.6989	0.7014
4	0.7804	0.6917	0.6875	0.6896
5	0.7868	0.7111	0.6724	0.6912
6	0.7858	0.6975	0.6996	0.6985

从参数选取的实验上看, 当关键句数量为 3 时有比较好的表现, 因此本模型将  $n$  设置为 3 来构成新闻文本摘要信息。从其它关键句数量对模型影响的结果来看, 可以看出关键句数量并非越多越好, 这可能与 BERT 模型本身不能很好地处理长文本序列有关。

#### 4.4.2. 消融实验

由于本文模型是在 BERT 模型基础上融入了外部新闻摘要信息。因此需要通过消融实验去与 BERT 模型进行对比, 从而验证本文模型的有效性。实验结果如表 4 所示。

**Table 4.** Results of ablation experiments

**表 4.** 消融实验结果

模型	准确率	精确率	召回率	F1 值
BERT	0.7744	0.6839	0.6777	0.6808
Our model	0.7887	0.7039	0.6989	0.7014

从实验结果可以看出, 与 BERT 相比四个指标均有提高, 说明了融入新闻摘要这一外部信息后可以提高观点句识别模型的性能。

#### 4.4.3. 对比实验

由于深度学习在观点句识别上的应用并不多, 因此本文与近年来流行的深度学习文本分类模型进行对比。基准模型采用了 TextCNN、TextRNN、DPCNN、TextRCNN 和 Transformer 模型。下面对五个基准模型进行介绍。

TextCNN [18], 由 Kim 等人于 2014 年提出。核心思想是利用多个不同大小的卷积核来提取句子中的局部特征, 从而能够更好地捕捉局部相关性。

TextRNN [19], 由 Liu 等人于 2016 年提出。核心思想是基于 RNN 设计三种不同的信息共享机制来进行模型训练, 在文本分类任务上取得较好结果。

DPCNN [20], 由 Johnson 等人于 2017 年提出, 主要思想是通过加深网络层来使模型能够捕捉长距离依赖关系。

TextRCNN [21], 由 Lai 等人于 2016 年提出, 主要思想是用 BiRNN 来得到文本的上下文表示, 并与词嵌入进行融合, 之后用 CNN 来获得整个文本的特征向量。同时弥补了双方的缺点。

Transformer [22], 由 Vaswani 于 2018 年提出。核心思想是在编码器-解码器的架构基础上, 利用多头注意力机制来表示输入和输出之间的全局依赖关系。以上所有对比模型均在同一实验数据集下进行, 并且超参数均相同。由于在小节 4.4.2 中的消融实验已经验证了模型融入新闻摘要信息的有效性, 因此基准模型均是在融入新闻摘要信息基础上进行的。实验结果如表 5 所示。

**Table 5.** Comparing experimental results  
**表 5.** 对比实验结果

模型	准确率	精确率	召回率	F1 值
TextCNN	0.7697	0.7485	0.7407	0.7446
TextRNN	0.6917	0.6654	0.6696	0.6675
DPCNN	0.7303	0.7042	0.6997	0.7019
TextRCNN	0.7500	0.7264	0.7186	0.7225
Transformer	0.7135	0.6853	0.6598	0.6723
本文模型	0.7980	0.7144	0.7180	0.7162

从对比实验结果中可以看出, 本文模型在准确率上取得最高结果, 在精确率、召回率和 F1 值上均不如 TextCNN 和 TextRCNN。可以看出本章模型在与 TextCNN 和 TextRCNN 模型对比上并没有明显优势, 经过分析可能原因在于 TextRank 算法是对句子得分进行从大到小排序, 再进行关键句抽取的, 其实得到的摘要信息并没有考虑到句子之间的顺序信息。但在四种评价指标上本章模型性能均优于 TextRNN、DPCNN 和 Transformer 模型, 依然说明了本文模型的有效性。

#### 4.4.4. 同一主题下的实验

为了说明模型的有效性, 本文还需要在某一新闻主题下对模型性能进行评估。本文选择在民事纠纷主题下对模型性能进行评估。实验数据集中符合民事纠纷主题的评论约有 4669 条, 对评论重新随机划分为训练集、验证集和测试集后, 实验在融入了关键句为 3 的新闻摘要信息前提下进行。实验结果如表 6 所示。

**Table 6.** Experiments under the theme of civil disputes  
**表 6.** 民事纠纷主题下的实验

模型	准确率	精确率	召回率	F1 值
TextCNN	0.7224	0.6872	0.6504	0.6683
TextRNN	0.6744	0.6147	0.5524	0.5819
DPCNN	0.7001	0.6680	0.6445	0.6560
TextRCNN	0.7472	0.7162	0.6952	0.7055
Transformer	0.6727	0.6110	0.5639	0.5865
Our model	0.7868	0.6723	0.6452	0.6585



#### 4.4.5. 公共数据集的实验

本文模型和基准模型也在 NLPCC&2012 微博观点句识别的公共数据集进行了实验, 虽然领域不同, 但可以用来验证模型的泛化能力, 实验结果如表 7:

**Table 7.** Experiments on public datasets  
**表 7.** 在公共数据集上的实验

模型	准确率	精确率	召回率	F1 值
TextCNN	0.7155	0.6917	0.6509	0.6707
TextRNN	0.6417	0.4897	0.4994	0.4945
DPCNN	0.6979	0.6656	0.6486	0.6570
TextRCNN	0.7002	0.6684	0.6481	0.6581
Transformer	0.6639	0.6281	0.6223	0.6252
本文模型	0.8038	0.8322	0.8718	0.8515

从结果上看, 本文模型在四个指标上均取得了最高结果, 证明了本文模型有一定的泛化能力, 其原因可能在于公共数据集的规模较小, 使得其它五个基准模型没有学习足够特征, 导致泛化能力较差。而本文模型是以 BERT 模型为基础的, BERT 模型本身在小规模数据集上就有不错的表现。因此本文模型的泛化能力相对较好。

## 5. 结语

本文提出了将 TextRank 算法与 BERT 相结合的新闻评论观点句识别方法, 利用 TextRank 算法从新闻文本中抽取摘要信息, 将摘要信息与评论文本一起送入 BERT 模型中去获得融合文本表示, 再传入全连接层计算是否为观点句的概率。通过实验证明了本文方法的有效性以及具有一定的泛化能力。未来会着力于解决数据集少和样本标签分布不平衡问题。另外也会对 TextRank 忽略文本顺序信息这一问题进行探讨。

## 基金项目

国家自然科学基金项目(61966020)和云南省基础研究计划面上项目(CB22052C143A)。

## 参考文献

- [1] Mihalcea, R. and Tarau, P. (2004) TextRank: Bringing Order into Text. *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona, July 2004, 404-411.
- [2] Devlin, J., Chang, M.W., Lee, K. and Toutanova, K. (2018) Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding, arXiv:1810.04805.
- [3] Kim, S.M. and Hovy, E. (2004) Determining the Sentiment of Opinions. *COLING 2004: Proceedings of the 20th International Conference on Computational Linguistics*, Geneva, 23-27 August 2004, 1367-es. <https://doi.org/10.3115/1220355.1220555>
- [4] Shen, Y., Li, S., Ling, Z., Ren, X. and Cheng, X. (2009) Emotion Mining Research on Micro-Blog. *2009 1st IEEE Symposium on Web Society*, Lanzhou, 23-24 August 2009, 71-75. <https://doi.org/10.1109/SWS.2009.5271711>
- [5] Pak, A. and Paroubek, P. (2010) Twitter as a Corpus for Sentiment Analysis and Opinion Mining. *Proceedings of the International Conference on Language Resources and Evaluation*, Valletta, 17-23 May 2010, 1320-1326.
- [6] Turney, P.D. (2002) Thumbs up or Thumbs down? Semantic Orientation Applied to Unsupervised Classification of

- Reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Philadelphia, 7-12 July 2002, 417-424. <https://doi.org/10.3115/1073083.1073153>
- [7] 刘培玉, 荀静, 费绍栋, 朱振方. 基于隐马尔可夫模型的主观句识别[J]. 中文信息学报, 2016, 30(4): 206-212.
- [8] Mihalcea, R., Banea, C. and Wiebe, J. (2007) Learning Multilingual Subjective Language via Cross-Lingual Projections. *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, Prague, 25-27 June 2007, 976-983.
- [9] Pang, B., Lee, L. and Vaithyanathan, S. (2002) Thumbs up? Sentiment Classification Using Machine Learning Techniques. *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing*, Vol. 10, PA, 6-7 July 2002, 79-86. <https://doi.org/10.3115/1118693.1118704>
- [10] 谢丽星, 周明, 孙茂松, 等. 基于层次结构的多策略中文微博情感分析和特征抽取[J]. 中文信息学报, 2012, 26(1): 73-83.
- [11] Hu, M. and Liu, B. (2004) Mining and Summarizing Customer Reviews. *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Seattle, 22-25 August 2004, 168-177. <https://doi.org/10.1145/1014052.1014073>
- [12] 刘荣, 郝晓燕, 李颖. 基于语义模式的半监督中文观点句识别研究[J]. 南京大学学报(自然科学), 2018, 54(5): 967-973.
- [13] Shi, H., Chen, W. and Li, X. (2013) Opinion Sentence Extraction and Sentiment Analysis for Chinese Microblogs. *2nd CCF International Conference on Natural Language Processing and Chinese Computing*, Chongqing, 15-19 November 2013, 417-423. [https://doi.org/10.1007/978-3-642-41644-6\\_41](https://doi.org/10.1007/978-3-642-41644-6_41)
- [14] Wang, G., Tian, X., Huang, D. and Zhang, J. (2016) Opinion Sentence Identification and Element Extraction in Chinese Micro Blogs. *Journal of Data Acquisition and Processing*, No. 1, 160-167.
- [15] 刘丹. 情感文本的识别与分类算法的研究与实现[D]: [硕士学位论文]. 北京: 北京交通大学, 2019.
- [16] 林思琦, 余正涛, 郭军军, 高盛祥. 融入多特征的汉越新闻观点句抽取方法[J]. 中文信息学报, 2019, 33(11): 101-106.
- [17] 王晓涵, 余正涛, 相艳, 郭贤伟, 黄于欣. 基于特征扩展卷积神经网络的案件微博观点句识别[J]. 中文信息学报, 2020, 34(9): 62-69.
- [18] Kim, Y. (2014) Convolutional Neural Networks for Sentence Classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, Doha, 25-29 October 2014, 1746-1751. <https://doi.org/10.3115/v1/D14-1181>
- [19] Liu, P., Qiu, X. and Huang, X. (2016) Recurrent Neural Network for Text Classification with Multi-Task Learning. arXiv:1605.05101.
- [20] Johnson, R. and Zhang, T. (2017) Deep Pyramid Convolutional Neural Networks for Text Categorization. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, Vancouver, July 2017, 562-570. <https://doi.org/10.18653/v1/P17-1052>
- [21] Lai, S., Xu, L., Liu, K. and Zhao, J. (2015) Recurrent Convolutional Neural Networks for Text Classification. *29th AAAI Conference on Artificial Intelligence*, Austin, 25-30 January 2015, 2267-2273.
- [22] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., et al. (2017) Attention Is All You Need. *31st Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 5998-6008.