

基于情感信息抽取与双向长短期记忆网络的情感分析研究

姜兴盛*, 张超, 毛义坪, 肖良成, 陈秋元, 任秋华

重庆对外经贸学院数学与计算机学院, 重庆

收稿日期: 2025年6月4日; 录用日期: 2025年6月27日; 发布日期: 2025年7月3日

摘要

网络舆情已成为当下不可忽视的研究领域, 利用深度学习技术对舆情文本进行情感分析, 可以通过分析结果对事件采取合适的决策。针对当前网络文本中因文本序列长度变化以及复杂的文本含义而导致的语义情感特征精准提取的问题, 本文提出了一种基于CNN-BiLSTM与多头注意力(CBLMAtt)的舆情分析模型。该模型首先通过对网络文本数据和方面级数据进行训练学习, 利用卷积神经网络(CNN)的变体对文本中的高级语义特征进行提取。其次将提取到的特征作为输入传递到双向长短期记忆网络(BiLSTM)层, 利用该层的特点对文中的上下文特征进行捕获表示。然后将BiLSTM层直接对输入数据进行计算。最后在这两种方法所得到的变量中, 输出隐藏表示被传递到多头注意力层, 使用softmax激活函数的输出层进行情感极性分类。我们在公开的数据集上评估了我们的模型, 在这些数据集上通过精确率的平均值和F1值的对比表明, 与传统的模型相比, 我们的方法相比这些模型都提升两个百分点以上。此外, 我们也进行了消融实验, 以显示不同文档级别权重对学习技术的影响。

关键词

注意力机制, 情感分析, CNN, BiLSTM

Sentiment Analysis Research Based on Emotional Information Extraction and Bidirectional Long Short-Term Memory Network

Xingsheng Jiang*, Chao Zhang, Yiping Mao, Liangcheng Xiao, Qiuyuan Chen, Qiuhua Ren

School of Mathematics and Computer Science, Chongqing College of International Business and Economics, Chongqing

*通讯作者。

文章引用: 姜兴盛, 张超, 毛义坪, 肖良成, 陈秋元, 任秋华. 基于情感信息抽取与双向长短期记忆网络的情感分析研究[J]. 数据挖掘, 2025, 15(3): 228-241. DOI: 10.12677/hjdm.2025.153019

Abstract

Online public opinion has become a research field that cannot be ignored nowadays, and the use of deep learning techniques for sentiment analysis of public opinion texts can be used to take appropriate decisions on events. This paper proposes a public opinion analysis model based on CNN BiLSTM and multi head attention (CBLMAtt) to address the issue of precise semantic and emotional feature extraction in current online texts due to changes in text sequence length and complex text meanings. The model trains and learns from both online text data and aspect level data. Firstly, a variant of Convolutional Neural Network (CNN) is used to extract advanced semantic features from text. Secondly, the extracted features are transferred as input to the bidirectional short-term memory network (BiLSTM) layer, and the context features in the text are captured and represented using the characteristics of this layer. Then, the BiLSTM layer is directly used to calculate the input data. Finally, among the variables obtained by the two methods, the output hiding representation is transferred to the multi head attention layer, and the output layer of the softmax activation function is used to classify the emotional polarity. We evaluated our model on publicly available datasets and compared the average accuracy and F1 values on these datasets, indicating that our method outperforms traditional models by more than two percentage points. In addition, we also conducted ablation studies to demonstrate the impact of different document level weights on learning techniques.

Keywords

Attention Mechanism, Sentiment Analysis, CNN, BiLSTM

Copyright © 2025 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

在全球化趋势下，互联网成为了人们了解世界的主要渠道，然而随着网民日益增多以及大数据时代的到来，互联网上的信息也开始剧增，人们在通过互联网获悉特定新闻事件时产生的观点也大相径庭。网络舆情可以看做是一把双刃剑，一方面它既可弘扬积极向上、促进社会稳定和谐发展的正能量，另一方面又可被居心叵测之人利用，随意大肆散播负面言论，企图误导公众，从中获取利益并影响社会和谐。且如今自媒体盛行，绝大多数网民在了解信息时或多或少都会接触到自媒体，越是关注人数越多的自媒体，其发表的言论越容易对广大网民产生引导性，而自媒体往往只想着扩大关注度和流量，导致在陈述新闻事件时带有一定的主观性和片面性，从而影响网民的情绪、意见和判断。

有关舆情情感分析的文本数据分析在国内自然语言领域中从广义上大致分为三类：基于规则情感词典[1]、基于机器学习[2]、基于深度学习[3]。

在基于规则情感词典方面主要是构建与之相匹配的情感词典来进行规则的设置，因而用以提升对文本数据中舆情情感分析的准确性。吴杰胜[4]等为追求微博文本情感分析的准确性，故而从舆情情感词典着手，进行改进和扩充，在其中加入了构造程度副词、微博领域词典和否定词词典。陈珂[5]等通过情感词典中的特征信息来对 Transformer 模型改善，在一定程度上提高了分类效果。虽然研究者们可以通过需求来增添规则词典内容，但是却需要大量精力去标注数据以及构造词典。在基于机器学习方面则是需要

通过选取特征进行词性标注,将原始数据集按照一定比例划分为训练集和测试集两部分,然后再利用朴素贝叶斯[6]、支持向量机[7]等分类器进行情感分析。为了有效提高情感情绪和强度计算的准确性,戚天梅等[8]致力于在四种机器学习算法的基础上设计情感倾向的计算方法。濮泽堃[9]提出了基于评论分析和集成分类的情感分析 ARAEC 模型,并且将 LS-SVM 模型通过多种算法优化调整最终改进为 SD-LS-SVM 情感分类算法,最后将两者进行融合,构建出基于机器学习的电商评论情感分析系统。在深度学习方面的舆情情感倾向分析可以在一定程度上弥补计算学习算法在上下文文本间关联性的缺陷。Lu 等人[10]通过方面门控图卷积网络模型利用句法信息和情感依赖关系进行基于方面的情感分析,在多个半实物数据集上进行实验分析,在精确度和宏 F1 方面比基线模型提高了 2.14%和 1.33%。在近几年的新冠疫情中,Nemes 和 Kiss [11]基于 COVID 19 (评论、标签、帖子、推文)对社交媒体进行情绪分析,虽然文献[12]和文献[13]所叙述的新冠疫情在全球范围产生了影响,但文献[11]的研究用递归神经网络进行分析,得出社交媒体上依然存有更多积极的推文。

神经网络在自然语言领域取得了非常不错效果,但是在关于舆情分析方面大多只是单纯针对文本主题内容或者文本中的语义特征进行分析研究,针对这一方面问题,本文提出一种基于卷积神经和双向长短时记忆神经的多头注意力模型(CBLMAtt)。首先对文本的词向量进行卷积操作,通过学习语义特征来提取文本语义情感特征向量,然后通过对语义情感特征进行长短时记忆神经层训练得到隐藏平均值,最后通过多头注意力进行权重划分与重组,对舆情文本进行分析。

本文主要贡献如下:

- 论文提出了一种结合情感特征和卷积神经网络的情感提取方法,用于对文本中的语义特征和情感特征提取并进行融合成语义情感向量,从而解决文本中语义和情感词性不能兼顾导致的语义情感偏差问题。
- 论文进一步构建了基于 CBLMAtt 的舆情分析模型,用于将提取到的语义情感向量经过双向长短时记忆层训练后和只由双向长短时记忆产生的隐藏向量的平均值结合,然后利用多头注意力机制的多组权重产生多个上下文向量,并对其拼接再线性变换输出,以此来解决文本向量中语义上下文对分析过程的影响。

2. 相关工作

2.1. 神经网络技术

如今深度学习技术发展迅速,很多研究人员将其应用领域转向自然语言方向,使得神经网络模型在文本方面得到极大应用。

目前部分在舆情情感分析方面的研究方法旨在使用文本嵌入能有效解决文本的顺序和可变序列长度等问题。Santos 等人[14]提出了一种被称作 CharSCNN 的深度学习模型,此模型含有对于短文本中情绪分析的字符级表示,这种字符级表示能够为作者采用的 CNN 模型的训练提供补充信息。除文本之外,也有学者偏向于研究深度模型体系结构。例如 J. Wang 等人[15]提出一种被称为区域 CNN-LSTM 的深度学习模型,这种模型在关注句子本身的局部特征信息的同时又兼顾了句子间的依赖关系,研究人员针对 CNN 和 LSTM 的优势精心设计了该模型的体系结构。Yoon 和 Kim [16]就舆情分析依次提出了结合 CNN 与 BiLSTM 的框架。他们先从词汇和情感词典角度入手,生成了一种多通道词嵌入。然后通过 CNN 和 BiLSTM 验证了该模型相对于多个基线的优越性。而 Wang 等人[17]提出的基于卷积神经和递归神经的情绪分析模型,通过顺序架构的 CNN 和 LSTM 层进行实验表明此模型在 Stanford Sentiment Treebank 数据集上获得相对最突出的效果。还有一些学者对于模型体系结构的调整进行研究,用以针对某些特定领域或文本中的情绪分析任务。如 A.A. Al Sallab 等人[18]提出的一种独特的深度学习模型,此模型针对的是

在阿拉伯语上进行舆情文本情感分析,其优势则是来源于发表前新近开发的阿拉伯语情感词库和其他标准词库。Van 等人[19]提出一种融合了卷积和递归神经的句子级情绪分析框架,用来减轻这两种模型的缺陷。Chen 等人[20]提出结合了视觉和文本信息作为两种不同的通道完成情绪分析任务,这种方法在多个数据集上都得到了不错的效果。

Lin 等人[21]利用术语频率-反向文档频率(TF-IDF)和深度学习技术,针对情感和心理健康预测的问题提出一种基于 CNN 的 COR-CNN 模型,并在新浪微博和 Twitter 舆情数据集上的结果表明,他们的模型高于其他三个基线模型。Sagnika S 等人[22]为了提高舆论中主观陈述意见挖掘的准确性,提出了一种基于 CNN 和 LSTM 的主观检测模型。在这项任务中,句子被表示成含有情感信息和部分语音的词嵌入,并且此模型在两个电影评论数据集得到了比其他先进方法更优良的效果。Priyadarshini I 等人[23]提出了一种基于网格搜索的 LSTM-CNN 模型用以分析舆论中的情绪,通过在多个数据集上对比其他一些基线模型发现他们的模型总体准确率更优。Zhang 等人[24]提出了 CNN-LSTM 模型,并依据设计的基于二叉树的情感计算算法对中文情感分类。通过对比不同的传统模型和分类器,最后结果表明其模型具有更大的优势,最后在实际的应急管理问题中检验了其方法拥有更高的分类性能。

2.2. 注意力应用技术

注意力机制在深度学习领域中也属于一个热门研究话题,近几年很多研究者都将其运用在自然语言领域当中,而最早把注意力机制运用在文本情绪分析方面的是 Wang, Y [25]等人,他们把注意力与 LSTM 结合,提出关于方面信息的两种方法。一种是将方面向量进行连接,组成句子级隐含向量来计算出注意力权重。第二种方法是把方向向量和输入字向量相结合。在 SemEval-2014 任务 4 的数据集上实验得出此方法比之前的方法有改进。Liu 和 Zhang [26]利用普通的基于 LSTM 的模型计算出整个句子中不同部分对句子的特定目标的注意权重。同时进一步改进模型,用以区分目标词的前后上下文。在两个基准数据集的结果显示有很大的提升。Ma 等人[27]采用了称为互注意力的机制对观点目标和各自的上下文进行建模。评估他们方法的标准是在 SemEval-2014 任务 4 的数据集的改进上进行的。Chen 等人[28]提出一种基于多注意力的框架来检测观点目标的情绪。此框架包含有五个模块,并在四个标准数据集上进行评估,其结果显示具有性能的提升。Fan 等人[29]研究出一种用来区别评论或评论情绪极性的多注意力机制的神经网络模型,并在四个标准数据集上显示出了效果。Huang 等人[30]提出一种基于神经网络的使用叠层注意力(attention-over-attention)机制学习的方法,此方法使得模型能够共同学习方面和句子的表示,并在两个标准数据集上进行了评估。Yang 等人[31]提出一种协同注意力(co-attention)机制,能够把目标和上下文的注意权重交替进行建模,用以关注目标关键词,以此学习更有效的上下文表示。他们还通过运用 LSTM 和另一种采用 multiple-hops 协同注意机制的 co-attention-MemNet 网络实现了这项技术,实验结果证明了其方法的有效。Zhang 等人[32]通过使用一个含有注意力机制的多层 CNN 模型来实现依赖于目标的情绪分析,这种模型可以对上下文建模表示,并在两个数据集上验证其效果。Gan 等人[33]提出一种基于稀疏注意力机制的扩张 CNN 模型,进行特定的情绪分析,识别情绪极性和处理复杂序列。通过在 Twitter 和评论数据集上进行验证发现此模型比各种标准和融合 LSTM 与 CNN 的模型更加优良。Lei 等人[34]通过运用情感信息增强的基于注意力机制的 LSTM 对数据集进行情绪分析,最后结果表明与其他类型的 LSTM 变体对比,此模型有更好的优势。

3. 模型描述

3.1. CBLMAtt 模型

在这部分中,对本文的模型的架构进行总体描述。此外我们还将介绍了基于 CNN-BiLSTM 与多头注

意力的舆情分析模型(CBLMAtt)各部分组成结构, CBLMAtt 模型基本架构如图 1 所示。

在对舆情文本的民众情感态度预测分析过程中, 使用的神经网络模型类型不同, 其具有的功能也不一样, 而文本的序列也拥有着重要的时序特征。因此本文设计 CBLMAtt 模型, 通过结合情感特征提取的卷积神经层获取序列重构相空间的带有情感的空间特征, 通过 LSTM 得到空间特征下的时空特征, 并采用注意力机制捕获混合神经网络提取到的关键时空特征。

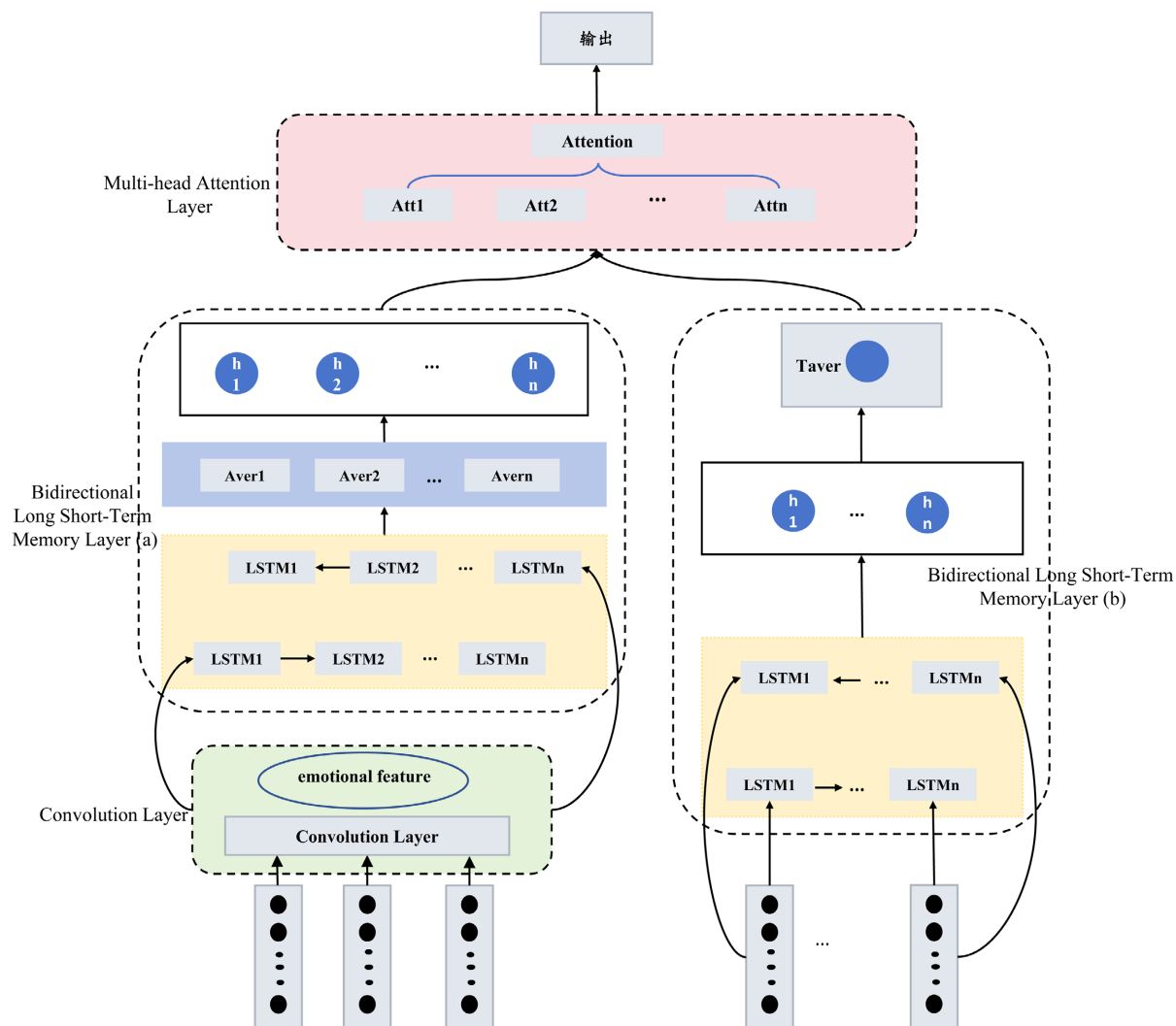


Figure 1. CBLMAtt model structure

图 1. CBLMAtt 模型结构

其各部分功能如下:

(1) 卷积神经网络层: 首先对文本中的情感特征进行提取保留, 然后通过对处理后的文本词向量处理网络输入的重构相空间域 D , 进行卷积操作, 获取带有情感的空间特征并输入到连接层;

(2) 双向长短期记忆层: 将连接层融合得到的张量输入到 BiLSTM 中, 网络根据 CNN 提取的空间特征获取时空特征, 然后将前向和后向传播的向量进行平均值计算;

(3) 注意力层: 计算 CNN 和 BiLSTM 获取的时空特征的平均值的权重, 获得序列的完整特征表示, 根据特征计算预测数值。

此外,在模型中设置了连接层和 Dropout 层,其中连接层是为了将卷积神经网络的输出,即序列的时序特征与原序列 T 进行特征融合。在深度学习模型中加入 Dropout 层其作用主要是为了防止模型训练时出现过拟合,进而提升模型的鲁棒性和泛化能力。

3.2. 卷积层

卷积网络又称卷积神经网络,是一种特殊的深度学习神经网络,常用于处理具有已知网格状拓扑的数据,在时间序列分析、计算机视觉和自然语言处理等领域有着广泛的运用。根据处理数据流的不同,CNN 可以分为一维卷积、二维卷积和三维卷积,其中一维卷积在时间序列分析和自然语言处理领域运用较多。

不管是何种 CNN,其组成主要包括基本的输入层和输出层,核心操作部分为卷积层、池化层(亦称采样层)和全连接层。在一维卷积中,卷积的作用可以理解为提取数据在某一方向的平移特征,在时间序列的分析中则表现为提取序列的特征,在这里卷积操作的本质是循环乘积与加和,设句子长度为 L ,句子情感特征向量 $X \in R^{L \times d}$,则 $X_i \in R^d$ 表示句子中 d 维的第 i 个单词向量。向量 $m \in R^{k \times d}$ 表示为卷积运算的过滤器,其中 k 为过滤器的大小。此外针对句子中从 j 位置起始的窗口向量 W_j 和连续 n 个输入的单词向量,其公式如(1)所示。

$$W_j = [X_j, X_{j+1}, \dots, X_{j+n-2}, X_{j+n-1}] \quad (1)$$

当过滤器 m 和窗口向量 W_j 中的每个位置的向量进行卷积运算后,生成特征图映射为 $C \in R^{L-n+1}$,而 W_j 的特征图中每个向量元素表示如公式(2)所示。

$$C_j = f(W_j * m + b) \quad (2)$$

其中 b 为偏置项, f 为非线性函数。由于模型采用不止一个过滤器进行特征映射,因此对于大小相同的 a 个过滤器,最终卷积生成的 a 个特征图排列成每个 W_j 的列向量组合特征表示如公式(3)所示。

$$W = [C_1; C_2; \dots; C_{a-1}; C_a] \quad (3)$$

其中 $C_i (1 < i \leq a)$ 表示为第 i 个过滤器卷积后的特征图表示。而此处的 $W \in R^{(L-n+1) \times a}$ 的每一行 W_j 都是由 a 个过滤器在位置 j 处窗口向量卷积成的新特征表示,最后得到的向量会传入下一层。

3.3. 双向长短时记忆层

双向长短时记忆模型是长短时神经网络模型的一种变体。首先,长短时又是根据循环神经网络改进而来。循环神经网络模型的长处是能很好地处理序列数据,而缺陷在于,模型无法处理因序列长度增加而产生的梯度消失和训练成本增加的问题,因此 RNN 很难捕获数据的长期依赖[35]。于是为了解决 RNN 的长期依赖问题,Hochreiter 和 Schmidhuber [36]提出了长短期记忆网络(LSTM),后来,在传统 LSTM 的基础上,又根据输入序列的向前向后输入的关系,由两个 LSTM 组成。同样在 BiLSTM 单元结构中一共有三个“门”:遗忘门、输入门和输出门。

在 BiLSTM 中,通过输入门、遗忘门、输出门以及一个记忆单元来实现长期记忆或遗忘信息,设当前时刻为 t ,则为前一时刻的记忆单元状态,为当前输入状态信息,和为当前 BiLSTM 的输出值。其中,LSTM 单元的计算公式如公式(4)~(9)所示。

$$\Gamma_f = \sigma(W_f [h^{t-1}, x^t] + b_f) \quad (4)$$

$$\Gamma_i = \sigma(W_i [h^{t-1}, x^t] + b_i) \quad (5)$$

$$\Gamma_o = \sigma(W_o[h^{t-1}, x^t] + b_o) \quad (6)$$

$$s'' = \tanh(W_s[h^{t-1}, x^t] + b_s) \quad (7)$$

$$s^t = \Gamma_{fs}^{t-1} + \Gamma_{fs'}^{t-1} \quad (8)$$

$$h^t = \Gamma_o \tanh(s^t) \quad (9)$$

式中 W_f , W_i , W_o 为各个门的神经元权重向量, b_f , b_i , b_o 为偏置向量, σ 为 sigmoid 函数, $\tanh$ 为非线性激活函数, S'' 为 t 时刻输入状态信息, S' 和 h' 分别为 t 时刻的记忆单元和隐藏状态输出值。为了从向前和向后两个方面捕获过去和未来的上下文关系, 以及考虑到本模型用到的输入序列, 最终得到的隐藏状态表示如公式(10)所示。

$$h^t = \mu(\bar{h}^t, \tilde{h}^t) \quad (10)$$

其中 μ 为两个方面得到的隐藏状态平均值。

3.4. 注意力层

人类大脑在观察某一事物时, 往往会将注意力集中到事物的某些部分, 这些注意力集中的部分往往也是从事物上获取信息的关键, 这些信息对于认知同类事物有着很强的指引作用, 而注意力机制(attention mechanism)就是一种模仿这一认知过程的特殊机制。注意力机制在计算机视觉和自然语言处理领域的运用都已经取得了较好的结果, 本文着重讲述在自然语言处理领域中有关舆情分析的注意力机制的运用。

首先, 前文的句子向量 X 中由单词向量 X_i 来表示, 当其直接进入 BiLSTM 中带有训练参数 θ , 用来捕获顺序和上下文信息, 输出 h_i 是输入的隐藏表示, 如公式(11)所示。

$$[h_1, h_2, \dots, h_n] = \text{BiLSTM}([X_1, X_2, \dots, X_n], \theta) \quad (11)$$

在舆情情感分析中, 先用 CNN 提取序列带有情感的空间特征, 在用 BiLSTM 根据空间特征提取时空特征后, 过多或非关键的特征会影响最后的预测结果。因此, 利用注意机制来提取序列的关键特征。注意力机制可以简单地理解为一个加权求和器, 也可以理解为一个关键特征提取器, 它主要进行的是加权求和操作。其中向量 c 为提取的关键特征, 其计算公式如(12)所示。

$$c = \sum_{i=1}^n \beta_i h_i \quad (12)$$

式中 n 为输入 LSTM 网络的时间步总和; h 为 BiLSTM 网络输出的隐藏特征向量, β 为 h 的权值, 为句子的每个单词分配权重。同时, 为了获取权值 β , 在注意力模型中加入一个小型的神经网络 α , 其输出层激活函数为 softmax, 其计算公式如(13)所示。

$$\beta_i = \frac{\exp(e_i)}{\sum_{k=1}^m \exp(e_k)} \quad (13)$$

式中的 e_i 由 $f_s(h_i, \text{Taver})$ 求得, 是根据隐藏状态 h_i 和 Taver 的语义关系和上下文关系计算句子中每个单词权重的函数, 而 Taver 则是隐藏特征表示的平均值, 其计算公式如(14)所示。

$$\text{Taver} = \frac{1}{n} \sum_{i=1}^n h_i \quad (14)$$

3.5. 模型训练

本文中针对模型训练部分时,神经网络最后的全连接层利用 softmax 进行概率输出,而多分类问题常采用的是交叉熵损失函数,因此本文中使用的损失函数为交叉熵损失函数,其计算公式如(15)所示。

$$H(p, q) = -\sum_{i=1}^n p(x_i) \log(q(x_i)) \quad (15)$$

其中 $p(x_i)$ 为真实概率分布, $q(x_i)$ 为预期的概率分布。此外,为了减轻过拟合,在训练过程中首先利用 dropout 技术随机失活一部分神经元,从而减少参数的共适应现象,降低了模型的复杂度,并增强模型对噪声和不同数据分布的适应性,对高次项的特征进行惩罚进而提高模型的泛化能力,本文采用 L_2 正则化进行训练,其计算公式如(16)所示。

$$L_2 = w_1^2 + w_2^2 + \dots + w_n^2 \quad (16)$$

其中 w 为需要训练的参数。

为了评估模型的性能,采用的基准测量分别是准确率和 Macro-F1 得分,其公式分别如(17)和(19)所示。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (17)$$

$$\text{F1} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (18)$$

$$\text{Macro-F1} = \frac{1}{n} \sum_{i=1}^m \text{F1}_i \quad (19)$$

式中 TP 代表样本为正,预测结果正确的真正例, TN 代表样本为负,预测结果正确的真反例, FP 代表样本为负,预测结果错误的假正例, FN 代表样本为正,预测结果错误的假反例。另外, Precision 代表预测样本为正,实际样本也为正的例数在所有预测样本为正的预测准确度。Recall 则代表所有正样本被正确预测的比例。

4. 实验及结果分析

4.1. 数据集

实验采用 SemEval-2014 任务 4 [37] (基于方面的情感分析,是 SemEval-2014 语义评测任务中的第四个任务,它又包含了四个子任务)、SemEval-2015 任务 12 [38] (是 SemEval-2014 任务 4 的延续)和 SemEval-2016 任务 5 [39] (是 SemEval-2015 的延续,增加新的测试数据之外还首次加入了除英语之外的其它多个语言)中四个涵盖餐厅和笔记本电脑领域的基准数据集。具体划分为 RES14 (SemEval-2014 任务 4 的餐厅域)、LAP14 (SemEval-2014 任务 4 的笔记本域)、RES15 (SemEval-2015 任务 12 的餐厅域)和 RES16 (SemEval-2016 任务 5 的餐厅域)。本文将极性标签显示为从 1 到 5 的等级,为了有一个统一的标签作为方面级数据,将这些评级转换为 3 级分类,评级 <3 、 $=3$ 和 >3 分别作为负面、中性和正面的极性。此外,通过随机抽样将方面级训练数据分割为 20% 的验证集和 80% 的训练集。数据集的统计数据如表 1 所示。

4.2. 实验设置

该模型使用 Tensorflow 作为技术支持。学习速率为 1×10^{-6} 。对模型进行正则化, Dropout 概率为 0.6。文档预训练模型的批处理大小为 8。在初始化深度学习模型时,通常采用基于词义和子词信息的预先训练语言模型。在本文预先训练的模型中,使用 300 维的 GloVe 嵌入来初始化单词。这些模型经过 8 个时期的训练以选出最佳结果。

Table 1. Statistics of experimental data
表 1. 实验数据统计

Dataset	Positive		Neural		Negative	
	Train	Test	Train	Test	Train	Test
Lap14	994	341	464	169	870	128
Rest14	2164	728	637	196	807	196
Rest15	912	326	36	34	256	182
Rest16	1240	469	69	30	439	117

4.3. 对比模型

本文为了验证本模型的性能，以下为本次实验对比的其他基于注意力机制的模型。

- ATAE-LSTM:ATAE-LSTM [25]通过将方向向量附加到输入的单词向量中来计算注意权重。
- IAN:IAN [27]通过交互注意网络分别建模意见目标及其上下文。
- RAM:RAM [28]提出了一个多注意框架来检测意见目标的情绪。
- MGAN:MGAN [29]开发了一种多注意神经网络模型，用于识别评论/评论中的情绪极性。
- AOA:AOA [30]引入了一种基于神经网络的方法，在学习中使用叠层注意力机制。

4.4. 实验结果及分析

表 2 为对比实验结果，通过表格可以得出，对比在 RES14 数据集上方法的性能，本文的方法达到了最先进的精度，宏观 f1 值分别为 81.96%和 73.16%。在 LAP14 数据集上，MGAN 的 Macro-F1 的值为 72.47%。在 RES15 数据集上，PABS-BiGRU 方法的准确率达到 80.88%，优于其他方法，但是本文的方法在 Macro-F1 值上却高于其他的对比方法。PABS-BiGRU 在 RES16 上的准确率最高，为 89.30%，而本文的方法在 Macro-F1 值上的准确率最高，为 69.35%。在笔记本电脑领域，我们的模型没有很好地泛化和执行低于一些比较算法。这是因为文档级数据来自通用电子领域，其中大多数都不是针对笔记本电脑的。值得注意的是，我们的方法在 RES15 和 RES16 数据集上取得了显著的优越的 Macro-F1 值。

Table 2. Results of model comparison
表 2. 模型对比结果

Method	RES14		LAP14		RES15		RES16	
	Acc	Macro-F1	Acc	Macro-F1	Acc	Macro-F1	Acc	Macro-F1
ATAE-LSTM	77.20	67.02	68.70	63.93	78.48	60.53	83.77	61.71
IAN	79.26	70.09	72.05	67.38	78.54	52.65	84.74	55.21
RAM	80.23	70.80	74.49	71.35	79.98	60.57	83.88	62.14
MGAN	81.25	71.94	75.39	72.47	79.36	57.26	87.06	62.29
AOA	79.97	70.42	72.62	67.52	78.17	57.02	87.50	66.21
Ours	81.96	73.16	76.53	69.63	80.07	63.68	87.30	69.35

4.5. 消融实验

为了研究本文提出的模型的不同构成部分对舆情分析结果性能的影响，特别进行了消融实验的研究，

其结果如表 3 所示。通过对比去掉不同部分的模型,可得知在各个数据集上所达到的结果中,单体的 CNN 和 Bilstm 在四个数据集上的表现相对来说最差,而在加入多头注意力机制后,准确率和 Macro-F1 都有了不同程度的提升,然而 CNN 和 Bilstm 组合后在未加入多头注意力机制后在四个数据集上的表现却有高有低。最后,本文所提出的模型在所有数据集上的表现都要优于前者。

Table 3. Results of ablation studies

表 3. 消融研究结果

Method	RES14		LAP14		RES15		RES16	
	Acc	Macro-F1	Acc	Macro-F1	Acc	Macro-F1	Acc	Macro-F1
CNN	77.77	68.50	67.71	65.10	76.68	61.38	83.88	65.65
Bilstm	78.75	67.66	71.94	68.22	78.04	60.13	86.55	66.02
CNN-MAtt	78.84	69.18	72.32	68.78	78.62	62.02	84.23	66.31
Bilstm-MAtt	80.18	71.40	75.63	69.04	78.66	60.53	86.82	66.39
CNN-Bilstm	81.06	72.42	74.96	68.93	79.17	62.32	85.50	66.01
Ours	81.96	73.16	76.53	69.63	80.07	63.68	88.14	72.40

此外,本文还通过对文档级权重的改变来探究不同文档级权重在不同数据集上会对模型产生的影响,其实验结果如表 4 所示。由表可知,在 0.1 到 1.0 之间随着权重提升模型的效果也越来越好,使用 1.0 的本模型的性能达到最佳。这个结果说明当使用的文档级权重不同,也会对模型的性能造成很大的影响,只有在不断测试,选出最优的文档级权重,模型效果才会达到最佳。

Table 4. Results of document weight ablation

表 4. 文档权重消融结果

f	RES14		LAP14		RES15		RES16	
	Acc	Macro-F1	Acc	Macro-F1	Acc	Macro-F1	Acc	Macro-F1
0.1	79.20	69.01	68.54	65.22	78.55	56.83	84.66	68.33
0.2	79.10	69.17	68.10	66.77	78.10	57.18	83.02	68.43
0.3	78.55	68.24	67.66	66.35	78.32	57.34	85.12	67.63
0.4	78.02	67.86	68.24	67.31	78.39	58.22	83.46	69.25
0.5	78.22	68.21	69.21	66.58	79.62	59.31	84.61	70.06
0.6	79.55	69.12	68.02	67.62	79.12	58.62	85.15	69.53
0.7	79.35	67.77	69.43	68.12	78.22	59.33	83.56	71.44
0.8	80.14	68.45	70.56	67.66	78.47	60.22	85.72	70.82
0.9	79.46	70.14	69.55	68.21	79.65	59.03	85.39	71.05
1.0	81.96	73.16	76.53	69.63	80.07	63.68	88.14	72.40

4.6. 卷积层数影响

本节内容主要是分析卷积层的数量可能对提出的模型的性能的影响,本文将卷积层的数目从 1 到 4 进行研究分析,其结果如图 2 所示。由实验结果得知,将卷积层的数目设为 2 时,得到的模型性能效果

比其余三个数目的卷积层数更好，故此次实验采取 2 层的卷积层。单层的卷积层可能是由于特征提取时的能力不足以将需要的特征完全提取出来，而 4 层的卷积在所有层数中的准确率和 F1 值最差，这就说明卷积层将不必要的特征也提取出来，造成了类似过拟合的情况。最后，由折线的趋势可以得知，卷积层数的增加虽然会增加模型的表达能力，使得提取的信息愈加复杂和抽象，但过度增加则会导致过拟合的产生，并不能给模型的性能带来更好的提升，反而会造成模型的学习效果变差。

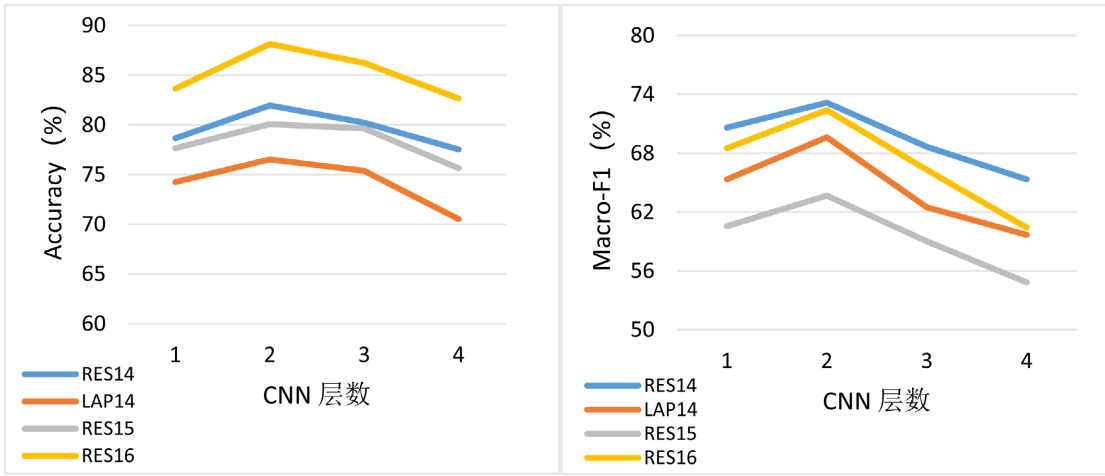


Figure 2. Effect of number of convolutional layers on accuracy and F1 value respectively
图 2. 卷积层的数量分别对准确率和 F1 值的影响

4.7. 案例分析

为了能够更加清晰地了解 CBLMAtt 模型的性能，在本节中通过与其他两个模型在三个测试案例中进行分析研究，其结果如表 5 所示。

Table 5. Case studies
表 5. 案例分析

Case study	Label		Prediction	Model
1. The quality of this person is really low , it's really disgusting .	Neg.	√	Neg.	IAN
		√	Neg.	MGAN
		√	Neg.	CBLMAtt
2. Although he looks fierce , he is indeed a good person .	Pos.	×	Neg.	IAN
		√	Pos.	MGAN
		√	Pos.	CBLMAtt
3. I think what he said is okay , but I can't find any evidence to support him.	Neu.	×	Pos.	IAN
		×	Neg.	MGAN
		√	Neu.	CBLMAtt

在三个案例中，具有情感特征的目标词用加粗显示，标签列显示为真实标签，三个模型的预测结果显示在预测列中，同时用“√”和“×”来对预测结果进行判断。在案例 1 中，消极的情感目标词情感色彩十分强烈，因此三个模型在进行分析预测后都得出了正确的结果。在案例 2 中消极的情感词和积极的

情感词同时出现, IAN 通过关注特定的上下文关注到了第一个消极的情感词, 故而给出了消极的标签, 与真实的标签相反, 而 MGAN 通过结合粗粒度和细粒度注意力捕捉到了上下文中词级别的信息, 给出了正确的结果, CBLMAtt 则充分利用语义情感特征提取的能力, 对目标情感词分析得出了正确的结果。而在案例 3 中加入了更多的情感目标词之后, CBLMAtt 通过提取到的更多的语义情感特征信息, 在后期进行分析时通过结合上下文得出了正确的标签结果。

5. 结论

本文着眼于特征融合的视角, 通过生成联合语义情感向量解决传统方法中语义与情感特征割裂导致的偏差问题, 并进一步构建了 CBLMAtt 模型, 利用双通道特征融合策略(双向 LSTM 隐藏层均值与多头注意力上下文向量拼接)增强文本上下文语义关联建模能力, 形成从底层特征融合到高层语义分析的全栈式解决方案, 有效提升了情感语义表征的完整性和舆情情感分析的准确性。最后通过对比实验证明本文提出的模型的有效性和适用性。

但是本文研究仍然存在一定的不足之处, 对于本文提出模型来说, 对数据的完整性和准确性要求较高, 如果针对即时性的数据来考量, 人工进行数据标注的工作可能避免不了, 所以说存在这方面的负担。未来的研究工作中可以着手于降低模型的复杂性, 但是与之同时也不能忽略情感分类任务的完成情况。然后减少数据标注工作的负担, 着眼于其他角度, 使用深度学习方法以期能更好地获取情感极性。

基金项目

重庆对外经贸学院科学研究项目(KYZK2024044)。

参考文献

- [1] 王家辉, 夏志杰, 王诣铭, 等. 基于句法规则和社会网络分析的网络舆情热点主题可视化及演化研究[J]. 情报科学, 2020, 38(7): 132-139.
- [2] 邓君, 孙绍丹, 王阮, 等. 基于 Word2Vec 和 SVM 的微博舆情情感演化分析[J]. 情报理论与实践, 2020, 43(8): 112-119.
- [3] 范涛, 吴鹏, 曹琪. 基于深度学习的多模态融合网民情感识别研究[J]. 信息资源管理学报, 2020, 10(1): 39-48.
- [4] 吴杰胜, 陆奎. 基于多部情感词典和规则集的中文微博情感分析研究[J]. 计算机应用与软件, 2019, 36(9): 93-99.
- [5] 陈珂, 谢博, 朱兴统. 基于情感词典和 Transformer 模型的情感分析算法研究[J]. 南京邮电大学学报(自然科学版), 2020, 40(1): 55-62.
- [6] Kononenko, I. (1989) Bayesian Neural Networks. *Biological Cybernetics*, **61**, 361-370. <https://doi.org/10.1007/bf00200801>
- [7] Anguita, D., Ridella, S. and Rovetta, S. (1998) Circuitual Implementation of Support Vector Machines. *Electronics Letters*, **34**, 1596-1597. <https://doi.org/10.1049/el:19981092>
- [8] 戚天梅, 过弋, 王吉祥, 等. 基于机器学习的外汇新闻情感分析[J]. 计算机工程与设计, 2020, 41(6): 1742-1748.
- [9] 濮泽堃. 基于机器学习的电商评论情感分析系统[D]: [硕士学位论文]. 南京: 南京邮电大学, 2020.
- [10] Lu, Q., Zhu, Z., Zhang, G., Kang, S. and Liu, P. (2021) Aspect-Gated Graph Convolutional Networks for Aspect-Based Sentiment Analysis. *Applied Intelligence*, **51**, 4408-4419. <https://doi.org/10.1007/s10489-020-02095-3>
- [11] Nemes, L. and Kiss, A. (2020) Social Media Sentiment Analysis Based on Covid-19. *Journal of Information and Telecommunication*, **5**, 1-15. <https://doi.org/10.1080/24751839.2020.1790793>
- [12] Priyadarshini, I., Mohanty, P., Kumar, R., Son, L.H., Chau, H.T.M., Nhu, V., et al. (2020) Analysis of Outbreak and Global Impacts of the Covid-19. *Healthcare*, **8**, Article No. 148. <https://doi.org/10.3390/healthcare8020148>
- [13] Dansana, D., Kumar, R., Das Adhikari, J., Mohapatra, M., Sharma, R., Priyadarshini, I., et al. (2020) Global Forecasting Confirmed and Fatal Cases of COVID-19 Outbreak Using Autoregressive Integrated Moving Average Model. *Frontiers in Public Health*, **8**, Article ID: 580327. <https://doi.org/10.3389/fpubh.2020.580327>
- [14] dos Santos, C.N. and Gatti, M. (2014) Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts.

- Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, Dublin, August 2014, 69-78.
- [15] Wang, J., Yu, L., Lai, K.R. and Zhang, X. (2016) Dimensional Sentiment Analysis Using a Regional CNN-LSTM Model. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, Volume 2, 225-230. <https://doi.org/10.18653/v1/p16-2037>
 - [16] Yoon, J. and Kim, H. (2017) Multi-Channel Lexicon Integrated CNN-BiLSTM Models for Sentiment Analysis. *Proceedings of the 29th Conference on Computational Linguistics and Speech Processing (ROCLING 2017)*, Taipei, 27-28 November 2017, 244-253.
 - [17] Wang, X.Y., Jiang, W.J. and Luo, Z.Y. (2016) Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts. *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, Osaka, December 2016, 2428-2437.
 - [18] Al Sallab, A., Hajj, H., Badaro, G., Baly, R., El Hajj, W. and Bashir Shaban, K. (2015) Deep Learning Models for Sentiment Analysis in Arabic. *Proceedings of the 2nd Workshop on Arabic Natural Language Processing*, Beijing, 30 July 2015, 9-17. <https://doi.org/10.18653/v1/w15-3202>
 - [19] Van, V.D., Thai, T. and Nghiem, M. (2017) Combining Convolution and Recursive Neural Networks for Sentiment Analysis. *Proceedings of the 8th International Symposium on Information and Communication Technology*, Nha Trang, 7-8 December 2017, 151-158. <https://doi.org/10.1145/3155133.3155158>
 - [20] Chen, X., Wang, Y. and Liu, Q. (2017) Visual and Textual Sentiment Analysis Using Deep Fusion Convolutional Neural Networks. 2017 *IEEE International Conference on Image Processing (ICIP)*, Beijing, 17-20 September 2017, 1557-1561. <https://doi.org/10.1109/icip.2017.8296543>
 - [21] Lin, H. and Bu, N. (2022) A CNN-Based Framework for Predicting Public Emotion and Multi-Level Behaviors Based on Network Public Opinion. *Frontiers in Psychology*, **13**, Article ID: 909439. <https://doi.org/10.3389/fpsyg.2022.909439>
 - [22] Sagnika, S., Mishra, B.S.P. and Meher, S.K. (2021) An Attention-Based CNN-LSTM Model for Subjectivity Detection in Opinion-Mining. *Neural Computing and Applications*, **33**, 17425-17438. <https://doi.org/10.1007/s00521-021-06328-5>
 - [23] Priyadarshini, I. and Cotton, C. (2021) A Novel LSTM-CNN-Grid Search-Based Deep Neural Network for Sentiment Analysis. *The Journal of Supercomputing*, **77**, 13911-13932. <https://doi.org/10.1007/s11227-021-03838-w>
 - [24] Zhang, W., Li, L., Zhu, Y., Yu, P. and Wen, J. (2022) CNN-LSTM Neural Network Model for Fine-Grained Negative Emotion Computing in Emergencies. *Alexandria Engineering Journal*, **61**, 6755-6767. <https://doi.org/10.1016/j.aej.2021.12.022>
 - [25] Wang, Y., Huang, M., Zhu, X. and Zhao, L. (2016) Attention-Based LSTM for Aspect-Level Sentiment Classification. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, 1-5 November 2016, 606-615. <https://doi.org/10.18653/v1/d16-1058>
 - [26] Liu, J. and Zhang, Y. (2017) Attention Modeling for Targeted Sentiment. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, Volume 2, 572-577. <https://doi.org/10.18653/v1/e17-2091>
 - [27] Ma, D., Li, S., Zhang, X. and Wang, H. (2017). Interactive Attention Networks for Aspect-Level Sentiment Classification. *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, Melbourne, 19-25 August 2017, 4068-4074. <https://doi.org/10.24963/ijcai.2017/568>
 - [28] Chen, P., Sun, Z., Bing, L. and Yang, W. (2017) Recurrent Attention Network on Memory for Aspect Sentiment Analysis. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Copenhagen, 9-11 September 2017, 452-461. <https://doi.org/10.18653/v1/d17-1047>
 - [29] Fan, F., Feng, Y. and Zhao, D. (2018) Multi-Grained Attention Network for Aspect-Level Sentiment Classification. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Brussels, 31 October-4 November 2018, 3433-3442. <https://doi.org/10.18653/v1/d18-1380>
 - [30] Huang, B., Ou, Y. and Carley, K.M. (2018) Aspect Level Sentiment Classification with Attention-over-Attention Neural Networks. In: Thomson, R., et al., Eds., *Social, Cultural, and Behavioral Modeling*, Springer International Publishing, 197-206. https://doi.org/10.1007/978-3-319-93372-6_22
 - [31] Yang, C., Zhang, H., Jiang, B. and Li, K. (2019) Aspect-Based Sentiment Analysis with Alternating Coattention Networks. *Information Processing & Management*, **56**, 463-478. <https://doi.org/10.1016/j.ipm.2018.12.004>
 - [32] Zhang, S., Xu, X., Pang, Y. and Han, J. (2019) Multi-Layer Attention Based CNN for Target-Dependent Sentiment Classification. *Neural Processing Letters*, **51**, 2089-2103. <https://doi.org/10.1007/s11063-019-10017-9>
 - [33] Gan, C., Wang, L., Zhang, Z. and Wang, Z. (2020) Sparse Attention Based Separable Dilated Convolutional Neural Network for Targeted Sentiment Analysis. *Knowledge-Based Systems*, **188**, Article ID: 104827. <https://doi.org/10.1016/j.knsys.2019.06.035>

-
- [34] Lei, Z., Yang, Y. and Yang, M. (2018) Sentiment Lexicon Enhanced Attention-Based LSTM for Sentiment Classification. *Proceedings of the AAAI Conference on Artificial Intelligence*, **32**, 8105-8106. <https://doi.org/10.1609/aaai.v32i1.12142>
 - [35] Pascanu, R., Mikolov, T. and Bengio, Y. (2013) On the Difficulty of Training Recurrent Neural Networks. *Proceedings of the 30th International Conference on Machine Learning*, Atlanta, 17-19 June 2013, 1310-1318.
 - [36] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
 - [37] Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I. and Manandhar, S. (2014) Semeval-2014 Task 4: Aspect Based Sentiment Analysis. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, Dublin, 23-24 August 2014, 27-35. <https://doi.org/10.3115/v1/s14-2004>
 - [38] Hercig, T., Brychcin, T., Svoboda, L. and Konkol, M. (2016) UWB at Semeval-2016 Task 5: Aspect Based Sentiment Analysis. *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, San Diego, 16-17 June 2016, 486-495. <https://doi.org/10.18653/v1/s16-1055>
 - [39] Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., AL-Smadi, M., *et al.* (2016) Semeval-2016 Task 5: Aspect Based Sentiment Analysis. *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, San Diego, 16-17 June 2016, 19-30. <https://doi.org/10.18653/v1/s16-1002>