

摘要

方面级情感分类问题是情感分析中的一项细粒度任务，其研究目标是判断给定方面词在句子中表现的情感极性。在现存的研究方法中，基于注意力机制的长短期记忆网络表现的性能较佳，但这些模型在处理微博话题数据集中偏口语化的语句时表现的性能不够理想。本文提出一种基于双向 LSTM 和注意力机制的网络模型，基本的设计思想是利用双向 LSTM 对语句序列的学习能力分别得到句子和方面词的隐含层，借助特定的注意力机制从方面词隐含层的平均值和句子的隐含层中生成句子的注意力向量，根据注意力向量对句子的隐含层进行加权计算，得到与方面词相关的句子表达向量，输入至 Softmax 层后选择概率最大的类别作为最终的预测结果。本文实验采用微博话题数据集验证模型的有效性，实验结果表明本文模型相较于几种 Baseline 模型在分类准确率上表现较优。另外，本文还采用两个可视化案例说明本文模型中的注意力机制的有效性。

关键词：方面级情感分类；情感分类；注意力机制；神经语言模型

ABSTRACT

Aspect-level sentiment classification problem aims to judge the sentiment polarity of a given aspect in a sentence, which is one of the basic tasks in sentiment analysis. In the recent research, the performance of the long short-term memory networks using the attention mechanism is pretty good, but the performance of these models is not ideal enough when dealing with spoken sentences in the microblog topic data set. This paper proposes a network model based on a bidirectional long short-term memory using attention mechanism to solve this problem. We use two bidirectional long short-term memory networks to obtain the hidden states of the sentence and the aspect respectively. The average of the hidden states of the aspect is used to influence the generation of the attention vector with our proposed attention mechanism to capture the emotional polarity information about the given aspect. According to the emotional polarity information, we can generate the final representation, which is inputted to the softmax function. We will select the category with the largest classification probability as the final prediction result. The topic data set verifies the validity of the model. The result of the experiments in this paper shows that the model performs better in classification accuracy than several baseline models. In addition, this paper also uses some visual case to illustrate the effectiveness of the attention mechanism in the model.

Keywords: aspect-level sentiment classification; sentiment classification; attention mechanism; neuro-language model

目录

第一章 导言	1
1.1 研究背景和意义	1
1.2 国内外研究现状	2
1.3 本文的工作	4
1.4 本文结构简介	5
第二章 相关技术	6
2.1 文本预处理	6
2.2 循环神经网络	6
2.3 注意力机制	8
第三章 基于双向 LSTM 和注意力的网络模型	10
3.1 问题的定义	10
3.2 模型介绍	10
3.2.1 词嵌入	11
3.2.2 双向长-短期记忆网络	12
3.2.3 注意力机制	13
3.2.4 预测模块	13
3.3 模型训练	14
3.4 本章小结	14
第四章 实验结果与分析	16
4.1 数据集介绍	16
4.2 实验环境及模型参数设置	16
4.3 实验算法流程	16
4.4 评价指标	17
4.5 模型比较	17
4.6 模型可视化	19
4.7 本章小结	20
第五章 总结与展望	21

参考文献.....	22
致谢	25
附录	26

第一章 导言

1.1 研究背景和意义

自然语言中的情感分类问题受到了研究者的广泛关注,其研究目的是通过文本分析展现人们对组织、个人、事件、产品、服务、主题等对象所体现的评价、看法、意见以及态度等主观情感^[1]。

随着网络技术的高速发展,社交网络日益发达,网络上有大量消费者对影视、手机、电脑、餐饮等各个行业的各种评论,从这些评论中不难看出消费者对各种服务和产品的某些方面的满意程度。根据 2010 年关于互联网产品评论的研究,Harris Poll 调查机构表明,在网上购物时会受到评论信息影响的调查者占 45%,对产品的评论信息毫不怀疑的调查者占 46%^[2]。因此,商家可依据消费者的评论了解自己的服务和产品中存在的优缺点,进而有的放矢地进行改进来提升产品的竞争力,更好地理解消费者的购买行为,为商家在以后的决策中提供强有力的支持。另一方面,人们在做选择和决策时,均会偏向于参考他人的意见^[3],消费者可根据以前享受过服务或购买过产品的消费者的评论中了解更多的信息,从而选择心仪的服务和产品,做出满意的购买决策。

根据中国互联网络信息中心(CNNIC)于 2019 年 8 月 30 日在京发布的第 44 次《中国互联网络发展状况统计报告》,截止 2019 年 6 月,我国网民规模达 8.54 亿,较 2018 年底增长 2598 万,互联网普及率达 61.2%,较 2018 年底提升 1.6 个百分点^[4]。随着网络用户数目的不断提高,网络上用户的评论数据呈爆炸式增长,这些评论数据包含了例如公众人物、个人经历、网络游戏、在线商品和服务等诸多主题,与此同时,这些数据广泛地分布于各大社交平台、新闻评论平台、电子商务评论平台等众多平台。面对数据庞大的评论文本,人工分类标注效率过低且不切实际,自动化进行文本情感分析已是大势所趋。情感分析的应用已经深入到各个领域,例如消费品行业、服务业、健康医疗和金融服务等^[5]。

按照分类的粒度层次来看,情感分类可分为“粗粒度情感分析”和“细粒度情感分析”,前者又可分为“篇章级情感分析”和“句子级情感分析”,后者则是基于给定的评价对象进行分析。对篇章级和句子级的情感分析的研究起步较早,

研究较为成熟,但当语句变得复杂时,二者的分析能力捉襟见肘,结果并不理想。为了适应多样复杂的语言环境并满足日益提高的分析需求,方面级情感分类问题逐渐成为了研究的重点。对于某些产品或服务的评论而言,整体的情感极性并不一定会与产品或服务在某些方面的情感极性保持一致。消极情感的文本并不代表用户对文中某些方面均保持消极的情感,因此精确分辨文本中各个方面的情感极性变得十分重要。

研究方面级情感分类问题的意义在于,帮助企业和消费者从海量的评论数据中快速地了解产品和服务在某些方面的情感极性,以便企业更好地做出未来的发展规划,以便消费者能够做出满意的购买决策。

1.2 国内外研究现状

方面级情感分类在许多领域得到广泛应用,学术界也对此特别关注。早期的研究采用例如情感词典、依赖信息和 n 元语法等人工定义特征,以构建分类器来进行情感分类的传统方法。Erik Boiy 等人在文献[6]中提出通过人工提取特征,例如词干、否定词、一元语法等,利用支持向量机 (Support Vector Machine)、朴素贝叶斯 (Naïve Bayes) 和最大熵 (Maximum Entropy) 模型进行分类,取得不错的效果。Long Jiang 等人在文献[7]中提出在给定搜索 (query) 条件下,根据 Twitter 对该搜索的情感是积极、消极还是中性,将该搜索语句进行情感分类,作者不仅采用 Luciano Barbosa 等人[8]提出的特征,而且利用语法依赖树扩展与目标相关的特征,最后使用支持向量机进行分类。综上所述,传统方法大同小异,只有在提取特征时存在不同。传统方法极度依赖耗时耗力的人工提取特征,而这正是传统方法的不足之处。

随着自然语言处理领域的神经网络模型和深度学习模型的不断发展,针对方面级情感分类问题,越来越多的研究人员采用深度学习模型来解决。Li Dong 等人在文献[9]中提出的自适应递归神经网络结构,利用依赖解析树找出与目标相互连接的词语来提取情感极性的信息,以帮助提高情感分类问题的准确率。之后的研究发现该模型在处理口语化文本时效果并不是特别理想,因为该模型特别依赖句子的句法关系,当出现句法分析错误时模型会受到较大影响。Duy-Tin Vo 等人在文献[10]中提出的模型不再采用句法分析的方法,该模型根据给定的目标将

句子分成语句左部、语句右部和目标三部分，接着采用分布式单词表示和神经池函数提取目标的相关特征。Duyu Tang 等人在文献[11]中将句子分为两个部分，即带目标的左部和带目标的右部，采用两个 LSTM 模型分别对句子的两个部分进行建模，然后结合二者的输出进行情感分类。Meishan Zhang 等人在文献[12]中提出的句子级神经网络模型克服池函数的局限性，采用双向门限神经网络连接单词以表示目标及其上下文，使用三向门限神经网络对目标和上下文建模。上述的几种基于 RNN 网络的模型均取得了较好的分类成绩，但 RNN 网络更为关注近期输入，对于那些复杂的语句，RNN 网络很难重视其中距离方面词较远的情感词与方面词之间的关系。因此，使用注意力机制可以克服这个局限。

注意力机制自从为解决机器翻译问题被提出以来，被广泛应用于自然语言处理领域。注意力机制的设计灵感来源于人对外界事务的感知，在自然语言处理领域中，判断方面词在句子中的情感极性时，注意力机制可使模型动态地关注输入，即模型将重视与方面词相关的情感词，而忽视与方面词无关的其他信息，从而能够在复杂的语句中提取出细节信息，做出更准确的判断。Yequan Wang 等人在文献[13]中提出的第一个模型即 Attention-based LSTM (AT-LSTM) 模型将 LSTM 网络的隐藏层向量与方面词嵌入向量拼接后，利用注意力机制得到每个词的权重，再使用隐藏层向量和权重计算得到句子的最终表示。第二个模型即 Attention-based LSTM with Aspect Embedding (ATAE-LSTM) 模型在 AT-LSTM 模型上做了改进，其将方面词的词向量和句子中每个词的词向量拼接，输入至 LSTM 网络中，再将方面词的词向量与隐藏层向量拼接，通过注意力机制生成最终的语句情感表达。Duyu Tang 等人在文献[14]中提出基于 multihop 注意力的记忆网络 (Memory Network, MemNet) 模型，模型中的每个 hop 包含注意力层和线性层，每个注意力层使用上一个 hop 的输出计算与句子中每个词的词向量的相似度，对词向量计算加权和得到句子的表示向量，再与线性层的输出求和得到 hop 的输出，多个相同结构的 hop 连接起来提取上下文特征，最终得到方面词的情感表达。Peng Chen 等人在文献[15]中提出的 Recurrent Attention Network on Memory (RAM) 模型对 MemNet 模型做了改进，RAM 模型采用双向 LSTM 对文本建模，使用 GRU 网络来更新注意力的抽象。Dehong Ma 等人在文献[16]中提出的 Interactive Attention Networks (IAN) 模型计算句子与方面词的交互注意力，利用方面词影

响句子注意力的生成进而产生句子的表示向量,利用句子影响方面词注意力的生成进而产生方面词的表示向量,再将二者拼接作为最终的分类特征。Binxuan Huang 等人在文献[17]中提出的 Attention-over-Attention (AOA) 模型也是基于注意力机制,该模型分别生成方面词到句子和句子到方面词的注意力,求方面词注意力对句子注意力的加权和得到句子级的注意力向量,再与句子的隐含层加权求和得到最终的句子表示。

上述的几个模型是近期在方面级情感分类问题中表现较好的模型,他们的相同点在于采用注意力机制得到有关方面词的句子表示向量,最终依据表示向量进行情感分类。

1.3 本文的工作

针对方面级情感分类问题,受 Dehong Ma 等人提出的交互式注意网络 (Interactive Attention Networks, IAN) 模型^[16]的启发,本文提出了一个基于双向 LSTM 和注意力机制的网络模型 (Bidirectional-LSTM based Attention Networks, BAN)。BAN 模型利用双向 LSTM 网络,以词嵌入向量为输入,分别捕获句子和方面词的隐含层。将方面词隐含层的平均值和句子的隐含层作为输入,借助特定的注意力机制生成句子的注意力向量,依据生成的注意力向量,对句子的隐含层进行加权计算,得到与方面词相关的句子表达向量。通过这种设计,可以使句子表达向量的生成受方面词的隐含层的平均值的影响。将句子表达向量输入至全连接网络,利用 Softmax 函数将句子表达向量转化为各个情感类别的概率。最后,选择概率最大的情感类别作为预测结果。在微博话题数据集上的实验表明, BAN 模型相较于 IAN 模型略有提升。

本文的贡献如下:

(1) 考虑到 LSTM 只能得到 t 序号词与 t 序号词之前的词的关系,而无法得到与 t 序号词之后的词的关系,因而采用双向 LSTM 分别获取句子和方面词的隐含层;

(2) 利用注意力机制使方面词的隐含层的平均值影响句子的注意力向量的生成,体现方面词在分类中的突出作用。

1.4 本文结构简介

本文共分成五个章节来阐述本人的研究工作：

第一章介绍方面级情感分类问题的研究背景和意义，分析其在社交网络迅速发展，尤其是电子商务网络快速发展，帮助消费者和商家做决策的重要作用；通过参考大量的文献，简要地介绍国内外在细粒度情感分析领域所涉及的不同任务的研究状况；概述本文的研究内容和主要贡献。

第二章介绍方面级情感分类的相关技术，主要包括文本预处理，循环神经网络以及 LSTM，最后介绍最早应用于自然语言处理领域的注意力机制。

第三章阐述方面级情感分类问题的定义，并详细地介绍本文提出的 BAN 模型，详细介绍模型的词嵌入、双向 LSTM、注意力机制和预测模块，同时还介绍模型中需要训练的参数以及采用的损失函数。

第四章详细介绍了本文实验时采用的微博话题数据集、实验环境、模型的超参数、评价指标、实验算法流程以及与其他 Baseline 模型的分类效果对比，并利用两个可视化案例来验证模型注意力机制的有效性。

第五章对本文工作进行分析和总结，分析存在的不足并提出未来研究的展望。

第二章 相关技术

2.1 文本预处理

在处理文本数据时，由于文本数据的格式和来源各不相同，例如本文使用的微博话题数据集是从微博网页爬取而来的，所以需要对爬取的数据进行预处理，以提取有效的文本信息，过滤无关信息，从数据源头处减少噪声的输入，提高模型的分类效果。本文对文本数据的预处理步骤如下：

(1) 对文本进行分词处理：在英文语句中，词与词之间有空格相间，而中文语句不同，语句中的词是相连的，没有分隔符作为间隔。为了使模型可以对其进行进一步处理，需要对中文语句进行分词处理。根据中文的语法规则，将中文语句分割成词和标点符号的序列。

(2) 处理文本中的停用词：中文文本中存在许多没有特定意义的虚词、代词、动词和名词等，这些词语一般只会带来噪声干扰，对情感分类毫无用处，所以往往选择将这些停用词从文本中去除。去除无用的停用词后，能够减小文本的长度，缩小模型输入的规模，提升模型预测的效率。

2.2 循环神经网络

Elman 等人^[18]提出的循环神经网络将同一个网络多次重复循环，每次循环时传递现有信息，使得信息持久化，结构如图 1 所示。

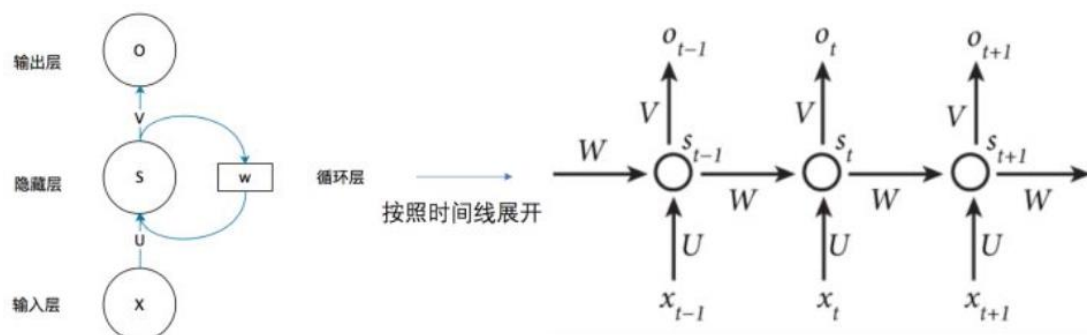


图 1 RNN 网络结构随时间展开示意图¹

图 1 中 x_t 表示 t 时刻的输入， s_t 表示 t 时刻的隐含层， o_t 表示 t 时刻的输出，

¹ 图片来源: <https://zhuanlan.zhihu.com/p/30844905>

它们的计算公式如下：

$$s_t = f(U \cdot x_t + W \cdot s_{t-1} + b_h) \quad (2.1)$$

$$o_t = g(V \cdot s_{t-1} + b_y) \quad (2.2)$$

上式中， U 表示输入与隐含层之间的连接权重， W 表示输入与时间步之间的连接权重， V 表示隐含层与输出之间的连接权重， b_h 和 b_y 均表示偏移量， f 和 g 均表示激活函数。图 1 显示的网络结构表明 RNN 可以处理序列数据。

RNN 能够依据以前的信息理解当前的信息，在 RNN 理解当前信息时，所依据的以前的信息与当前信息相比有一定的时间差，这就会导致长期依赖的问题。当时间差逐渐增大时，RNN 不能学习距离较远的信息，这就会出现梯度消失和梯度爆炸。采用梯度修剪的方法可以解决梯度爆炸的问题，也就是说，当梯度向量超过设定的阈值时，可以缩减该梯度向量。然而，梯度消失依然是一个难题。

针对 RNN 存在的梯度消失和梯度爆炸的难题，Hochreiter 等人在文献[19]中提出 RNN 的变种 LSTM，结构如图 2 所示。

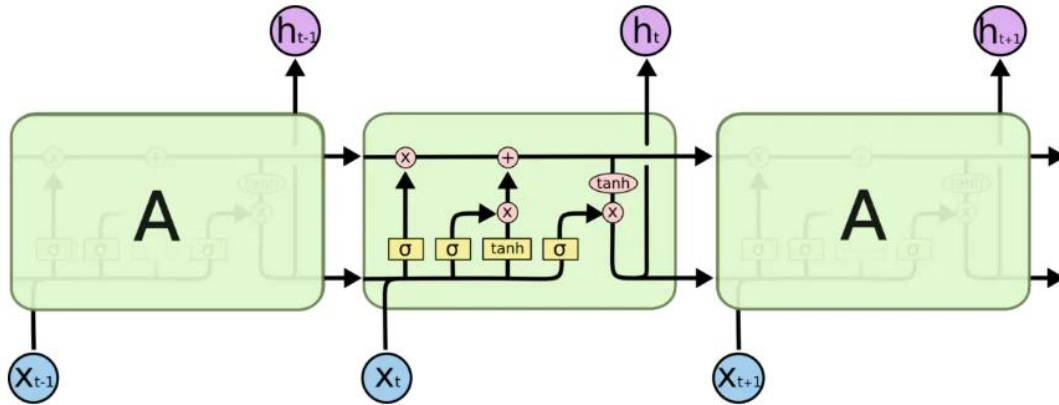


图 2 LSTM 的重复模块示意图²

LSTM 中不断被重复的新模块名为记忆细胞，记忆细胞在每个时刻都有一个对应的状态，以此存储该时刻的信息。LSTM 使用选择性通过信息的输入门、遗忘门和输出门控制并保护细胞的状态。

对于每个 LSTM 的记忆细胞，在词嵌入向量 v_t 、前一记忆状态 c_{t-1} 和前一隐含层 h_{t-1} 给定的情况下，输入门 i_t 的更新公式为

$$i_t = \sigma(W_i^v \cdot v_t + W_i^h \cdot h_{t-1} + b_i) \quad (2.3)$$

² 图片来源: <https://www.jianshu.com/p/9dc9f41f0b29>

遗忘门 f_t 的更新公式为

$$f_t = \sigma(W_f^v \cdot v_t + W_f^h \cdot h_{t-1} + b_f) \quad (2.4)$$

输出门 o_t 的更新公式为

$$o_t = \sigma(W_o^v \cdot v_t + W_o^h \cdot h_{t-1} + b_o) \quad (2.5)$$

记忆状态 c_t 的更新公式为

$$\hat{c}_t = \tanh(W_c^v \cdot v_t + W_c^h \cdot h_{t-1} + b_c) \quad (2.6)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \hat{c}_t \quad (2.7)$$

隐含层 h_t 的更新公式为

$$h_t = o_t \odot \tanh(c_t) \quad (2.8)$$

在上述更新公式中， σ 表示 sigmoid 函数， \odot 表示按位乘运算， \cdot 表示矩阵乘运算， W_i^v 、 W_f^v 、 W_o^v 、 W_c^v 、 W_i^h 、 W_f^h 、 W_o^h 、 W_c^h 分别表示对应的权重矩阵， b_i 、 b_f 、 b_o 、 b_c 分别表示对应的偏移量。

2.3 注意力机制

注意力机制在自然语言处理领域被广泛应用，之前的机器翻译方法基于编码器-解码器结构，编码器将输入的整个源句子编码成一个固定长度的向量，而这存在两个缺点，一是该向量可能无法完整表示整个源句子的语义信息，二是后输入到网络的信息会覆盖先前输入到网络的信息，输入序列越长，这一弊端越明显。这两个缺点使解码器不能获得足够的语义信息，从而解码器的准确率大大折扣。

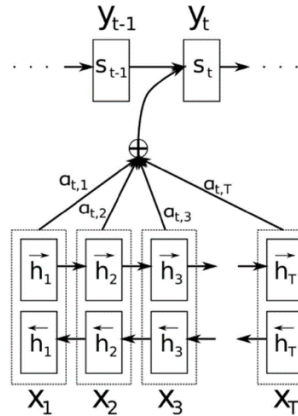


图 3 基于注意力机制的机器翻译模型示意图^[19]

Dzmitry Bahdanau 等人在文献[19]中首次将注意力机制应用到机器翻译上，模型如图 3 所示。编码器将输入的源句子编码成一个向量序列，通过注意力机制从该向量序列中选择与目标词最相关的向量序列的子集，解码器基于该子集和之前的目标词来生成当前的目标词，这样在生成每个目标词时都能够充分利用输入的源句子中的信息。在图 3 中， $\alpha_{t,i}$ 表示解码器在生成目标词 y_t 时，分配给向量序列中第 i 个向量的权重，以此来表示该向量对应的词的重要性， $\alpha_{t,i}$ 的计算公式如(2.9)所示：

$$\alpha_{t,i} = \frac{\exp(a(s_{t-1}, h_i))}{\sum_{k=1}^T \exp(a(s_{t-1}, k))} \quad (2.9)$$

其中 a 表示一个评分函数，采用的是前馈神经网络。

第三章 基于双向 LSTM 和注意力的网络模型

3.1 问题的定义

作为情感分析中的一项细粒度任务，方面级情感分类的目的是确定句子描述的方面词在上下文中表现的情感是积极、中立、还是消极。假设一个句子包含 n 个词，表示为

$$C = [w_{c_1}, w_{c_2}, w_{c_3}, \dots, w_{c_n}] \quad (3.1)$$

其中每个词表示为 w_{c_i} 。假设句子中的方面词包含 m 个词，表示为

$$T = [w_{t_1}, w_{t_2}, w_{t_3}, \dots, w_{t_m}], \quad (0 < m \leq n) \quad (3.2)$$

其中每个词表示为 w_{t_i} 。

比如说，句子 C 为“#吴亦凡# [二哈]#吴亦凡公益正能量# 这个超话有毒吧 @Mr_凡先生”，方面词 T_1 为“吴亦凡公益”，方面词 T_2 为“超话”。方面级情感分类问题的输入就是句子 C 、方面词 T_1 和方面词 T_2 分别对应的词嵌入矩阵 V_C 、 V_{T_1} 和 V_{T_2} ，任务是判断 T_1 在 C 中的情感极性为积极， T_2 在 C 中的情感极性为消极，也就是说，若将情感极性中的积极、中性和消极分别设为 2、1 和 0，那么任务就是当模型的输入为 V_C 和 V_{T_1} 时，输出应为 2，当模型的输入为 V_C 和 V_{T_2} 时，输出应为 0。

从上面的例子可以看出，句子描述了“吴亦凡公益”和“超话”两个方面，整个句子体现了积极和消极两种相反的极性。如果忽视方面词，我们将很难确定情感极性。如此，方面级情感分类的提出克服了篇章级别情感分类在处理一个句子出现多个方面词时情感分类的局限性。

3.2 模型介绍

借鉴交互式注意网络 (IAN) 模型^[16]的设计思想，本文提出了基于双向 LSTM 和注意力机制的网络模型 (BAN)，该网络模型由四部分组成，第一部分将样本中的句子和方面词转换为词嵌入向量，第二部分利用双向 LSTM 网络，以词嵌入

向量为输入，分别捕获句子和方面词的隐含层，第三部分将方面词隐含层的平均值和句子的隐含层作为输入，借助特定的注意力机制生成句子的注意力向量，依据生成的注意力向量，对句子的隐含层进行加权计算，得到与方面词相关的句子表达向量。通过这种设计，可以使句子表达向量的生成受方面词的隐含层的平均值的影响。第四部分将句子表达向量输入至全连接网络，利用 Softmax 函数将句子表达向量转化为各个情感类别的概率。最后，选择概率最大的情感类别作为预测结果。总体架构如图 4 所示。

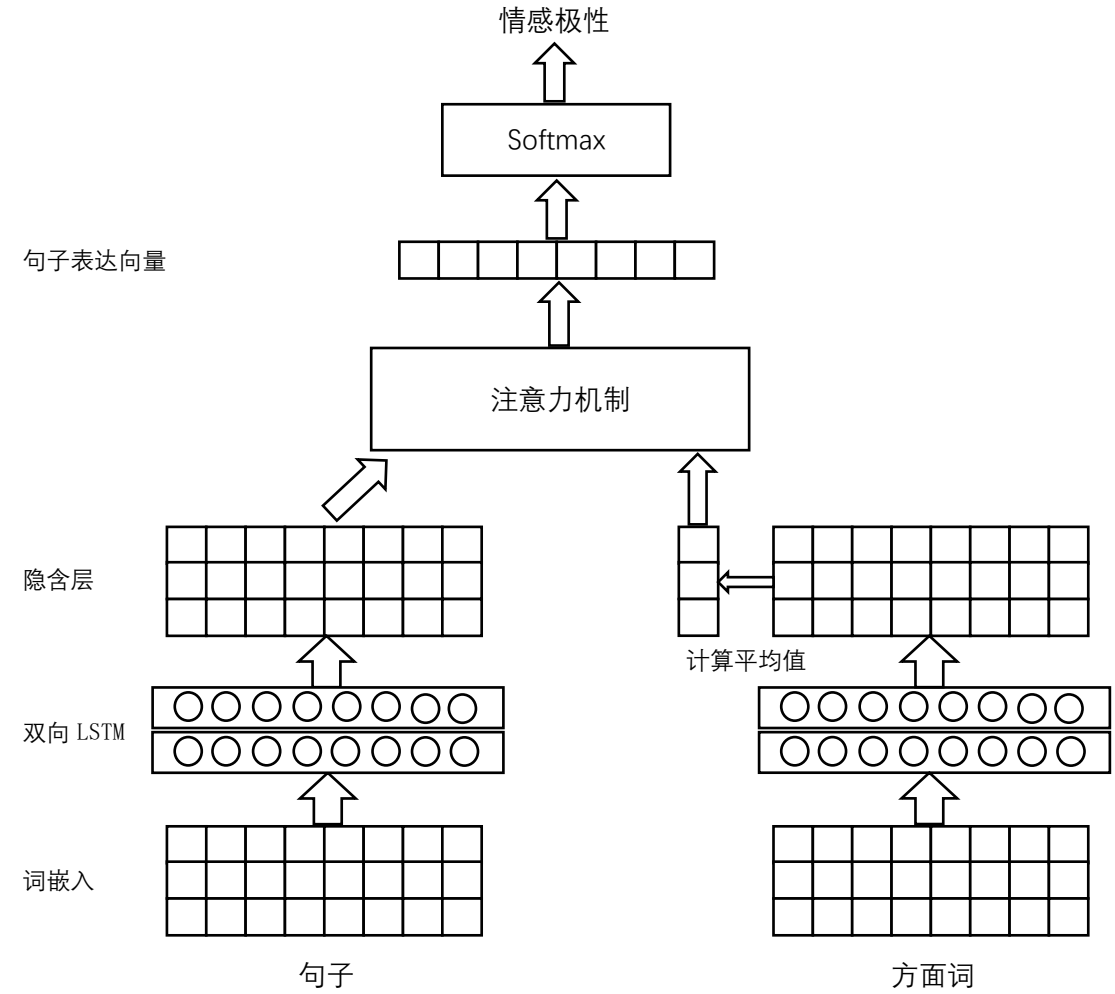


图 4 BAN 模型的总体架构

3.2.1 词嵌入

词嵌入向量是本文模型的输入，因为模型不能处理维度为所有词的数量的数据，所以将句子或方面词中的每个词映射到一个低维实值向量来表示。词嵌入向

量可在整个模型的训练过程中作为神经网络的参数训练得到,也可以从适当的语料库中采用其他的语言模型预训练获得。

本文模型采用的是 Shen Li 等人^[21]训练好的词嵌入矩阵。利用该词嵌入矩阵 $M^{d \times |V|}$, $|V|$ 表示词的总数, d 表示词嵌入向量的维度, 针对每个词 w_i 可以获得对应的词嵌入向量 $v_i \in R^d$, 也就是说, 每个词 w_i 在词嵌入矩阵 $M^{d \times |V|}$ 中存在对应的一个列向量 v_i 。如此, 句子 C (3.1) 可表示为

$$V_C = [v_{c_1}, v_{c_2}, v_{c_3}, \dots, v_{c_n}] \quad (3.3)$$

其中 $V_C \in R^{d \times n}$, n 表示句子 C (3.1) 的长度。同样, 方面词 T (3.2) 可表示为

$$V_T = [v_{t_1}, v_{t_2}, v_{t_3}, \dots, v_{t_m}] \quad (3.4)$$

其中 $V_T \in R^{d \times m}$, m 表示方面词 T (3.2) 的长度。

3.2.2 双向长-短期记忆网络

通过上述方法分别得到句子和方面词的词嵌入向量后, 将这两组词嵌入向量分别输入至两个不同的双向 LSTM 网络中, 以分别捕获句子和方面词的语义, 即得到二者的隐含层。本文模型使用的这两个不同的双向 LSTM 网络均是由前向 LSTM 和后向 LSTM 叠加得到的, 与普通 RNN 相比, LSTM 不仅擅长处理序列变化的数据, 还能够解决在长序列训练过程中出现的梯度爆炸和梯度消失的问题。

假设将句子或方面词中的词从左至右依次编号, 前向 LSTM 可捕获 t 序号词与 t 序号词之前的词的关系, 而无法捕获与 t 序号词之后的词的关系。同样, 后向 LSTM 可捕获 t 序号词与 t 序号词之后的词的关系, 而不能捕获与 t 序号词之前的词的关系。所以, 本文模型使用的双向 LSTM, 既可以捕获 t 序号词与 t 序号词之前的词的关系, 又可以捕获 t 序号词与 t 序号词之后的词的关系。

将句子 V_C (3.3) 输入至前向 LSTM 以获得句子的语义, 即得到句子的前向隐含层:

$$\vec{H} = [\vec{h}_1, \vec{h}_2, \dots, \vec{h}_n], \vec{H} \in R^{n \times d_h} \quad (3.5)$$

输入至后向 LSTM 以获得句子的语义, 即得到句子的后向隐含层:

$$\bar{H} = [\bar{h}_1, \bar{h}_2, \dots, \bar{h}_n], \bar{H} \in R^{n \times d_h} \quad (3.6)$$

然后，将两者拼接后作为句子的最终语义，即句子的隐含层：

$$H_C = [(\vec{h}_1, \bar{h}_1), (\vec{h}_2, \bar{h}_2), \dots, (\vec{h}_n, \bar{h}_n)] = [h_{c_1}, h_{c_2}, \dots, h_{c_n}], H_C \in R^{n \times 2d_h} \quad (3.7)$$

同样，将方面词 V_T (3.4)输入至双向 LSTM 得到方面词的最终语义，即方面词的隐含层：

$$H_T = [h_{t_1}, h_{t_2}, \dots, h_{t_m}], H_T \in R^{m \times 2d_h} \quad (3.8)$$

在上述的式子中， d_h 表示隐含层的维度。

3.2.3 注意力机制

本文模型采用特定的注意力机制，利用方面词对句子表达向量的生成的影响，挖掘句子中更多与情感极性相关的信息。

对方面词的隐含层 H_T (3.8)求平均值，得到方面词的语义向量 t_{avg} ，即

$$t_{avg} = \frac{1}{m} \sum_{i=1}^m h_{t_i}, t_{avg} \in R^{1 \times 2d_h} \quad (3.9)$$

注意力机制将方面词的语义向量 t_{avg} (3.9)和句子的隐含层 H_C (3.7)中每个隐藏层输出 h_{c_i} 作为输入，计算得到与 h_{c_i} 对应的权重 α_i ，即

$$\alpha_i = \frac{\exp(\tanh(h_{c_i} \cdot W_\alpha \cdot t_{avg}^T + b_\alpha))}{\sum_{j=1}^n \exp(\tanh(h_{c_j} \cdot W_\alpha \cdot t_{avg}^T + b_\alpha))} \quad (3.10)$$

其中 $h_{c_i} \in R^{1 \times 2d_h}$, t_{avg}^T 表示 t_{avg} 的转置， W_α 和 b_α 分别表示权重矩阵和偏移量。

3.2.4 预测模块

根据权重 α_i (3.10)对句子的隐含层 H_C (3.7)中每个隐藏层输出 h_{c_i} 进行加权求和，可以得到句子的表达向量 s_r ，即

$$s_r = \sum_{i=1}^n \alpha_i h_{c_i} \quad (3.11)$$

将得到的句子表达向量 s_r (3.11)输入至全连接层，全连接层中神经元的数目

为 3，因为该数目就是情感类别的个数，即积极、消极和中性三种。在全连接层中使用 Softmax 函数将句子的表达向量 s_r (3.11)转化为各个情感类别的概率，计算公式如(3.12)所示。

$$\hat{y}_i = \frac{\exp(s_r \cdot w_i)}{\sum_{j=1}^C \exp(s_r \cdot w_j)} \quad (3.12)$$

其中， w_i 表示第 i 个情感类别与句子表达向量的连接权重， C 表示情感类别数目，显然 3 个类别的概率之和为 1：

$$0 \leq \hat{y}_i \leq 1 \quad (3.13)$$

$$\sum_{j=1}^C \hat{y}_j = 1 \quad (3.14)$$

最后，选择概率最大的情感极性作为最终的预测结果。

3.3 模型训练

本文模型含有众多参数，双向 LSTM 中的 $[W_i^v, W_i^h, W_f^v, W_f^h, W_o^v, W_o^h, W_c^v, W_c^h, b_i, b_f, b_o, b_c]$ ，注意力机制中的 $[W_a, b_a]$ ，Softmax 函数中的 w_i ，将这些参数记为 Φ ，它们均需要利用误差反向传播算法计算梯度并更新优化，即

$$\Phi = \Phi - \lambda_l \frac{\partial J(\Phi)}{\partial \Phi} \quad (3.15)$$

其中， λ_l 表示学习率。

模型使用监督学习的训练方式，为了防止过拟合采用 dropout 策略，随机丢弃 LSTM 单元的部分输入，并使用带 L2 正则化的交叉熵损失函数，即

$$loss = - \sum_{i=1}^C y_i \log(\hat{y}_i) + \sum_{\theta \in \Phi} \lambda \theta^2 \quad (3.16)$$

其中， i 表示类别的索引， $y_i \in R^C$ 表示实际情感类别， $\hat{y}_i \in R^C$ 表示对应的预测类别， λ 表示 L2 正则项的系数。

3.4 本章小结

本章详细介绍了本文提出的 BAN 模型。首先阐述了方面级分类问题的具体定义，接着针对该问题，本文提出基于双向 LSTM 和注意力机制的网络模型。然

后详细说明模型中双向 LSTM 的原理，与 LSTM 相比，双向 LSTM 既可得到 t 序号词与 t 序号词之前的词的关系，又可得到 t 序号词与 t 序号词之后的词的关系，所以模型可以获取更多的信息，提高分类的准确率。本章还详细介绍模型中使用的注意力机制，考虑了方面词对生成句子注意力向量的影响。最后阐述了模型中需要训练的参数、更新参数的算法以及采用的损失函数。

第四章 实验结果与分析

4.1 数据集介绍

本文的实验数据集是从微博话题中自行爬取的 6000 个话题数据，人工标注它们的方面词和情感极性，得到 3050 个在情感极性方面多人标注完全一致的话题数据，其中情感积极的有 1189 个，情感中立的有 1461 个，情感消极的有 400 个。选取 2000 个话题数据作为训练集，剩下的作为测试集，统计结果如表 1 所示。

表 1 微博话题数据集统计

	Train	Test
Positive	922	267
Neural	806	655
Negative	272	128
Total	2000	1050

4.2 实验环境及模型参数设置

实验环境为 Windows 10 操作系统、Intel® Core™ i5-6300HQ 4 核 CPU、8G 内存、模型搭建使用 torch-1.4.0+cpu。

实验采用 jieba 中文分词器，在微博话题数据集中，句子的最大词语个数设置为 80，对于包含词语数大于 80 的句子，从后截断使其满足设定值；对于词语数少于 80 的句子，则通过在句末补 0 的方式使其达到设定值。实验使用 Shen Li 等人^[21]基于微博语料库预训练的 300 维词向量，双向 LSTM 隐藏层的维度设置为 300，学习率设置为 0.0001，L2 正则项设置为 0.01，dropout 设置为 0.3，模型中的所有权重矩阵采用均匀分布 $U(-0.01, 0.01)$ 随机初始化，所有偏移量设置为 0。实验使用 Adam 优化算法和小批量梯度下降算法，每个批次 8 个样本。

4.3 实验算法流程

实验的具体操作流程如下所述：

a) 从数据集中提取出语句、语句中的方面词和方面词在语句中的情感极性，以此生成对应的数据样本，即语句、方面词、情感极性。

b) 采用 jieba 中文分词器将数据样本中的语句和方面词进行分词，依据 Shen Li 等人^[21]的预训练词嵌入向量将样本中的文本转换为词嵌入矩阵，情感极性标签采用 one-hot 编码，如此获得转换后的数据样本。

c) 根据分词后的数据样本和词嵌入矩阵，可得到每条语句的词向量矩阵，将词向量矩阵作为模型的输入。

d) 使用微博话题数据集进行实验。

4.4 评价指标

本文实验根据式(4.1)的计算结果作为指标，以评价不同模型整体性能。

$$Accuracy = \frac{T}{N} \quad (4.1)$$

其中 *Accuracy* 表示准确率，*T* 表示模型在测试集中准确预测的样本数目，*N* 表示测试集的样本总数。一般而言，整体性能越强的模型往往具有越高

4.5 模型比较

为了全面评估 BAN 模型的性能，将其与几种 Baseline 方法进行比较，如下所示：

a) Majority: 最基本的方法，把训练集中最多的情感极性作为测试集中每个样本的预测结果。

b) LSTM: 只使用一个 LSTM 网络，将句子作为输入，得到每个词的隐藏状态后，计算所有隐藏状态的平均值作为最终表示向量，输入至 Softmax 函数计算各个情感极性的概率。

c) ATAE-LSTM^[13]: 将方面词的嵌入向量拼接每个词的嵌入向量中表示句子的嵌入矩阵，输入至 LSTM 网络中得到单词的隐含层，再将其与方面词的嵌入向量相结合生成注意力向量，进而生成用于情感分类的最终表示向量，输入至 Softmax 函数计算各个情感极性的概率。

d) IAN^[16]: 使用 LSTM 分别对句子和方面词建模，再在句子和方面词之间

相互采用注意力机制，分别生成句子的表示向量和方面词的表示向量，将二者拼接成用于情感分类的最终表示向量，输入至 Softmax 函数计算各个情感极性的概率。

e) AOA^[17]: 采用双向 LSTM 分别对句子和方面词建模，生成二者的交互矩阵，通过列式 Softmax 和行式 Softmax 分别得到句子和方面词交互的注意力向量，将方面词注意力向量求平均，再对句子注意力向量计算加权和得到最终的句子注意力向量，利用该注意力向量对句子的隐藏层语义状态计算加权和，得到最终的分类型特征，使用 Softmax 函数计算各个情感极性的概率。

表 2 显示了在微博话题数据集上各个模型的结果。

表 2 模型结果对比

Methods	准确率
Majority	0.254
LSTM	0.643
ATAE-LSTM	0.694
AOA	0.745
IAN	0.764
BAN	0.797

从表中可以看到最基本的 Majority 模型效果最差，LSTM 模型在微博话题数据集上相比于 Majority 模型性能有巨大的提升，ATAE-LSTM 模型在 LSTM 模型的基础上结合方面词向量并采用注意力机制，使得 ATAE-LSTM 模型的分类准确率比 LSTM 模型提高了 5.1%。AOA 模型和 IAN 模型也是基于 LSTM 模型，两者也都在句子和方面词之间交互采用注意力机制，从而使得分类准确率分别提高了 10.2%和 12.1%。

与表中的其他模型相比，本文提出的 BAN 模型取得最好的分类结果，本文模型的设计灵感来源于 IAN 模型，但相比 IAN 模型的改进之处在于采用双向 LSTM 模型，对句子和方面词学习时可以使模型捕获更多的语义信息，在微博话题数据集上相比 IAN 模型的准确率提高了 3.3%。

4.6 模型可视化

为说明模型的有效性，下面将举两个可视化的例子予以说明。随机从微博话题数据库中抽取两条句子作为测试对象，使用训练好的模型对其建模，绘制模型中注意力机制输出的注意力权重向量，如图 5、图 6 和图 7 所示，图中方块颜色越深表示对应的词所占的权重越高，与方面词的情感分类越相关。

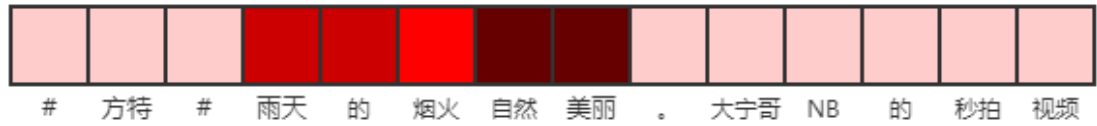


图 5 对“#方特#雨天的烟火自然美丽。大宁哥 NB 的秒拍视频”的可视化分析

对于图 5 中的话题数据“#方特#雨天的烟火自然美丽。大宁哥 NB 的秒拍视频”，针对方面词“烟火”，注意力机制关注到与“烟火”相关的词“雨天的”、“自然”和“美丽”，并且对“自然”、“美丽”的关注程度高于“雨天的”，而注意力机制忽略了句子中的标点符号和其他与方面词不相干的词。



图 6 对“#百度#百度是中国互联网最无耻的一个公司。其它公司都慢慢转好。百度从未变好”的可视化分析

对于图 6 中的话题数据“#百度#百度是中国互联网最无耻的一个公司。其它公司都慢慢转好。百度从未变好”，针对方面词“百度”，注意力机制关注到与“百度”相关的词“最无耻”和“从未变好”，而没有关注与“百度”无关的词“慢慢转好”。



图 7 对“#百度#百度是中国互联网最无耻的一个公司。其它公司都慢慢转好。百度从未变好”的可视化分析

对于图 7 中的话题数据“#百度#百度是中国互联网最无耻的一个公司。其它公司都慢慢转好。百度从未变好”，针对方面词“其他公司”，注意力机制关注到与“其他公司”相关的词“慢慢转好”，而并未关注与“其他公司”无关的词“最无耻”和“从未变好”。

从以上的案例分析可以看出，本文模型的注意力机制能够从句子中提取出与

方面词相关的重要词汇信息，并且能够忽略毫不相干的词汇和标点符号，以此可说明本文模型引入注意力机制取得不错的效果。

4.7 本章小结

本章详细介绍了本文的实验部分。对实验中采用的数据集进行了统计分析以了解它的特点，接着介绍了实验的环境和本文模型参数的设置，然后阐述了本文实验的算法流程。本章还说明了实验结果的评价指标，即准确率。通过与几种 **Baseline** 模型实验结果的对比分析，验证文本模型的有效性。本章通过两个可视化的案例直观地说明模型的注意力机制着重关注与方面词有关的情感词，印证了模型注意力机制的有效性。

第五章 总结与展望

本文研究方面级情感分类问题,在参考大量文献以及对现有工作分析的基础上,针对 LSTM 网络只能得到 t 序号词与 t 序号词之前的词的关系,而无法得到与 t 序号词之后的词的关系这一缺陷,提出一个基于双向 LSTM 和注意力机制的网络模型。本文模型利用双向 LSTM 网络,以词嵌入向量为输入,分别捕获句子和方面词的隐含层。将方面词隐含层的平均值和句子的隐含层作为输入,借助特定的注意力机制生成句子的注意力向量,依据生成的注意力向量,对句子的隐含层进行加权计算,得到与方面词相关的句子表达向量。通过这种设计,可以使句子表达向量的生成受方面词的隐含层的平均值的影响。将句子表达向量输入至全连接网络,利用 Softmax 函数将句子表达向量转化为各个情感类别的概率。最后,选择概率最大的情感类别作为预测结果。本文模型在微博话题数据集中与其他几种 Baseline 模型相比分类准确率较高。本文还采用两个可视化的案例说明模型的注意力机制能够从句子中提取出与方面词相关的重要词汇信息。

在未来的工作中可以考虑将方面词的词向量和句子中每个词的词向量拼接后作为模型的输入,使模型能够捕捉更多的信息,同时在模型中加入位置信息、词性、情感词典和情感词极性额外信息来辅助提高模型的分类准确率。此外在以后的工作中可以尝试改进注意力机制的计算方式,进一步提高注意力的准确性,进而提高模型的性能。

参考文献

- [1]. 何炎祥, 孙松涛, 牛菲菲, 李飞. 用于微博情感分析的一种情感语义增强的深度学习模型 [J]. 计算机学报, 2017, 40(04): 773-790.
- [2]. 黄胜. Web 评论文本的细粒度意见挖掘技术研究[D]. 北京理工大学, 2014.
- [3]. Bing Liu. Sentiment analysis and subjectivity [J]. Indurkha N, and Damerau FJ. Handbook of Natural Language Processing, Second Edition. CRC Press, Taylor & Francis Group, 2010: 627-666.
- [4]. 于朝晖. CNNIC 发布第 44 次《中国互联网络发展状况统计报告》[J]. 网信军民融合, 2019(09): 30-31.
- [5]. 唐晓波, 刘广超. 细粒度情感分析研究综述 [J]. 图书情报工作, 2017, 61 (05): 132-140.
- [6]. Erik Boiy, Marie-Francine Moens. A machine learning approach to sentiment analysis in multilingual Web texts [J]. Information retrieval, 2009, 12(5): 526-558.
- [7]. Long Jiang, Mo Yu, Ming Zhou, Xiaohua Liu, Tiejun Zhao. Target-dependent Twitter Sentiment Classification [J]. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics. Portland, Oregon. 2011: 151-160.
- [8]. Luciano Barbosa, Junlan Feng. Robust Sentiment Detection on Twitter from Biased and Noisy Data [C]. Proceedings of International Conference on Computational Linguistics: Posters. Association for Computational Linguistics, 2010: 36-44.
- [9]. Li Dong, Furu Wei, Chuanqi Tan, Duyu Tang, Ming Zhou, Ke Xu. Adaptive Recursive Neural Network for Target-dependent Twitter Sentiment Classification [C]. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, Maryland, USA. 2014:49-54.
- [10].Duy-Tin Vo, Yue Zhang. Target-Dependent Twitter Sentiment Classification with Rich Automatic Features [C]. Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence. 2015:1347-1353.
- [11].Duyu Tang, Bing Qin, Xiaocheng Feng, Ting Liu. Effective LSTMs for target-dependent sentiment classification [C]. Proceedings of the 26th International

- Conference on Computational Linguistics: Technical Papers. Osaka, Japan, 2016: 3298-3307.
- [12].Meishan Zhang, Yue Zhang, Duy-Tin Vo. Gated Neural Networks for Targeted Sentiment Analysis [C]. Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence. Phoenix, USA, 2016: 3087-3093.
- [13].Yequan Wang, Minlie Huang, Li Zhao, Xiaoyan Zhu. Attention-based LSTM for Aspect-level Sentiment Classification [C]. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, Texas, USA, 2016: 606-615.
- [14].Duyu Tang, Bing Qin, Ting Liu. Aspect Level Sentiment Classification with Deep Memory Network [C]. Proceeding of the 21st Conference on Empirical Methods in Natural Language Processing. Austin, USA, 2016: 214-224.
- [15].Peng Chen, Zhongqian Sun, Lidong Bing, Wei Yang. Recurrent Attention Network on Memory for Aspect Sentiment Analysis [C]. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen, Denmark, 2017: 452-461.
- [16].Dehong Ma, Sujian Li, Xiaodong Zhang, Houfeng Wang. Interactive Attention Networks for Aspect-Level Sentiment Classification [C]. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. Melbourne, Australia, 2017: 4068–4074.
- [17].Binxuan Huang, Yanglan Ou, Kathleen M. Carley. Aspect Level Sentiment Classification with Attention-over-Attention Neural Networks [C]. Proceedings of the 11th International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation. Washington, DC, USA, 2018: 197-206.
- [18].Zachary C. Lipton, John Berkowitz, Charles Elkan. A Critical Review of Recurrent Neural Networks for Sequence Learning[J]. arXiv preprint arXiv:1506.00019, 2015.
- [19].Sepp Hochreiter, Jurgen Schmidhuber. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.

- [20].Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio. Neural Machine Translation by Jointly Learning to Align and Translate[C]. Proceedings of the 3rd International Conference on Learning Representations, San Diego, California, 2015.
- [21].Shen Li, Zhe Zhao, Renfen Hu, Wensi Li, Tao Liu, Xiaoyong Du. Analogical Reasoning on Chinese Morphological and Semantic Relations [J]. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, Australia, 2018: 138-143.

致谢

附录