

# 注意力增强的双向 LSTM 情感分析\*

关鹏飞<sup>1</sup>, 李宝安<sup>1</sup>, 吕学强<sup>1</sup>, 周建设<sup>2</sup>

(1. 北京信息科技大学 网络文化与数字传播北京市重点实验室, 北京 100101; 2. 首都师范大学, 北京成像技术高精尖创新中心, 北京 100048)

**摘要:** 在基于深度学习的情感分析工作中, 传统的注意力机制主要以串行的方式作为其他模型的下一层, 用于学习其他神经网络模型输出的权重分布。该文在探究使用深度学习进行句子级情感分析任务的基础上, 提出一种注意力增强的双向 LSTM 模型, 模型使用注意力机制直接从词向量的基础上学习每个词对句子情感倾向的权重分布, 从而学习到能增强分类效果的词语, 使用双向 LSTM 学习文本的语义信息, 最终通过并行融合的方式提升分类效果。通过在 NLPCC2014 情感分析语料上进行测试, 该模型的结果优于其他句子级情感分类模型。

**关键词:** 注意力机制 双向 LSTM 情感分析

**中图分类号:** TP391.1

**文献标志码:** A

## Attention to strengthen Bi-directional LSTM for sentiment analysis

Guan Pengfei<sup>1</sup>, Li Baoan<sup>1</sup>, Lv Xueqiang<sup>1</sup>, Zhou Jianshe<sup>2</sup>

(1. Beijing Information Science & Technology University, Beijing Key Laboratory of Internet Culture & Digital Dissemination Research, Beijing 100101, China; 2. Capital Normal University, Beijing Advanced Innovation Center for Imaging Technology, Beijing 100048, China)

**Abstract:** In the study of based on deep learning sentiment analysis, traditional attention mechanism is mainly in the form of serial as the next layer of other model, is used to study other weight distribution of neural network model output. Based on exploring how to use deep learning to solve sentiment analysis of sentence level task, this paper proposes a model of attention enhanced Bi-directional LSTM, use attention mechanism directly learn every word weight distribution of sentiment tendency from the word vector, so as to learn words that can enhance the classification effect, use Bi-directional LSTM learning text semantic information, eventually improve classification by means of parallel fusion effect. Through the test on corpus NLPCC2014 sentiment analysis, the results of the model is superior to the other sentence level sentiment classification model.

**Key word:** Attention mechanism, Bi-directional LSTM, sentiment analysis

## 1 引言

随着互联网的发展, 网民数量在近几年急剧上升, 人们在信息交互的过程中产生了大量的对诸如人物、事件、产品等有价值的评论信息。这些信息表达了人们的各种感情色彩和情感倾向性, 通过对情感信息的挖掘, 可以更好的理解用户行为, 从而预测出事件的发展方向或趋势<sup>[1]</sup>。但随着信息规模的极具膨胀, 仅靠人工已经不可能完成情感分析工作, 所以使用计算机进行高效准确的情感分析工作有着重要的意义。

目前情感分析技术主要分为 3 类: 基于情感词典的方法、基于特征的方法和基于深度学习的方法。基于情感词典的方法主要根据文本中的情感词来判断文本的情感倾向, 需要人工构建情感词典, 结合情感词典和人工设置规则实现对文本的情感分析<sup>[2][3]</sup>。这种方法基于人类

\* 收稿日期:

定稿日期:

**基金项目:** 国家自然科学基金项目 (项目编号: 61671070); 国家语委重点项目 (ZDI135-53); 北京成像技术高精尖创新中心项目 (项目编号: BAICIT-2016003)

**作者简介:** 关鹏飞 (1993-), 男, 河北省唐山人, 硕士研究生, 主要研究方向为自然语言处理, 北京市北四环中路北京信息科技大学健翔桥校区图书馆五层, perfect\_guan@foxmail.com; 李宝安 (1965-), 男, 硕士, 教授, 主要研究方向为社会计算与服务计算; 吕学强 (1970-), 男, 博士, 教授, 主要研究方向为中文与多媒体信息处理; 周建设, 首都师范大学教授, 博士, 博导, 主要研究方向为语言学及应用语言学

语言的表述方式，透过情感词可以反映人的情感倾向，但是没有考虑到上下文的语义信息。基于特征的方法是采用统计学知识，从大量语料中选取特征，使用这些特征对文本进行表示，然后使用决策树、支持向量机（SVM）等机器学习算法进行分类<sup>[4][5]</sup>。该方法对经验要求较高，特征的选取直接影响分析结果。基于深度学习的方法，可以分别对词语、句子和篇章进行向量化表示，学习文本的深层语义信息<sup>[6]</sup>。该方法有强大的特征学习能力，省去了特征选取和规则制定等步骤。常见的深度学习模型有：卷积神经网络（Convolutional Neural Networks, CNN）、循环神经网络（Recurrent Neural Networks, RNN）等。

本文以短文本为情感分析的研究对象，提出一种注意力增强的双向 LSTM 模型，双向 LSTM（Bi-directional Long Short-Term Memory, BiLSTM）模型用于获取文本的语义信息，注意力机制（Attention mechanism）用于获取文本中的关键词信息。与传统的使用注意力机制方式不同，本文不再使用注意力机制学习 CNN、LSTM 等模型的输出，而是直接在词向量的基础上使用注意力模型来发现文本中重要信息。实验表明该模型在情感分类上取得了较好的效果。

本文将在下一节介绍使用深度学习技术进行情感分析的相关工作，在论文第三节介绍注意力增强的双向 LSTM 模型的结构及原理，第四节将对实验数据进行介绍，并在数据集上使用不同模型进行对比，最后一节对本论文的工作进行总结并提出展望。

## 2 相关工作

自 2006 年 Hinton 等提出深度学习后，深度学习在计算机视觉方面取得了巨大的成就，越来越多的学者也开始将深度学习应用到自然语言处理领域。由于短文本包含的信息量较少，同时传统方法不能发现文本中的深层语义信息，所以深度学习成为情感分析研究的主流方法。

基于深度学习的情感分析通常是使用神经网络将文本表示成一定长度的向量，所以基于深度学习的情感分析可以理解为自动学习文本编码然后分类的过程。Kim 等<sup>[7]</sup>在多个英文语料上利用卷积神经网络实现句子级别的分类任务，并在相关电影评论数据集上讨论了词向量预训练及参数调整等细节工作。Cao 等<sup>[8]</sup>利用迁移学习的思想，将 CNN 提取的文本特征使用 SVM 进行分类，该方法将传统机器学习模型和深度学习模型相融合，在情感分析任务上取得不错的效果。冯兴杰等<sup>[9]</sup>使用 CNN 和注意力机制在词向量的基础上表示文本，相比与 CNN 网络方法有明显的提高，该模型结构对本文的工作有很大的启发。Qian 等<sup>[10]</sup>通过改变损失函数的方式，把情感词典，否定词和程度副词等现有语言规则融入到句子级 LSTM 情感分类模型中，在没有增大模型复杂度的情况下，有效的利用语言学规则，在实验数据集上取得了较好效果。张仰森等<sup>[11]</sup>采用双向 LSTM 网络分别对微博文本和文本中包含的情感符号进行编码，该方法通过将注意力模型和常用网络用语的微博情感符号库相结合，有效增强了对微博文本情感语义的捕获能力，提高了微博情感分类的性能。胡朝举等<sup>[12]</sup>通过双向 LSTM 将主题词向量和文本词向量进行训练，将得到主题特征和文本特征进行特征融合，经过深层注意力机制的处理，使模型加强了对主题特征和情感信息的有效关注。

现阶段基于深度学习的情感分析是以发掘文本语义信息为主的，但是在情感分析工作中，句中每个词对于整体情感倾向的影响是不同的，尤其是一些情感词，这些词语往往能够直接反映出人情感倾向。文<sup>[10]</sup>虽然使用了注意力机制学习权重分布，但其是在双向 LSTM 编码的基础上进行的，没有直接从文本中挖掘重点信息。本文提出一种注意力增强的双向 LSTM 模型，使用注意力机制直接从词向量的基础上学习每个词对情感倾向的权重分布。本模型通过使用双向 LSTM 学习文本的语义信息，最终使用重点词增强的方式提升分类效果。

## 3 注意力增强的双向 LSTM 结构

本文采用注意力机制与双向 LSTM 结合的方式，使用双向 LSTM 学习文本语义信息，使用注意力机制加强对重点词的关注，整体结构如图 1 所示。首先对输入的句子利用预训练好词向量进行表示，然后分别经过双向 LSTM 模型和注意力模型学习表示，将两部分表示后的向

量拼接，最终通过分类器完成文本情感分析的工作。

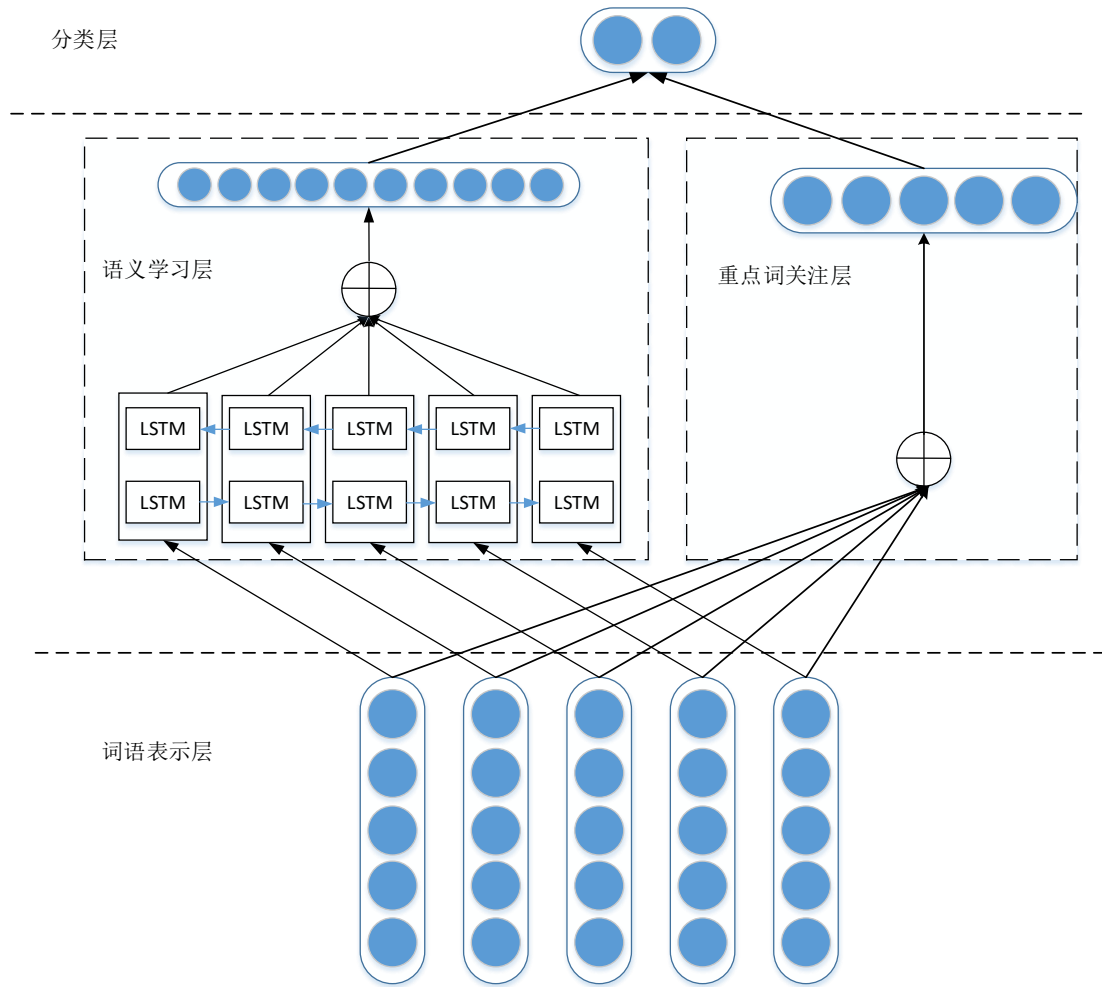


图1 词向量注意力机制的双向 LSTM 模型结构

### 3.1 词语表示层

词作为该模型处理的基本单元，第一步是对词进行符号化表示。与传统的表示方法不同，本文使用连续的稠密向量作为模型的输入。随着 Mikolov 等人<sup>[13][14]</sup>提出了 Word2vec 模型，词向量可以在低维空间中更好的学习词语蕴含语义信息。词语表示层会在相关语料上预训练出一个 $R^{N \times d}$ 规模的词典， $N$ 表示词典中词的个数， $d$ 表示词向量的维度。在进行词语表示时，用 $x_t$ 表示文本中的第  $t$  个词， $x_t \in R^d$ 。若文本长度为 $T$ ，则输入文本表示为：

$$S = [x_1; x_2; \dots; x_T] \in R^{T \times d}$$

### 3.2 语义学习层

自然语言的词语之间存在时序关系，为了让模型可以学习词语间的语义依赖关系，本文采用双向 LSTM 模型对句子的语义信息进行编码。双向 LSTM 由正反两个 LSTM 模型组成，长短期记忆网络（Long Short-Term Memory, LSTM）<sup>[15]</sup>是循环神经网络的一种，它有较强的长距离语义捕获能力，图 2 是 LSTM 的基本结构。

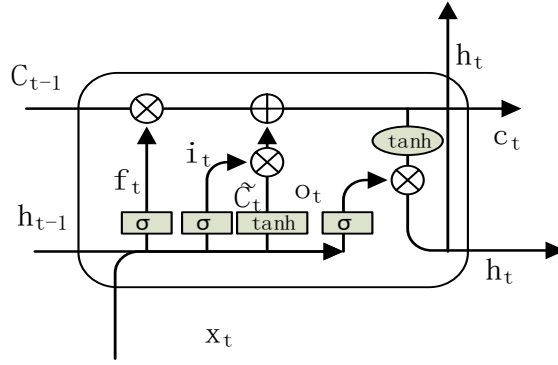


图 2 LSTM 结构

$x_t$ 为  $t$  时刻 LSTM 单元的输入数据,  $h_t$ 是  $t$  时刻输出,  $C$ 是不同时刻记忆单元的值。LSTM 主要有三个门结构控制模型中信息的加工。遗忘门 $f_t$ 决定记忆信息的通过量, 该门将 $x_t$ 和上一时刻输出 $h_{t-1}$ 作为输入, 输出值在 0 和 1 之间, 值用来描述每个部分通过量的多少。 $f_t$ 的计算公式为:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

式中 $\sigma$ 为 sigmoid 函数,  $W_f$ 是遗忘门权重,  $b_f$ 为遗忘门偏置。输入门 $i_t$ 控制当前的输入数据对记忆单元状态的影响,  $\tilde{C}_t$ 为要被添加到记忆单元的候选值。 $i_t$ 和 $\tilde{C}_t$ 的更新公式分别为:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C)$$

$W_i$ 为输入门权重,  $b_i$ 是输入门偏置,  $\tanh$ 为双曲正切函数,  $W_C$ 为更新后候选值权重,  $b_C$ 为更新候选值偏置。接下来更新记忆单元的状态, 由状态 $C_{t-1}$ 至 $C_t$ 状态, 原来状态 $C_{t-1}$ 根据遗忘门的值丢弃要屏蔽的信息, 根据输入门的值添加新的信息。 $C_t$ 更新公式为:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

输出门 $o_t$ 控制输出信息, 该门通过 $x_t$ 和上一时刻输出决定从当前状态中输出的信息量, 状态 $C_t$ 通过  $\tanh$  函数得到区间在-1 和 1 的值, 该值乘以 $o_t$ 作为本时刻的输出值。 $o_t$ 和 $h_t$ 的公式分别为:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

式中 $W_o$ 为更新输出值的权重,  $b_o$ 是更新输出值偏执,  $h_t$ 为最终输出值。

传统的 LSTM 模型能学习到文本中的正向语义信息, 但忽略了未来的上下文信息, 因此本文使用双向 LSTM 模型增加对文本逆向语义的学习。连接正反两个方向的 LSTM 模型输出向量 $\vec{h}_t$ 、 $\overleftarrow{h}_t$ 作为  $t$  时刻双向 LSTM 的输出 $B_t$ :

$$\vec{h}_t = \overrightarrow{\text{LSTM}}(h_{t-1}, x_t, C_{t-1}), \overleftarrow{h}_t = \overleftarrow{\text{LSTM}}(h_{t+1}, x_t, C_{t+1})$$

$$B_t = [\vec{h}_t, \overleftarrow{h}_t]$$

对每个双向 LSTM 单元的输出做和取平均作为语义学习层输出, 公式为:

$$\text{output}_{\text{sema}} = \frac{\sum_1^T B_t}{T}$$

### 3.3 重点词关注层

在句子级情感分析中, 由于句子中包含词数较少, 也就难以获取很多的语义信息。我们通过对语料的分析, 见表 1, 在句子中的情感词(如: 喜欢、失望等)往往更能体现用户的情感倾向。为了加强这些情感词在分类时的作用, 本文采用注意力机制学习句子中不同词的权重分布。

表 1 重点词分析

评论	重点词
很 <b>喜欢</b> ，这个价格，质量比想像中 <b>好</b> 很多，很 <b>喜欢</b> !	好、喜欢
希望以后普通版不要那么 <b>寒碜</b> 了，买的很 <b>不值</b> 。	寒碜、不值
整个系列都是 <b>非常好</b> 的书	非常好
音质 <b>不好</b> 。有的还可以，有的就像是录的音似的。 <b>失望</b> !	不好、失望

注意力机制最早在计算机视觉领域提出来的，模仿人类的注意力机制，给图像不同的局部赋予不同的权重。后来 Bahdanau 等人<sup>[16]</sup>将注意力机制应用到机器翻译，这也是注意力机制在自然语言处理领域的首次应用。但是机器翻译中的注意力机制是将前一时间步输出的隐含状态与当前时间步输入的隐含状态进行对齐的模式，而在情感分析任务中使用对当前输入自适应加权的自注意力机制（Self-attention）更合适。Wang 等人<sup>[17]</sup>把自注意力模型应用到方面级（aspect-level）情感分析任务上，将自注意力机制与 LSTM 结合在一起，通过自注意力机制去获取对不同方面更重要的上下文信息，来解决某个给定的方面情感分类的问题。自注意力机制通常也不会使用其他额外的信息，它会自动从所给数据中学习权重分布：

$$v_t = \tanh(Wx_t + b)$$

$$a_t = \frac{\exp(v_t A)}{\sum_{t=1}^T \exp(v_t A)}$$

其中  $a_t$  表示第  $t$  个词对于当前文本的重要程度， $v_t A$  作为一种评分制度由模型从语料中自动学习， $A$ 、 $W$  均为权重矩阵， $b$  为偏置。得到每个词的权重后，假设句中词数为  $T$ ，将词向量按权重求和作为重点词语关注层的输出：

$$output_{att} = \sum_{t=1}^T a_t x_t$$

### 3.4 分类层

文本情感分析本质上是一个分类任务，所以模型的最后一层是分类层，分类层将语义学习层和重点词语关注层的结果连接作为输入：

$$input_{classify} = [output_{sema}, output_{att}]$$

输出每类的概率为  $p_c$ ：

$$V = w_{classify} input_{classify} + b_{classify}$$

$$p_c = \frac{\exp(V_c)}{\sum_{j=1}^C \exp(V_j)} \quad \text{for } c \text{ in } (1, 2, \dots, C)$$

$w_{classify}$  为  $L \times C$  的权重矩阵， $L$  为输入向量的维度， $C$  为类数， $b_{classify}$  为分类层的偏置向量。在得到预测的概率分布后，本文采用交叉熵损失函数来衡量真实分布和预测分布之间的差距，从而利用反向传播对模型中的参数进行更新。

## 4 实验与分析

### 4.1 实验数据

本文的实验数据为 NLPC2014 情感分析（NLPC2-SCDL）评测任务中文数据集，该数据集中共收集 12500 条中立、负面均衡的网购商品评论。从语料随机选取中立、负面数据 1250 条作为测试数据（共 2500 条），其余作为训练数据。实验语料采用 Jieba 分词工具进行分词，分词后句子长的分布如图 3 所示

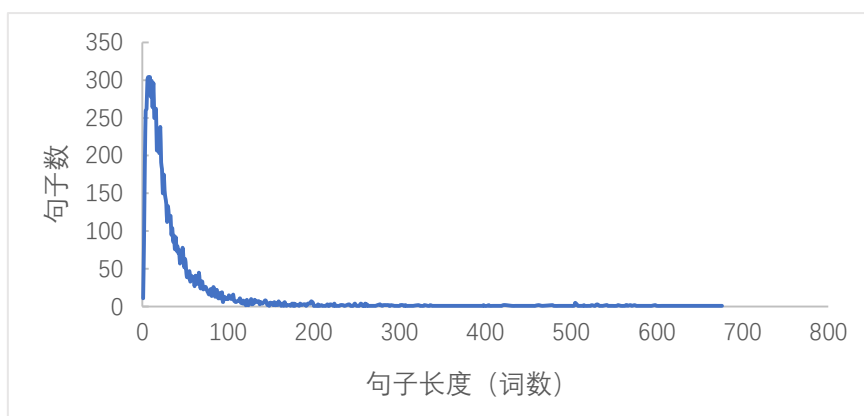


图3 语料句子长度分布

本文选用维基百科的中文语料作为预训练词向量的语料库,因为中文维基百科语料中含有大量的繁体字,所以需要先对语料进行繁简转换,再使用 Jieba 进行分词。本文利用 Google 开源的 word2vec 工具预训练词向量,选择 cbow 模型,上下文窗口大小设置为 5,词向量维度大小设为 200,采样值大小设为  $1e-3$ ,其他参数使用模型的默认值。在使用预训练的词向量时采用随机初始化<UNK>来表示未登录词,使用全零初始化<PAD>补齐句子,并在训练神经网络时对词向量进行微调。

#### 4.2 评价标准

本实验采用正确率(Precision)、召回率(Recall)、F1 值(F1-measure)准确率(Accuracy)作为评价标准。

TP: 正确分类中的中立条数

FP: 错误分类中的中立条数

TN: 正确分类中的负面条数

FN: 错误分类中的负面条数

$$\text{Precision}_{neu} = \frac{TP}{TP+FP}, \text{Precision}_{neg} = \frac{TN}{TN+FN}$$

$$\text{Recall}_{neu} = \frac{TP}{TP+FN}, \text{Recall}_{neg} = \frac{TN}{TN+FP}$$

$$\text{F1}_{neu} = \frac{2 * \text{Precision}_{neu} * \text{Recall}_{neu}}{\text{Precision}_{neu} + \text{Recall}_{neu}}$$

$$\text{F1}_{neg} = \frac{2 * \text{Precision}_{neg} * \text{Recall}_{neg}}{\text{Precision}_{neg} + \text{Recall}_{neg}}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

#### 4.3 实验对比模型

**NBOW:** Mikolov 提出的神经网络词袋模型,该模型将句中的词向量的平均值作为句子表示,在文本分类任务中的效率很高。

**CNN:** 基于 kim 等人提出的 CNN 分类模型,使用卷积核大小为 1、2、3 分别提取文本特征,经过最大池化后的使用 softmax 进行情感分类。

**CNN+SVM:** Cao 等人提出的使用 CNN 提取文本特征,使用迁移学习的思想,利用 SVM 进行文本分类,该模型在 NLPCC-SCDL 任务上取得了最好的效果。

**CNN+EMB\_ATT:** 冯兴杰等人提出的基于词向量注意力机制的卷积神经网络模型。

**LSTM:** 使用 LSTM 进行语义学习,将每个输入对应的输出取平均作为文本表示。

**LSTM+ATT:** 使用 LSTM 进行语义学习,使用自注意力机制学习 LSTM 的输出,最后按权重做和

作为句子的表示。

**BiLSTM:** 使用双向 LSTM 学习文本语义，将每个输入对应的输出取平均作为文本表示。

**BiLSTM+ATT:** 使用双向 LSTM 进行语义学习，使用自注意力机制学习双向 LSTM 的输出，最后按权重做和作为句子的表示。

**BiLSTM+EMB\_ATT:** 本文提出的注意力增强的双向 LSTM 模型。

#### 4.4 实验参数

因为模型输入需要一定的长度限制，根据 4.1 节中的语料句子的长度分布，将模型输入文本的最大长度限制在 100 个词，超出 100 个词的句子截取前 100 个词，不足的用<PAD>补齐。LSTM 模型和自注意力机制的隐藏层设置为 200。每轮参数更新的学习率衰减为上轮学习率的 0.8 倍。当准确率在 5 轮内没有超过当前最高的准确率时，模型提前终止学习。模型中使用的主要超参数及参数值如表 2。

表 2 模型中的参数设置

参数名	参数值
句子最大长度	100
词向量维度	200
隐藏层大小	200
批处理个数	64
学习率	0.01
学习率的衰减率	0.8
提前终止	5
L2 正则系数	0.001

#### 4.5 实验结果及分析

本实验在服务器上进行，服务器 CPU 为 2 颗英特尔至强 (R) 处理器 E5-2603 v4，GPU 为 NVIDIA Tesla K40M。在训练阶段，将训练数据分为 5 份做交叉验证，最终实验结果取 5 次实验的平均值。实验结果如表 3。

表 3 实验结果

模型	中立			负面			Acc	训练用时
	P	R	F1	P	R	F1		
NBOW	76.92	74.40	75.64	75.21	77.68	76.43	76.04	2min52s
CNN	76.42	78.56	77.48	77.94	75.76	76.84	77.16	2min43s
CNN+SVM	75.73	75.12	75.42	75.32	75.92	75.62	75.52	3min11s
CNN+EMB_ATT	77.68	78.80	78.24	78.49	77.36	77.92	78.08	6min18s
LSTM	77.58	78.08	77.83	77.94	77.44	77.69	77.76	27min5s
LSTM+ATT	78.32	77.76	78.04	77.92	78.48	78.20	78.12	20min59s
BiLSTM	<b>79.35</b>	76.24	77.76	78.14	<b>80.16</b>	78.62	78.20	36min19s
BiLSTM+ATT	77.65	77.28	77.47	77.39	77.76	77.57	77.52	33min
BiLSTM+EMB_ATT	78.54	<b>79.04</b>	<b>78.79</b>	<b>78.90</b>	78.40	<b>78.65</b>	<b>78.72</b>	43min44s

从表中可以看出，本文提出的 BiLSTM+EMB\_ATT 模型在除中立准确率、负面召回率两个指标上均取得最好的结果。因为自然语言存在时序性的特点，LSTM 模型中的记忆单元有效记录了文本时序信息，本文采用双向 LSTM 结构学习文本的语义信息，为模型中增加了逆向文本的语义信息，加强了模型对文本上下文的语义学习能力。但 LSTM 模型只是理论上可以持续记忆文本信息，在实际效果上看 LSTM 依然存在着不足，而自注意力机制在情感分析的

任务需求上,可以通过学习文本规律,以分配不同词语相应权重的方式自动捕获影响文本情感倾向重要信息,为了便于观察文本中的注意力分布,本文将权重输出,如图4。所以可以使用注意力机制作为辅助的方式从文本中得到更多的信息。通过在 BiLSTM 模型上加入基于词向量的注意力机制,结果得到了提升,验证了可以使用注意力机制加强模型直接从词向量中学习文本信息的想法。

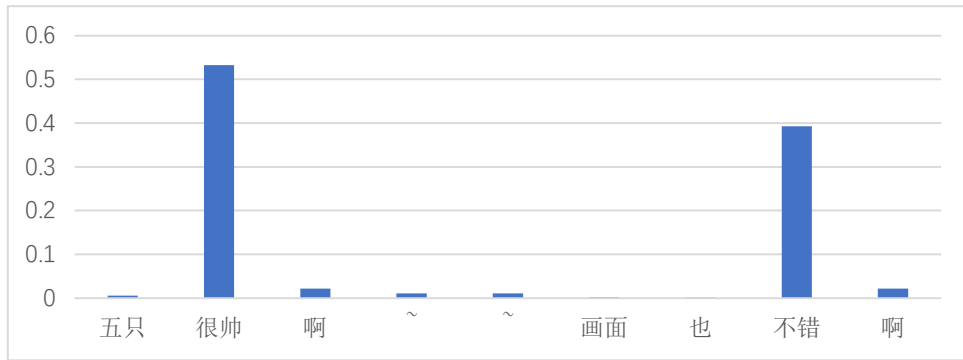


图4 注意力分布图

另外,根据实验结果还可以得出结论:(1) CNN 和 CNN+SVM 的结果说明了在该实验语料上,使用 SVM 代替 softmax 分类器的方式并不能提高结果,因为 softmax 分类会使用一层全链接网络对特征进行学习,在最小化损失学习后,softmax 可以让概率分布更接近标准结果一些。(2) 在 LSTM 模型和 BiLSTM 模型上直接加上注意力机制的结构训练用时更短,说明了注意力机制虽然增加了模型的规模,但是其关注关键信息的特性使得模型的收敛速度更快。

(3) 在 LSTM 和 BiLSTM 模型上增加注意力机制并没有达到相同的效果,这说明盲目加入注意力机制有时不会使结果更好,相反,我们在 CNN 上加入基于词向量的注意力机制的 CNN+EMB\_ATT 模型和 BiLSTM 上加入基于词向量的注意力机制的 BiLSTM+EMB\_ATT 模型的效果都得到了提高,这也验证了正确的加入注意力机制有助于模型效果的提升。

## 5 总结与展望

本文提出了一种注意力增强的双向 LSTM 情感分析模型,使用双向 LSTM 学习文本的语义信息,使用在词向量上建立的自注意力机制加强对句中情感关键词的关注度。与传统双向 LSTM 模型加注意力机制不同,本文采用的词向量注意力机制与双向 LSTM 为并行结构,实验表明,本文提出的模型表现出优越的性能,并在多个指标上超过了已知的最好模型。

在深度学习流行之前,前人们总结了很多如 HowNet 等知识库,这些知识是自然语言处理领域的巨大宝藏,现在很多学者都在尝试如何利用这些知识库,在接下来的研究中,我们也将深入研究如何用现有的语言学规则和知识库提高注意力增强的双向 LSTM 模型的效果。

## 参考文献

- [1] 李然,林政,林海伦,等. 文本情绪分析综述[J]. 计算机研究与发展, 2018, 55(1): 30-52.
- [2] 吴江,唐常杰,李太勇,等. 基于语义规则的 Web 金融文本情感分析[J]. 计算机应用, 2014, 34(2): 481-485.
- [3] Taboada M, Brooke J, Tofiloski M, et al. Lexicon-based methods for sentiment analysis[J]. Computational Linguistics, 2011, 37(2): 267-307.
- [4] 李婷婷,姬东鸿. 基于 SVM 和 CRF 多特征组合的微博情感分析[J]. 计算机应用研究, 2015, 32(4): 978-981.
- [5] Pu X, Wu G, Yuan C. Exploring overall opinions for document level sentiment classification with structural SVM[J]. Multimedia Systems, 2017(1-2): 1-13.
- [6] 梁军,柴玉梅,原慧斌,等. 基于深度学习的微博情感分析[J]. 中文信息学报, 2014, 28(5): 155-161.



- [7] Kim Y. Convolutional Neural Networks for Sentence Classification[J]. Eprint Arxiv, 2014.
- [8] Cao Y, Xu R, Chen T. Combining Convolutional Neural Network and Support Vector Machine for Sentiment Classification[C]// Chinese National Conference on Social Media Processing. Springer, Singapore, 2015: 144-155.
- [9] 冯兴杰, 张志伟, 史金钊. 基于卷积神经网络和注意力模型的文本情感分析[J/OL]. 计算机应用研究, 2018, 35(05): 1434-1436.
- [10] Qian Q, Huang M, Lei J, et al. Linguistically Regularized LSTM for Sentiment Classification[C]// Meeting of the Association for Computational Linguistics. 2017: 1679-1689.
- [11] 张仰森, 郑佳, 黄改娟, 等. 基于双重注意力模型的微博情感分析方法[J]. 清华大学学报(自然科学版) 2018, 58(02): 122-130.
- [12] 胡朝举, 梁宁. 基于深层注意力的 LSTM 的特定主题情感分析[J]. 计算机应用研究, 2019, (05): 1-3.
- [13] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space[J]. Computer Science, 2013.
- [14] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality [C] // Proc of the NIPS. 2013: 3111-3119.
- [15] Hochreiter S, Schmidhuber J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [16] Bahdanau D, Cho K, Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate[J]. Computer Science, 2014.
- [17] Wang Y, Huang M, Zhu X, et al. Attention-based LSTM for Aspect-level Sentiment Classification[C]// Conference on Empirical Methods in Natural Language Processing. 2016: 606-615.