

# 注意力机制在评论文本情感分析中的应用研究

申静波, 李井辉, 孙丽娜

(东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318)

**摘要:**长短期记忆神经网络(long short term memory, LSTM)在文本情感分类的准确率方面拥有优秀的表现,能够解决基于长文本序列的模型训练过程中的梯度消失和梯度爆炸等问题。针对传统的LSTM分类模型不能突出体现输出的某个词语对分类类别的贡献程度和重要性的现象,在循环神经网络(RNN)变体长短期记忆人工神经网络(LSTM)隐藏层和输出层之间引入注意力机制,其主要目的是在模型做最后的极性分类预测时,在重要的句子成分加上权重,加大了对最后分类的预测概率的影响因素。实验证明LSTM与注意力机制的融合可突出模型产生动态变化的背景向量以表现出不同输入词语对于输出词语分类的重要性,从而有效提高了分类速率和准确率。

**关键词:**评论文本;情感分析;长短期记忆网络;Seq2Seq模型;注意力机制

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2020)07-0169-05

doi:10.3969/j.issn.1673-629X.2020.07.036

## Research on Application of Attention Mechanism in Comment Text Emotional Analysis

SHEN Jing-bo, LI Jing-hui, SUN Li-na

(School of Computer and Information Technology, North East Petroleum University, Daqing 163318, China)

**Abstract:** Long short term memory (LSTM) has an excellent performance in the accuracy of text emotional classification, and can solve the problems of gradient disappearance and gradient explosion in the process of model training based on long text sequences. In view of the phenomenon that the traditional LSTM classification model cannot incarnate the output word's contribution to the classification category degree and the importance, we introduce the attention mechanism between hidden layer and output layer in long short term memory (LSTM) variated from circulatory neural network (RNN), aiming to add weights to the important sentence components when the model makes the final prediction of polarity classification and increasing the influence factors on the prediction probability of the final classification. Experiment shows that the integration of LSTM and attention mechanism can highlight the background vector of dynamic changes in the model to show the importance of different input words to the output words classification, thus effectively improving the classification rate and accuracy.

**Key words:** comment text; sentiment analysis; LSTM; Seq2Seq model; attention mechanism

## 0 引言

情感分析是自然语言处理 NLP 领域最受欢迎的应用之一,挖掘出来的情感信息可以反映发布者当时的情感状态,在个性化推荐、舆情控制、基于社会调查的政策制定等方面蕴含着极大的应用价值<sup>[1-2]</sup>。利用情感极性,即情绪的正值或负值量化,判定文本情感倾向并抓住文本情感趋势<sup>[3-4]</sup>。LSTM<sup>[5-7]</sup>应用于情感分析领域,结合 Word2Vec 词嵌入技术<sup>[8]</sup>来进行实验可取得不错的情感分类效果。注意力机制最早是在计算机视觉领域被提出的,应用在图像处理任务中,其目的是利用注意力机制的特征针对性使网络模型在训练时

重点关注重要信息。传统的 Seq2Seq 模型是不含注意力机制的解码-编码模型,模型训练过程中只利用唯一的背景变量,效果不明显<sup>[9]</sup>。文中介绍了应用于机器翻译任务的编码-解码器的工作原理,同时揭示了其生成的固定长度的背景变量在长序列文本处理中的局限。而注意力机制是从编码-解码器的内部表征结构中衍生出来的一种方法,注意力机制可以有选择地关注输出与输入样本的相关性<sup>[10-11]</sup>。文中引入注意力机制与神经网络相结合,处理文本分类问题,不但降低了高维数据的计算难度,而且直观地表现出了分类结果与相关极性词语的关联性,提升了实验的准确率。

收稿日期:2019-06-13

修回日期:2019-10-16

基金项目:国家青年基金(61702093);黑龙江省高等教育教学改革项目(SJGY20170044)

作者简介:申静波(1974-),女,硕士,副教授,研究方向为自然语言处理、人工智能等。

## 1 注意力机制方法研究

### 1.1 Seq2Seq 模型

注意力机制在解决序列学习任务上有巨大的提升意义。与循环神经网络不同,RNN 的输入是不定长序列,而输出为定长序列,例如情感分类模型的词语。然而注意力机制的提出最早应用于机器翻译工作中,机器翻译的输入是不定长,输出也是不定长的<sup>[12]</sup>。最典

型的模型为编码-解码模型,该模型的实质结构是两个 RNN 网络的链接。一个 RNN 模型作为编码器,将输入的不定长序列转化为定长的背景向量  $c$ ,这个背景向量相当于整个模型的隐藏层的输出;另一个 RNN 模型作为解码器,它的作用是接受编码器产生的背景向量,同时结合之前生成的历史信息生成当前时刻的单词<sup>[13]</sup>。结构示意图如图 1 所示。

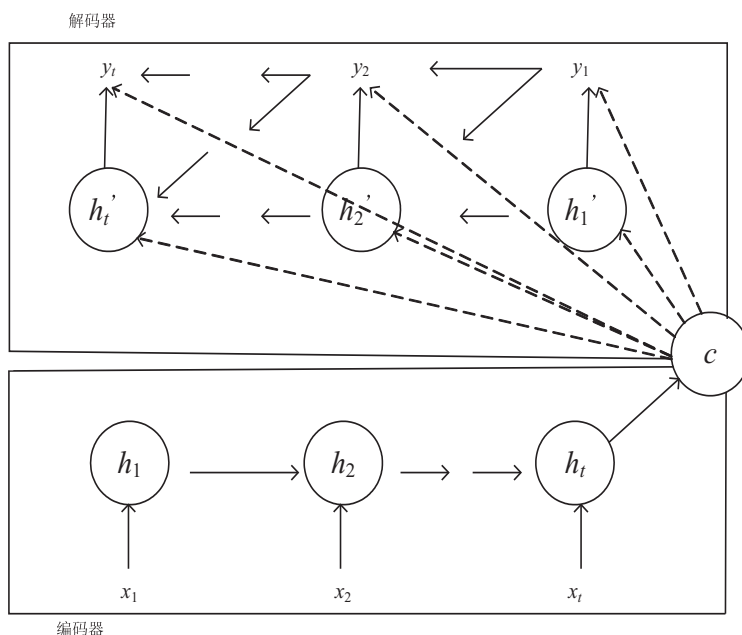


图 1 Seq2Seq 模型结构

$c$  为编码器产生的定长的背景变量,得到背景变量最简单的方式是将编码器中的最后一个隐藏层直接作为背景变量。输出层的  $y_i$  是每一时刻编码词生成的单词,由图可以看出生成过程,如公式:

$$y_1 = f(c) \quad (1)$$

$$y_2 = f(c, y_1) \quad (2)$$

$$y_t = f(c, y_1, y_2) \quad (3)$$

由此可以看出在生成单词的时候背景向量都是  $c$ ,每个单词并没有区别。这就意味着在编码器输入的句子中的任意一个单词对于解码器中生成的  $y_i$  贡献率是一样的。所以传统的 Seq2Seq 模型是不含注意力机制的解码-编码模型。在下一节将介绍加入注意力机制的模型原理。

### 1.2 自注意力机制

注意力机制的基本功能可以理解为,当人们在观察一样东西的时候,不可能将这个东西的全部特征同时记载下来,而关注点只在当前观察的东西的每一个局部。或者说,当人们的视线在转移的同时注意力也在发生转移,这就意味着当人们观察某个物品或者是某个背景,该目标内的每一处空间位置上的注意力分布是不一样的。将注意力机制引入到自然语言处理任务中时,可以理解为,当人们去描述一件事情或者转述

一个消息的时候,最先表达的单词和句子正是这件事情中某个片段的最相关部分<sup>[14]</sup>。针对 Seq2Seq 模型只依赖于唯一一个背景向量的局限及软注意力机制的编解码结构一般是应用在机器翻译、图像标注和语音识别上,这种任务对应的输入和输出都是序列,并表示了输入输出之间的相似对齐模型。但是对于文本情感分析这种只有输入是序列而输出是二分类的任务来说,编解码结构是不适应的,为此,对于所需的任务引入自注意力机制(self attention)。

抽象地介绍自注意力机制的本质思想:将输入序列 Source 中的构成元素拆分为<Key, Value>即关键词和有用信息的二元组表示,并将此刻的查询元素 Query 看作是 Source 中的某个词。此刻将进行 Query 对 Source 的相似度计算,得到每一个关键词对有用信息的相关性权重系数,然后对于得到的有用信息即 Value 进行加权求和计算:

$$f(x_i, x_j) = W^T \sigma(W_1 * x_i + W_2 * x_j) \quad (4)$$

$$\text{Attention}_j = \sum_{i=1}^n \text{softmax}(f(x_i, x_j)) * x_i \quad (5)$$

其中,  $x_j$  表示 Query 查询序列,  $x_i$  表示  $t$  时刻与  $x_j$  呈对齐关系的 Source 源句中的单词,  $f(x_i, x_j)$  表示对应关系的相似度,也就是注意力机制的权重。自注意力机

制的结构如图2所示。

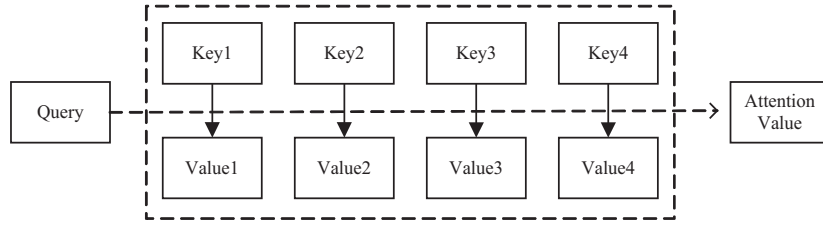


图2 自注意力机制结构

图中Query查询序列就代表了Source源句中的单词,并对句子中的每个词计算相关性。自注意力机制可以捕获同一句子中单词之间的一些语义语法的相关性依赖关系,它解决了LSTM按序列逐步计算的局限。对于LSTM来说,远距离相互依赖特征要经过若干时间步的信息累积才能将两者联系起来,显然距离越远,有效捕捉可能性越小。而自注意力机制在计算过程中会直接将句子中的任意两个单词的联系通过一个计算步骤直接联系起来,所以远距离依赖特征之间的距离被极大地缩短了,有利于有效利用这些特征来提高情感分类的准确率。

## 2 基于LSTM方法与注意力机制的结合

研究者们发现注意力机制的动机主要来自人类对于语言描述存在的侧重性,在分析文本情感时,需要对

文本语义进行理解。以语句“我今天非常高兴”为例,可以看出,文本前半段的“我今天”对本次情感分析的贡献性较小,因为“我今天”这三个字(或者是“我”和“今天”构成的两个词组)并不传达特殊的情感信息。相反,“非常高兴”显然传达出强烈的积极情感。因此,分类模型应该能够认识到这种表意能力的区别,为“非常高兴”部分添加更大的权重,在决定分类结果时重点考虑该部分的激活值。受此启发,文中提出中文场景下的情感分析对应的注意力机制。

通过对编码器所有时间步的隐藏状态做加权求和来得到背景变量。模型每一时间步调整这些权重,即注意力权重,从而能够在不同时间步分别关注输入序列中的不同部分并编码进相应时间步的背景变量。文中设计的LSTM模型与注意力机制的结合示意图如图3所示。

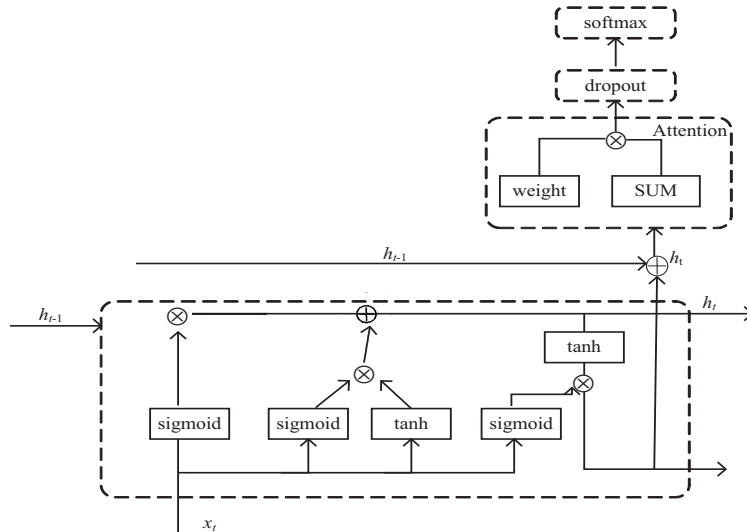


图3 LSTM中的注意力机制示意图

注意力机制整体设计如下:

(1)计算背景变量。首先,模型根据时间步1的隐藏状态及其在各个时间步的隐藏状态计算一个softmax运算的输入。然后,softmax运算输出该概率分布并对模型各个时间步的隐藏状态做加权求和,从而得到背景变量,加权求和公式如下:

$$c_t = \sum_{i=1}^T \alpha_{[it]} h_i \quad (6)$$

其中, $c$ 为背景变量, $\alpha$ 为权重, $h$ 为原隐含层状态。

(2)更新隐藏层的状态。在得到背景变量后,利用其对隐藏层的状态进行更新,提出将隐藏层的状态修改为:

$$\hat{s}_t = z_{[t]} \odot s_{[t-1]} + (1 - z_{[t]}) \odot \tilde{s}_{[t]} \quad (7)$$

上式中的忘记门、更新门和候选隐含状态分别被更新为:

$$r_{[t]} = \sigma(w_{[yr]} y_{[t-1]} + w_{[sr]} s_{[t-1]} + w_{[cr]} c_{[t-1]} + b_r) \quad (8)$$

$$z_{[i]} = \sigma(w_{[yz]}y_{[i-1]} + w_{[sz]}s_{[i-1]} + w_{[cz]}c_{[i]} + b_z) \quad (9)$$

$$\tilde{s}_{[i]} = \tanh(w_{[ys]}y_{[i-1]} + w_{[ss]}(s_{[i-1]} \odot r_{[i]}) + w_{[cs]}c_{[i]} + b_s) \quad (10)$$

其中,  $W$  和  $b$  分别为 LSTM 中门控循环单元的权重和偏置参数。

LSTM 具有序列中每个字的输出向量。使用 LSTM 的最常用方法是将序列中最后一个单词的输出向量作为整个序列的表示。

(3) 只需使用最后一个单词的向量提供了一种将可变长度序列转换为固定长度向量的简单方法。易于使用密集层进行分类等。

(4) 假定递归神经网络使当前字的输出向量取决于序列的先前字。LSTM 应该将此功能扩展到很长的序列。因此最后一个字的输出矢量编码来自整个序列的信息。

实际上, 将整个序列中的信息编码到单个向量中是不合理的, LSTM 也不能为过多的时间步保证良好的依赖性, 因为这试图在有限的空间内压缩过多的信息。可以通过如下方式解决这个问题: 不使用最后一个单词的输出向量, 而是使用所有单词的输出向量。然而, 这不容易实现, 因为现在处理的是一个可变维数矩阵而不是固定长度的矢量, 即存在一个序列的每个字的向量。因此, 更好的方法是基于某些上下文将单词向量聚合成固定长度向量。以图 2 中神经机器翻译的例子来更好地理解, 其中位置  $t$  处的译文取决于输入序列的所有输出矢量的聚合。通过这种方式, 网络可以选择哪些单词是重要的, 并且仅使用来自那些单词的信息来构造固定长度向量。

注意力模型是进行上述将神经网络隐层的输出聚合的一种方式。它提供了一种方法, 用于根据某些上下文将每个单词的输出向量聚合为单个向量。

### 3 基于 LSTM 方法与注意力机制的实验设计

LSTM 实际工作机理是通过识别输入序列对接受的所有特征进行相同的操作, 由于这种的无差别对待, LSTM 模型的输出结果不能直观地体现出每一个输入单词对于分类结果的重要程度。例如文本情感判断中, 一个句子中决定情感极性的只是带有感情色彩的词语, 这样只与特定目标相关的描述信息才是最重要的。针对文中所要解决的任务, 对于微博评论文本的情感分类进行了实验设计, 使注意力机制与 LSTM 相结合, 首先根据 LSTM 输出向量和上下文向量计算权重系数, 其中第一个阶段根据 LSTM 输出向量和上下

文向量计算两者的相似性或者相关性<sup>[15]</sup>; 第二个阶段对第一阶段的原始分值进行归一化处理; 然后, 根据权重系数对构成元素进行加权求和, 得到最终的表示。

具体步骤如下:

(1) 将 LSTM 输出向量与上下文向量混合并得到中间状态:

$$e_{[ij]} = f(h_i, c_j) \quad (11)$$

通常,  $f$  是具有 2 个全连接层的前馈网络, 向量  $c_j$  表示上下文,  $h_i$  表示位置  $i$  处的词的 LSTM 输出向量。

(2) 生成 Attention 的权重与偏置项, 以对中间状态进行非线性映射:

$$u_i = \tanh(W_w e_{[ij]} + b_w) \quad (12)$$

其中,  $W_w$  与  $b_w$  为 Attention 的权重与偏置项。

(3) 计算每个单词向量  $h_i$  的权重, 经过 softmax 进行归一化就得到了符合概率分布取值区间的注意力分配概率分布数值(也就是注意力权重), 这对应着不同的源句子单词的注意力分配概率分布:

$$\alpha_i = \frac{e^{e_i}}{\sum_{i=1}^n e^{e_i}} \quad (13)$$

这里, 每个单词将有  $n$  个中间状态, 中间状态的 softmax 操作为每个单词向量生成权重, 权重的总和为 1(这符合概率分布的特征)。这里的加权机制为模型提供了更好的解释性, 可以决定在文本分类中哪些词或句子更重要。

(4) 对构成元素加权求和, 计算最终的固定长度的向量。用  $i$  时刻的隐层节点状态  $h_i$  去一一和输入句子中每个单词对应的 LSTM 隐层节点状态进行对比, 获得向量  $v$  和每个输入单词对应的对齐可能性。定义一个条件概率如下:

$$p(y_i | y_1, y_2, \dots, y_{i-1}, X) = g(y_{i-1}, \alpha_i) \quad (14)$$

(5) 在合并这些输出向量时, 希望可以自适应地将注意力集中在那些对当前任务更重要的向量上, 也就是给它们都分配一个权值, 将所有的输出向量加权求和。合并后的表示为:

$$v = \sum_{i=1}^n \alpha_i h_i \quad (15)$$

这里假设  $h_i$  为输出向量,  $\alpha_i$  为权值, 不同输出向量的注意力权重是作为训练的一部分学习到的。此项计算成为网络计算图的一部分。要学习的参数是函数  $f$  的参数。在该流程中, 计算每个上下文  $c_j$  的注意力, 使得到的特征可以更好地表征文本, 也使训练出的模型具有更好的表现。

## 4 实验结果与分析

### 4.1 数据集选取

使用斯坦福的大规模电影评论数据集 (Stanford's



large movie review dataset,IMDb<sup>[3]</sup>)作为文本情感分析实验的数据集。该数据集从IMDB电影评论网站收集了50 000条评论,每部电影不超过30条评论。在该数据集中,标签为“正面情感”和“负面情感”的评论数量相等,因此随机猜测将产生50%的准确性。该数据集只保留了高度两极化的评论,负面评价的得分 $\leq 4$ 分,总分为10分,正面评价的得分 $\geq 7$ 分。中性评价不包括在数据集中。该数据集分为训练和测试两个集合,分别包含25 000条评论(数据来源:<http://www.andrew-maas.net/data/sentiment>)。

## 4.2 注意力机制引入的实验结果分析

注意力机制的引入是文中对传统深度学习网络对于情感分析方法的一个改进。注意力机制的主要目的就是在模型做最后的极性分类预测时,在重要的句子成分加上权重,加大对最后分类的预测概率的影响因素。在注意力机制引入的LSTM网络结构的实验结果如表1所示,表中结合展示了特征提取两种方法的对比。

表1 评价标准对比

方法	精确率	准确率	召回率
Word2Vec+LSTM	87.8	84.6	85.9
Word2Vec+LSTM+Attention	89.6	86.8	87.4
FastText+LSTM	91.3	89.7	90.4
FastText+LSTM+Attention	93.4	90.9	92.4

表中数据展示了在微博数据集中,文中提出的加入注意力机制的网络结构模型的分类效果以及准确率的对比,可以明显看出,通过加入注意力机制,网络可以选择哪些单词是重要的,并且仅使用来自那些单词的信息来构造固定长度向量。考虑相关场景下计算序列到序列网络的注意力的方式,字符序列由编码器部分转换成矢量(编码表示),并且解码器接收该输入并产生输出字。这种编码表示向量只不过是编码器中LSTM层的最后隐藏状态的输出。当计算注意力向量时,使用编码器的所有隐藏状态。因此,解码器可以捕获某些全局信息,而不仅仅是基于正常网络中的一个隐藏状态进行推断,因为文中使用了所有的隐藏状态。

## 5 结束语

主要介绍了注意力机制的作用与应用目的,通过对不含注意力机制的传统Seq2Seq模型的原理讲解,以及加入注意力机制的模型,Soft Attention模型的工作原理的对比,突出表现随时间变化的背景向量对于生成单词的贡献率的不同,并着重引入了Self Attention针对于只有输入数据的结构,对文中的文本

情感分析做出阐述,加入注意力机制的模型更能表现出输入词语对于输出词语的相关性和重要程度。并通过与机器翻译进行对比,进一步解释了注意力机制应用于分类模型的应用原理。同时,设计了注意力机制的结构,并主要研究了注意力机制与长短期记忆模型的结合方法与算法构架的设计。

## 参考文献:

- [1] 唐慧丰,谭松波,程学旗. 基于监督学习的中文情感分类技术比较研究[J]. 中文信息学报,2007,21(6):88-94.
- [2] 周世妍. 新浪微博用户及其微博特征分析[J]. 新媒体研究,2017,3(21):8-9.
- [3] MAAS A L,DALY R E,PHAM P T,et al. Learning word vectors for sentiment analysis[C]//Proceedings of the 49th annual meeting of the association for computational linguistics; human language technologies-volume 1. Washington D C: Association for Computational Linguistics,2011:142-150.
- [4] SIEGELMANN H T, SONTAG E D. On the computational power of neural nets[J]. Journal of Computer and System Sciences,1995,50(1):132-150.
- [5] HOCHREITER S,SCHMIDHUBER J. Long short-term memory[J]. Neural Computation,1997,9(8):1735-1780.
- [6] 赵晓伟. 基于LSTM模型的文本情感分析技术的研究[D]. 北京:华北电力大学,2018.
- [7] GREFF K,SRIVASTAVA R K,KOUTNÍK J,et al. LSTM: a search space odyssey[J]. IEEE Transactions on Neural Networks & Learning Systems,2015,28(10):2222-2232.
- [8] MIKOLOV T,CHEN K,CORRADO G,et al. Efficient estimation of word representations in vector space[C]//ICLR 2013. [s.l.]:[s.n.],2013.
- [9] 唐善成,马付玉,张璞月,等. 采用Seq2Seq模型的非受限词义消歧方法[J]. 西北大学学报:自然科学版,2019,49(3):351-355.
- [10] BAHDANAU D,CHO K,BENGIO Y. Neural machine translation by jointly learning to align and translate [C]//ICLR 2015. [s.l.]:[s.n.],2015.
- [11] XU K,BA J,KIROS R,et al. Show,attend and tell:neural image caption generation with visual attention[C]//ICLR 2015. [s.l.]:[s.n.],2015.
- [12] 赵庆东,郭中华. 基于模型Seq2Seq机器翻译的研究[J]. 电子世界,2019(8):69-70.
- [13] 王 钧,张 鹏,袁 帅. 基于股票预测的Seq2Seq RNN和LSTM模型比较[J]. 时代金融,2018(35):381-382.
- [14] 尹 鹏,周 林,郭 强,等. 基于短语级注意力机制的关系抽取方法[J]. 计算机技术与发展,2019,29(9):24-30.
- [15] TANG D,QIN B,FENG X,et al. Effective LSTMS for target-dependent sentiment classification[C]//Proceedings of COLING 2016,the 26th international conference on computational linguistics. Osaka,Japan:[s.n.],2016:3298-3307.