学号\_2018302110379\_\_

密级\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

武汉大学本科毕业论文

基于图卷积网络的情感依赖模型在方面级别情感分类中的应用

院（系）名 称：计算机学院

专 业 名 称 ：软件工程

学 生 姓 名 ：陈威

指 导 教 师 ：刘进 教授

二○二二年三月

**郑 重 声 明**

本人呈交的学位论文，是在导师的指导下，独立进行研究工作所取得的成果，所有数据、图片资料真实可靠。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确的方式标明。本学位论文的知识产权归属于培养单位。

本人签名： 日期：

摘要

伴随着技术发展，各大社交平台的用户流入，各式各样的文本数据等随着智能移动终端的代代升级流入互联网，文本数据也呈现了爆炸式的增长，对于这些数据，内层都包含着人们对于事情的看法等，于是由计算机帮助处理文本数据的研究项目也逐渐成为了热点，情感分类也被学者们认为是数据分析的有效办法之一。

情感分类是指在对带有感情色彩的主观文本进行分析时，去推理、运算的过程，就是说去推断人对事物的主观看法，它是倾向于正面、反面抑或是中性的，在传统的文本主题分类中，是将文本的客观内容进行分类，而情感分类则异于这一点，在实际应用中用来分析具体事物对不同人群所产生的看法。

本文基本图卷积网络依赖模型对方面级情感进行分类且应用，所采用的研究方法是SDGCN[1]，该研究方法有区别于其他同类型的图卷积神经网络，首先，它使用了一种携带了位置编码信息的双向注意力机制去捕捉特定方面的表达，且提出了一种新异的多方面情感分类框架，该框架可以利用GCN来有效地捕捉句子中不同方面级之间的情感依赖关系。该方法提供了两种不同的方法来实现，本文仅使用GloVe模型来做具体讨论。

关键字**：**情感分类；深度学习；图卷积神经网络；情感依赖；GloVe

# 绪论

* 1. 研究背景与意义

近年来随着信息时代互联网时代的到来，社交网络以及电影评论网站像雨后春笋一样接踵而至，大量的文本数据涌入互联网当中，像是Twitter、微博、抖音、RottenTomatoes等为首的各种社交软件中， 各式各样的人们在其中留下自己的主观看法，此类应用很大程度上改变了我们的生活方式提供了我们很多的参考数据去观察，通常我们会去查看用户留下来的评论文本数据来获取信息以决定选择，这些大量的主观文本数据随着人们手中的智能终端的技术提升而逐渐产生爆炸式的增长。其中包含的用户情绪观点，蕴含了大量的主观数据可供分析，而利用人工的方法去提取其中的主观含义无异于无稽之谈，于是，由计算机的强大算力所支持的机械学习方法是提出文本中主观观点的不二之选。

基于方面级的情感分类是为了分析文本中某个实体在个体的主观评价下所产生的主观情感，譬如于“The service is reasonable although the price is expensive”在这句中“service”与“price”为两个方面级词语，对于这两个方面级词语来说，“service”明显是正向的，而“price”则是负面的，这是方面级情感分类的优势，那么对于这样的简单句来说，只有“service”和“price”这两个方面级词汇，我们可以建立一个简单的情感依赖图，我们可以简易的推断出，“service”和“price”的情感是相异的，对于这样的情感依赖图来说，推断另外的方面级词汇的情感是有益的。

从本质上讲，其他人的观点对我们自身的行为有着根本性的影响，我们对这个世界的认知和观念在一定程度上取决于其他人对这个世界的认知和观念，这不仅体现在个人，也体现在一个集体上，就好比消费者的购物需求在一定程度上取决于企业的市场营销，而企业生产营销策略也需要根据市场需求制订。因此，情感分析能对数据的情感倾向做出解析，对人类决策具有重要意义。

情感分析属于一类特殊的文本分类任务，通过对文本数据进行处理和分析，从而预测其中蕴含的情感趋势，如正/负/中立或者其他多个类别。更进一步讲，根据文本数据的不同类型，情感分析还可分为多种粒度：文章级别的情感分析，句子级别的情感分析和更细粒度的方面级别的情感分析。情感分析还可包含许多自然语言处理的其他任务，比如在句子级别的情感分析中，主观性检测、讽刺检测有着重要的作用；在细粒度的方面级别的情感分析中，还需要进行命名实体识别、方面抽取、同义词聚类等等；对某些数据量较小的小众语言，可以利用较多的英文数据，进行跨语言的情感分析。

得益于计算机硬件性能的飞速提升和网络中数据规模的急剧增长，深度学习算法凭借其在特征学习上的优势，在计算机视觉，语音处理等领域取得了长足发展，全连接神经网络、卷积神经网络、循环神经网络等深度学习模型在不同应用条件下展现出不俗的性能。由Tomas Mikolov在2013年提出的Word2Vec算法，使得文本从离散的符号表示，转化为自然包含语义信息的连续的向量表示，为自然语言处理与深度学习的结合奠定了基础，由此也推动了深度学习技术在情感分析中的应用与发展[1] 。

基于英文的情感分析，早在千禧年就成为自然语言处理、文本挖掘和信息检索等研究领域中的热门方向。从2014年开始，国际语义评测比赛（SemEval）提供了情感分析的较大规模数据集，更是促进了深度学习算法在方面级的情感分析上蓬勃发展。而中文情感分析在近十年才获得研究与应用人员关注，且可用的公开数据集也有待丰富，因此基于深度学习的中文情感分析尚有待进一步的研究。

* 1. 国内外研究现状

情感分析最初的研究是基于英文文本，国外对于基于深度学习的英文情感分析的研究已趋于成熟，在各类任务的公开数据集上皆取得了较好的效果。根据文本类型，可将情感分析分为文章级别、句子级别和方面级别三种不同的粒度，以下分别简述各个粒度下，英文情感分析的研究现状与发展趋势。

文章级别的情感分析通常假设文章中包含对某一事物的单一情感的观点，需要通过各种数据挖掘与分析算法，分析出其中蕴含的观点以及其情感类别。由于文章中包含了连续的文本内容，因此还可以利用语法或语义信息提升效果，比如传统的Part-of-Speech tagging、语法树，和深度学习中常用的attention机制等[2] 。Zhang等人对近年来的文章级情感分析进展进行了简要概述[3] 。

句子级别的情感分析目标是挖掘单个句子中，对某一事物的情感倾向。无论是文章还是句子的情感分析，文本序列的连续表示都有着重要作用，这里的文本序列可以是文章、段落或者句子。文献[4]基于文献[1]的思想，将文本序列的离散符号表示，转换为了连续向量表示，在得到句子的连续表示的同时，也自动包含了该句子的语义，从而有利于句子级别的情感分析。Kim将深度学习图像处理中常用的卷积神经网络（CNN）应用到文本序列中，进行文本分类，获得了良好的效果[5]。循环神经网络（RNN）及其变式GRU、LSTM以其对变长序列的处理优势在自然语言处理中广受欢迎，在句子级别的情感分析中也取得了不俗的效果[6]。不同于常见的语言模型学习方式和RNN序列式的处理方式，Devlin等人基于文献[7]提出的Masked LM和文献[8]的Transformer建立了新的语言表示模型BERT，发表时刷新了包括情感分析在内的各项任务的纪录[9]。总的来讲，句子级别的情感分析与语言表示模型的效果有紧密联系，语言表示模型的语义表示能力越强，情感分析的效果越好。

不同于文章或句子级别，方面级别的情感分析的要求更加严格，其目标是预测一个句子中，包含的对某一事物的某一方面的情感倾向，如：“这台电脑的屏幕很不错。“这句话中，”电脑“是事物本身，”屏幕“则是”电脑“的一个方面。方面级别的情感分析还可细分为四个子任务，情感词项的提取、情感词项的倾向分析、情感词类的提取、情感词类的倾向分析，这四个任务也是2014年SemEval中方面级别情感分析的子任务。

由于attention机制在处理方面特征上具有优势，Wang等人结合LSTM与attention机制，建立了情感词项的倾向分析模型[10]。Bao等人对文献[10]做出改进，在计算attention vector时加入了方面的向量表示，从而将深度学习模型与传统的词典信息结合，提升了情感分析的性能[11]。为了更深层次的结合句法信息，Zhang等人[12]运用图卷积网络（GCN）在目标句子对应的解析树上进行情感词项的倾向分析，方面信息在经GCN处理后，由attention机制加入[12, 13] 。在GCN的基础上同时引入attention机制后即为图注意力网络（GAT）[14]，Huang等人采用了该模型，在方面级别的情感分析任务上，超越了基于BERT的微调模型[15]。综上所述，深度学习模型在方面级别的情感分析上的应用逐渐成熟，并有集成其他传统自然语言处理方法的趋势。

与英文情形相比，基于深度学习的中文情感分析发展尚不够成熟，其主要原因有两点。一是发展时间尚短，文献[16]提到，中文情感分析仅在近十年才得到发展。二是数据集的不足，由文献[17]可知，截至2018年，尚未建立公开的方面级别的情感分析中文数据集。因此，基于深度学习的中文情感分析研究仍有较为广阔的发展空间。

Che等人基于条件随机场（CRF）算法，设计了保留情感信息的句子压缩模块，该模块可作为情感词项提取的预处理模块，也可用于其他情感分析任务[18]。Zeng等人将CNN和LSTM结合，设计了可同时进行情感词类提取和情感词类的倾向分析的模型[19]。Cui等人基于全词遮罩（Whole Word Masking）的思想提出了适用于中文的BERT-wwm[20]，促进了中文文本处理的发展。

* 1. 本文的主要研究工作

本文的主要研究工作是探究基于深度学习模型的中文情感分析算法，对不同的深度学习算法，在句子级别中文情感分析任务上的性能进行实验和分析，主要内容可分为以下几个部分。

（1）建立基于卷积神经网络的TextCNN模型，作为中文情感分析的基本模型，将其性能作为其他模型的性能参考。选用卷积神经网络模型作为参考，是因为该模型在深度学习领域广受青睐，在图像处理领域大现身手，而将其应用于文本处理领域，既能引入深度学习模型在特征学习上的优势，又能结合自然语言处理传统模型，n-gram模型的思想。

（2）由于循环神经网络（RNN）在变长数据处理中具有天然优势，结合文本的序列化特征，选取其中具有稳定优异性能的双向长短期记忆网络（Bi-LSTM），建立了基于该模型的中文情感分析模型。由于Bi-LSTM模型生成的句子向量表示存在对句子末端单词的偏向性，因此，其文本的特征表示能力有待提高。因此，参考RCNN的池化机制[21]，对模型进行改进，建立了基于RCNN的中文情感分析模型，使得其性能有了稳定且显著的改善。

（3）预训练模型推动了众多自然语言处理研究任务的发展，因此，本文也探究了预训练模型对情感分析的影响，建立了基于BERT的中文情感分析模型，并对其结果进行分析，探讨其性能优异的原因，以及句子级别情感分析的局限性。

* 1. 本文的结构与安排

第一章绪论，主要介绍了情感分析的研究背景与意义，简要概述了国内外情感分析的研究进展，并阐述了本文对基于深度学习的中文情感分析研究路线，最后，对本文内容安排进行了说明。

第二章相关技术简介，首先介绍了神经网络模型的基本结构及其原理，并由此对深度学习进行简单介绍；随后阐述了卷积神经网络模型、循环神经网络的基本原理；最后，简要介绍了attention技术，并据此说明了Transformer的encoder的结构，为BERT原理进行铺垫。

第三章基于深度学习的中文情感分析，首先根据卷积神经网络的原理，结合中文分词的词向量，建立了基于CNN的情感分析模型，TextCNN；其次基于循环神经网络的原理，搭建基于Bi-LSTM的情感分析模型，并参考RCNN，引入池化机制对Bi-LSTM中存在的偏差问题以及文本的特征表示进行改进，建立了基于RCNN的情感分析模型；最后搭建用于情感分析的BERT模型，对其结构进行了概述。

第四章对比实验与结果分析，针对基于深度学习的中文情感分析算法进行实验，选取句子级别的中文情感分析数据集，根据不同模型输入的需求，介绍预处理方法。随后，对TextCNN、Bi-LSTM、RCNN和BERT模型在数据集上的准确率进行对比和分析。其次，对BERT进行了错误分析，最后探究了其Multi-head attention机制的优点，以及句子级别情感分析的局限。

第五章总结，对本文工作进行了总结，并对有待进一步研究的工作进行简要了介绍。

# 相关技术简介

1. 1. 情感分析概述

如前所述，情感分析是通过对文本数据进行处理，从而提取出其中蕴含的情感倾向。虽然情感分析有篇章级、句子级、方面级多种粒度，但本质上讲，还是一种分类任务，因而也特别适合深度学习算法应用。



图 2.1 情感分析流程图

总的来说，中文情感分析任务的流程如图2.1所示。将中文文本数据输入到深度学校算法中，经过处理后得到情感类别的预测。文本数据根据任务粒度不同，有不同的输入形式，如句子级别是一段句子，而篇章级别则是一篇完整的文章。中文情感分析算法由研究人员设计与实现，通常包括分词、词向量嵌入、算法处理，具体分词规则和词向量嵌入方式根据算法输入需要而定。情感类别通常是人为实现设定，可以是正负两类情感极性，也可以是多种情感极性。

* 1. 深度学习基础

2. 2. 1. 深度学习概述

神经网络模型是深度学习主流模型的基础，在介绍深度学习之前，首先对神经网络模型简要说明。

神经网络是一种模仿生物大脑神经网络的模型，并在其结构进行了计算上必要的简化，从结构，实现机理和功能上实现了生物神经网络的模拟[22]。其基本计算单位是神经元，多个神经元之间通过可学习的权值进行连接，如同生物神经网络中，神经元通过树突和轴突连接。各个神经元收到的信息通过连接的权值线性组合得到输入，将神经元输入通过一个非线性的激励函数计算得到该神经元的输出，模拟了生物神经元中，神经信号的传递过程，以及突触对神经信号的激活或抑制，其数学模型如(2.1)式所示。

其中，为输入，为权值，。为该神经元的bias项，为非线性激励函数。基本结构如图2.2所示。



图 2.2 神经元结构示意图

不同于生物大脑中无规则团块状的网络结构，为了简化计算，神经网络模型通常以层的形式组织神经元，其中，最常见的是全连接层，该相邻层间的神经元相互连接，层内的神经元无连接，其结构示意图如图2.3所示。



图 2.3 全连接神经网络

将输入数据视为神经网络的输入层，中间的全连接层为隐含层，前一层的输出作为后一层的输入，激励信号不断向前传播，最终由最后一层全连接层作为输出层输出整个模型的处理结果。通过事先定义的目标函数衡量模型输出与事实结果之间的误差，计算整个模型损失，再通过反向传播算法有效解决权值更新问题，最终构成了一个完整的神经网络模型。可以将神经网络视为一个通用的函数拟合模型，将其应用到机器学习任务中，理论上，只要有足够的数据和神经元，就能拟合到足够复杂的函数，换句话说，参数越多，模型的容量就越大。

虽然理论上已经证明，两层的神经网络已有拟合任意非线性函数的能力[23]，但其特征学习效率和特征表示能力远远不足，深度学习就是用于更进一步的特征表示的学习[22]。为了自动学习到问题中涉及的特征信息，构建深层次的特征转换模型往往更具有优势。这是因为在深度学习模型中，更低（更靠近数据输入）的层更容易学习到局部特征，而高阶的层则用于合理组合底层学习到的局部信息，从而更加高效、合理的学习到特征信息。而神经网络模型应用到深度学习中，具有可扩展性和基于反向传播算法的优化方法等优势，此外，基于神经网络的深度学习模型通常还具有端到端学习的优点，因此成为主流。

然而仅仅依靠简单的全连接神经网络无法完成深度学习任务，因为其全连接的神经元连接方式存在不足，3到4层之后就失去了进一步提取特征的能力，本文后续将对用到的深度学习模型进行简要介绍，并以此建立基于深度学习的情感分析模型。

* + 1. 卷积神经网络简介

卷积神经网络（CNN）是一种具有局部连接、权重共享等特性的深层神经网络，早期主要用于处理图像信息[24]。全连接神经网络在处理图像数据时，存在参数过多和无法处理图像的局部不变性特征的问题。参数过多将导致整体模型训练效率大幅度降低，且极易过拟合。局部不变性特征是指图像中的特征经过缩放、平移、旋转等操作，其语义信息不受影响，全连接网络较难处理这一特征。而CNN的局部连接，权重共享以及深层连接的特点则较好的解决了这两个问题。

CNN是受视觉神经系统中的视觉皮层的启发而提出的，视觉皮层中的神经元仅接受视网膜上对应的感受野区域的信号激励，由此提出了具有视觉皮层特性的卷积层。一个典型的CNN模型主要由卷积层、池化层、全连接层组成，其结构示意图如图2.4所示。



图 2.4 经典卷积神经网络结构示意图[3]

卷积层由个卷积核组成，每个卷积核对应一个感受野，由可学习的权值组成，设为，即感受野大小为的像素区域。设网络模型输入为的图像像素，其中c为图像通道，图像高和宽均为16。每个卷积核在图像上滑动，并对图像上对应的感受野进行卷积，即将像素值与卷积核权值对应位置相乘后再相加，得到输入图像在该感受野区域，对当前卷积核的激励值。将激励值按原图像上的顺序排列可组成一个二维数组，得到某一卷积核在该图像上的激励，所有卷积核卷积结果可组合在一起，称为特征图（feature map），作为整个卷积层的输出，特征图又可以作为新的图像像素输入到下一卷积层，或通过其他方式进行处理。设滑动步伐为1个像素，则上图中卷积层输出的特征图大小为。

池化层也称作降采样层，用于降低图像分辨率，减少模型参数，提升模型训练的效率同时防止过拟合。池化层与卷积层的图像处理方式类似，但不含可学习的权值，每层预设了固定大小的降采样区域，设为，该区域在输入图像的每个通道上滑动，对该区域进行降采样，从而形成该区域输出，每个区域按原顺序可组成一个更小的特征图。设滑动步骤与采样区域相同，采样区域间不覆盖，上图中池化层输出的特征图大小为。

原始图像经过卷积和池化之后，得到了一个高度抽象的图像特征表示，可将其展开为一维向量，输入到全连接层，用于特定的任务。

以上对CNN模型在图像处理中的应用方式进行了分析，对卷积网络中的关键技术进行了介绍，该模型在文本中的应用与图像应用类似，3.1节将详细介绍基于CNN的情感分析模型。

* + 1. 循环神经网络简介

循环神经网络（RNN）能通过维持一个状态信息，来处理序列化的输入。不同于其他神经网络模型的无回路结构，RNN每层内部存在有向回路，使得其神经元不仅可以接受其他神经元的输入信息，还能接受自身的状态信息，使得RNN运算的每一时刻不仅依赖于当前元素，也依赖于状态信息，即之前所有输入元素的信息[22]。和单一的前向神经网络相比，RNN更符合生物学上的神经网络结构。

图2.5展示了RNN模型的基本结构，左侧为实际结构，右侧为按时刻展开的结构，每一时刻，RNN处理输入序列中的一个元素。



图 2.5 RNN结构示意图[3]

其中，垂直的矩形框表示一个RNN单元，其中包含了多个神经元。对时刻t，RNN输入为输入序列中的第t个元素和t-1时刻的输出，即状态，二者分别通过可学习的权值矩阵和通过矩阵乘法相加和一个非线性函数，得到t时刻的状态。每一时刻的状态都可以再通过一个可学习的变换矩阵进行维度转换从而得到输出，其数学模型如(2.2)和(2.3)式所示。

值得注意的是，RNN中的权值矩阵都是重复使用的，由此得到了内存上高效的序列处理模型，可以处理任意长度的输入序列，且理论上，每一时刻都可以通过状态信息获得该序列当前时刻之前的所有信息。然而在实际实践过程中，RNN存在一些问题，一是由于其序列化特征，RNN在训练时间上并不高效，失去了深度学习优化中至关重要的并行特性；二是存在梯度爆炸或消失问题，使得模型训练效果不稳定，且对较长的输入序列，并不能真正获取序列中较早处理的信息。

为克服这些问题，对RNN提出了许多变种模型，本文选取双向长短期记忆网络（Bidirectional Long Short Term Memory），建立了基于Bi-LSTM的情感分析模型，3.2节对其结构进行详细介绍。

* 1. Attention机制

当人在阅读一个句子，对句子中某些结构有较高的注意力，而另一些结构的注意力则较低，这称为注意力（attention）机制，是一种较为高级的认知能力，它使得人脑在接受大量视觉、听觉、触觉等多种来源信息时，能自动选取有用的信息，而过滤掉不重要的信息。下面介绍基本的attention机制。

设输入序列为个向量，存在一个任务相关的查询向量，为了从输入向量中获取与特定任务相关的信息，首先需要计算每个输入向量对查询向量的相关程度，其计算公式可表示为：

函数为打分函数，用于评判输入向量对的相关程度，softmax函数使得软性表示了第i个输入向量对查询向量相关的概率，即attention的分布。最后，(2.5)式采用加权平均的方式计算的注意力向量：

以上是基于单个查询向量的基本attention机制描述。随着NLP技术发展，attention机制发展出了众多变体，更有研究人员提出了仅基于attention机制的Transformer[8]，在高效处理变长序列的同时，解决了RNN无法并行的问题，而Transformer中的encoder为BERT模型的提出奠定了基础[9]。下面对encoder结构进行简要介绍。

每个encoder由multi-head attention层、全连接层两个子层组成，每个子层的输出又经过了残差连接和Layernorm处理[25]，其结构示意图如图2.6所示。



图 2.6 Encoder层结构图[8]

Encoder中最重要的结构为Multi-head Attention层。前一节介绍了单个查询向量的attention机制的原理，为了计算句子中每个单词对句子中其他句子的重要性，可以将每个单词作为查询向量，计算出对应的attention向量，作为该句子特征转换后的向量表示，因此也称之为self-attention。

设一个句子经Embedding得到词向量后，拼接成为的二维词向量，则self-attention将同时视为查询向量与输入向量，其打分函数为缩放的点积运算（scaled dot product），因此，其attention的软性概率分布为：

其中，为一个的二维矩阵，每一行表示对应单词在当前句子中的重要性。计算出attention分布后，就可以对句子进行特征转换：

为了进一步从文本中获取语义信息，Multi-head Attention将，，投影到了不同的向量空间，再做self-attention，各向量空间称为head。将各个向量空间的转化结果拼接后，作为句子新的特征表示，其数学模型如(2.8)所示。

其中，

通过、、，将X投影到了不同的语义空间。由图2.6可知，因为采用了残差连接，因此各模块输出的维度应当与输入一致，的作用就是将拼接后的高维向量转换到低维空间，与encoder其他模块连接。

此外，由于attention没有考虑到输入序列的位置关系，因此，对序列中的每个词向量加上一个表示位置信息的位置编码（Positional Encoding），Transformer中的编码方式是对词向量每一个维度加上一个与单词在序列中的位置相关的三角函数。

# 基于深度学习的中文情感分析

1. 1. 基于CNN的情感分析

为了将深度学习模型用于文本数据，进而建立情感分析模型，首先需要采用word2vec算法将离散的文本序列转换为连续向量形式[1]。为获取更具有泛化能力的词向量，本文选择Li等人提供的基于中文分词的预训练词向量sgns.renmin.bigram-char[26]，作为基本词向量，根据数据中的单词建立用于本文实验的初始词向量。各模型具体词向量建立方式与out-of-vocabulary（OOV）的处理方式见4.2节。

建立词向量后，每一个单词存在一个对应的向量，将一个句子所有单词的向量组成一个的二维数组，其中为句子中的单词数，为词向量的维度，则可以将该二维数组视为一张单通道的二维图像，由此，可建立基于卷积神经网络的情感分析模型TextCNN，本模型的基本结构如图3.1所示。



图 3.1 TextCNN结构示意图

设一个句子经过分词后分成了个单词，每个单词对应一个维的向量，则该句子可表示为二维向量。为尽力抽取每个单词的语义信息，设卷积层中，卷积核的大小为，为单词数，为词向量维度，则该卷积核输出为

其中，为bias，是该卷积核的一个实数参数，为非线性激励函数，为第i到第i+h-1个单词对应的h个词向量卷积输出。该卷积核对该句子输出为向量。

上面描述了一个卷积核对一个句子生成的二维向量进行的操作，实际卷积层包含多种尺度的卷积核，即卷积核的大小可变，但不变，这对应了n-gram模型中，选用不同的gram尺度，从而获取不同程度的上下文信息，而通过词向量，则能更好从语义层面挖掘情感相关的信息。每种尺度的卷积核也可以设置多个，普通卷积层中，卷积核的数目表示输出特征图的通道数，TextCNN中，多个相同尺度不同参数的卷积核则使模型学习到相同尺度下，不同的文本特征，从而更能提升数据挖掘的能力。

然而由于尺度不同的原因，将导致不同卷积核输出的向量长度不同，因此，使用max-pool池化层，对每个卷积核的输出进行最大值池化，即选取中最大值作为输出，从而解决输出长度问题。

将每个卷积核池化后的输出拼接起来，将得到该句子经过卷积层提取后的特征，将其输入到全连接层，进行维度转换，即可对该句子的情感倾向进行预测。

* 1. 基于Bi-LSTM的情感分析

为充分利用文本的序列化特征挖掘情感相关信息，本文选取多层双向LSTM建立基于Bi-LSTM的情感分析模型，其基本结构如图3.2所示。



图 3.2 Bi-LSTM结构示意图

其中，底层表示模型输入，每一圆点表示LSTM单元的某一时刻，中间每一层对应模型的LSTM单元层，实线表示正向的状态流动，虚线表示反向的状态流动。

与基本RNN不同的是，LSTM增加了门控机制，其状态的计算规则为：

通过、、三个可学习的门和中间信息单元对状态的生成进行控制，有效解决了基本RNN模型中存在的梯度爆炸或消失问题，将LSTM单元的数学模型表示为(3.7)式。

为了进一步结合句子的上下文挖掘情感相关信息，本模型采用了多层双向LSTM叠加的结构，第一层双向LSTM的数学模型为：

由上式可知，t时刻的状态，由LSTM从两个方向对输入学习到的状态拼接而成，而第二层及之后的数学模型为：

即第i层的输入为第i-1层输出的状态，通过多层双向LSTM的叠加，使得模型能够从文本序列与上下文的角度挖掘其中的情感特征信息，最后，将LSTM两个方向的最后状态拼接为单个向量，作为文本的向量表示，输入到全连接层，即可进行情感倾向的分析。

* 1. 基于RCNN的情感分析

虽然通过拼接，Bi-LSTM能有效获取上下文信息，但拼接后的状态能否作为原句子的有效表示却有待商榷，比如，通过正向LSTM处理的句子，将其最终状态作为句子表示，则句末的单词相对句首的单词的影响更多，简单的拼接或许将损失情感相关特征信息。因此，参考Lai等人提出的RCNN模型[21]，对RNN模型进行两点改进。

首先，对LSTM单元的状态的表示进行改进。对每一时刻的状态，除了双向的状态输出，还考虑了该时刻对应的词向量，即，每个时刻的输出状态表示为：

其次，对一个句子的特征表示进行改进。设一个句子经过LSTM处理后，输出状态为的二维向量，其中，为句子单词数，为状态向量的维度，为得到句子的向量表示，采用max-pool池化层，对句子每一维度个值进行最大值池化，得到维向量，作为句子的特征表示。

通过上述改进，可以更好的获取句子中的语义信息，从而进一步挖掘整个句子的情感信息，实验证明，能明显提高循环神经网络的情感特征学习能力，4.4.2节对改进后的池化循环神经网络和基本循环神经网络的性能进行了对比。

* 1. 基于BERT的情感分析

2.5节介绍了Transformer中的encoder的结构，而BERT模型则采用了多层encoder叠加[9]，将attention的序列处理能力与深度学习特征转换能力结合，深入提取了句子的语义信息。

BERT的优势还在于预训练，通过MLM和连续语句预测对模型进行无监督训练，使模型充分学习到了语料库中的语义信息，研究表明，该预训练方式语义学习能力明显高于浅层的双向RNN状态拼接，或者CNN的局部特征转换。此外，通常情况下，模型参数越多，训练越容易在小规模的数据集上产生过拟合，然而，通过在大规模语料库中进行预训练，即使在小规模的数据集上也能达到较好的效果，该模型最大的亮点是，经过大量无监督任务预训练后，仅需针对下游任务进行微调即可超越众多任务相关的模型，也同样适合本次的情感分析任务。本文选用了Cui等人提供的中文BERT预训练模型BERT-wwm-ext[20]，在其基础上进行下游任务微调，使其适用于本文的情感分析任务。

设encoder的模型输出为句子的二维向量表示，为获取句子情感相关特征的向量表示，在句首增加一个特殊符号“[CLS]”，其经BERT处理后的k维向量作为适用于情感分析的句子特征表示，将其输入到全连接层进行情感倾向的预测。

# 对比实验与结果分析

1. 1. 数据集简介

本文根据网络搜集到的ChnSentiCorp，建立了句子级别的中文情感分析数据集。该数据集是一个综合领域的数据集，包含了酒店、书籍、数码用品三个领域的用户评论，每一条样本数据由文本和标签字段组成，文本即用户的中文评论句子，标签为用户的情感极性，1表示正向评论，0表示负向评论。训练集、验证集、测试集的数据统计信息如表4.1所示：

表 4.1 数据集统计信息

|  |  |  |  |
| --- | --- | --- | --- |
|  | 数据量 | 正向 | 负向 |
| 训练集 | 9600 | 4799 | 4801 |
| 验证集 | 1200 | 593 | 607 |
| 测试集 | 1200 | 608 | 592 |

该数据集的句子长度适中，且长短皆有，情感极性分布均匀，因此符合本文任务要求。

* 1. 数据预处理

数据的预处理工作主要是根据训练集分词结果建立词典，再从预训练的词向量中根据词典建立适用于各模型的词向量，以及根据各模型输入需求准备输入。

为了保证实验结果无偏差，仅在训练集上建立词典。本文采用开源工具jieba0F[[1]](#footnote-1)进行中文分词，将训练集中每条文本分割成由单词组成的列表，再对整个训练集中出现的单词进行统计，实验表明共有25662个不同的单词，由此建立词典。除了训练集中出现的单词，在词典中，还预先加入了“[PAD]”和“[UNK]”两个特殊符号，分别用于文本的补齐和对未知单词的处理。

本实验选用的预训练中文词向量sgns.renmin.bigram-char包含了356053个单词，词向量维度为300。按词典中的单词顺序从中抽取对应的词向量，将之组合成一个25662x300的二维数组，即为实验用到的初始词向量，而“[PAD]”和“[UNK”则分别初始化为全零向量和25662个向量的平均值。

以上得到的基于中文分词的词向量可用于TextCNN、Bi-LSTM、RCNN三个模型的词向量初始化，而BERT模型有自有的分词技术及词向量。本文选用的中文BERT模型BERT-wwm-ext，沿用了谷歌提供的BERT-base（Chinese）的字分割，即将每条中文文本分割成单个字符的列表，预训练模型将字向量作为模型参数的一部分，一起训练，因此其自带的字向量已较好的学习到了无监督训练中语料库的语义信息。为了符合模型预训练的输入模式，对于每个句子，还需要在句子首尾分别加入“[CLS]”和“[SEP]”，“[CLS]”经BERT特征转换后，将作为句子的情感特征表示，用于情感倾向预测。

此外，通过对比前述训练集上的单词数和预训练中单词数可知，训练集的规模远小于预训练语料库，因此，在训练时，为充分保留词向量中的语义信息，词向量将不参与训练。

* 1. 训练参数配置

本节对各个模型实验训练配置进行说明。本实验数据集情感倾向为正负两类，因此各模型输出为2维向量，用以表示句子在两个情感极性上的置信度，由此选择Cross Entropy作为损失函数。

选择F1分数作为模型评价的指标。在介绍F1分数前，首先介绍混淆矩阵。以本文任务为例，每个样本的真实值为正负两类，而模型的预测值不一定与真实相同，可分为四类：真实值为正，预测值也为正（True Positive, TP）；真实值为正，预测值为负（False Negative, FN）；真实值为负，预测值为正（False Positive, FP）；真实值为负，预测值为正（True Negative, TN），则混淆矩阵如表4.2所示。

表 4.2 混淆矩阵示意图

|  |  |  |  |
| --- | --- | --- | --- |
|  | | 真实值 | |
| True | False |
| 预测值 | Positive | TP | FP |
| Negtive | TN | FN |

在四类基本结果的基础上，可分别计算查准率（Precision）和查全率（Recall），查准率是预测值为正向样本中，真实值为正向的数量，其计算公式如式(4.1)所示。

而查全率是真实值为正向的样本中，预测值为正向的数量，计算公式如式(4.2)所示。

基于查全率和查准率，得到F1分数的计算规则，如(4.3)式所示。相比简单的准确率，F1分数更加全面了评价了模型的性能。

模型采用AdamW对模型进行优化[27]，TextCNN、Bi-LSTM、RCNN的初始学习率为0.001，BERT的初始学习率设为。为防止过拟合，设置L2正则参数设置为0.05，且设置early stop为10个epoch，即验证集上的F1分数经过10个epoch未上升，则停止训练。通过合理的超参数调整，保证各模型结构在本实验任务上达到最优，各模型超参数的详细调参结果见附录。每个模型参数随机初始化，连续训练10次，记录验证集和测试集上的最好和平均结果，并保存测试集上F1得分达到最高的模型。

* 1. 模型评估

2. 4. 1. 实验结果与分析

本节对基于深度学习的各个情感分析模型在本实验数据集上的最优性能进行对比。首先，对TextCNN、Bi-LSTM、RCNN三个非预训练模型进行超参数搜索，选择各模型在本次任务中的最优结构。

图 4.1 TextCNN实验结果

TextCNN主要的超参数有卷积核的尺度和每种尺度卷积核的数目，该模型调参结果如图4.1所示，各模型详细数值结果见附录A。其中，不同的柱体表示不同卷积核尺度，横坐标表示每种尺度卷积核数目，如“(2 – 4)”和128对应的模型中，有2、3、4个词向量尺度的卷积核，每个卷积核有128维。

由图4.1可知，TextCNN本数据集上最优的超参数配置为采用2和3尺度的卷积核，每个卷积核128个。该结果从一定程度上说明数据集中情感相关的局部信息用2到3个单词就能概况。

Bi-LSTM的超参数配置主要是Bi-LSTM单元的层数和单向LSTM状态的维度，该模型调参结果如图4.2所示，各模型详细数值结果见附录B。其中各柱体表示不同的Bi-LSTM层数，横坐标表示不同的中间状态维度。

图 4.2 Bi-LSTM实验结果

实验表明，Bi-LSTM的最优结构为2层LSTM单元，中间状态维度为512，且各模型的结果相差不大，仅在4层512维的Bi-LSTM出现了较强的衰退，结合循环神经网络的原理，可初步断定为参数过多，出现了梯度爆炸或消失。

RCNN的超参数配置与Bi-LSTM类似，主要考虑基本双向LSTM单元的层数，和中间状态的维度，该模型调参结果如图4.3所示，详细数值结果见附录C。其中每个柱体表示不同的LSTM单元层数，而横坐标则对应不同的中间状态维度。

图 4.3 RCNN实验结果

实验结果表明，在512维单层的Bi-LSTM结合池化机制，能在本任务上达到最优结果。

由于BERT为预训练模型，因此其结构不需修改，本次选用的预训练模型由12层encoder组成，词向量及中间状态维度为768维。

表 4.3 各模型最优F1分数

|  |  |  |
| --- | --- | --- |
| 模型 | 验证集（最高 / 平均） | 测试集（最高 / 平均） |
| TextCNN | 0.9207 / 0.9146 | 0.9356 / 0.9295 |
| Bi-LSTM | 0.9226 / 0.9169 | 0.9354 / 0.9264 |
| RCNN | 0.9273 / 0.9231 | 0.9408 / 0.9348 |
| BERT | 0.9528 / 0.9482 | **0.9550 / 0.9500** |

最终，各模型最优结构下的F1分数如表4.3所示。实验结果有个明显特征是，测试集上的F1分数明显普遍高于验证集，因此相对于验证集，该数据集的测试集分布更接近于训练集。

就测试集上的最优F1分数来看，BERT、RCNN、Bi-LSTM、TextCNN的性能依次递减，从各个模型的基本原理分析，可认为深度学习模型在句子级别中文情感分析任务上的性能与模型的语义挖掘能力密切相关，模型的语义表示能力越强，则越能对文本中的情感倾向做出准确的预测。

TextCNN采用卷积形式获取了文本中符合卷积核参数的局部情感特征，但由于文本的变长限制，卷积层无法叠加，对于情感特征较复杂，分隔较远的文本特征无法有效解决。

Bi-LSTM模型虽然在序列化计算上具有优势，但在对文本单词的语义学习能力上与TextCNN相差无几，即使经过多层双向的LSTM进行特征转换，简单的状态拼接得到的句子向量表示，对句子的情感信息抽取仍有不足。

RCNN通过加入池化层，最大限度的获取了句子特征中的情感信息，从而使得模型性能有所提升。另外，从attention的角度看，RCNN采用的池化也是一种打分函数，即选取每个特征维度上，最大的值，作为整个句子在该维度上的值，而普通的attention机制，则是所有输入向量基于attention分布的加权平均。

BERT的语义表示能力来源于其multi-head attention结构和预训练过程。Multi-head attention使该模型从不同语义空间学习到文本中的情感信息；通过MLM和连续语句预测，预训练模型及其词向量充分获取了语料库中的语义信息，因此仅需在下游任务的数据集上对模型参数进行微调，即可获取较好的情感倾向预测能力。

* + 1. RCNN优化与分析

将不同超参数的RCNN与Bi-LSTM在测试集上的F1分数进行对比，结果如图4.4所示。

图 4.4 Bi-LSTM与RCNN测试集上F1分数对比

其中，H表示单向状态的维度，而L表示LSTM单元的层数。实验表明，RCNN特征转换所需的参数，即RNN单元层数少于RNN，且其性能更加稳定。因此，RCNN在提升模型性能的同时，降低了参数量要求，提升了模型训练和特征转换的效率。

此外，值得注意的是，随LSTM层数增加，RCNN性能有所下降，这可能是因为本文实验所用的RCNN模型，跟Lai等人提出的RCNN模型并非完全一致[21]，该模型每层LSTM状态训练时均使用了前一层的输入。因此仅在1层LSTM单元时，本文的RCNN模型等同于该模型，且能达到较优的效果，但总之，RCNN引入的池化机制使其优于Bi-LSTM模型。

* + 1. BERT错误分析

为了进一步研究BERT模型的优势，对BERT在测试集中预测错误的样本，进行人工分析，可以将这些错误样本大致分成三类。

一类是标签错误，该类错误中，原本是正向含义的样本，标签为负向，或原本是负向语义的样本，标签为正向。例如，“3999的时候抢购的 运气真好 这个价格还有什么号说的 品牌和价格都很不错”这个句子表明用户对商品十分满意，但标注为负向，而BERT能正确预测为正向。

一类是语义错误，例如，测试集中有这样一个样本：“住这个酒店实在是太享受了,不仅可以使用五彩缤纷的白毛巾,还可以免费听赏别人KTV包厢里的高音演奏”。字面上，似乎对文本中的主体“酒店”含有称赞的意思，然而结合常识仔细分析便知，这句话含有讽刺意味，表达的是完全相反的意思。本文搭建的BERT模型对该句子情感倾向的理解便产生了偏差，因此将这类错误归为语义错误。

最后一类错误是歧义错误，一个句子中包含多个情感极性或者整体呈中性，带有歧义，则无法从中判别出评论用户对整个句子中的主体的情感倾向，例如，“周边环境较差，服务的速度较慢，态度还可以，价格太高”，这一句子的标签为负向，BERT模型将其判定为了正向，然而无论是人工还是机器，都无法猜测用户对这一句子中的主体到底是什么态度。

BERT模型在测试集上的错误数目为55条，人工分析得到三类错误样本统计数据如表4.4所示。

表 4.4 BERT错误样本分类统计

|  |  |
| --- | --- |
| 错误类型 | 错误数目 |
| 标签错误 | 23 |
| 语义错误 | 14 |
| 歧义错误 | 18 |

由表4.4可知，难以判定的歧义错误较多，可将其视为本文的句子级别情感分析任务中，处于两个情感极性间的噪声。标签错误表明数据集中的标注存在问题，然而模型经过充分训练后，能深入挖掘文本中的情感信息，因此仍能对其做出客观而正确的预测，这体现了BERT模型有较强的语义学习能力和抗过拟合的能力。语义错误则是该数据集中的难点，难度较大，但数量不多，表明模型的语义学习能力较强。

需要注意的是，也可能存在模型分类正确，而实际语义错误的样本，可将其视为噪声，但从分类正确这一角度看，也可说明模型从某种角度学习到了“错误”本身的特征。

* + 1. Multi-head Attention分析

本节从BERT模型的multi-head attention机制的角度，对其处理歧义错误的能力进行分析。

考虑到attention机制中，能生成具有概率分布性质的attention分布，因此可对BERT模型中，多层encoder中的多个head生成的attention进行可视化分析。本文选用的预训练模型有12层encoder，每层有12个head，因此，一个样本的情感分析将生成144个attention分布，每个attention都生成于不同的语义空间。针对带有歧义的样本：“周边环境较差，态度还可以”，选取生成的具有代表性的attention分布可视化结果，如图4.5和4.6所示。

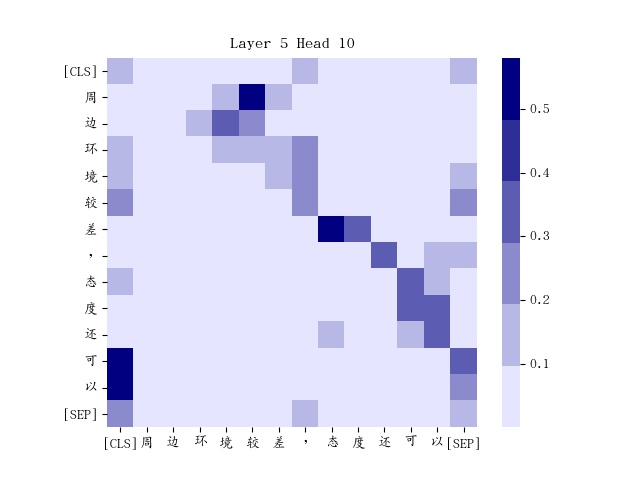


图 4.5 第5层第10个attention分布

图4.5中颜色较深的单元表明，在该语义空间中，attention机制学习到了对某些单词较为重要的语义信息。具体来讲，“环境”与“较差”有一定的关联，“态度”和“可以”也有类似的attention分布，该语义空间对评价对象有一定的认知能力，然而由于在句子粒度上进行情感分析，无法对用户情感进行精准的刻画，存在无法解决的局限性，采用更细粒度的方面级情感分析，对句子中的评价实体“酒店”的这两个“方面”的情感倾向分别进行预测，可以解决这一问题。

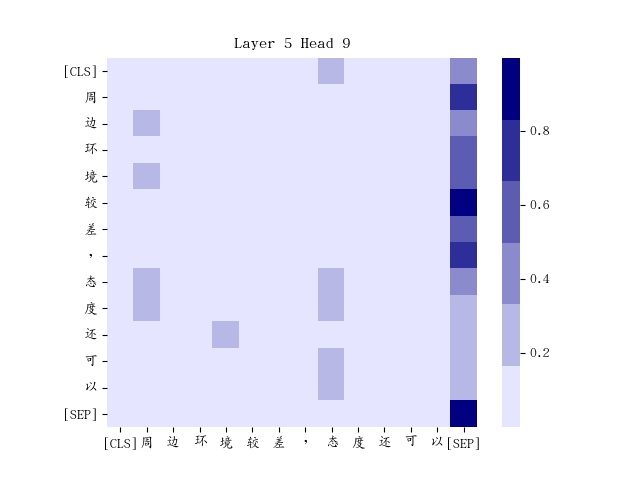


图 4.6 第5层第9个attention分布

图4.6中总体较浅的颜色分布则表示，该语义空间对句子在情感分析任务中的贡献较小。推广到其他基于单个attention机制模型，若该模型对应的语义空间生成的attention分布类似图4.6，则其结果就有可能受到不利影响。而在BERT中，虽然也可能存在这种影响，但因为多层encoder的叠加和多个语义空间attention拼接，这种影响在一定程度上得到消减。

# 总结

1. 1. 工作总结

本文对基于深度学习的句子级别中文情感分析算法进行了研究。采用不同深度学习模型建立了情感分析模型，并从模型原理角度，对比分析了各个模型在情感分析任务中的优缺点。

基于卷积神经网络，建立了中文情感分析模型TextCNN，该模型结合了深度学习的特征学习能力和传统n-gram模型的思想，通过卷积的方式探索了文本中的局部情感特征。但由于文本的变长特性限制，卷积层无法叠加，因此无法提取高层抽象特征，分隔较远的情感相关特征无法得到有效处理。

根据文本的序列特征，建立了基于Bi-LSTM的中文情感分析模型，充分利用了文本的序列信息，对文本的语义进行进一步的发掘。然而，仅采用简单的句子首尾LSTM状态拼接作为对于文本的特征表示，将存在单词顺序上的偏差，进而降低文本表示能力。

参考RCNN，建立了基于RCNN的中文情感分析模型，对Bi-LSTM特征转化后的文本表示进行池化处理，从而在减少模型参数的同时，最大限度的利用了文本表示每一维度的特征信息，使得情感分析性能进一步得到提升。从attention机制的角度来讲，池化也是一种attention打分函数，且实验表明，该机制在本次任务中，相比于加权平均机制更有效。

最后建立了基于BERT的中文情感分析模型，结合BERT预训练的语义表示能力，该模型达到了本次任务中最高的准确率。为进一步分析BERT模型的优越性，对BERT进行了错误分析和attention可视化分析，实验表明，BERT具有较好的语义表示能力，能充分发掘句子中的情感相关信息。

* 1. 研究展望

从深度学习算法性能的角度看，BERT模型的multi-head attention机制相比于单次Attention占有极大优势，然而实际模型训练过程中，由于模型参数较多，临时生成的attention分布占用内存较大，因此时间、空间效率都有待提高。而根据Bi-LSTM和RCNN性能对比发现，RCNN通过池化特征向量，同样能较好的提取语义信息，同时不需要生成attention分布。结合RCNN的原理，或许能解决上述问题。

从情感分析的角度看，实验对BERT模型存在的错误，进行了分类，分析发现歧义错误占有一定比例，而通过BERT的attention可视化发现，BERT对歧义文本的评价对象的各个方面仍有一定的认知能力，但句子级别的情感分析存在局限，无法对文本总体做出情感倾向判定，更适合采用细粒度的方面级情感分析。

参考文献

[1] Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space[A]. International Conference on Learning Representations(ICLR)[C], 2013.

致谢

感谢我的指导老师，刘进教授对我这次毕业设计的支持与帮助。刘进老师为我提供了良好的学习氛围，以及更优秀的科研平台，开阔了我的眼界。在刘进老师的指导下，我领略到了科研工作的乐趣，体会到了知识的魅力，让我对未来的生活更加憧憬，我将更加积极的面对今后的学习生活。

感谢实验室各位师兄，帮助我解决了实验和论文中的许多问题，对于一些细节上的难题总能给出很好的建议，让我少走弯路，他们的渊博学识是我今后追赶的目标。特别感谢谢志文师兄对我在实验思路上的指点，以及赵焜松师兄在论文工作上不辞劳苦的为我解惑。

感谢武汉大学的各位同学，老师毕设期间乃至整个本科四年对我的帮助，四年时光匆匆过去，有许多值得铭记的人和事，无论何时何地，我都心怀感激。特别感谢我的女友，闫世奇女士的陪伴。

最后，还要感谢我的家人，他们的支持是我最坚实的后盾，也是我一路走来，成长的动力。

附录

1. TextCNN模型结构调参结果

1. https://github.com/fxsjy/jieba [↑](#footnote-ref-1)