

基于深度学习的微博情感分析*

梁军¹, 柴玉梅¹, 原慧斌², 咎红英¹, 刘铭¹

(1.郑州大学 信息工程学院, 河南 郑州 450001; 2.中国核科技信息与经济研究院 北京 10048)

摘要: 中文微博情感分析旨在发现用户对热点事件的观点态度。已有的研究大多使用 SVM、CRF 等传统算法根据手工标注情感特征对微博情感进行分析。本文主要探讨利用深度学习来做中文微博情感分析的可行性, 采用递归神经网络来发现与任务相关的特征, 避免依赖于具体任务的人工特征设计, 并根据句子词语间前后的关联性引入情感极性转移模型加强对文本关联性的捕获。本文提出的方法在性能上与当前采用手工标注情感特征的方法相当, 但节省了大量人工标注的工作量。

关键词: 深度学习; 微博情感分析; 递归神经网络; 自编码

Deep Learning for Chinese Micro-blog Sentiment Analysis

Liang Jun¹, Chai Yumei¹, Yuan Huibin², Zan Hongying¹, Liu Ming¹

(1.School of Information Engineering, Zhengzhou University, Zhengzhou, Henan 450001, China;

2. China Institute of Nuclear Information & Economics, Beijing 10048, China)

Abstract: Chinese micro-blog sentiment analysis aims to discover users' attitude towards hot events. Existing studies mostly analyzed the micro-blog sentiment based on hand-engineered features by using other traditional algorithms such as SVM, CRF. This paper explores the feasibility of performing Chinese micro-blog sentiment analysis by deep learning. We try to avoid task-specific features, and use recursive neural networks to discover relevant features to the tasks. We propose a novel model - sentiment polarity transition model - based on the relationship between neighboring words of a sentence to strengthen the text association. The proposed method achieved close to state-of-the-art performance with the current method based on the hand-engineered features, thus saving a lot of manual annotation work.

Key words: deep learning; micro-blog sentiment analysis; recursive neural networks; autoencoder

1 引言

随着社交网络的不断发展, 人们更愿意通过微博、博客社区来表达自己的观点, 发表对

基金项目: 本文承国家自然科学基金项目(60970083, 61272221)、国家社会科学基金项目(14BYY096)、国家高技术研究发展 863 计划(2012AA011101)、河南省科技厅科技攻关计划项目(132102210407)、河南省科技厅基础研究项目(142300410231, 142300410308)、河南省教育厅科学技术研究重点项目(12B520055, 13B520381)资助。

作者简介: 梁军(出生年 1990—), 男, 研究生, 主要研究领域为机器学习、自然语言处理, Email: zhengdaxg@163.com, 通讯作者; 柴玉梅(1964—), 女, 教授, 主要研究领域为机器学习、数据挖掘和软件工程; 原慧斌(1966—), 女, 中国核科技信息与经济研究院保密测评中心副主任, 理学士, 研究员级高工, 主要研究领域为中文全文检索技术、信息安全技术、涉密信息系统安全保密测评; 咎红英(1966—), 女, 教授, 主要研究领域为自然语言处理、文本分类和情感计算; 刘铭(1990—), 男, 研究生, 主要研究领域为自然语言处理。

热点事件的评论。从而使得通过微博、博客、影评以及产品评价等来了解社交网络用户的情感倾向得到了学术界的广泛关注。根据微博数据进行情感分析是一个具有挑战性的任务，并且近年来引发了研究学者极大的兴趣^[1]。

目前，情感分析的主要研究方法还是一些基于机器学习的传统算法，如 SVM、信息熵、CRF 等。这些方法归纳起来有三类：有监督学习、无监督学习和半监督学习。而当前大多数基于有监督学习的研究都取得了不错的成绩^[2]，但是由于有监督学习依赖于大量人工标注的数据，使得基于有监督学习的系统需要付出很高的标注代价。相反的，无监督学习不需要人工标注数据训练模型，是降低标注代价的解决方案，但由于其完全依赖算法学习结果，往往效果不佳，难以达到实际要求。而半监督学习则是采取综合利用少量已标注样本和大量未标注样本来提高学习性能的机器学习方法，它兼顾了人工标注成本和学习效果，被视为一种折中方案。本文围绕半监督学习方法基于自编码算法提出一种新的分析中文微博情感倾向的深度学习算法。

已有研究所采用的方法大多数都基于词袋模型，而这种模型无法捕获到很多有关情感倾向性分析的语言现象特征。例如，“反法西斯联盟击溃了法西斯”和“法西斯击溃了反法西斯联盟”这两个词组拥有相同的词袋模型表示方法，而前一个带有积极的感情色彩，后一个带有消极的感情色彩。除此之外，还有很多研究者使用人工标注的数据（情感词典及句法分析等），虽然采用这些方法可以有效的提高情感分析的准确性，但由于需要较多的人工标注数据，从而限制了这些方法在其他领域以及跨语言的推广。

本文采用基于深度学习方法，不使用任何人工标注的情感词典以及句法分析结果，仅使用少量标注训练集和测试集，使用递归神经网络进行情感分类。本文首次将句子中每个词语的标签关联考虑在内提出情感极性转移模型，通过实验发现加入该模型可明显提高算法准确率。

2 相关工作

2.1 深度学习

深度学习是机器学习研究中的一个新的领域，其动机在于建立、模拟人脑进行分析学习的神经网络，它模仿人脑的机制来解释数据，例如图像，声音和文本。随着深度学习方法在图像处理和语音识别方向的成功应用，越来越多的深度学习的方法也被应用于自然语言处理方向。Bengio et al.^[3] 2003 年提出用神经网络构建二元语言模型的方法。随后，Ronan Collobert 和 Jason Weston^[4] 于 2008 年推出 SENNA 系统，使用词向量方法去完成自然语言处理中的各种任务，如词性标注、命名体识别、短语识别、语义角色标注等。Andriy Mnih 和 Geoffrey Hinton 于 2008 年提出一种层次的思想训练语言模型^[5]，Mikolov^[6-8] 最终使用“Log-Bilinear”模型成功将深度学习的模型降低到可接受范围内，并随后在谷歌推出 word2vec 将词语转换

成词向量的工具^[9]。Socher^[10]于 2012 年提出基于递归自编码 (Recursive AutoEncoder, RAE) 的树回归模型用来分析句子的情感倾向性, 但该模型无法准确捕获文本中的情感极性转移现象, 本文采用 RAE 这种深度学习模型, 引入词语间前后的关联性对其进行改进使得更适合于中文微博情感分析任务。

2.2 情感分析

情感分析自从 2002 年由 Bo Pang 提出之后, 获得了很大程度的关注, 特别是在在线评论的情感倾向性分析上获得了很大的发展。本文主要关注无监督的情感分析方法, 由于不需要大量标注语料, 无监督情感分析方法一直受到许多研究者的青睐, 但同时效果也低于有监督的情感分析方法。Turney^[11]首次提出给予种子词 (excellent, poor) 的非监督学习方法, 使用 “excellent” 和 “poor” 两个种子词与未知词在搜索网页中的互信息来计算未知词的情感极性, 并用以计算整个文本的情感极性。后续的非监督情感分析方法大都是基于生成或已有的情感词典或者相关资源进行情感分析。例如, Kennedy 和 Inkpen^[12]考虑文本中词的极性转移关系并基于种子词集合进行词计数决定情感倾向。朱嫣岚^[13]等人将一组已知极性的词语集合作为种子, 基于 HowNet 对未知词语与种子词进行语义计算, 从而判别未知词的极性。Lin et al.^[14] 采用 LSM 模型、JST 模型、Reverse-JST 模型构建了三种无监督的情感分析系统。但是由于深层情感分析必然涉及到语义的分析, 以及文本中情感转移现象的经常出现, 所以基于深层语义的情感分析效果并不理想, 本文针对中文文本中经常出现的情感转移现象提出情感极性转移模型, 提高了深层语义情感分析的分析效果。

3 基于 RAE 的深度学习模型

为了实现用于微博情感分析的深度学习模型, 本文提出了一个基于 RAE 的情感极性转移模型。该模型首先将文本数据转为为低维实数向量表示, 建立表示文本特征的矩阵, 然后将其作为基于 RAE 的情感极性转移模型的输入, 最后使用 LBFGS 算法多次迭代生成最终的模型。该模型可对低维实数向量表示的文本进行情感分类输出其情感极性。

3.1 使用词向量表示词语

不同于传统的词袋模型, 在使用神经网络解决自然语言处理问题时, 通常采用一个低维的实数向量来表示词语以避免数据维数灾难 (Bengio et al., 2003; Collobert and Weston, 2008)。例如, 使用向量 $[0.3 \ 0.1 \ 0.6]^T$ 表示 “深度”, 向量 $[0.2 \ 0.5 \ 0.7]^T$ 表示 “学习”。然后将这些实数向量作为神经网络的输入。通常情况下, 将每一个词语映射到 n 维实数向量空间上, 即 $x \in \mathbb{R}^n$, 将这些向量堆放在一起构成一个词向量矩阵 $L \in \mathbb{R}^{n \times |V|}$, 其中 $|V|$ 是词表的大小。

给定句子 $c_{[1:n]}$ ，它包含 n 个词 c_i ， $1 \leq i \leq n$ 。其中 c_i 的特征向量为矩阵 L 的第 k_i 列，可表示为

$$x_i = L e_{k_i} \in \mathbb{R}^n \quad (1)$$

其中 $e_{k_i} \in \mathbb{R}^{|V|}$ 是 $\mathbb{R}^{|V|}$ 中的第 k_i 个单位向量（除第 k_i 个分量为 1 外其他分量均为 0）。这样该查询操作可视为一个简单的映射层。

3.2 递归自编码

自编码的目的是学习输入数据中隐含着的特定结构，本部分将具体介绍如何使用递归自编码获取非叶子结点的低维向量表示。例如“深度”的向量表示为 $[0.3 \ 0.1 \ 0.6]^T$ ，“学习”的向量表示为 $[0.2 \ 0.5 \ 0.7]^T$ ，那么该如何表示“深度 学习”呢？假定“深度 学习”是“深度”、“学习”的父节点 p ，“深度”是第一个子节点 c_1 ，“学习”是第二个子节点 c_2 ，那么 p 可由函数 f 从 c_1 、 c_2 映射得到：

$$p = f(w^{(1)}[c_1; c_2] + b^{(1)}) \quad (2)$$

其中 $[c_1; c_2] \in \mathbb{R}^{2n}$ ，由向量 c_1 、 c_2 联接得到； $w^{(1)} \in \mathbb{R}^{n \times 2n}$ 是一个参数矩阵； $b \in \mathbb{R}^n$ 为偏置项；编码器函数 $f(\cdot)$ 在这里选取双曲正切函数 $\tanh(\cdot)$ 。

按照公式(2)可以得到父节点的 n 维向量表示 $[0.8 \ 0.2 \ 0.7]^T$ 。对整个句子递归使用这种神经网络结构，使用贪心算法每次选取两个节点结合之后父节点得分最高的节点组合，最终可以得到整个句子的 n 维向量表示，如图 1 所示。

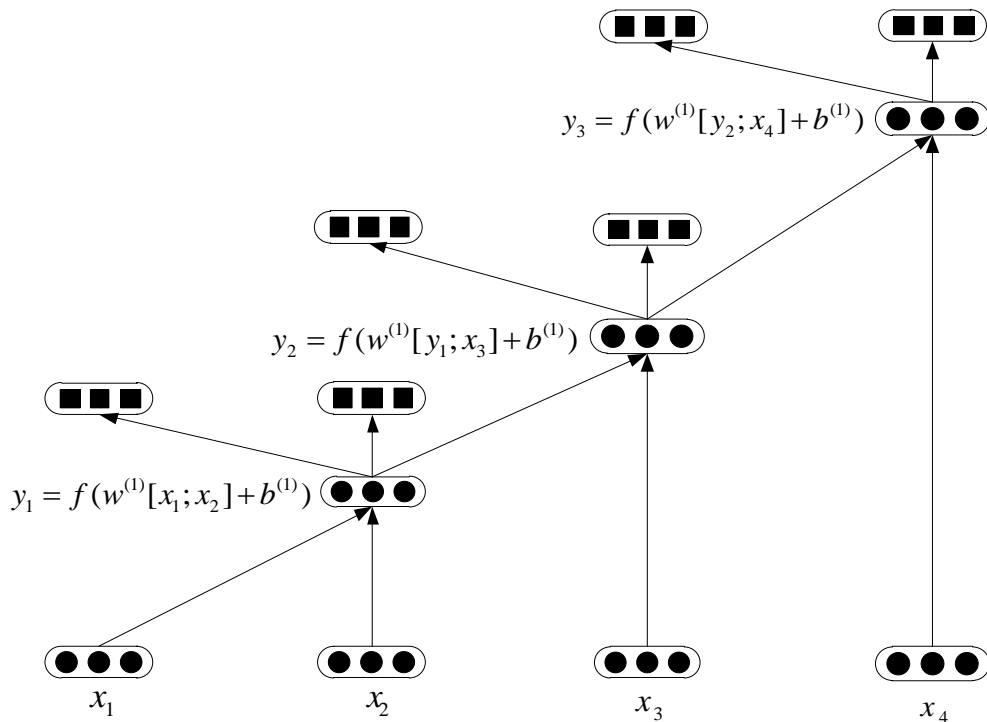


图 1 递归自编码结构图，圆形节点表示原始节点，矩形节点表示重建节点

为了验证父节点 p 对子节点 c_1 、 c_2 的表示度，在父节点 p 上建立一个重建层，重建两个子节点：

$$[\dot{c}_1; \dot{c}_2] = g(w^{(2)}p + b^{(2)}) \quad (3)$$

其中 $\dot{c}_1, \dot{c}_2 \in \mathbb{R}^n$ 是重建子节点； $w^{(2)} \in \mathbb{R}^{n \times 2n}$ 是参数矩阵； $b^{(2)} \in \mathbb{R}^n$ 是偏置项；解码器函数 $g(\cdot)$ 这里取恒等函数。类似地，将自编码网络递归地应用于二叉树的每一层。为了评估重建节点 \dot{c}_1 、 \dot{c}_2 与原始节点的相似度，使用重构误差函数进行描述：

$$E_{rec} = \|c_1 - \dot{c}_1\|^2 + \|c_2 - \dot{c}_2\|^2 \quad (4)$$

这种神经网络被称为递归自编码^[10]，图 1 展示了应用于一个二叉树的递归自编码网络，其中圆形节点表示原始节点，矩形节点表示重建节点。对二叉树的每一个节点使用自编码器，那么该二叉树最终可由一个三元组（如： $(p_1 \rightarrow c_1 c_2)$ ， p 是父节点， c_1 、 c_2 是子节点）集合表示。图 1 中的二叉树可表示为： $(y_1 \rightarrow x_1 x_2)$ ， $(y_2 \rightarrow y_1 x_3)$ 和 $(y_3 \rightarrow y_2 x_4)$ 。

仔细分析发现，按照上面的自编码方式进行训练可能出现两个问题：①节点的向量表示为零向量，这会导致该节点的重建误差 E_{rec} 为零，事实上这是没有意义的；②节点 c_1 和节点 c_2 可能包含的叶子节点数有很大差异，计算重建误差时会导致不平衡。

为了解决问题①，可以使用单位向量来表示节点 p (p 表示任一节点) 即：

$$p = \frac{f(W^{(1)}[x_1; x_2] + b^{(1)})}{\|f(W^{(1)}[x_1; x_2] + b^{(1)})\|} \quad (5)$$

通过对父节点的向量进行正则化，可以有效避免优化算法因最小化重建误差而导致父节点向量为零向量。

为了解决问题②，可以在计算重建误差 E_{rec} 时，为每个子节点添加权重，权重的设定根据每个子节点所包含叶子节点数计算，假定节点 c_1 的叶子节点数为 n_1 ，节点 c_2 的叶子节点数为 n_2 ，重新定义重建误差 E_{rec} ：

$$E_{rec} = \frac{n_1}{n_1 + n_2} \|c_1 - \dot{c}_1\|^2 + \frac{n_2}{n_1 + n_2} \|c_2 - \dot{c}_2\|^2 \quad (6)$$

通过节点权重的添加，可以使得计算重建误差时更多地偏向于子节点数目更多的结点，从而达到平衡误差的效果。

3.3 情感极性转移模型

使用递归自编码可以得到二叉树中每个节点的 n 维向量表示，根节点则是整个句子的 n 维向量表示。为了得到句子的情感倾向性，可以再加一层输出层：

$$s(p; \theta) = \mu(w^{(3)}p + b^{(3)}) \quad (7)$$

其中 $w^{(3)} \in \mathbb{R}^{k \times n}$ (k 是情感标签的数量，本文仅关注负向和正向两类，即 $k=2$)； $b^{(3)} \in \mathbb{R}^k$ 为偏置项；激活函数 $\mu(\cdot)$ 选择 $\text{softmax}(\cdot)$ 函数。如图 2 所示，其中三角形节点表示输出层，输

出句子的情感倾向性。

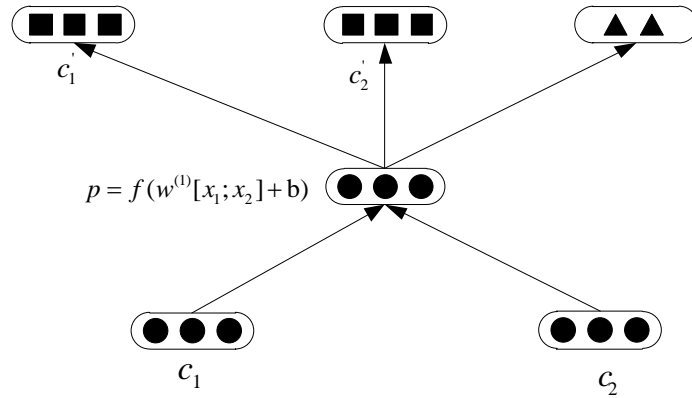


图2 加入输出层的递归自编码模型，其中三角形节点表示输出层

定义情感标注集 $T = \{\text{正}, \text{负}\}$ ，这样得到的 $s(p; \theta)$ 中的每一个分量 s_i 就对应该句的情感倾向为 T_i 的得分。为了验证该得分的准确性，采用交叉熵作为代价函数（假定 t_i 表示句子情感倾向性为 T_i 的真实概率）：

$$E_{ce} = -\sum_{k=1}^K t_k \log s_k(p; \theta) \quad (8)$$

对于情感分析这样的自然语言处理任务来说，一个句子中的各个词语之间存在很强的依赖性。例如，“不 喜欢”这样的子句，“不”的情感倾向为负，“喜欢”的情感倾向为正，而“不 喜欢”子句的情感倾向性显然为负。当连续两个词语的情感标签由 T 中的第 i 个 t_i 变为第 j 个 t_j 时，引入相应的转换分数 A_{ij} ($A_{ij} \in \{A_{01}, A_{00}, A_{10}, A_{11}\}$)，如图3所示。

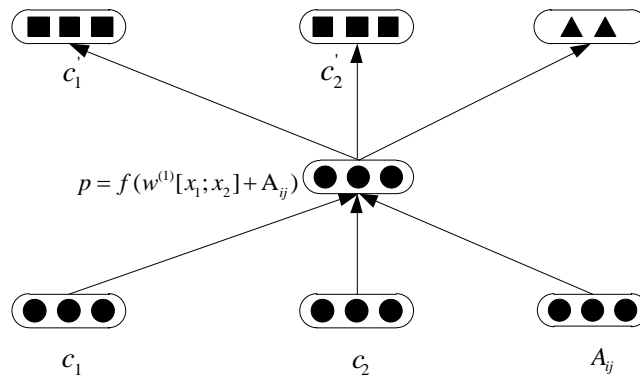


图3 加入情感转换因子的递归自编码模型， A_{ij} 为 x_1 到 x_2 的极性转换因子

根据 3.2 节得到的重建误差函数和 3.3 节得到的标签误差函数，可以得到情感极性转移模型的代价函数为（数据集假定为 S ，每个数据对象包含一个句子 c 和对应的情感标签 t ）：

$$J = \frac{1}{N} \sum_{(c,t) \in S} E(c,t; \theta) + \frac{\lambda}{2} \|\theta\|^2 \quad (9)$$

其中， $E(c,t; \theta) = \alpha E_{rec} + (1-\alpha) E_{ce}$ (α 为可调参数)， λ 也为可调参数。具体模型的算法

描述:

训练情感分析模型

输入: 训练语料 S 及其对应标签 t

输出: $\theta = \{w^{(1)}, w^{(2)}, w^{(3)}, A, b^{(2)}, b^{(3)}\}$

1) 使用高斯分布初始化训练语料的词向量表示;

2) **while** 不收敛 **do**

$\nabla J = 0$

for all $\langle s, t \rangle \in S$ **do**

 使用贪心算法生成句子二叉树结构

 计算 $\nabla J_i = \partial J(s, t) / \partial \theta$

 更新 $\nabla J = \nabla J + \nabla J_i$

end for

 更新 $\theta = \frac{1}{N} \nabla J + \lambda \theta$

end while

4 实验

为了保证结果的可靠性, 本文选用两个数据集来验证模型的有效性。①COAE2014 微博数据集, 采用人工标注的方式, 标注了一万条数据, 其中带有情感色彩的有 2740 句, 带有消极情感的有 1608 句, 带有积极情感的有 1132 句。②COAE2014 微博数据集完整版, 共计 40000 条数据, 其中包含已标注过的一万条数据。

表 1 COAE2014 数据集样例

积极	消极
奥迪 Q3 SUV, 不错	保险公司很无语啊!
我的三星被我摔了好几次, 摔坏了好几个壳, 手机还是完好无损哈。	我同意, 外资的保险公司似乎缺少了狼性。但友邦有待考量。
蒙牛那个点滴的广告还不错	三星要是能领导行业那才见鬼了
过六天测试。奥迪 Q5 不错。	手机真害人不浅啊!

本文共设计三个实验来验证情感极性转移模型的有效性。实验一: 可调参数选择, 用来确定可调参数对算法的影响并选择一组最优的参数; 实验二: COAE2014 数据集封闭测试, 用来对比在相同实验集上未使用其他人工标注数据的情感极性转移模型与其他传统方法的效果; 实验三: 对比实验, 比较本文提出的改进模型与 RAE 模型的优劣。

4.1 可调参数选择

本实验用来测试可调参数（具体包括算法迭代次数、正则惩罚项系数、重建误差权重），对算法性能的影响。实验数据采用人工标注 2740 句带有感情色彩的句子进行十折交叉验证（即随机抽取 2466 个句子作为训练集，274 个句子作为开发集）。

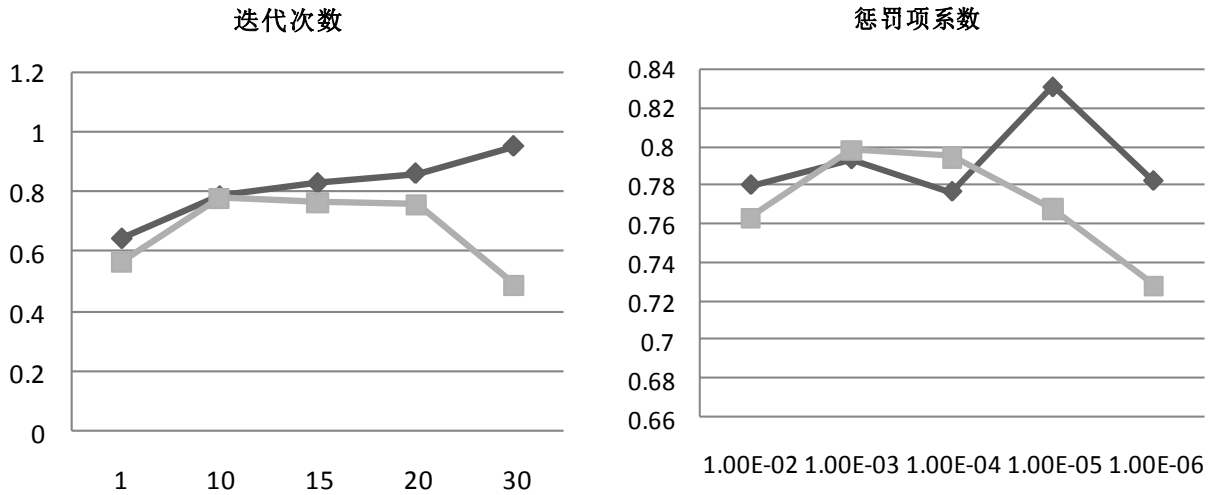


图 a 迭代次数选择

图 b 惩罚项系数选择

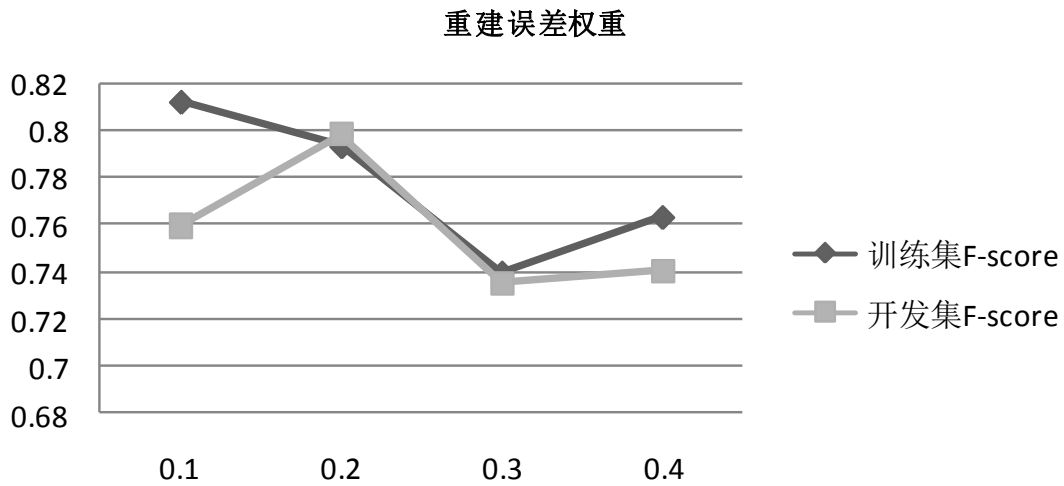


图 c 重建误差权重选择

图 4 迭代次数、惩罚项系数和重建误差权重对算法的影响

通过实验可以发现，随着迭代次数的增加模型对训练数据的拟合越来越好，F 值可以达到 95%，带随之而来的就是过拟合问题，导致模型的泛化能力下降，对比发现当迭代次数为 10 次数在训练数据和测试数据上的 F 值都能达到 80% 左右。观察惩罚项系数对算法性能的影响曲线可知：当惩罚项系数为 0.001 时，对于训练数据和测试数据均可以达到较好的结

果，随着惩罚系数的减小，过拟合问题越来越严重（测试数据准确率下降）。根据图 4 结果，可以发现，当重建误差占误差比 20%时算法的性能达到最好。

4.2 COAE2014 数据集封闭测试

本实验数据采用 COAE2014 微博数据集完整版（与实验一的训练集、开发集不同，该数据集共计 40000 条数据，包含实验一采用的 2740 句数据）作为测试集，进行完全封闭实验，不采用任何评测外数据集，并与其他应用于该评测数据集的结果作比较。

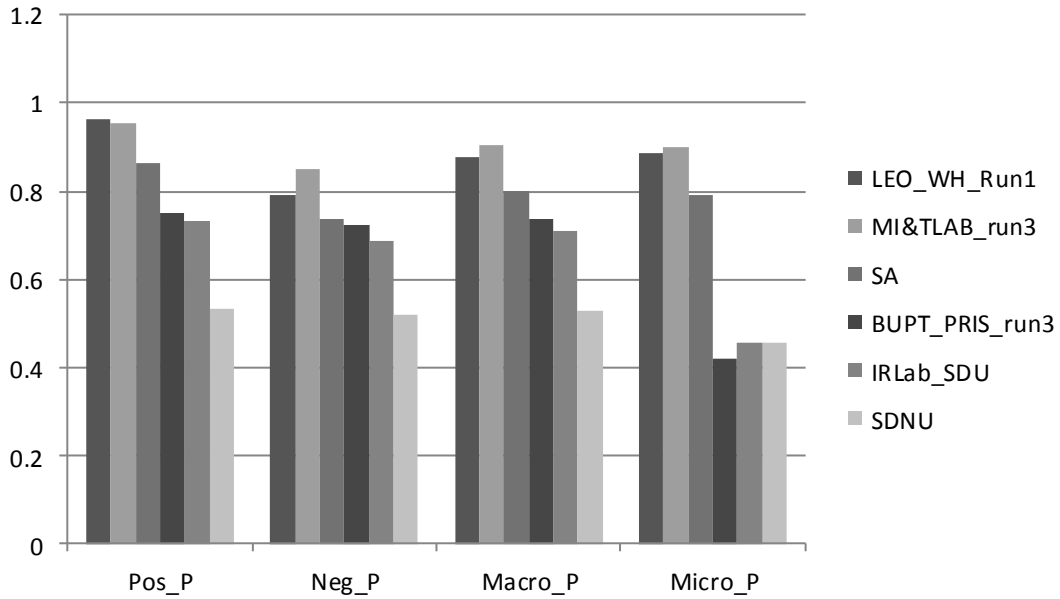


图 5 比较不同系统在 COAE 数据集上的性能

SA 为本文系统，与 COAE2014 上其他系统相比，在准确率上均与最好水平接近，且通过分层训练模型大大降低了深度学习的训练时间复杂度，达到了可以应用的水平。

4.3 对比试验

本实验选用除标注数据外的三万条数据作为测试数据，在实验数据上分别使用递归自编码模型和情感极性转移模型，对比情感极性转移模型相对于传统递归自编码模型的优劣。

表 2 可调参数的设置

可调参数	值
词向量维数	50
迭代次数	10
正则惩罚项系数	1e-03
重建误差与交叉熵误差权重比	0.2

按照表 2 设定可调参数,在相同条件下对比情感极性转移模型相对于传统递归自编码模型,结果如下:

表 3 两个模型对比结果

系统	准确率	召回率	F-Score
Recursive AutoEncoder	0.787546	0.784394	0.778786
情感极性转移模型	0.805861	0.805918	0.797973

通过对比,发现在加入情感极性转移因子后,该模型明显有了较大程度的提升。因为句子中的情感发生转移在中文句子中是很常见的,而传统 RAE 模型没有考虑该因素。

5 结语

本文提出了一种基于 RAE 的情感极性转移模型,在不增加神经网络复杂度的前提下,提高了模型对中文微博情感分析的准确性,算法的表现已经很接近当前采用许多手工标注特征的传统算法的性能。

深度学习方法在情感分析上的研究还存在很多问题需要进一步探讨。例如,从上面的实验结果可以看出,神经网络模型由于参数数量较多在训练模型时很容易出现过拟合的情况。下一步工作将深入研究如何降低模型的结构风险,从而寻找到更适合情感分析的深度学习算法。

6 参考文献

- [1] B. Pang and L. Lee. 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In ACL, pages 115-124.
- [2] 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, 6(2):88-94
- [3] Y. Bengio, R. Ducharme, P. Vincent, and C. Janvin. 2003. A neural probabilistic language model. Journal of Machine Learning Research, 3:1137-1155.
- [4] Collobert R, Weston J. A unified architecture for natural language processing: Deep neural networks with multitask learning[C]//Proceedings of the 25th international conference on Machine learning. ACM, 2008: 160-167.
- [5] Mnih A, Hinton G E. A Scalable Hierarchical Distributed Language Model[C]//NIPS. 2008: 1081-1088.
- [6] Mikolov T, Karafiát M, Burget L, et al. Recurrent neural network based language model[C]//INTERSPEECH. 2010: 1045-1048.
- [7] Mikolov T, Kombrink S, Burget L, et al. Extensions of recurrent neural network language

- model[C]//Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. IEEE, 2011: 5528–5531.
- [8] Kombrink S, Mikolov T, Karafiát M, et al. Recurrent Neural Network Based Language Modeling in Meeting Recognition[C]//INTERSPEECH. 2011: 2877–2880.
- [9] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space[J]. arXiv preprint arXiv:1301.3781, 2013.
- [10] Richard Socher, Jeffrey Pennington, Eric Huang, Andrew Y. Ng, and Christopher D. Manning Conference on Empirical Methods in Natural Language Processing (EMNLP 2011, Oral) Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions.
- [11] Turney P D. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews[C]//Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics, 2002: 417–424.
- [12] Kennedy A, Inkpen D. Sentiment classification of movie reviews using contextual valence shifters[J]. Computational Intelligence, 2006, 22(2): 110–125.
- [13] 朱嫣岚, 闵锦, 周雅倩, 等. 基于 HowNet 的词汇语义倾向计算 [J]. 中文信息学报, 2006, 20(1): 14–20.
- [14] Lin C, He Y, Everson R. A comparative study of Bayesian models for unsupervised sentiment detection[C]//Proceedings of the Fourteenth Conference on Computational Natural Language Learning. Association for computational linguistics, 2010: 144–152.