

基于深度学习的中文电商评论情感分析

王宏生, 金相宇

(沈阳工业大学信息科学与工程学院, 辽宁 沈阳 110870)

摘要: 随着在线评论和推荐系统可用性的不断增加, 情感分析逐渐成为学术研究中的一个重要的任务, 近年来得到了许多研究人员的关注。传统的解决文本情感分析问题的方法主要是基于情感词典或浅层学习, 利用回归、分类等方案实现特征的提取及分类。但由于评论语料内容简短, 特征稀疏, 含有大量未登录词使得上述这些方法存在数据稀疏问题, 且忽略了词的语序问题。以这类方法为起点, 本文采用深度学习的方法对评论文本进行情感分析, 通过组合实验对比, 找到可以深度理解人类情感表达的模型。

关键词: 情感分析; 深度学习; 评论文本; 情感词典

中图分类号: TP391

文献标识码: A

文章编号: 1673-1131(2018)03-0051-03

Sentiment Analysis of Chinese Ecommerce Reviews Based on Deep Learning

Wang Hongsheng, Jin Xiangyu

(College of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China)

Abstract: With the increasing availability of online commentary and recommendation systems, sentiment analysis has gradually become an important task in academic research. In recent years, it has attracted the attention of many researchers. The traditional methods to solve sentiment analysis are mainly based on sentiment dictionaries or shallow learning. The methods of regression and classification are used to extract and classify features. However, due to the shortness of the review corpus, sparse features and the large number of unregistered words, these methods have the problem of data sparseness and ignore the word order. Starting from this kind of method, we use the method of deep learning to analyze the sentiment of the comment texts, and find the model that can deeply understand the human emotion expression through the combination of experiments.

Key words: sentiment analysis; deep learning; commenting text; sentiment dictionaries

0 引言

近些年, 随着社交媒体如论坛, 微博, 电子商务等的兴起, 用户对于产品的评论, 评分等内容不断增加, 能够自动的过滤并抽取其中有价值信息就成了企业出售商品, 发现新的市场机遇的关键挑战。这引起了情感分析研究的激增。文本的情感分析又称意见挖掘, 是对人的观点进行分析, 识别出其情感趋向——高兴、伤悲等, 或得出其观点是“赞同”还是“反对”, 它是对带有情感色彩的主观性文本进行分析、处理、归纳和推理的过程^[1]。

目前主流的情感分析的方法有三类, 一类是基于情感词典的方法, 这类方法主要依赖于情感词典, 并通过人为设计的规则进行情感分类, 第二类是传统的基于机器学习的方法, 这类方法需首先挖掘出词语的特征, 再使用如支持向量机(Support Vector Machine, SVM)、朴素贝叶斯(NB)、随机森林等分类算法对文本进行情感倾向性的判断。最后一类则为基于深度学习的方法, 即使用不同的神经网络模型, 将文本映射到向量空间得到词语的数值表示, 再将向量输入到分类器中。使用神经网络的好处在于可以避免繁琐的特征选择过程, 可捕捉到复杂的特征。因此, 本文利用深度学习的方法对中文电商评论文本进行情感分析。首先对文本进行预处理, 再对文本进行向量化表示, 最后引入 GRU 深度学习模型进行情感倾向性判断。

1 相关工作

随着深度学习在语音识别及图像处理飞速发展, 近几年越来越多的学者将其引入到情感分析中。在国外, 考虑到词

语的上下文信息, Kim^[2]提出使用卷积神经网络进行句子建模, 解决情感分析任务, 在多个数据集上均取得了很好的结果。Socher 在^[3]中相继提出 RNN、MRNN、RNTN 等多个递归神经网络(Recursive Neural Network)模型, 将文本的句法结构考虑到神经网络模型中, 在句子级别、短语级别得到的结果均远高于基准系统。另外, 考虑到文本中的词之间的时序信息, Tai 等^[4]使用更加复杂的长短期记忆模型(Long-Short Term Memory)来解决情感分析问题。Lai 等人使用循环神经网络的架构来获得上下文的信息, 相对于传统的基于窗口的神经网络在性能上获得了很大的提升^[5]。

在国内, 梁军等人首次将句子中每个词语的标签关联考虑在内提出一个基于 RAE 的情感极性转移模型^[6]。该模型首先将文本数据转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。该模型可明显提高算法准确率。Xiao SUN 基于微博提出了一种内容扩展方法^[7], 将文本和其评论相结合构成一个用于情感分析的微博对话。利用一种新的卷积自动编码器, 从该微博对话中提取上下文情感信息。此外, 河炎祥^[8]等将常用的表情符号引入到情感分类任务中, 将其作为情感特征输入到 MCNN 模型进行分类, 该模型多个情感分类实验中取得了较好的效果。目前已经有人将 CNN, LSTM 等深度神经网络引入到中文情感分析这一领域, 但由于 CNN 无法很好的解决时序问题, LSTM 虽可以很好处理上下文关系, 但是需训练过多的参数。因此, 本文将 GRU(Gated Recurrent Unit)引入到中文情感分析这一问题中, 用 GRU 实现情感

分析这一任务。

2 情感分析模型

2.1 使用词向量表示词语

为了利用深度学习模型对文本进行情感分析,需首先通过训练将文本中的每一个词映射成一个固定长度的向量,作为深度神经网络模型的输入层。基于传统词向量表示的深度神经网络模型在情感分析任务上取得了很好的效果,但由于传统的词向量仅考虑了文本中的语义或语法信息,未考虑词语本身的情感极性,往往导致一些具有用法相似,情感极性却相反的词语被映射到相邻的位置,这对于情感分析这一任务是致命的错误。因此本文在传统词向量的基础上融合了情感信息。

使用谷歌研究人员 Mikolov^[9]提出的 word2vec 模型框架中的 Continuous Bags of Words (CBOW) 来训练词向量 X_{cbow}。CBOW 根据上下文来预测目标词,表示为:

$$P(W_t | W_{t-n}, W_{t-n-1}, \dots, W_{t+n}, W_{t+n+1}) \quad (1)$$

其中 W_t 为语料库中词语。CBOW 通过与 W_t 相邻上下文的窗口大小为 n 的词来预测 W_t 出现的概率。

X_{cbow} 是根据词语的共现语境训练得到的,只包含词的语义信息。为了获取情感信息,本文将利用情感词典中词语的情感标签获取情感信息向量 X_{sem} 。首先结合清华大学李军的中文褒贬义词典和台湾大学 NTUSD 中文情感词典构成最终的情感词典,表 1 为情感词典的词语分布情况。

表 1 情感词典

情感词典	积极	消极	总计
中文褒贬义词典	5567	4468	10035
NTUSD 中文情感词典	2810	8276	11080
本文情感词典	7407	10982	18389

在这其中设定,对于给定的长度为 l 的输入句子 s ,将句子中每个词语 W_t , $1 \leq t \leq l$, 映射到一个 k 维的向量空间, k 代表本模型考虑到的词语本身具有的属性的特征个数,每一维用 0, 1 来表示, 0 表示词语不具有此特征, 1 表示词语具有此特征,令 X 为情感词向量,则可表示为:

$$X = X_{cbow} + X_{sem} \quad (2)$$

2.2 模型结构

由于待解决的文本类型为电商评论,这类文本中存在大量转折型句式。而传统的深度学习模型不能够解决时序问题,本文设计了可获取时序信息的 GRU 深度学习模型,来进行情感极性分类。整体的模型架构如图 1 所示,模型包括四层,分别为情感向量输入层,GRU 层,语义表达层,Softmax 情感分类层。

第一层为输入层:从预训练好的词向量中查找待分类文本对应的词向量,将其拼接起来作为输入矩阵。

第二层为 GRU 层:GRU 是由 Cho 在 2014^[10]年提出的,它能够使每个单位适应性地捕捉不同时间尺度的依赖关系能够有效的处理时序信息。GRU 具有调节单元内部信息流的门控单元。其中有两个门单元分别为更新门和重置门。更新门用于控制前一时刻的状态信息被带入到当前状态中的程度,更新门的值越大说明前一时刻的状态信息带

入越多。

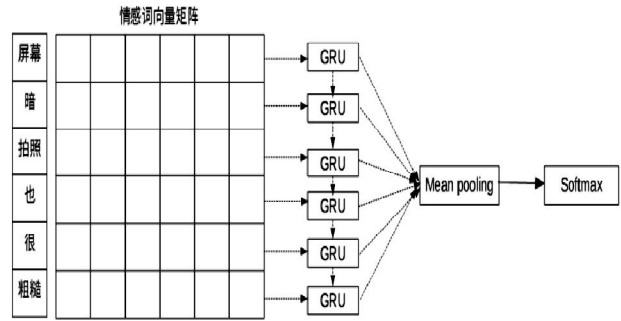


图 1 情感分析模型图

对于第 j 个 GRU 神经元,在 t 时刻更新门的计算公式为:

$$z_t^j = \sigma(W_z x_t + U_z h_{t-1}^j) \quad (3)$$

重置门用于控制忽略前一时刻的状态信息的程度,重置门的值越小说明忽略得越多。在 t 时刻重置门的计算公式为:

$$r_t^j = \sigma(W_r x_t + U_r h_{t-1}^j) \quad (4)$$

则时间 t 处 GRU 的输出 h_t^j 是前一时刻输出 h_{t-1}^j 与候选输出 \tilde{h}_t^j 之间的线性内插:

$$h_t^j = (1 - z_t^j) h_{t-1}^j + z_t^j \tilde{h}_t^j \quad (5)$$

其中,候选输出 \tilde{h}_t^j 的计算方法公式为:

$$\tilde{h}_t^j = \tanh(W x_t + U(r_t \odot h_{t-1}^j)) \quad (6)$$

第三层为语义表达层:上一层输出的句子的向量表示作为 Mean pooling 层的输入向量,Mean pooling 层对所有时间单位上的 h_t 求平均,得到向量 h ,则 h 为该句子的语义表达。

最后一层为 Softmax 情感分类层:由语义表达层的输出向量 h 作为 Softmax 层的输入,求得情感分类的预测概率,同时根据训练数据的实际分类标签,对 GRU 的参数采用反向传播算法进行梯度更新。

$$P(y | h, W, b) = \text{softmax}(W \cdot h + b) \quad (7)$$

其中 $y \in \{+1, -1\}$,代表情感积极或消极; W 代表参数; b 为偏执项。

3 实验结果及分析

3.1 数据集

实验中的数据集是从淘宝网上抓取的有关手机、书籍、热水器等商品评论语料,经过去重处理,选取了约 21101 条评论作为语料。根据其评论星级及人工校对,将其分为积极,消极两种类别。最终积极评论为 10627 条,消极评论为 10429 条。因为数据集的数量较小,我们随机按照 80%, 20% 的比例将数据分为训练集和测试集,并进行多次重复实验,实验结果取平均值。

表 2 评论数据样例

积极评论	消极评论
很好的手机,从3米高的地方摔下去都没坏,质量非常好	电池充完了电连手机都打不开.简直烂的要命.真是金玉其外,败絮其中!连5号电池都不如
质量不错,是正品,安装师傅也很好,才要了83元材料费	屏幕较差,拍照也很粗糙
东西非常不错,安装师傅很负责人,装的也很漂亮,精致,谢谢安装师傅	手机质量太差了,黑心店家,赚黑心钱,以后再也不用买了

3.2 数据处理

实验需对这 21101 条数据进行预处理,首先进行繁简体转化。与其他传统方式不同的是实验中没有去除停用词,因为有些停用词可能含有某些情感信息,如“不”,同样的,本文对一些特殊符号如“!、?”也进行了保留。预处理结束后,使用 Python 中 jieba 分词将每条评论文本转化为以空格分隔的词语序列,来训练词向量。其中,词向量的维度设为 50,迭代次数为 100 次。

3.3 实验结果与讨论

本文在传统词向量基础上加入了情感信息进行了情感分析,为了验证方法的准确性,将其与以下模型进行对比:

(1)Sword2vec+CNN: 使用文中训练出的情感词向量作为特征矩阵输入卷积神经网络模型中,证明 GRU 模型对于情感分析这一任务的优势。

(2)Sword2vec+LSTM: 使用文中训练出的情感词向量训练出的词向量作为特征矩阵输入 LSTM 中,证明 GRU 模型对于情感分析这一任务的优势。

(3) word2vec+GRU: 使用 word2vec 中 CBOW 训练出的词向量作为特征矩阵输入 GRU 中,证明融合了情感信息的词向量在情感分析这一任务上的优势。

(4) word2vec+SVM: 在相同的数据集上,使用传统的机器学习的方法来证明 GRU 在情感分析这一任务上的优势。使用 word2vec 中 CBOW 训练词向量,取词向量的均值作为 SVM 的特征,通过构建一个超平面,对训练样本进行正确分类。

采用准确率和召回率来评估方法的性能,表 3 中列出了不同方法进行情感极性分类这一任务的结果。

表 3 情感分析结果

方法	准确率	召回率
Sword2vec+CNN	0.8706	0.7564
Sword2vec+LSTM	0.9024	0.8856
word2vec+GRU	0.8661	0.8365
word2vec+SVM	0.8096	0.8309
Sword2vec+GRU (本文方法)	0.9150	0.8979

实验结果表明,情感分析这一任务上使用深度学习模型的方法在普遍优于 SVM 的方法,将 3、5 方法对比发现,在保证 GRU 网络模型不变的情况下,融合了情感信息的词向量能够有效提升模型的准确率。而比较 CNN,LSTM,GRU 这几种深度学习模型后,发现 CNN 准确率明显低于 LSTM,GRU 的准确率,这是由于 CNN 在训练的过程中摒弃了上下文间的关联,不能精确的分析如转折句式这类数据。LSTM,GRU 模型的准确率不相上下。

表 4 不同模型完成一次迭代的训练时间

模型	时间
Sword2vec+CNN	29
Sword2vec+LSTM	114
Sword2vec+GRU (本文方法)	80

为了分析不同模型在相同条件下的时间性能,本文在相同的硬件配置和网络框架下完成所有的神经网络实验。同

时,所有实验使用相同的词向量和数据集。表 4 给出了不同网络模型在淘宝评论文本上完成一次迭代的训练时间对比结果。

虽然 LSTM,GRU 的准确率不相上下,但是从表中 4 可以看出,GRU 的训练速度要明显快于 LSTM,因此若拥有大量训练数据,使用 GRU 可更有效的降低模型训练时间且得到更精确的分类结果。

4 结语

本文在词向量学习过程中融合词语的情感信息,解决了传统的基于深度神经网络训练出的词向量只含有语义信息而忽略了情感极性表达,不能准确的进行情感分析任务这一问题。并且将 GRU 这一深度神经网络引入到情感分析这一任务,设计了多组对比实验。实验结果证明,GRU 这一深度学习模型在情感分析这一任务上更具优势。

参考文献:

- [1] Whitelaw C,Garg N,Argamon S.Using Appraisal Groups for Sentiment Analysis. In: Proceedings of the ACM SIGIR Conference on Information and Knowledge Management (CIKM),2005,625-631.
- [2] Kim Y.Convolutional neural network for sentence classification[C].EMNLP.2014 1746-1751.
- [3] Socher R, Manning C D, Ng A Y. Learning Continuous Phrase Representations and Syntactic Parsing with Recursive Neural Networks[J]. 2010:1-9.
- [4] Tai K S, Socher R, Manning C D. Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks[J]. Computer Science, 2015.
- [5] LAI S, XU L, LIU K.Recurrent convolutional neural networks for text classification[C].Twenty-Ninth AAAI Conference on Artificial Intelligence.2015.
- [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] Sun X, Gao F, Li C, et al. Chinese microblog sentiment classification based on convolution neural network with content extension method[C].International Conference on Affective Computing and Intelligent Interaction. IEEE Computer Society, 2015:408-414.
- [8] 何炎祥, 孙松涛, 牛菲菲,等. 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(4): 773-790.
- [9] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space[J]. Computer Science, 2013.
- [10] Cho K, Merrienboer B V, Bahdanau D, et al. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches[J]. Computer Science, 2014.

作者简介:王宏生(1964-),男,硕士,副教授,主要研究方向:智能搜索引擎,人工智能;金相宇(1992-),女,硕士,主要研究方向:智能信息处理。